UNIVERSIDADE DE TRÁS-OS-MONTES E ALTO DOURO

IDENTIFICAÇÃO DE PATOLOGIAS PULMONARES EM EXAMES DE TOMOGRAFIA COMPUTORIZADA

TESE DE DOUTORAMENTO EM ENGENHARIA ELETROTÉCNICA E DE COMPUTADORES

VERÓNICA MARIA MARQUES DO CARREIRO SILVA VASCONCELOS

Orientador

Professor Doutor João Manuel Pereira Barroso

Coorientador

Professor Doutor José Silvestre Serra da Silva

Vila Real, 2017

UNIVERSIDADE DE TRÁS-OS-MONTES E ALTO DOURO

IDENTIFICAÇÃO DE PATOLOGIAS PULMONARES EM EXAMES DE TOMOGRAFIA COMPUTORIZADA

TESE DE DOUTORAMENTO EM ENGENHARIA ELETROTÉCNICA E DE COMPUTADORES

VERÓNICA MARIA MARQUES DO CARREIRO SILVA VASCONCELOS

Orientador

PROFESSOR DOUTOR JOÃO MANUEL PEREIRA BARROSO

Coorientador

PROFESSOR DOUTOR JOSÉ SILVESTRE SERRA DA SILVA

Vila Real, dezembro 2017

IDENTIFICAÇÃO DE PATOLOGIAS PULMONARES EM EXAMES DE TOMOGRAFIA COMPUTORIZADA

Tese de doutoramento apresentada por Verónica Maria Marques do Carreiro Silva Vasconcelos à Universidade de Trás-os-Montes e Alto Douro para o cumprimento dos requisitos necessários à obtenção do grau de Doutor em Engenharia Electrotécnica e de Computadores, sob a orientação do Professor Doutor João Manuel Pereira Barroso e coorientação do Professor Doutor José Silvestre Serra da Silva.

COMPOSIÇÃO DO JÚRI

PRESIDENTE:

Doutor José Boaventura Ribeiro da Cunha

Professor Associado com Agregação da Universidade de Trás-os-Montes e Alto Douro

VOGAIS:

Doutor Paulo Jorge Santos Gonçalves Ferreira

Professor Catedrático da Universidade de Aveiro

Doutor Manuel José Cabral dos Santos Reis

Professor Associado com Agregação da Universidade de Trás-os-Montes e Alto Douro

Doutor João Manuel Pereira Barroso

Professor Associado com Agregação da Universidade de Trás-os-Montes e Alto Douro

Doutor António Manuel Jesus Pereira

Professor Coordenador do Instituto Politécnico de Leiria

Doutor João Miguel Fernandes Rodrigues

Professor Adjunto do Instituto Superior de Engenharia da Universidade do Algarve

RESUMO

A medicina moderna é cada vez mais baseada na evidência da imagem, sendo o recurso a técnicas complementares de diagnóstico uma prática comum nos centros hospitalares das sociedades modernas. O elevado volume de imagens gerado pelos novos equipamentos impulsionou o desenvolvimento de sistemas de Diagnóstico Assistido por Computador (DAC). Estes sistemas são, cada vez mais, uma ferramenta essencial, auxiliando os imagiologistas na deteção, quantificação, caracterização e acompanhamento de patologias.

A Tomografia Computorizada de Alta Resolução (TCAR) é uma técnica imagiológica de referência no estudo de patologias pulmonares, nomeadamente da Doença Intersticial Pulmonar (DIP). Em imagens de TCAR a presença da DIP caracteriza-se por padrões imagiológicos específicos cujo reconhecimento é determinante para a elaboração de um diagnóstico correto. Em colaboração com o Centro Hospitalar e Universitário de Coimbra foram reunidos exames de pacientes a partir dos quais foram construídas duas bases de dados de padrões do tecido pulmonar associados ao enfisema, à DIP e ao pulmão saudável, tendo sido desenvolvida para o efeito a aplicação *Lung Study Tool*.

Foram realizados testes comparativos para a diferenciação de regiões com padrão normal de regiões com enfisema, recorrendo a diversos métodos de análise de textura e considerando várias hipóteses relativamente às regiões de análise e ao algoritmo de aprendizagem automático considerado, o *Support Vector Machines*. Foi proposto e validado um sistema de DAC que permite a classificação de padrões do tecido pulmonar com base na análise da sua textura. Dada a heterogeneidade dos padrões pulmonares em estudo, verificou-se que a integração no sistema da análise de textura multiescala, recorrendo à lacunaridade diferencial, potenciou uma melhoria global no desempenho do sistema e aumentou a sua estabilidade. O sistema permitiu a diferenciação entre o padrão normal e padrões associados à DIP com uma exatidão de 96.4±1.3% e a categorização de todos os padrões pulmonares com uma exatidão global de 95.6±1.2%.

Palavras-Chave: Doença Intersticial Pulmonar, Tomografia Computorizada de Alta Resolução, Análise de Textura, Lacunaridade, *Support Vector Machines*, Diagnóstico Assistido por Computador.

ABSTRACT

Modern medicine is increasingly based on the evidence of the image, being the use of diagnostic techniques a common practice in hospitals of modern societies. The high number of images generated by new equipment boosted the development of Computer-Aided Diagnosis (CAD) systems that are an essential tool, supporting radiologists in the detection, quantification, characterization, and follow-up of diseases.

High-resolution computed tomography (HRCT) is a reference imaging technique for the study of lung diseases, namely Interstitial Lung Disease (ILD). In HRCT images, the presence of ILD is characterized by specific lung tissue patterns whose recognition is decisive to a correct diagnosis. In collaboration with Coimbra Hospital and Universitary Center, images with lung tissue patterns associated with emphysema, ILD, and normal were gathered, and organized in two databases, having been developed the application Lung Study Tool for this purpose.

Comparative tests were performed with the objective to differentiate lung regions with emphysema pattern from healthy pattern, using different methods of texture analysis and considering several hypotheses regarding the regions of analysis and the machine learning algorithm considered, the Support Vector Machines. Was proposed and validated a CAD system that allows the classification of lung tissue patterns based on texture analysis. Due to the heterogeneity of the lung patterns in study, it was verified that the integration of a multiscale texture analysis, using differential lacunarity, allowed an overall improvement in system performance and increased its stability. The system allowed the differentiation between normal and ILD patterns with an accuracy of 96.4 ± 1.3% and the categorization of all the lung patterns with an overall accuracy of 95.6 ± 1.2%.

Key Words: Interstitial Lung Disease, High-Resolution Computed Tomography, Texture Analysis, Lacunarity, Support Vector Machines, Computer-Aided Diagnosis.

AGRADECIMENTOS

Este trabalho não seria possível sem o contributo de diversas pessoas e instituições que não posso deixar de destacar. Assim, quero agradecer:

Em primeiro lugar ao Professor João Manuel Pereira Barroso, meu orientador, pelo acompanhamento do trabalho, pela disponibilidade, amizade e generosidade reveladas ao longo destes anos de trabalho, assim como pelas críticas, correções e sugestões efetuadas. Ao meu coorientador Professor José Silvestre Serra da Silva pela oportunidade de trabalhar na área da imagem médica e pelo apoio na revisão de todos os documentos.

Ao Dr. Miguel Seco e à Dr.ª Luísa Andrade pela disponibilidade demonstrada na análise e marcação de exames de Tomografia Computorizada e à Dr.ª Luísa Teixeira e ao Professor Doutor Casaleiro Alves por tornarem possível o acesso ao Departamento de Imagiologia do Centro Hospitalar e Universitário de Coimbra.

Ao Instituto Superior de Engenharia de Coimbra em particular ao Departamento de Engenharia Eletrotécnica pelo excelente ambiente de trabalho que me proporcionam e pela motivação em terminar este trabalho. A todos os meus colegas e técnicos pelo apoio demonstrado.

Num plano pessoal não posso deixar de expressar a minha gratidão à minha amada família. Aos meus pais, à minha filha e à minha irmã pelo seu amor incondicional e por compreenderem as minhas inúmeras ausências. Ao meu marido por caminhar sempre a meu lado e pelo seu insubstituível apoio.

ÍNDICE

COMPOSIÇÃO DO JÚRI vii

RESUMO ix

ABSTRACT xi

AGRADECIMENTOS xiii

ÍNDICE xv

LISTA DE FIGURAS xix

LISTA DE TABELAS xxi

LISTA DE ABREVIATURAS E SIGLAS xviii

CAPÍTULO 1 - INTRODUÇÃO 1

1.1 Enquadramento 1

1.2 Motivações e Objetivos 2

1.3 Principais Contribuições 4

1.4 Estrutura do Documento 6

CAPÍTULO 2 - REVISÃO BIBLIOGRÁFICA 9

2.1 Breve Resenha Histórica 9

2.2 Estado da Arte 15

CAPÍTULO 3 - PATOLOGIAS PULMONARES EM TC DE ALTA RESOLUÇÃO ... 37

3.1 TC de Alta Resolução do Tórax 37

3.2 Estrutura Pulmonar ao Nível do Lóbulo Pulmonar Secundário 44

3.3 Enfisema Pulmonar 47

3.4 Doença Intersticial Pulmonar 49

3.5 Considerações Finais 52

CAPÍTULO 4 - EXTRAÇÃO E SELEÇÃO DE CARACTERÍSTICAS 53

4.1 Estatística de Níveis de Cinzento de Primeira Ordem 53

4.2 Método de Dependência Espacial de Níveis de Cinzento 55

4.3 Método de Comprimento de Primitivas de Níveis de Cinzento 59

4.4 Método de Diferenças de Níveis de Cinzento 63

4.5 Lacunaridade Diferencial 65
4.6 Seleção de Características 72
4.7 Considerações Finais 76

CAPÍTULO 5 - CLASSIFICAÇÃO DE PADRÕES 79

5.1 Metodologia de Aprendizagem 79
5.2 Teoria da Aprendizagem Estatística 81
5.3 Algoritmo SVM 85
5.3.1 SVM para Dados Linearmente Separáveis 85
5.3.2 SVM para Dados Não Linearmente Separáveis 89
5.3.3 SVM Não Lineares 91
5.3.4 Exemplos de Fronteiras de Decisão 94
5.3.5 SVM Multiclasse 95
5.4 Avaliação de Desempenho 96
5.4.1 Técnicas de Validação 96
5.4.2 Métricas 100
5.5 Considerações Finais 102

CAPÍTULO 6 - SISTEMA DE DAC PROPOSTO 105

6.1 Aplicação *Lung Study Tool* 105
6.1.1 Modo de Funcionamento do Separador "SaveROI" 107
6.1.2 Modo de Funcionamento do Separador "OpenROI" 109
6.2 Caracterização das Bases de Dados 110
6.2.1 Base de Dados I - Enfisema *versus* Normal 113
6.2.2 Base de Dados II - Padrões Intersticiais 114
6.3 Testes Experimentais 114
6.3.1 Características das ROIs e Vetores de Características 115
6.3.2 Parâmetros Ótimos do Modelo de Classificador 120
6.4 Categorização de Padrões Intersticiais em Imagens TC de Alta Resolução 124
6.4.1 Seleção de FH-ROIs 124
6.4.2 Processamento das ROIs 124
6.4.3 Extração de Características 126
6.4.4 Conjunto de Treino e Teste 127
6.4.5 Treino do Classificador e Seleção de Características 128

6.4.6 Resultados 129
6.4.6.1 Normal *versus* Patológico 130
6.4.6.2 Categorização de Padrões Intersticiais 131
6.5 Considerações Finais 137
CAPÍTULO 7 - CONCLUSÕES E TRABALHO FUTURO 139
7.1 Conclusões 139
7.2 Trabalho Futuro 142
REFERÊNCIAS BIBLIOGRÁFICAS 145
ANEXO A - APLICAÇÃO DA LACUNARIDADE NA ANÁLISE DE IMAGENS MÉDICAS 155

LISTA DE FIGURAS

Figura 2.1: Sistema ImageChecker™ CT. 11

Figura 2.2: Sistema de DAC genérico para deteção, categorização e quantificação de patologias. 12

Figura 2.3: Modelo genérico de um sistema de DAC CBIR. 13

Figura 2.4: Fluxograma genérico de um sistema de DAC para DIP. 14

Figura 3.1: Aquisição helicoidal volumétrica. a) Com uma linha de detetores. b) Com várias linhas de detetores. 39

Figura 3.2: Escala de Hounsfield. 42

Figura 3.3: Imagem de TC do tórax. a) Como se encontra guardada no ficheiro DICOM. b) Após a aplicação da janela pulmonar. 43

Figura 3.4: Estrutura pulmonar. a) Representação anatómica da periferia da árvore brônquica. b) Anatomia de um lóbulo pulmonar secundário sem patologia. 44

Figura 3.5: Lóbulo pulmonar secundário e identificação dos tipos de tecidos intersticiais e sua localização 45

Figura 3.6: Regiões do parênquima pulmonar extraídas de imagens de TCAR e respetivos histogramas. a) e c) região normal. b) e d) região com enfisema centrolobular. .. 47

Figura 3.7: Imagens de TCAR do tórax de pacientes com enfisema. a) Enfisema centrolobular. b) Enfisema paraseptal. c) Enfisema panlobular. d) Enfisema bolhoso. 49

Figura 3.8: Padrões do tecido pulmonar associados à DIP em imagens TCAR do tórax. Na coluna da esquerda uma vista de pormenor do padrão e à direita a respetiva imagem. a) Favo de mel; b) Vidro despolido; c) Normal. 51

Figura 4.1: Exemplo do cálculo do histograma normalizado da imagem $I(x,y)$. 54

Figura 4.2: Matrizes $\Omega(1,0^{o})$ e $\Omega(1,45^{o})$, correspondentes à imagem $I(x,y)$. 57

Figura 4.3: Matrizes de primitivas *run length* para as direções 0^{o} e 90^{o}, correspondentes à imagem $I(x,y)$..... 59

Figura 4.4: Exemplo do cálculo dos histogramas $HD(1,0^{o})$ e $HD(1,45^{o})$ para a imagem $I(x,y)$. 64

Figura 4.5: Algoritmo *differential box-counting* ($w = 9$, $r = 3$). 70

Figura 4.6: Cálculo da lacunaridade diferencial. a) Janela de 4×4 pixels. De b) a e) As quatro posições possíveis da caixa deslizante. 71

Figura 4.7: Métodos de filtragem em classificação. 74

Figura 4.8: Métodos *wrapper* em classificação. 75

Figura 5.1: Abordagem supervisionada em classificação de padrões. 80

Figura 5.2: Conjunto de treino e três fronteiras de decisão diferentes. 82

Figura 5.3: Ilustração da minimização do erro estrutural. 83

Figura 5.4: Conceito de margem. À esquerda um hiperplano com uma pequena margem de separação entre classes e à direita um hiperplano com uma margem maior. 84

Figura 5.5: Possíveis soluções para um problema de classificação de dados linearmente separáveis. 85

Figura 5.6: Hiperplano separador ótimo 86

Figura 5.7: Hiperplano separador ótimo para dados não linearmente separáveis. 89

Figura 5.8: a) Padrões não separáveis linearmente em $\Re^2$. b) Mapeamento dos padrões para $\Re^3$, onde é possível a separação linear das classes. 91

Figura 5.9: Exemplos de fronteiras de decisão. SVM lineares a) e b) e não lineares c) e d). 94

Figura 5.10: A técnica *holdout* com $p = 0.7$. 97

Figura 5.11: A técnica de validação cruzada *k-fold* ($k = 4$) 98

Figura 5.12: Validação cruzada LOPO para $k = N_p = 4$ pacientes. 100

Figura 6.1: Ecrã inicial da aplicação *Lung Study Tool* 106

Figura 6.2: Ecrã da aplicação *Lung Study Tool* em modo de funcionamento "SaveROI". 107

Figura 6.3: a) Opções disponíveis para caracterizar o padrão/patologia pulmonar. b) Nível anatómico pulmonar da área selecionada. 108

Figura 6.4: Ecrã da aplicação *Lung Study Tool* em modo "OpenROI" 110

Figura 6.5: a) Visão geral de uma sala do Departamento de Imagiologia do CHUC. b) *Scanner* da General Electric Healthcare. 111

Figura 6.6: Exemplo de uma região selecionada pelo imagiologista (contorno a azul) e a grelha que permite obter ROIs de 40×40 pixels. 112

Figura 6.7: Gráfico de contornos correspondente à exatidão média (%) considerando o conjunto de características GLDM e o *kernel* RBF 122

Figura 6.8: Principais etapas do sistema de DAC proposto. 125

Figura 6.9: Curvas da lacunaridade média normalizadas 127

LISTA DE TABELAS

Tabela 2.1: Resumo das características dos trabalhos descritos. 31

Tabela 5.1: Matriz de confusão para um cenário de *c* classes..... 101

Tabela 6.1: Resumo do conteúdo da Base de Dados II – Padrões Intersticiais..... 114

Tabela 6.2: Valores médios da exatidão para as várias combinações de ROIs e vetores de características extraídas através do SGLDM..... 117

Tabela 6.3: Valores médios da exatidão para as várias combinações de ROIs e vetores de características extraídas através do GLRLM..... 118

Tabela 6.4: Valores médios da exatidão para as várias combinações de ROIs e vetores de características extraídas através do GLDM..... 119

Tabela 6.5: Resultados obtidos para os três vetores de características SGLDM, GLRLM e GLDM e os *kernels* linear, RBF e polinomial..... 123

Tabela 6.6: Resultados obtidos para o cenário normal *versus* patológico, considerando os dois conjuntos de características *Set 1* e *Set 2*. Valores apresentados em percentagem. 130

Tabela 6.7: Exemplos de ROIs mal classificadas entre as duas classes Normal e Patológica. 131

Tabela 6.8: Resultados obtidos para o cenário multiclasse com o conjunto de características *Set 1*. Valores apresentados em percentagem..... 131

Tabela 6.9: Resultados obtidos para o cenário multiclasse com o conjunto de características *Set 2*. Valores apresentados em percentagem..... 132

Tabela 6.10: Exemplos de ROIs mal classificadas entre as classes Vidro Despolido e Normal. 134

Tabela 6.11: Exemplos de ROIs mal classificadas entre as classes Favo de Mel e Vidro Despolido..... 135

Tabela 6.12: Exemplos de ROIs mal classificadas entre as classes Normal e Enfisema..... 136

LISTA DE ABREVIATURAS E SIGLAS

ALARA - *As Low As Reasonably Achievable*

AMFM - *Adaptive Multiple Feature Method*

AUC - *Area Under the Curve*

BEEMD - *Bidimensional Ensemble Empirical Mode Decomposition*

BI-RADS - *Breast Imaging - Reporting and Data System*

CBIR - *Content Based Image Retrieval*

CFS - *Correlation Feature Selection*

CHUC - *Centro Hospitalar e Universitário de Coimbra*

CNN - *Convolutional Neural Network*

DAC - Diagnóstico Assistido por Computador

DCT - *Discrete Cosine Transform*

DICOM - *Digital Imaging and COmmunications in Medicine*

DIP - Doença Intersticial Pulmonar

DL - *Deep Learning*

DMO - Densitometria Mineral Óssea

DPOC - Doença Pulmonar Obstrutiva Crónica

EMD - *Earth Mover's Distance*

EPC - Ênfase em Primitivas Curtas

EPCNCA - Ênfase em Primitivas Curtas com Níveis de Cinzento Altos

EPCNCB - Ênfase em Primitivas Curtas com Níveis de Cinzento Baixos

EPL - Ênfase em Primitivas Longas

EPLNCA - Ênfase em Primitivas Longas com Níveis de Cinzento Altos

EPLNCB - Ênfase em Primitivas Longas com Níveis de Cinzento Baixos

EPNCA - Ênfase em Primitivas com Níveis de Cinzento Altos

EPNCB - Ênfase em Primitivas com Níveis de Cinzento Baixos

ESTI - *European Society of Thoracic Imaging*

FDA - *Food and Drug Administration*

FDG - *FuoroDeoxyGlucose*

FFP - Fração de Falso Positivos

FH-ROI - *Free Hand* - ROI

FLAIR - *Fluid-Attenuated Inversion Recovery*

FOS - *First Order Statistics*

FVP - Fração de Verdadeiros Positivos

GLDM - *Gray Level Difference Method*

GLRLM - *Gray Level Run Length Method*

GUI - *Graphical User Interface*

HRCT - *High-Resolution Computed Tomography*

IMF - *Intrinsic Mode Functions*

KKT - *Karush - Kuhn - Tucker*

KPCA - *Kernel Principal Component Analysis*

***k*-NN** - *k-Nearest Neighbors*

LBP - *Local Binary Patterns*

LDA - *Linear Discriminant Analysis*

LOPO - *Leave-One-Patient-Out*

MID - Momento Inverso da Diferença

MIP - *Maximum Intensity Projection*

mIP - *Minimum Intensity Projection*

MKL - *Multiple Kernel Learning*

NIG - *Nakagami-generalized Inverse Gaussian*

NUCP - Não Uniformidade do Comprimento das Primitivas

NUNC - Não Uniformidade dos Níveis de Cinzento

OMP - *Orthogonal Matching Pursuit*

PACS - *Picture Archiving and Communication Systems*
PASA - *Patch-Adaptive Sparse Approximation*
PP - Percentagem de Primitivas
PET - *Positron Emission Tomography*
Pixel - *Picture Element*
PNN - *Probabilistic Neural Network*
PTA(*l*, *r*) - *Plus l-Take Away r*

QDA - *Quadratic Discriminant Analysis*
QWF - *Quincunx Wavelet Frames*

RBF - *Radial Basis Function*
RBM - *Restricted Boltzmann Machine*
RF - *Random Forest*
RIG-LBP - *Rotation-Invariant Gabor - Local Binary Pattern*
RLM - Regressão Logística Multinomial
RM - Ressonância Magnética
RN - Redes Neuronais
RNP - Redes Neuronais Probabilísticas
ROC - *Receiver Operating Characteristic*
ROI – *Region of Interest*

SBS - *Sequential Backward Selection*
SDA - *Stepwise Discriminant Analysis*
SFBS - *Sequential Floating Backward Selection*
SFFS - *Sequential Floating Forward Selection*
SFS - *Sequential Forward Selection*
SGLDM - *Spacial Gray Level Dependence Method*
SMA - Segundo Momento Angular
SR - *Sparse Representation*
SVM - *Support Vector Machines*

TC - Tomografia Computorizada

TCAR - Tomografia Computorizada de Alta Resolução

UH - Unidades Hounsfield

VC - Vapnik - Chervonenkis

VO - *Volume Overlap*

VOI - *Volume of Interest*

Voxels - *Volume Elements*

VS - Vetores de Suporte

Capítulo 1
INTRODUÇÃO

Neste capítulo introdutório é efetuado o enquadramento do trabalho realizado, apresentadas as motivações que levaram ao seu desenvolvimento e objetivos principais que se pretenderam atingir. São enunciadas as contribuições do projeto e, por fim, realiza-se um resumo dos capítulos que constituem o documento.

1.1 Enquadramento

A imagiologia médica desempenha um papel fulcral na medicina moderna. A aquisição de imagens/vídeos é um procedimento rápido que possui um excelente binómio custo-efeito. As diferentes modalidades imagiológicas e as novas técnicas de processamento e análise de imagem digital tornaram-se num instrumento imprescindível na visualização e inspeção de estruturas anatómicas, assim como, na deteção, quantificação, diagnóstico e acompanhamento de patologias. É também cada vez mais frequente o recurso à aquisição de imagens em tempo real como apoio à realização de biópsias, cirurgias e outros procedimentos médicos que exijam precisão e tomadas de decisão no momento.

Ao longo das últimas décadas, a evolução tecnológica dos equipamentos, aliada a computadores com cada vez maior capacidade de processamento, deram origem à aquisição de uma elevada quantidade de imagens, originárias de diferentes técnicas de apoio ao diagnóstico. A título ilustrativo, podemos referir que em 2009, no Departamento de Radiologia do Hospital Universitário de Genebra se verificou um armazenamento médio diário nos seus arquivos de imagem PACS (*Picture Archiving and Communication Systems*) de 114 000 imagens. Tendo em conta a aumento exponencial que se tem verificado nos últimos anos, prevê-se que este número atinja os 3 milhões de imagens em 2020 (Depeursinge, 2010). Estas imagens

necessitam de ser analisadas, interpretadas, anotadas e armazenadas para posteriores consultas. Consequentemente, o volume de trabalho diário dos imagiologistas aumentou consideravelmente nos últimos anos, tendo sido o grande impulso para o desenvolvimento de sistemas automáticos que minimizem a intervenção humana e diminuam o tempo de realização do diagnóstico e posterior acompanhamento dos pacientes, sem, no entanto, diminuir a qualidade do diagnóstico.

Uma das técnicas imagiológicas mais comuns é a Tomografia Computorizada (TC), uma evolução lógica e natural da convencional imagem radiográfica, amplamente utilizada desde o início do século XX. A TC de Alta Resolução (TCAR) permite obter imagens com elevado detalhe morfológico do parênquima pulmonar sendo consensualmente aceite como a técnica imagiológica de referência no apoio ao diagnóstico de doenças difusas do parênquima pulmonar, como é o caso da Doença Intersticial Pulmonar (DIP) (Verschakelen, De Wever, & Matamoros, 2008; Pisco, 2009; Anthimopoulos, Christodoulidis, Christe, & Mougiakakou, 2015).

1.2 Motivações e Objetivos

As grandes unidades hospitalares e os centros de diagnóstico por imagem são confrontados diariamente com um elevado número de imagens/vídeos resultantes de diferentes meios complementares de diagnóstico. Esta realidade tornou evidente a necessidade do desenvolvimento de sistemas automatizados que ajudem o imagiologista na sua rotina diária.

Dada a capacidade dos atuais equipamentos de TC helicoidais multicorte, que permitem a aquisição de várias secções em simultâneo e em poucos segundos, a quantidade de imagens adquiridas por cada exame realizado aumentou substancialmente. Estas imagens devem ser cuidadosamente examinadas por parte dos imagiologistas para a elaboração precisa do diagnóstico ou para avaliar a evolução de patologias já previamente diagnosticadas. O processo de análise e interpretação de padrões imagiológicos é monótono, cansativo, muito consumidor de tempo e requer bastante experiência, sendo realizado muitas vezes em ambientes bastante stressantes. Para além destas razões, existe uma grande variabilidade na interpretação das mesmas regiões por observadores diferentes, assim como pelo mesmo observador em instantes temporais distintos. No caso de imagens de TCAR de pacientes com DIP a interpretação e

quantificação dos padrões imagiológicos presentes nem sempre é fácil, mesmo para imagiologistas experientes. Os padrões nem sempre são específicos, podendo surgir misturados e sofrer alterações ao longo da evolução da doença.

A realidade descrita nos parágrafos anteriores tornou evidente a necessidade do desenvolvimento de sistemas automatizados que auxiliem o imagiologista no seu trabalho diário. Estes sistemas, genericamente conhecidos como Diagnóstico Assistido por Computador (DAC), proporcionam ao imagiologista uma segunda opinião, objetiva e reprodutível. Podem também ser utilizados como uma primeira triagem, aumentando a produtividade e libertando o imagiologista para situações mais complexas. Os sistemas DAC podem ser também uma mais-valia em centros hospitalares onde não existe pessoal especializado ou que possui pouca experiência. Nestas circunstâncias, consideram-se os resultados procedentes de um sistema DAC como uma primeira triagem, no sentido em que o paciente seja encaminhado para um centro especializado caso os resultados assim o justifiquem.

Os aspetos atrás referidos constituíram as principais motivações para a realização do trabalho descrito neste documento. Por outro lado, o facto do desenvolvimento de sistemas de DAC ser uma área muita ativa de investigação, de grande multidisciplinariedade e que procura dar resposta a problemas atuais e concretos, constituíram motivações adicionais para a escolha do tema.

Os objetivos norteadores deste trabalho foram:

- A conceção e desenvolvimento de uma aplicação informática *user-friendly* que permita reunir dados relevantes extraídos de exames de TC com vista ao teste de algoritmos de processamento e análise de imagens e à compreensão dos dados recolhidos. Assim, a aplicação deve possibilitar a visualização de exames de TC, a seleção de regiões de interesse (ROIs - *Regions of Interest*), a sua anotação e armazenamento. A posterior visualização das ROIs selecionadas, enquadradas na imagem de origem deve também ser possível. Uma vez que as regiões selecionadas e anotadas podem constituir material de apoio pedagógico que auxilie os estudantes de medicina, e a comunidade científica em geral, no estudo de patologias e padrões pulmonares em imagens TCAR, a aplicação deve englobar funcionalidades que contemplem uma vertente pedagógica.

- A construção de bases de dados de regiões do parênquima pulmonar extraídas a partir de imagens de TCAR do tórax. Estas bases de dados devem ser representativas de padrões do tecido pulmonar de pacientes com DIP e de pacientes com pulmão saudável. A análise das imagens, a seleção de áreas de interesse e respetiva caracterização devem ser realizadas por imagiologistas com experiência em patologias pulmonares.

- O desenvolvimento de um sistema automático de auxílio ao diagnóstico, baseado em imagens de TCAR do tórax, que permita a diferenciação de regiões do parênquima pulmonar de pacientes com DIP de regiões de pacientes saudáveis. Numa fase posterior o sistema deve ser capaz de categorizar as regiões do parênquima pulmonar nos diferentes padrões imagiológicos usualmente associados a pacientes com suspeita ou diagnosticados com DIP. O problema que se coloca pode ser dividido em duas etapas principais:

    - A caracterização dos referidos padrões pulmonares através de um conjunto de medidas que expressem a sua textura;

    - O desenvolvimento de estratégicas que permitam a construção de um modelo de classificador e a sua avaliação de desempenho.

1.3 Principais Contribuições

A realização do trabalho proposto deu origem a um conjunto de contribuições, das quais destacamos as seguintes:

- Desenvolvimento de uma aplicação informática com duas vertentes distintas:

    - A primeira vertente possui as funcionalidades base inerentes à visualização de sequências de imagens de acordo com a norma DICOM[1] (*Digital Imaging and COmmunications in Medicine*). A estas funcionalidades foram associadas um conjunto de opções que permitem guardar dados num formato apropriado à sua utilização em aplicações que recorram a técnicas de reconhecimento de padrões. De uma forma *user-friendly* é possível a seleção de regiões de uma imagem, a sua

[1] http://medical.nema.org/, consultado em julho de 2017.

caracterização e o seu armazenamento, conjuntamente com os dados que caracterizam a região selecionada.

- A segunda vertente da aplicação é de carácter pedagógico, sendo dirigida ao meio científico e académico. A informação recolhida pode ser visualizada no seu contexto original, constituindo material de estudo para todas as pessoas que tenham como foco de interesse patologias pulmonares em imagens de TCAR.

- Construção de duas bases de dados, em colaboração com o Departamento de Imagiologia do Centro Hospitalar e Universitário de Coimbra (CHUC), que reúnem dados das regiões do parênquima pulmonar extraídas de exames de TCAR do tórax. Para este efeito os exames de pacientes com diagnóstico confirmado de enfisema, de DIP e sem patologia pulmonar foram analisados por imagiologistas do referido departamento. As regiões que possuíam padrões imagiológicos típicos foram selecionadas e devidamente anotadas. Os dados recolhidos foram organizados em duas bases de dados:

  - Base de dados I – Normal *versus* Enfisema que reúne dados de pacientes com enfisema pulmonar e de pacientes com pulmão normal. Uma vez que o enfisema é uma patologia bastante comum, este conjunto de dados foi o primeiro a ser recolhido, permitindo a realização de estudos preliminares;

  - Base de dados II – Padrões Intersticiais que integra padrões imagiológicos associados à DIP e ao pulmão normal.

- Realização de estudos comparativos da influência do tamanho das ROIs e dos níveis de quantização dessas regiões na diferenciação das regiões de pulmão normal em relação às regiões enfisematosas. A textura dos padrões imagiológicos foi descrita através de diferentes conjuntos de características extraídas recorrendo a métodos estatísticos de análise de textura. Foram também realizados estudos com o algoritmo de aprendizagem *Support Vector Machines* (SVM) considerando várias funções de *kernel* e sua importância na distinção entre regiões de padrão pulmonar normal e regiões com enfisema.

- Estudo da potencialidade da lacunaridade diferencial na categorização de padrões imagiológicos associados à DIP e associados ao pulmão normal. De acordo com as

pesquisas efetuadas, à data, a análise da lacunaridade nunca foi utilizada no estudo da textura do parênquima de pacientes com DIP.

- Desenvolvimento de um sistema de DAC inspirado no sistema de visão do imagiologista, capaz de categorizar padrões pulmonares associados à DIP e ao pulmão normal em imagens de TCAR do tórax baseado na informação da textura.

1.4 Estrutura do Documento

Este documento está organizado em sete capítulos, que serão apresentados de forma sucinta nos parágrafos seguintes.

No presente capítulo é efetuado o enquadramento ao tema da tese, expostas as principais motivações e objetivos para o desenvolvimento do trabalho. São ainda destacadas as principais contribuições. De seguida, é efetuado um resumo dos capítulos que integram este documento.

A área de investigação onde se enquadra o trabalho descrito nesta tese é muito ativa em termos de publicações científicas. No Capítulo 2 apresenta-se um resumo histórico dos primeiros sistemas de DAC, sendo também descritos os principais tipos destes sistemas. Posteriormente, são apresentados os trabalhos relevantes na área da categorização de padrões pulmonares associados à DIP e à Doença Pulmonar Obstrutiva Crónica (DPOC) em imagens de TC. De acordo com a informação disponibilizada pelos autores, foram realçados os objetivos do sistema descrito, identificados os padrões pulmonares em análise, os métodos de extração e o tipo de características utilizadas na descrição dos padrões, a metodologia e os algoritmos de aprendizagem automática empregues na criação do modelo do classificador, a metodologia de validação dos resultados e uma avaliação do desempenho do sistema.

No Capítulo 3 são descritas as duas abordagens genéricas disponíveis para a aquisição das imagens de TCAR: a axial e a volumétrica. É efetuado o enquadramento a nível anatómico da doença intersticial e das regiões pulmonares afetadas. São descritos os padrões intersticiais considerados neste trabalho e as suas principais características quando visualizados em imagens de TCAR.

No Capítulo 4 são apresentados os métodos de análise de textura utilizados na extração de características das regiões pulmonares de interesse. Dada a complexidade da textura do parênquima pulmonar são reunidas características através de métodos diferentes. O conjunto de características visa conjugar informação determinante para a categorização dos diferentes padrões. Na última secção é dada uma visão geral dos três grupos de métodos de seleção de características, tendo em consideração a interação das estratégias de pesquisa do subconjunto de características ótimo com o algoritmo de aprendizagem automático.

No Capítulo 5 são introduzidos os conceitos de aprendizagem estatística e descrito o algoritmo de aprendizagem automática utilizado para a categorização dos padrões imagiológicos em estudo – o SVM. São considerados os classificadores SVM para dados linearmente e não linearmente separáveis, considerando cenários binários e multiclasse. Por fim, são expostas as técnicas de validação que permitem o treino e teste do algoritmo de aprendizagem e métricas de avaliação de desempenho do mesmo.

No Capítulo 6 é descrita a aplicação *Lung Study Tool* e apresentadas as suas principais funcionalidades. Posteriormente, são caracterizadas as duas bases de dados construídas e que foram utilizadas ao longo do trabalho para a realização de testes e avaliação de desempenho do sistema de DAC proposto. A primeira base de dados construída engloba regiões pulmonares extraídas de exames de pacientes com enfisema e de pacientes com pulmão saudável. A segunda base de dados reúne regiões de pacientes com DIP e de pacientes com pulmão saudável. São descritos e apresentados os resultados de estudos comparativos realizados tendo em consideração vários métodos de análise de textura, várias hipóteses de ROIs e várias opções do algoritmo de aprendizagem SVM. Posteriormente, são descritas as várias etapas do sistema de DAC proposto que visa a categorização de padrões pulmonares, finalizando com a apresentação e discussão dos resultados obtidos.

Por fim, o Capítulo 7 apresenta as conclusões e tópicos para trabalho futuro.

Capítulo 2
REVISÃO BIBLIOGRÁFICA

Neste capítulo, após uma breve visão global dos sistemas de DAC, apresenta-se o estado da arte referente aos trabalhos de investigação na área do trabalho proposto. A pesquisa centrou-se em sistemas DAC baseados em imagens de TCAR que visam a deteção e classificação de padrões pulmonar associados à DIP e à DPOC. Genericamente, estes sistemas de DAC são constituídos por duas grandes etapas: a extração de características relevantes e a deteção e classificação dos padrões em análise, recorrendo a técnicas de inteligência artificial. No estado da arte foram realçados: o objetivo do sistema DAC, identificados os tipos de padrões pulmonares, os métodos de extração de características utilizados na descrição dos padrões, o método de classificação e a avaliação do desempenho do sistema.

2.1 Breve Resenha Histórica

A medicina baseada na "evidência" tem como consequência imediata o aumento exponencial do volume de imagens adquiridas diariamente nos centros de imagiologia e hospitais, resultantes da utilização das diferentes técnicas de apoio ao diagnóstico baseado na imagem. A quantidade e variedade de informação visual que nos dias de hoje os médicos têm ao seu dispor, o formato digital das imagens médicas e a uniformização no seu armazenamento e transmissão, nomeadamente através da norma DICOM, deram origem a uma área de investigação muito ativa, os sistemas de DAC. Esta área de investigação é muito abrangente pois integra vários domínios do conhecimento, como por exemplo: medicina, processamento de imagem, inteligência artificial, matemática, física e estatística. Um dos objetivos dos sistemas de apoio ao diagnóstico é efetuar uma "primeira leitura", libertando os imagiologistas de tarefas exaustivas e muito consumidoras de tempo, permitindo o enfoque das equipas de imagiologia

em regiões, imagens ou séries de imagens informativas do ponto de vista da deteção e diagnóstico da doença. Para além disso, estudos realizados mostram que o recurso a sistemas computorizados dedicados aumenta a precisão do diagnóstico médico, reduz a percentagem de falhas na deteção devido à fadiga e excesso de trabalho dos imagiologistas e diminui a variabilidade intra e interobservador (Uppaluri et al, 1999b; Fujita et al., 2008).

Os primeiros estudos baseados na análise computacional de imagens médicas tiveram início nos anos sessenta. Em 1963, Lodwick et al. estudaram a utilização de métodos automáticos no diagnóstico de tumores ósseos (Lodwick, Haun, Smith, Keller, & Robertson, 1963). Em 1964, Meyers e a sua equipa propuseram um sistema para a distinção automática de radiografias do tórax normais *versus* anormais através da razão cardiotorácica (Meyers et al., 1964, referenciado por Fujita et al., 2008). Nos finais dos anos sessenta e inícios dos anos setenta do século XX surgiram os primeiros sistemas para a análise automática de mamografias (Winsberg, Elkin, Macy, Bordaz, & Weymouth, 1967; Ackerman, & Gose, 1972, referenciados por Fujita et al., 2008). No entanto, as limitações da capacidade de processamento dos computadores e a dificuldade no processamento e análise das imagens colocaram limitações severas a estes primeiros sistemas.

Os primeiros sistemas de DAC para a deteção de patologias pulmonares em imagens TC datam dos anos oitenta e baseavam-se na variação quantitativa das alterações da densidade do histograma (Gilman, Laurens Jr, Somogyi, & Honig, 1983; Müller, Staples, Miller, & Abboud, 1988).

O ano de 1998 foi um marco na história dos sistemas de DAC uma vez que houve a transferência de tecnologia da investigação académica, realizada nos laboratórios da Universidade de Chicago, para a prática clínica. Nesse ano deu-se a aprovação do sistema ImageChecker^{TM} (R2 Technology, Inc., Sunnyvale, CA adquirida em 2006 pela Hologic, Inc.) pela *Food and Drug Administration* (FDA) para fins comerciais. O ImageChecker^{TM} analisa mamografias digitais para a deteção de regiões suspeitas que possam indiciar a presença de cancro da mama. Atualmente, os sistemas de DAC são usados em mais de 90% das mamografias realizadas nos Estados Unidos (Lin, Huang, & Cheng, 2016).

Figura 2.1: Sistema ImageChecker™ CT[2].

O primeiro sistema de DAC pulmonar comercial, designado por CAD RapidScreen™, a obter a aprovação da FDA data de 2001. Este sistema desenvolvido pela Deus Technologies, LLC baseia-se em radiografias torácicas para a deteção do cancro do pulmão em estágios inicias, associados a nódulos pulmonares isolados (9 mm a 30 mm de tamanho). Em 2004 surgiu o ImageChecker™ CT, o primeiro sistema de DAC aprovado pela FDA para imagens de TC. Na Figura 2.1 pode observar-se uma imagem do ecrã do sistema em funcionamento. Este sistema tinha como principal objetivo a deteção precoce de nódulos pulmonares sólidos (entre 4 mm e 30 mm de diâmetro) em imagens TC do tórax. O sistema DAC analisa todas as imagens de TC obtidas durante o exame, assinala nódulos candidatos e calcula o seu diâmetro, volume e densidade (Doi, 2005; Doi, 2007;Fujita et al., 2008).

Ao longo dos últimos anos outros sistemas DAC foram sendo aprovados pela FDA e introduzidos no mercado em diversas modalidades imagiológicas e aplicados a diferentes órgãos (Fujita et al., 2010; Lin et al., 2016).

Na literatura são propostos quatro cenários de sistemas de DAC baseados em imagens, que podem surgir associados:

- Deteção de patologias;

[2] http://www.cancernetwork.com/, consultado em julho de 2017.

- Quantificação de regiões com patologia;
- Categorização da patologia;
- Recuperação de imagens similares com diagnóstico conhecido (CBIR - *Content–Based Image Retrieval*).

Qualquer um dos cenários baseia-se no conteúdo visual da imagem para extrair informação. O DAC baseado na imagem é um conceito amplo que pode incluir a deteção, a classificação de várias patologias ou padrões patológicos, e a quantificação da doença que pode ser aplicada na análise volumétrica, progressão da doença ou resposta a tratamentos. Um esquema genérico de um sistema DAC encontra-se apresentado na Figura 2.2, adaptada de (Depeursinge, 2010).

Figura 2.2: Sistema de DAC genérico para deteção, categorização e quantificação de patologias.

O cenário CBIR refere-se a sistemas de DAC que consistem na recuperação de imagens baseada em conteúdos. Quando o imagiologista recebe o exame de um paciente, os sistemas CBIR exploram o conteúdo visual das imagens armazenadas nos PACS, organizam, classificam e disponibilizam ao imagiologista exames semelhantes, com diagnóstico confirmado, e que são relevantes para a elaboração do diagnóstico, de acordo com a proximidade com o exame original. Geralmente, esta proximidade é estabelecida a partir de um conjunto de características visuais que descrevem o conteúdo das instâncias (sejam elas vídeos, séries de imagens, ROIs, etc.). O imagiologista pode consultar os exames disponibilizados, os relatórios e anotações efetuadas anteriormente para casos semelhantes, aumentado a precisão e rapidez do seu diagnóstico, especialmente para imagiologistas pouco experientes em situações de diagnóstico difícil. Estes sistemas visam gerir os extensos repositórios de dados existentes nos centros hospitalares e auxiliar o médico na elaboração do diagnóstico. Para além destas aplicações, dado o elevado volume de imagens guardadas nos PACS, os sistemas CBIR podem também ser usados no domínio do ensino e pesquisa científica. Uma revisão da literatura no domínio dos sistemas CBIR pode ser consultada em (Müller, Michoux, Bandon, & Geissbuhler, 2004; Long, Antani, Deserno, & Thoma, 2009; Depeursinge, 2010; Zhou, Li, & Tian, 2017). A maioria dos

sistemas CBIR possuem uma arquitetura semelhante na forma como arquivam, pesquisam e indexam as instâncias, como integram métodos para extração de características visuais e para o cálculo de medidas de similitude. Usualmente, os sistemas CBIR possuem uma interface gráfica (GUI - G*raphical User Interface*) de fácil utilização. A arquitetura genérica destes sistemas pode ser representada através do esquema da Figura 2.3, adaptada de (Müller et al., 2004).

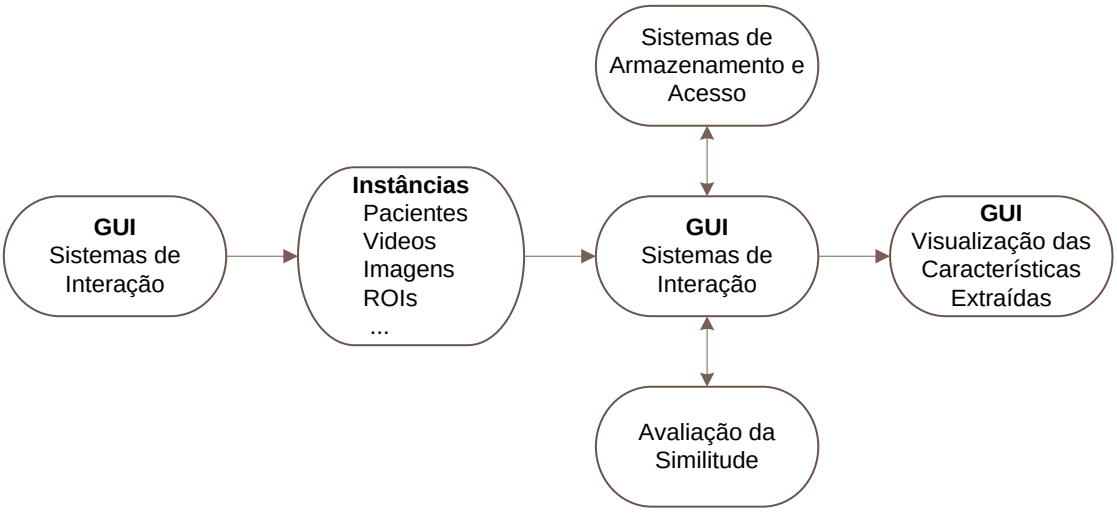


Figura 2.3: Modelo genérico de um sistema de DAC CBIR.

A conceção e desenvolvimento de novos sistemas de DAC são efetuados por equipas multidisciplinares que devem incluir imagiologistas experientes. Existem equipas de investigação a desenvolver sistemas de DAC nas mais variadas modalidades imagiológicas: radiografia digital, ultrassom, Ressonância Magnética (RM), TC e nas técnicas mais recentes de medicina nuclear. Os sistemas anatómicos abrangidos são também muito vastos: pulmonar, cardíaco, geniturinário, digestivo, neurológico, esquelético, mamário e vascular.

Os sistemas de DAC têm por objetivo auxiliar as equipas médicas na tomada de decisão ao longo do processo de acompanhamento do doente, com enfoque na elaboração do diagnóstico. Estes sistemas colocam ao dispor do médico uma segunda "opinião", baseada na análise quantitativa de imagens médicas, possuindo um enorme potencial, uma vez que permitem aumentar a precisão, reduzir o tempo de análise de cada imagem e aumentar a consistência do diagnóstico, diminuindo a variabilidade entre o mesmo observador e entre observadores diferentes. Enquanto a capacidade de análise de um imagiologista depende de fatores como a sua experiência naquele domínio específico, cansaço e muitos outros aspetos que se prendem com a natureza humana, um sistema de DAC é imune a estas perturbações.

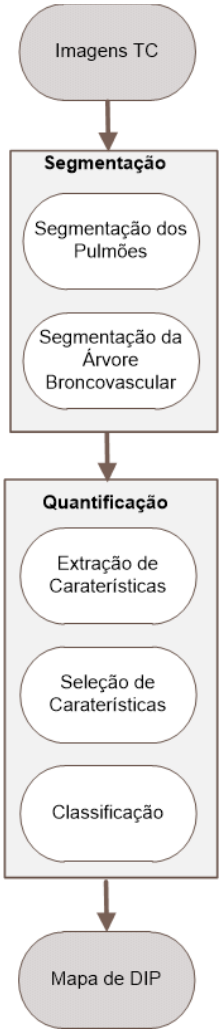


Figura 2.4: Fluxograma genérico de um sistema de DAC para DIP.

Os sistemas computorizados modernos possuem um conjunto de potencialidades que tornam o seu desempenho superior aos humanos em muitas aplicações, como ser 100% reproduzíveis. No entanto, em tarefas de reconhecimento de padrões, como por exemplo, o reconhecimento de texturas, objetos, vozes, caracteres escritos ou paisagens, o sistema de visão humano é extraordinariamente rápido e preciso. É um sistema complexo que resulta da interação de numerosos componentes desde os olhos, em especial da retina, ao córtex visual (Depeursinge, 2010). Uma aproximação comum implementada em sistemas de DAC baseados em imagens é inspirada pelo sistema de visão do imagiologista. Este, quando confrontado com imagens de um paciente, observa-as e extrai delas as características que considera relevantes, que podem tomar formas muito diferentes. Baseado na informação visual recebida, no seu conhecimento, na experiência adquirida ao longo dos seus estudos e prática clínica é capaz de, na maioria das vezes, proceder à identificação e classificação de padrões ou estruturas

anómalas. Em reconhecimento de padrões, estas duas etapas podem ser identificadas como: extração de características e aprendizagem supervisionada. No caso concreto do objetivo do sistema de DAC ser a classificação do tecido pulmonar em duas ou mais classes, a tarefa de aprendizagem envolvida é uma tarefa de classificação de padrões.

2.2 Estado da Arte

Na literatura especializada são vários os sistemas de DAC propostos para diagnóstico de DIP em imagens de TC do tórax. Alguns sistemas propõem a deteção e quantificação dos padrões associados à DIP, outros vão mais além propondo a classificação dos padrões. Na Figura 2.4 encontra-se representado um fluxograma genérico de um sistema de DAC para a DIP completamente automático (Anthimopoulos et al., 2015). Nele podem ser identificadas duas etapas: a segmentação pulmonar, que pode incluir a segmentação da árvore broncovascular, e a deteção e quantificação dos padrões associados DIP.

Em imagem médica o processo de segmentação consiste em separar órgãos, tecidos ou patologias para que possam ser analisados individualmente. A segmentação pulmonar é uma fase de pré-processamento que consiste na identificação do contorno do tecido pulmonar, na identificação dos dois pulmões e, por vezes, na localização dos lobos pulmonares. Em algumas situações, a segmentação da árvore broncovascular também é efetuada de forma a isolar o parênquima pulmonar e facilitar a sua análise. Os primeiros sistemas de DAC propunham uma metodologia de segmentação manual ou semiautomática, o que requer muito tempo e um grande esforço, especialmente nos novos *scanners* TC multidetetor que produzem muitos dados por exame. Tornou-se essencial a automatização desta etapa de pré-processamento em sistemas de DAC que envolvam patologias pulmonares. Dado que o tecido pulmonar apresenta, normalmente, uma densidade menor que os tecidos que o envolvem muitos métodos de segmentação pulmonar baseiam-se em técnicas de *thresholding*, seguidas por operações morfológicas e análise de ligação entre componentes para refinar os resultados obtidos (Sluimer, Schilham, Prokop, & van Ginneken, 2006; Anthimopoulos et al., 2015). No entanto, estas técnicas que se baseiam fundamentalmente nos níveis de intensidade dos tecidos falham na presença de patologias com densidade alta como é o caso de algumas DIP, com uma distribuição que é muitas vezes periférica e basal, dando origem a regiões pulmonares subsegmentadas. Estas situações podem dar origem a erros nas fases seguintes dos sistemas de

DAC. Revisões bibliográficas que incluem segmentação pulmonar e suas estruturas podem ser encontradas em (Sluimer et al., 2006; Sonka et al., 2007; Anthimopoulos, Christodoulidis, Ebner, Christe, & Mougiakakou, 2016).

Após a etapa de segmentação, os padrões DIP em estudo são identificados através da análise e classificação das regiões consideradas. Uma vez que a presença de DIP se caracteriza por alterações da textura do parênquima pulmonar muitos sistemas de DAC recorrem a métodos de análise de textura 2D ou 3D para detetar, quantificar e categorizar padrões radiológicos tipicamente associados à DIP. Na literatura especializada, os trabalhos de investigação existentes mostram que a análise de textura é uma abordagem correta para o estudo de alterações do parênquima pulmonar. Usualmente, as características que descrevem as regiões são extraídas através de métodos distintos dando origem a um conjunto de diferentes tipos de características, obtidas de forma cumulativa. Deste conjunto é selecionado um subconjunto de características relevantes que permitam a maior diferenciação entre as classes, com a menor dimensionalidade possível. Recorrendo a algoritmos de aprendizagem automática, geralmente supervisionados, cada instância é classificada de acordo com os padrões pulmonares DIP considerados. As regiões identificadas podem ser posteriormente analisadas pelo imagiologista ou por sistemas computorizados para a obtenção de um diagnóstico diferencial.

Num trabalho de revisão do estado da arte é difícil estabelecer uma comparação da avaliação de desempenho dos sistemas descritos em diferentes trabalhos uma vez que: (i) os objetivos dos sistemas DAC são diferentes, (ii) o conjunto de imagens e respetivas anotações são muito diferentes entre si, (iii) o número e tipo de instâncias (ROI, imagem, volume de interesse (VOI - *volume of interest*), entre outras) consideradas também diferem bastante, assim como os métodos de avaliação do desempenho do sistema.

Em 1997, Delorme et al. (Delorme, Keller-Reichenbecher, Zuna, Schlegel, & van Kaick,1997) propuseram o primeiro sistema de DAC completamente automático para a classificação de padrões pulmonares (normal, enfisema, vidro despolido, fibrose intralobular, vasos sanguíneos e brônquios). As imagens são extraídas de três regiões anatómicas distintas: 3 cm abaixo e acima do nível da carina, assim como, aproximadamente, 2 cm acima da mais alta cavidade diafragmática. Após uma primeira fase correspondente à segmentação pulmonar, são usadas três pirâmides de resolução de modo a minimizar a influência de pequenas estruturas

pulmonares na análise da imagem. Cada ROI, em cada pirâmide, é caracterizada através de um conjunto de descritores de textura estatísticos de primeira ordem (FOS – *First Order Statiscs*), extraídos com base no histograma de níveis de cinzento, a partir do Método de Dependência Espacial de Níveis de Cinzento da região (SGLDM - *Spatial Gray Level Dependence Method*) e do Método de Comprimento de Primitivas de Níveis de Cinzento (GLRLM – *Gray Level Run Length Method*). A categorização das ROIs em seis classes é efetuada através de uma Análise Discriminante passo-a-passo (SDA – *Stepwise Discriminant Analysis*) após uma pré-seleção das características baseadas no valor médio de Hounsfield da ROI. Os resultados mostram que a análise de textura é válida no reconhecimento da maioria das lesões associadas à DIP, com classificações corretas que vão de 68.7% a 80.7%, mas demonstrou dificuldade em classificar áreas de enfisema, veias e brônquios, com valores entre 16.2% e 53.7%. O treino do sistema é efetuado num conjunto de ROIs selecionadas e a avaliação em todas as ROIs. Assim, o método de validação utilizado não garante que uma ROI integre apenas o conjunto de treino ou o conjunto de teste.

Uppaluri et al. (Uppaluri, Mitsa, Sonka, Hoffman, & McLennan, 1997) propuseram o método AMFM (*Adaptive Multiple Feature Method*) para distinguir regiões pulmonares saudáveis de regiões com enfisema severo, cujos pacientes se encontravam em avaliação para eventual cirurgia de redução pulmonar. O método incorpora características estatísticas (FOS, SGLDM e GLRLM) e a dimensão fractal geométrica para descrever a textura pulmonar. Após a segmentação pulmonar automática, as imagens são processadas através de *edgmentation*, uma técnica de crescimento de regiões que une pixels adjacentes em que a diferença de níveis de cinzento é inferior a um limiar, sendo atribuída à área construída a média dos níveis de cinzento que integram a região. Sobre estas áreas, da imagem original, são calculadas as características de primeira ordem e a dimensão fractal e nas imagens processadas são calculadas as restantes características. As ROIs são divididas aleatoriamente no conjunto de treino e teste. A seleção do conjunto de características ótimo é efetuada nas ROIs do conjunto de treino, usando medidas de divergência das características conjugada com a correlação entre elas.

Em (Uppaluri, Hoffman, Sonka, Hunninghake, & McLennan, 1999a) o método AMFM foi estendido para a discriminação de quatro classes: normal, enfisema, fibrose pulmonar idiopática e sarcoidose. Numa análise global, as ROIs são constituídas por toda a região pulmonar, tendo sido obtida uma exatidão global de 81%, calculada num conjunto de imagens de teste

aleatoriamente selecionadas. Foi ainda estudada a correlação entre os testes de função pulmonar e o método proposto. Os resultados apresentados nos dois estudos são superiores aos obtidos com métodos clássicos de avaliação do enfisema pulmonar como a densidade pulmonar média e a análise do histograma.

Em (Uppaluri et al., 1999b) o método AMFM é aplicado na categorização de seis padrões pulmonares associados à DIP. São apenas consideradas as ROIs onde houve a concordância acerca do padrão presente por três dos quatro imagiologistas envolvidos no estudo, existindo uma sobreposição de 50% entre as ROIs. Neste estudo, a árvore pulmonar é modelada como um fractal geométrico (autossimilaridade geométrica em diferentes escalas) e estocástico (autossimilaridade estatística em diferentes escalas), tendo sido adicionadas cinco características ao conjunto original do método AMFM baseadas no modelo fractal estocástico. É ainda avaliada a concordância intra e interobservador, sendo o desempenho obtido pelo método AMFM semelhante ao de observadores experientes após o conhecimento prévio do diagnóstico do paciente. É de referir que as métricas de avaliação de desempenho do sistema são calculadas sobre um conjunto de teste onde podem existir ROIs de pacientes que integraram o conjunto de treino do sistema.

Em (Uchiyama et al., 2003) foi introduzida a utilização de medidas geométricas baseadas na transformação morfológica *top-hat* para a deteção e caracterização automática de sete padrões associados à DIP (normal, vidro despolido, opacidades reticulares e lineares, opacidades nodulares, favo de mel, enfisema e consolidação). A primeira etapa consiste na segmentação dos pulmões através de operações morfológicas e técnicas de *thresholding*. Inicialmente é aplicada a operação morfológica de abertura, para retirar as pequenas regiões de elevada intensidade correspondentes aos vasos sanguíneos. Posteriormente, o histograma de níveis de cinzento é construído e o *threshold* considerado para proceder à binarização das imagens é aquele que maximiza a separação entre os dois picos do histograma. Este método de segmentação falha em imagens com a presença do padrão consolidação, sendo, nestes casos, efetuada uma segmentação pulmonar manual. Em imagens de regiões anatómicas específicas, a área segmentada dos pulmões é dividida em ROIs contíguas, marcadas por três imagiologistas experientes. As ROIs que possuem padrões específicos de DIP e onde houve concordância na marcação de todos os imagiologistas são consideradas o *gold standard*. De cada uma das ROIs são extraídas seis medidas físicas, três baseadas na distribuição dos níveis de cinzento – média,

desvio padrão e fração da área da ROI que possui componentes da densidade do ar (estabelecido com valores entre -910 UH (Unidades Hounsfield) e -1000 UH) e três medidas geométricas usadas para a caracterização das componentes nodulares, lineares e multiloculares. O recurso a transformação morfológica *white top-hat* e *black top-hat* seguidas de *thresholding* permite a diferenciação entre os padrões nodular, lineares e reticulares, assim como, a eliminação de estruturas morfológicas naturais. A avaliação do sistema é efetuada apenas no conjunto de treino. A principal dificuldade encontrada pelos autores foi a diferenciação entre o padrão normal e o nodular.

Chabat e seus colaboradores (Chabat, Yang, & Hansell, 2003) propuseram um sistema de DAC com o objetivo de categorizar regiões pulmonares de pacientes com DPOC e saudáveis. É efetuada a segmentação automática das grandes veias pulmonares recorrendo a morfologia matemática. Este pré-processamento é importante porque a dimensão das estruturas segmentadas aproxima-se do tamanho das ROIs utilizadas, 22 pixels de raio. Em cada imagem são extraídas quatro ROIs circulares de cinco níveis anatómicos predefinidos. A textura de cada ROI é caracterizada através de um vetor constituído por características estatísticas de primeira ordem, extraídas do SGLDM e do GLRLM. A classificação das regiões é efetuada através do classificador de Bayes, sendo o treino e teste do sistema efetuados recorrendo a grupos distintos de sujeitos.

Em 2003, Sluimer et al. (Sluimer, van Waes, Viergever, & van Ginneken, 2003) propuseram um sistema de DAC que distingue padrões patológicos de padrões normais de pacientes com DIP. As regiões a classificar são selecionadas manualmente da mesma região anatómica (zona anterior dos pulmões, à altura do arco da aorta) sob a supervisão de um imagiologista. As imagens são processadas através da aplicação de bancos de filtros que integram os filtros simétricos Laplaciano e Gaussiano e a sua primeira e segunda derivada parcial. Todos os filtros são aplicados segundo quatro escalas diferentes, sendo os filtros derivativos aplicados segundo seis ângulos. Das imagens convolvidas são extraídas características correspondentes aos quatro momentos centrais das imagens filtradas: média, desvio padrão, simetria e curtose. Um subconjunto de características é obtido através do método *Sequential Forward Selection* (SFS). São considerados vários algoritmos de classificação tais como Análise Discriminante Linear (LDA - *Linear Discriminant Analysis*), Análise Discriminante Quadrática (QDA - *Quadratic Discriminant Analysis*), *k-Nearest Neighbor* (*k*-NN) e SVM. No caso do classificador SVM são

utilizadas a totalidade das características. Os melhores resultados foram obtidos para os classificadores *k*-NN e SVM com 0.86 obtido recorrendo à análise das curvas ROC (*Receiver Operating Characteristic*), desempenho semelhante à avaliação das ROIs efetuada por dois imagiologistas (0.88 e 0.89). O método de validação utilizado reflete o desempenho do sistema em ambiente clínico, uma vez que é efetuada com base nos pacientes, apesar de ser aplicado apenas a regiões pulmonares da região anterior do pulmão.

De modo semelhante à metodologia usada por imagiologistas na interpretação de imagens, Zrimec e seus colaboradores (Zrimec, & Wong, 2007) propõem um sistema multimodal que incorpora o conhecimento anatómico e a aparência dos padrões para aumentar o desempenho do sistema de DAC aplicado ao padrão favo de mel. Na fase de pré-processamento é realizada a segmentação pulmonar recorrendo a *thresholding* adaptativo, operadores morfológicos e contornos ativos. São considerados dois tamanhos de ROIs (7x7 e 15x15 pixels) para captar as características dos diferentes tamanhos de quistos, a partir das quais são extraídas características de primeira e segunda ordem que são selecionadas através do método *Correlation Feature Selection* (CFS). A validação do sistema é efetuada de duas formas diferentes: através de validação cruzada permitindo o treino e teste com ROIs do mesmo paciente e o teste em dados novos. Nesta última situação a sensibilidade diminui consideravelmente o que mostra a importância da metodologia de validação. O conhecimento da predominância periférica do padrão favo de mel (*honeycombing*) levou ao desenvolvimento de uma estratégia de classificação em que o sistema só classifica ROIs que se encontram na periferia. Destas, as que forem classificadas como favo de mel constituirão sementes para encontrar ROIs com favo de mel vizinhas, através de crescimento de regiões. Esta estratégia permitiu melhorar a exatidão devido à diminuição do número de falsos positivos.

Kim et al. (Kim et al., 2009) mostram que para além da textura, a forma dos padrões é importante na diferenciação de padrões associados a regiões de baixa atenuação associadas à DPOC. ROIs de diferentes tamanhos (16×16, 32×32 e 64×64) foram marcadas por dois imagiologistas experientes. Áreas com atenuação entre -1024 UH e -400 UH foram segmentadas para isolar os grandes vasos sanguíneos. A textura é descrita através de características baseadas em estatísticas de primeira ordem, do SGLDM, do GLRLM, do gradiente. São consideradas as transformações *white* e *black top-hat* e zonas de atenuação anormalmente baixas (inferiores a -950 UH) para a extração de descritores de forma. Foram

investigados métodos de redução do número de características, nomeadamente: SFS, *Sequential Backward Selection* (SBS), *Plus l-Take Away r* (PTA(*l*, *r*)), *Sequential Floating Forward Selection* (SFFS) e *Sequential Floating Backward Selection* (SFBS). Não houve diferenças significativas no desempenho do sistema de DAC entre os algoritmos de redução de características testados, pois na realidade todos os eles se baseiam em processos iterativos semelhantes. O sistema é testado através de validação cruzada de todas as ROIs, repetida vinte vezes. A junção da informação da forma permitiu ao classificador de Bayes e SVM melhorar a sensibilidade, em especial em ROIs de pequenas dimensões, o que é um aspeto relevante em termos de prática clínica.

Em 2009, Lee e colaboradores (Lee et al., 2009) compararam o desempenho dos classificadores redes neuronais (RN), SVM, Bayes e naïve Bayes na diferenciação de padrões associados à DPOC. Os padrões são descritos através de características estatísticas de primeira ordem, do SGLDM, do gradiente e do GLRLM calculadas para ROIs de diversos tamanhos, previamente marcadas por dois imagiologistas. O método é testado através de validação cruzada baseada no paciente sendo repetida vinte vezes. O classificador SVM mostrou possuir maior sensibilidade e especificidade na predição das classes em análise para as ROIs de dimensão 32×32 e 64×64, assim como a melhor exatidão global. O SVM foi também o classificador com menor tempo de teste. Os resultados mostraram a importância do tamanho da ROI no desempenho global do sistema.

Em (Sørensen, Shaker, & de Bruijne, 2010) é apresentado um sistema para quantificação de enfisema que se baseia em padrões binários locais (LBP - *Local Binary Patterns*) invariantes à rotação. Através dos LBP é possível detetar microestruturas como, por exemplo, contornos, pontos e cantos. Uma vez que por conceção o LBP é invariante a transformações de intensidade monotónicas, verificou-se que ao adicionar a informação do histograma de intensidades ao histograma LBF a classificação melhorou significativamente, aumentando a exatidão de 79.2% para 95.2%. Este valor é similar ao obtido considerando bancos de filtros derivativos baseados na função de Gauss, invariantes à rotação (Sluimer et al., 2003), através dos quais se obteve uma exatidão de 94%. A classificação das ROIs é efetuada através do *k*-NN sendo a dissimilaridade do histograma a medida da distância. O treino do classificador é efetuado através de uma estratégia *leave-one-patient-out* (LOPO).

O modelo proposto por Vo e Sowmya (Vo, & Sowmya, 2010) para a classificação de quatro padrões DIP baseia-se em características obtidas através de *discrete wavelet frames*, transformada *contourlet* e filtros derivativos direcionais de Gauss. A classificação é efetuada através de aprendizagem com recurso a múltiplos *kernels* (MKL - *Multiple Kernel Learning*) multiclasse com base numa validação cruzada de todas as ROIs, repetida dez vezes. Os três espaços de características foram avaliados individualmente verificando-se que o método baseado em *contourlet* supera os outros dois. A combinação dos três grupos permitiu obter uma sensibilidade e especificidade médias de 94.2% e 98.7%, respetivamente.

Em (Depeursinge et al., 2012a) é proposto um conjunto de descritores de textura baseados em propriedades específicas da transformada de *wavelet* com subamostragem Quincunx. Os autores obtiveram a isotrópica *Quincunx Wavelet Frames* (QWF) que é invariante à translação, rotação e covariante à escala. A segmentação do volume pulmonar é efetuada de forma semiautomática através de crescimento de regiões a partir de uma semente escolhida pelo utilizador e finalizada com operações morfológicas de fecho. Ajustes manuais à segmentação são necessários em casos pontuais para a exclusão da traqueia e em aéreas extensas de consolidação. Dois imagiologistas desenham e anotam regiões do pulmão em cinco classes diferentes, a partir de exames de 85 pacientes, sendo cada uma destas regiões dividida em blocos sobrepostos de 32×32 pixels. Cada bloco é caracterizado por um vetor que inclui os coeficientes da QWF, modelados por uma mistura de duas Gaussianas com escala diferente e igual média, o histograma de intensidades e o número de pixels de ar, com valores inferiores ou iguais a -1000 UH. A classificação é efetuada através do classificador SVM recorrendo à validação cruzada LOPO. Foi obtida uma exatidão média de 76.9%. Este resultado é consequência de classificação errada entre, essencialmente, os padrões: normal e micronódulos e fibrose e vidro despolido.

Anthimopoulos et al. (Anthimopoulos, Christodoulidis, Christe, & Mougiakakou, 2014) propõem um método para a descrição e classificação de seis classes de padrões pulmonares. A extração de características baseia-se na análise do espectro local recorrendo à Transformada Discreta do Cosseno 2D (DCT - *Discrete Cosine Transform*). A DCT descreve um sinal discreto como combinações lineares de funções cosseno com diferentes frequências. Após a convolução das imagens com os 25 bancos de filtros (5x5 DCT) são calculados os 10-quantis de cada imagem filtrada a partir dos histogramas acumulados e concatenados num vetor de

características. A este vetor de características é adicionado o histograma de níveis de cinzento da imagem original. O classificador usado na discriminação das seis classes foi o *Random Forest* (RF). O RF resulta da conjugação de várias árvores de decisão, em que cada árvore é "alimentada" por vetores de características aleatoriamente amostrados. A classe atribuída a um novo vetor de características é obtida através da maioria dos votos de todas as árvores de decisão. Segundo os autores, comparativamente ao SVM e RN, o RF é mais rápido, especialmente na fase de predição, podendo gerir bases de dados com elevado número de variáveis de entrada, e ser completamente paralelizável. O procedimento proposto é avaliado em cerca de 2500 regiões com uma sobreposição máxima de 33%, recorrendo à validação cruzada em cinco conjuntos de teste e treino.

Hamzah et al. propõem um procedimento que pode ser considerado uma triagem inicial na deteção de DIP (Hamzah, Kasim, Yunus, Rijal, & Noor, 2017). A base de dados é constituída por três tipos de pacientes: saudáveis, com patologias não DIP e com DIP. Todas as imagens selecionadas encontram-se na região 1 a 2 cm do diafragma. Para cada imagem são calculadas sete medidas de textura (Energia, Energia Média, Entropia, Contraste, Homogeneidade, Desvio Padrão e Desvio Padrão da Energia) e a correlação entre pares de medidas. As correlações são convertidas em medidas de distância que através de técnicas de *clustering* permitem distinguir dois grupos de medidas, usados para criar três modelos de regressão linear. O modelo de regressão mais discriminatório é dado por Desvio Padrão da Energia = α_0Contraste + α_1. A estimação dos parâmetros (α_0, α_1) para os três modelos de dados é efetuada através de validação cruzada. O método distingue dados de pacientes com DIP de pacientes com patologia não DIP com uma taxa de sucesso de 73% e distingue pacientes com DIP de pacientes saudáveis com uma taxa de sucesso de 67%.

Alguns sistemas de DAC exploram a informação volumétrica adquirida pelos modernos tomógrafos multidetetor que permitem a aquisição de voxels submilimétricos, praticamente isotrópicos. Muitos dos trabalhos procedem à expansão de conjuntos de características 2D propostos e propõem uma análise tridimensional.

Em (Xu, Sonka, McLennand, Guo, & Hoffmanc, 2005; Xu, Sonka, McLennan, Guo, & Hoffman, 2006a; Xu et al., 2006b) o conjunto de características do método AMFM foi adaptado à análise 3D da textura em exames TCAR volumétricos. O método foi aplicado na classificação

de padrões pulmonares de pacientes com DPOC e padrão normal de indivíduos fumadores e não fumadores, num total de cinco classes (Xu et al., 2005; Xu et al. 2006a). A fase de pré-processamento das imagens consiste na segmentação das vias áreas e veias, *edgmentation* e segmentação dos pulmões. A exclusão da árvore broncovascular diminui o "ruído" pois corresponde a zonas hipodensas e hiperdensas que afetam as características e podem ser considerados falsos positivos relativamente a algumas patologias, por exemplo, entre as vias aéreas e o padrão enfisema. A árvore brônquica é segmentada a partir de crescimento de regiões baseado numa VOI semente. As VOIs da imagem são comparadas com a semente e se o valor de similitude for superior a um valor estabelecido a região é ligada à árvore. Posteriormente, o "esqueleto" da árvore obtido é criado para identificar os segmentos centrais e os pontos de bifurcação que serão usados para o refinamento da segmentação (Tschirren, Hoffman, McLennan, & Sonka, 2005). A árvore vascular é segmentada através do método proposto por Shikata et al. (Shikata, Hoffman, & Sonka, 2004) que realça o aspeto tubular dos vasos através de um filtro de linha baseado em combinações dos valores próprios da matriz Hessiana. Após a aplicação de um *threshold* é usado um algoritmo de *tracking* para eliminar falhas e descontinuidades. A segmentação automática dos pulmões é efetuada através do método proposto por Hu et al. (Hu, Hoffman, & Reinhardt, 2001) que possui três fases: a extração da região dos pulmões através da aplicação de um *threshold* iterativo que permite acomodar as pequenas variações existentes nas imagens correspondentes a diferentes indivíduos, a separação dos dois pulmões através da identificação das junções entre eles e por fim a suavização dos contornos ao longo do mediastino através de sequências de operações morfológicas. A extensão das características de 2D para 3D e a segmentação das estruturas áreas e vasculares permitiram aumentar significativamente a diferenciação entre patologias associadas ao tabagismo. Os testes foram realizados de modo realista recorrendo à validação LOPO através um classificador de Bayes. Os autores mostram também que a análise 3D permite a perceção de estruturas diferentes entre os padrões pulmonares normais correspondentes a indivíduos não fumadores e a indivíduos fumadores. Em (Xu et al., 2006b) o sistema 3D AMFM foi aplicado com sucesso na classificação de cinco padrões pulmonares associados à DPOC e à DIP. A classificação das VOIs foi efetuada recorrendo ao classificador de Bayes e SVM através de validação cruzada (k = 10) e ao teste McNemar. Verificou-se que os dois classificadores mostraram desempenhos similares na classificação de padrões associados à DIP.

Em (Zavaletta, Bartholmai, & Robb, 2007) as VOIs são descritas através das assinaturas dos seus histogramas de *bins* adaptativos e classificadas através da *Earth Mover's Distance* (EMD) das assinaturas canónicas das classes. São consideradas VOIs de cinco classes (normal, vidro despolido, reticular, favo de mel e enfisema) obtidas a partir de exames de pacientes com fibrose pulmonar idiopática ou pneumonia intersticial. Cada base de dados é pré-processada o que inclui a segmentação dos pulmões e da árvore broncovascular. A segmentação pulmonar baseia-se em *thresholding*, crescimento de regiões e aplicação de operadores morfológicos. Uma vez que estão presentes padrões de elevada atenuação, a segmentação da árvore vascular torna-se mais complexa que a árvore brônquica. A traqueia e os brônquios centrais são segmentados através de um processo iterativo de crescimento de regiões em pequenas vizinhanças de 6-8 pixels sendo aplicados diferentes *thresholds*. É efetuada a segmentação dos vasos sanguíneos superiores a 3 mm de diâmetro através da aplicação de filtros 3D de linha baseados na matriz Hessiana seguido de *thresholding* manual. O histograma adaptativo é construído recorrendo ao algoritmo *K-means clustering* e a sua assinatura é definida pelos pares de valores $\{(\mu_i, w_i), \dots (\mu_k, w_k)\}$, sendo μ_i o centróide do *cluster* e w_i o respetivo número de VOIs. A assinatura canónica de cada classe é calculada a partir das assinaturas de cada VOI readaptando os *k clusters* à nova distribuição. A criação de assinaturas canónicas para cada classe permite otimizar o *matching* entre assinaturas, não sendo necessário o cálculo da distância entre todas as assinaturas das VOIs do conjunto de treino e todas as VOIs do conjunto de teste, mas apenas entre estas e as assinaturas canónicas de cada classe. Os valores obtidos são altamente condicionados pela baixa sensibilidade do padrão reticular devido à sua semelhança com o padrão vidro despolido.

Também Korfiatis et al. (Korfiatis, Karahaliou, Kazantzi, Kalogeropoulou, & Costaridou, 2010) propõem a identificação e quantificação volumétrica automática de padrões associadas à pneumonia intersticial, um subconjunto da DIP, baseado na análise de textura 3D. Na fase de pré-processamento é efetuada a segmentação automática dos pulmões e da árvore vascular. O método de segmentação 3D dos pulmões é adaptado aos padrões associados à pneumonia intersticial. Na primeira etapa a segmentação é obtida através de *threshold* após o realce dos contornos recorrendo a *wavelets*. O volume pulmonar segmentado é deficitário devido à presença de padrões patológicos nas regiões periféricas. O seu refinamento é efetuado através da etiquetagem recursiva dos pixels vizinhos dos contornos iniciais através do classificador SVM baseado na textura extraída de ROIs centradas em todos os pontos dos contornos. Para

isso é aplicado um filtro 3D de linha em várias escalas baseado nos valores próprios da matriz Hessiana para realçar as veias/artérias e as suas bifurcações. A classificação dos voxels (*volume elements*) em três classes de padrões (normal, vidro despolido e reticular) é efetuada com base na média e desvio padrão de cada uma das 13 características extraídas das matrizes de coocorrência 3D, segundo 13 direções. São consideradas cinco distâncias entre VOIs o que perfaz um total de 130 características para descrever a textura de cada VOI. A redução da dimensionalidade é efetuada através de uma SDA, sendo posteriormente usado o classificador *k*-NN para a classificação de cada voxel. O desempenho do método baseia-se nas métricas: sobreposição do volume (VO - *Volume Overlap*) entre a marcação do imagiologista e os resultados do computador, fração de verdadeiros positivos (FVP) e fração de falsos positivos (FFP) tendo como referência os contornos manuais delineados por um imagiologista. O treino e teste do método é efetuado em exames de pacientes diferentes. Os resultados obtidos mostram a importância da etapa de refinamento da segmentação, especialmente no padrão reticular.

Após um processo de pré-processamento das imagens semelhante ao descrito por (Korfiatis et al., 2010), que resulta na segmentação dos pulmões e árvore vascular, Mariolis et al. (Mariolis et al., 2010) avaliam de uma forma sistemática conjuntos de características extraídas através do GLRLM 3D e *features Laws* 3D na capacidade de diferenciar VOIs de quatro tipos de padrões associados à DIP (normal, vidro despolido, reticular e favo de mel) quando conjugados com os classificadores *k*-NN, *probabilistic neural network* (PNN), Bayes e regressão logística multinominal (RLM), sendo avaliados 12 DACs diferentes. As primitivas de comprimento 3D são calculadas em 13 direções e delas são extraídas a média e desvio padrão de 11 características. Através da filtragem das VOIs com 125 máscaras *laws* 3D são extraídas cinco características de primeira ordem, num total de 625. O número de características é reduzido através do método SDA. Os parâmetros correspondentes ao tamanho das VOIs, número de níveis de cinzento e os parâmetros de cada um dos classificadores são escolhidos baseados numa grelha de pesquisa, sendo escolhidos aqueles que correspondem à maior taxa de exatidão global. O desempenho dos vários classificados é avaliado através de validação cruzada, com $k = 10$. Apesar dos dois conjuntos de características mostrarem um alto poder de diferenciação dos padrões, com taxas de exatidão superiores a 96%, a sua conjugação permitiu melhorar os resultados para todos os classificadores considerados com valores de exatidão superiores a 98.6%.

Os trabalhos descritos anteriormente recorrem às designadas características *hand-craft* para descrever padrões pulmonares em imagens de TCAR. A limitação destas características é que por vezes não se adaptam a novos dados ou padrões. Têm surgido trabalhos que implementam métodos de aprendizagem para a extração de características adaptadas ao conjunto de treino, baseados em *bag-of-words* (por vezes designado por *bag-of-features*) (Gangeh at al., 2010; Xu, Hirano, Tachibana, & Kido, 2013) ou em modelos de representação esparsa (SR - *Sparse Representation*) (Zhao, Xu, Hirano, Tachibana, & Kido, 2013; Zhang et al., 2013; Song, Cai, Zhou, & Feng, 2013; Zhao, Xu, Hirano, Tachibana, & Kido, 2015).

Em (Gangeh et al., 2010) é proposto um sistema de análise de textura para a classificação de subtipos de enfisema baseado em *textons* de intensidade. Esta aproximação pode ser decomposta em três fases: a construção do *codebook* de *textons* recorrendo a um algoritmo de *clustering*, o treino do modelo através da criação dos histogramas de *textons* para as imagens do conjunto de treino e, por fim, a classificação do conjunto de teste através da criação dos seus histogramas de *textons* e comparando-os com os histogramas criados durante a fase dois, de forma a encontrar o padrão mais próximo. Neste trabalho, para a construção do *codebook* de *textons* que representará as três classes (normal, enfisema centrolobular e enfisema paraseptal), foram extraídas aleatoriamente *patches* de todas as ROIs de cada classe que integram o conjunto de treino. Estas pequenas áreas são agrupadas recorrendo ao algoritmo *k-means* de forma a criar o *codebook*, que guarda os centros dos *clusters* criados. Para cada ROI são extraídas regiões sobrepostas do mesmo tamanho das da fase anterior, através de uma janela deslizante, e de acordo com a similitude (distância Euclidiana) com o *codebook* é criado um histograma que caracteriza cada ROI. Os histogramas construídos são usados como características para treinar o classificador SVM. Para a classificação de cada ROI do conjunto de teste é necessário a criação dos seus histogramas, de forma semelhante ao descrito para o conjunto de treino. O classificador já treinado é usado para atribuir uma classe a cada ROI de teste. Para além dos parâmetros do classificador SVM – RBF é necessário otimizar o tamanho do *patch* e o número de *clusters*. O desempenho do sistema foi avaliado através de uma estratégia de validação LOPO. Os resultados da classificação da textura baseada em *texton* é melhor que a obtida em sistemas baseados nos momentos do histograma da resposta a bancos de filtros de Gauss e ligeiramente melhor que a proposta baseada no LBP (Sørensen et al., 2010).

Em (Xu et al., 2013) é proposto um método baseado em *bag-of-words* para a classificação de seis padrões pulmonares. As características locais 3D têm em consideração os níveis de intensidade e a forma dos padrões, através dos valores próprios da matriz Hessiana. As regiões marcadas nos exames obtiveram a concordância de três imagiologistas e a partir delas são extraídas VOIs não sobrepostas. Cada VOI é amostrada regularmente através de uma grelha e sobre esses pontos são considerados quatro cubos, três obtidos através dos valores próprios da matriz Hessiana e um contendo os valores de intensidade originais da imagem. De cada um são extraídas a média, desvio padrão, simetria e curtose. Assim, cada característica local é composta por 16 medidas. As VOIs são separadas em dois conjuntos distintos sendo que as VOIs de um paciente se encontrem apenas num dos conjuntos. O *codebook* é construído no conjunto de treino. Posteriormente, cada VOI é caracterizada por um histograma de acordo com o número de características locais associadas a cada *cluster*. O conjunto de todos os histogramas do conjunto de treino serão as características que irão treinar o classificador. Os parâmetros ótimos do sistema foram obtidos através de uma estratégia de validação cruzada (k = 20) no conjunto de treino. A otimização é efetuada para o tamanho do *patch* 3D, o número de *cluster* no algoritmo *k-means* e o *kernel*, e respetivos parâmetros, do classificador SVM. O método proposto é comparado com quatro métodos: (Gangeh et al., 2010) e a sua adaptação a 3D, (Zavaletta et al., 2007) e (Uchiyama et al., 2003). Em todos os casos foi considerada a metodologia proposta por Xu et al. (Xu et al., 2013). O método proposto é melhor que os quatro métodos testados. Este estudo comparativo mostrou a superioridade dos métodos 3D e a importância da informação da forma, para além da informação da intensidade, especialmente nos padrões normal, favo de mel e nodular.

Os trabalhos baseados na representação esparsa têm como ideia principal a aproximação dos exemplos ou amostras através da combinação linear esparsa de um pequeno número de características chave, que neste contexto se designam por átomos, selecionadas a partir de um dicionário de átomos, garantindo um pequeno erro de reconstrução. Uma vez que se trata de um *overcomplete dictionary* (o número de átomos do dicionário é muito superior à dimensionalidade do exemplo de entrada) cada exemplo pode ser representado por mais do que uma combinação de átomos.

O método proposto por (Zhao et al., 2013) combina as características locais, estratégia de treino e validação propostos por (Xu et al., 2013) com um modelo de representação esparsa para a

classificação do tecido pulmonar normal e cinco padrões associados à DIP, através do classificador SVM. O modelo de representação esparsa possui três etapas: a aprendizagem do dicionário, a codificação esparsa dos exemplos de entrada e o *pooling* espacial médio. A primeira fase consiste na extração de um elevado número de características locais a partir das VOIs. A aprendizagem do dicionário e a codificação esparsa faz-se a partir dos vetores de características do conjunto de treino através dos algoritmos *k*-SVD e OMP (*Orthogonal Matching Pursuit*) que são executados iterativamente até o erro de reconstrução ser menor que um limiar de paragem. Posteriormente, a representação esparsa das características locais originais é calculada de acordo com o dicionário. Por fim, os vetores de coeficientes de representação esparsa obtidos da mesma VOI são combinados para obter um único vetor que será usado como entrada do classificador. Recorrendo a uma validação cruzada, os parâmetros a otimizar são: o tamanho do *patch* 3D, o número de *átomos*, a esparsidade e o parâmetro de regularização *C* do classificador SVM. O método proposto apresenta resultados superiores aos obtidos em (Xu et al., 2013), sendo obtida uma exatidão superior a 95%. Em (Zhao et al., 2015) é proposta a substituição do método k-SVD pelo método *k-means* o que reduz o tempo de treino do dicionário em 98.2%. É proposto um método para a codificação esparsa que simplifica o método OMP reduzindo o tempo de teste de uma VOI em 55.2%. O desempenho do sistema melhorou, tendo sido obtida uma exatidão de 96.4%

Na análise visual das imagens TCAR pulmonares efetuada pelos imagiologistas, observa-se que a textura, intensidade e a distribuição do gradiente dos tecidos pulmonares pertencentes a uma área de interesse são bastante informativos e discriminatórios das diferentes categorias de tecidos pulmonares. Song et al. (Song et al., 2013) estruturaram um conjunto de características para descrever as áreas da imagem de acordo com esta trilogia. É proposto o descritor de textura *Rotation - Invariant Gabor* - LBP (RIG-LBP) que consiste em extrair características LBF invariantes à rotação de um conjunto de imagens convolucionadas com filtros de Gabor, para diferentes escalas e orientações. Para que as imagens filtradas sejam invariantes à rotação são somadas as imagens obtidas para a mesma escala, para as diferentes rotações. Para as imagens de cada escala é criado um histograma baseado no LBP invariante à rotação. A informação da intensidade é incluída através do histograma de níveis de cinzento. A informação do gradiente é obtida através do histograma dos gradientes orientados calculados num sistema de multicoordenadas radial. Para cada classe é construído um *overcomplete dictionary* que reúne os vetores de características de todas as regiões dessa classe, em que o número de linhas

(dimensão dos vetores de características) é inferior ao número de colunas (número de regiões de cada classe). A classificação é efetuada através do método proposto *Patch-Adaptive Sparse Approximation* (PASA) que se baseia no critério da discrepância mínima. A classe a atribuir corresponde àquela que produz menor discrepância entre o vetor original e as aproximações esparsas obtidas para todos os dicionários. As métricas de desempenho foram calculadas através de uma validação LOPO em regiões anotadas por dois imagiologistas.

Também através da representação esparsa de um dicionário de *textons* extraídos diretamente dos níveis de intensidade dos pixels, Zhang et al. propõem um método para a diferenciação do tecido pulmonar normal de três tipos de enfisema (Zhang et al., 2013). A construção do dicionário de *textons* é efetuado a partir de 10% da totalidade das ROIs, a partir das quais são extraídos *patches.* É aplicado um filtro passa-alto para excluir os *patches* pouco informativos. A aprendizagem do dicionário é efetuada através dos algoritmos *k*-SVD e OMP. A classificação da textura é feita considerando os coeficientes esparsos conjuntamente com o histograma de intensidades.

A Tabela 2.1 reúne um conjunto de características dos trabalhos resumidos nos parágrafos anteriores. Os trabalhos encontram-se na tabela por ordem cronológica.

Mais recentemente têm sido adotadas técnicas de *Deep Learning* (DL) na análise e classificação de padrões do tecido pulmonar, com especial ênfase para as *Convolutional Neural Networks* (CNNs) (Li et al., 2014; Gao et al., 2016; Anthimopoulos et al. 2016). As CNNs fazem a aprendizagem das características e treino de uma RN simultaneamente, através da minimização do erro de classificação. Apesar do conceito DL implicar várias camadas de aprendizagem os trabalhos desenvolvidos em imagens de TC pulmonar recorrem a arquiteturas pouco profundas (Anthimopoulos et al., 2016). O recurso a *Restricted Boltzmann Machine* (RBM) é usado em (Tulder, & de Bruijne, 2014; Tulder & de Bruijne, 2016) tanto na aprendizagem das características como na classificação do tecido pulmonar. As RBM são RN generativas capazes de captar e reproduzir a estrutura estatística dos dados.

Tabela 2.1: Resumo das características dos trabalhos descritos.

Estudo	**Padrões/Patologia**	**Características**	**Base de Dados**	**Classificação**	**Desempenho**
(Delorme et al., 1997)	Normal, Enfisema, Vidro Despolido, Fibrose Intralobular, Vasos Sanguíneos, Brônquios	FOS, SGLDM, GLRLM	Sujeitos: 10 ROIs: 1022, 5x5	SDA	**Exatidão**:70.7%
(Uppaluri et al., 1997)	Normal *versus* Enfisema	FOS, SGLDM, GLRLM, Dimensão Fractal	Sujeitos:19 ROI: todo o pulmão, ROIs geradas por *edgmentation*	Bayes	**Exatidão** = **Sensibilidade** = **Especificidade**:100%
(Uppaluri et al., 1999a)	Normal, Enfisema, Fibrose Pulmonar Idiopática, Sarcoidose	FOS, SGLDM, GLRLM, Dimensão Fractal	Sujeitos: 72 Imagens: 288	Bayes	**Exatidão**: 81%
(Uppaluri et al., 1999b)	Normal, Enfisema Severo, Vidro Despolido, Favo de Mel, Nodular, Broncovascular	FOS, SGLDM, GLRLM, Características Fractais	Sujeitos: 72, Imagens: 288 ROIs: 899, 31x31	Bayes	**Exatidão**: 93.5%
(Uchiyama et al., 2003)	Normal, Vidro Despolido, Opacidades Reticulares e Lineares, Opacidades Nodulares, Favo de Mel, Enfisema, Consolidação	FOS, Densidade de ar, Medidas Geométricas (Transformações *top-hat*)	Sujeitos:105 Imagens: 315 ROIs: 32×32, 96×96	RN	**Sensibilidade** Vidro Despolido: 99.2% Opacidades Reticulares e Lineares: 100% Opacidades Nodulares: 88% Favo de Mel: 100% Enfisema: 95.8% Consolidação: 100% **Normal *versus* Patológico** Sensibilidade: 97.4% Especificidade: 88.0%

Tabela 2.1: Resumo das características dos trabalhos descritos. (Continuação).

Estudo	**Padrões/Patologia**	**Características**	**Base de Dados**	**Classificação**	**Desempenho**
(Chabat et al., 2003)	Normal, Enfisema Centrolobular, Enfisema Panlobular, Bronquiolite Obliterante	FOS, SGLDM, GLRLM	Sujeitos:77 Imagens: 308 ROIs: 2 420, raio 22	Bayes	**Sensibilidade**: 73.6% **Especificidade**: 91.2%
(Sluimer et al., 2003)	Normal *versus* Patológico	Momentos centrais das imagens, após aplicação de banco de filtros: Gaussiano, Laplaciano, Derivativos de primeira e segunda ordem	Sujeitos: 116 ROIs: 657, raio 40	SVM, *k*-NN	**AUC**: 0.86
(Xu et al., 2005) (Xu et al., 2006a)	Enfisema em DOPC severa (EC), Enfisema em DOPC média (EM), Normal em DOPC (NC), Normal não Fumadores (NN), Normal Fumadores (NS)	FOS, SGLDM, GLRLM, Características Fractais (3D AMFM)	Sujeitos: 34 VOIs: 1 865, 31x31x31	Bayes	**Exatidão**: EC: 84.9% MC: 89.8% NC: 87.5% NN: 100%
(Xu et al., 2006b)	Enfisema, Vidro Despolido, Favo de Mel, Normal Não Fumador, Normal Fumador	FOS, SGLDM, GLRLM, Características Fractais (3D AMFM)	Sujeitos: 20 VOIs: 1 865, 21x21x21	SVM, Bayes	**Exatidão**: Bayes: 86.2% SVM: 83.8%
(Zavaletta et al., 2007)	Normal, Vidro Despolido, Favo de Mel, Reticular, Enfisema	Assinaturas dos Histogramas Adaptativos	Sujeitos: 14 VOIS: 909, 15x15x15	EMD entre assinaturas canónicas de cada classe	(**Sensibilidade**, **Especificidade**) Normal (92%, 95%) Vidro Despolido (75%, 89%) Reticular (22%,92%) Favo de Mel (74%,91%), Enfisema (94%, 98%)

Tabela 2.1: Resumo das características dos trabalhos descritos. (Continuação).

Estudo	**Padrões/Patologia**	**Características**	**Base de Dados**	**Classificação**	**Desempenho**
(Zrimec et al., 2007)	Normal *versus* Favo de Mel	FOS, SGLDM	Sujeitos: 8 Imagens: 42 ROIs: 18 407, 7x7 e 15x15	Bayes	**Exatidão**: 87.2% **Sensibilidade**: 97.4% **Especificidade**: 85.5%
(Kim et al., 2009)	Normal, Enfisema Centrolobular Severo, Enfisema Centrolobular Médio, Bronquiolite Obliterante	FOS, Gradiente, SGLGM, GLRLM, Transformadas *top-hat*, Forma	Sujeitos: 82 Imagens: 265 ROIs: 64x64	SVM	**Sensibilidade**: 93.5± 1.0%
(Lee et al., 2009)	Normal, Enfisema Centrolobular Severo, Enfisema Centrolobular Médio, Bronquiolite Obliterante	FOS, Gradiente, SGLGM, GLRLM	Sujeitos: 92 Imagens: 265 ROIs: 236, 64x64	SVM	**Exatidão**: 83.1± 1.5%
(Sørensen et al., 2010	Normal, Enfisema Centrolobular, Enfisema Paraseptal	Histograma LBP Histograma de intensidade	Sujeitos: 39 Imagens: 117 ROIS: 168, 31x31	*k*-NN	**Exatidão**: 95.2%
(Gangeh et al., 2010)	Normal, Enfisema Centrolobular, Enfisema Paraseptal	Histograma de *Textons* de Intensidade	Sujeitos: 25 Imagens: 75 ROIS: 168, 50x50	SVM	**Exatidão**: 96.43%
(Korfiatis et al., 2010)	Normal, Vidro Despolido, Reticular	SGLDM (3D)	Sujeitos: 13 VOIs: 21x21x21	*k*-NN	(**FVP**, **FFP**, **VO**) Vidro Despolido (0.6, 0.5, 0.7); Reticular (0.9, 0.2, 0.8)

Tabela 2.1: Resumo das características dos trabalhos descritos. (Continuação).

Estudo	**Padrões/Patologia**	**Características**	**Base de Dados**	**Classificação**	**Desempenho**
(Mariolis et al., 2010)	Normal, Vidro Despolido, Reticular, Favo de Mel	GLRLM (3D), *Laws* (3D)	Sujeitos: 30 VOIs : 1 173, dimensão variável	PNN, *k*-NN	**Exatidão**: 99.8%
(Vo et al., 2010)	Normal, Enfisema, Favo de Mel, Vidro Despolido	*Wavelet*, *Contourlet*, Filtros derivativos de Gauss	Sujeitos:38 Imagens: 89 ROIs: 73 000, 32x32	MKL - SVM	**Sensibilidade**: 94.2% **Especificidade**: 98.7%
(Depeursinge et al., 2012a)	Normal, Enfisema, Vidro Despolido, Fibrose, Micronódulos	Histograma de intensidades, Número de pixels de ar, Q*uincunx Wavelet Frames*	Sujeitos: 85 ROIs:17 848, 32x32	SVM	**Exatidão**:76.9%
(Xu et al, 2013)	Normal, Consolidação, Vidro despolido, Favo de Mel, Enfisema, Nodular	Histograma de *textons* de intensidade e forma	Sujeitos: 117 VOI:3 009, 32x32x32	SVM	**Exatidão**: 93.18%
(Zhao et al., 2013)	Normal, Consolidação, Vidro despolido, Favo de Mel, Enfisema, Nodular	Descritores esparsos, obtidos a partir dos momentos do histograma de intensidade e forma	Sujeitos: 117 VOI:2 360, 32x32x32	SVM	**Exatidão**: 95.4%
(Song et al., 2013)	Normal, Enfisema, Vidro Despolido, Fibrose, Micronódulos	RIG-LBP, Histograma de níveis de cinzento, Histograma dos gradientes orientados em sistema de multicoordenadas	Sujeitos: 95 Imagens:112 ROIS: 3 731, 31x31	Critério da Discrepância Mínima	***F-score*** Normal: 84.0%; Enfisema: 75.3% Vidro Despolido:78.2%, Fibrose: 84.1% Micronódulos: 85.7%

Tabela 2.1: Resumo das características dos trabalhos descritos. (Continuação).

Estudo	**Padrões/Patologia**	**Características**	**Base de Dados**	**Classificação**	**Desempenho**
(Zhang et al., 2013)	Normal, Enfisema Centrolobular, Enfisema Paraseptal, Enfisema Panlobular	Descritores esparsos obtidos a partir do histograma de intensidades	Sujeitos: 14 ROIs: 470, 40x40	*k*-NN	**Exatidão**: 88%
(Anthimopoulos et al. 2014)	Normal, Vidro Despolido, Favo de mel, Consolidação Reticulação, Reticulação/ Vidro Despolido	Quantis do histograma acumulado das imagens filtradas com DCT, Histograma de intensidades	Sujeitos: 113 ROIs: 2503, 21x21	RF	***F-score*** = 89%
(Hamzah et al., 2017)	DIP *versus* Patologia não DIP DIP *versus* Normal	Energia, Energia média, Entropia, Contraste, Homogeneidade, Desvio Padrão, Desvio Padrão da Energia	Sujeitos: 45 Imagens: 225	*k*-NN	**Exatidão**: 73% **Exatidão**: 67%

Capítulo 3
PATOLOGIAS PULMONARES EM TC DE ALTA RESOLUÇÃO

Atualmente a TCAR constitui a técnica imagiológica principal no estudo de doenças que afetam o parênquima pulmonar, como o enfisema, a DIP, e as estruturas respiratórias de pequena dimensão. Neste capítulo são referidos os protocolos de TC atualmente utilizados na aquisição de imagens de Alta Resolução e descritos os padrões imagiológicos associados ao enfisema pulmonar e à DIP estudados neste trabalho.

3.1 TC de Alta Resolução do Tórax

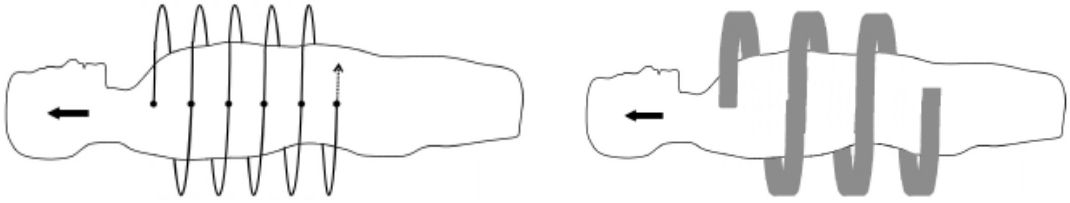
A radiografia e a TC do tórax constituem as técnicas radiográficas mais importantes no estudo imagiológico dos pulmões, sendo complementadas, em situações concretas pela angiografia, RM e ecografia. A radiografia do tórax é habitualmente a primeira técnica imagiológica prescrita na abordagem clínica das doenças do aparelho respiratório, uma vez que é uma técnica de fácil acesso, baixo custo e baixa dose de radiação. Apesar da interpretação das imagens radiológicas constituírem uma fonte de informação importante, na maioria das situações não conduz a um diagnóstico definitivo. Dado o baixo contraste e a sobreposição de órgãos e estruturas, inerente à técnica de aquisição da radiografia, a ausência de alterações numa radiografia não garante a ausência de patologia. Por outro lado, a presença de lesões não significa que estas se encontrem ativas. A radiografia do tórax é considerada normal em 10% a 15% de pacientes com DIP e apenas fornece um diagnóstico correto em 23% das doenças pulmonares em geral (Stark, 2008; Depeursinge, 2010). Os testes pulmonares funcionais permitem detetar alterações globais, no entanto, possuem falta de sensibilidade para uma deteção precoce, não permitindo isolar as regiões afetadas nem estabelecer uma relação entre os resultados e as alterações nas estruturas abrangidas.

O desenvolvimento da TC permitiu ultrapassar as limitações da radiografia dita convencional, evitando a sobreposição dos órgãos e estruturas, sendo atualmente a principal técnica imagiológica na deteção e avaliação de patologias pulmonares. No caso de doenças que afetam o interstício pulmonar como a DIP, ou as pequenas vias respiratórias a TCAR é considerada o *gold standard*, especialmente nas primeiras fases da doença. Esta técnica possui um enorme potencial na análise da distribuição, quantificação e caracterização de regiões afetadas com DIP, de uma forma não invasiva. As imagens TCAR dos pacientes afetados com DIP possuem padrões específicos cuja distribuição e análise do seu conteúdo visual são particularmente relevantes na elaboração de um diagnóstico diferencial entre as diversas patologias englobadas na DIP (Sluimer et al., 2003; Kazerooni et al. , 2014).

A TCAR consiste na utilização de imagens de TC de espessura fina (entre 0.625 - 1.5 mm), recorrendo a algoritmos de reconstrução de alta frequência espacial, como por exemplo, o *bone algorithm* para o estudo pulmonar. Nas imagens TCAR é possível discriminar estruturas submilimétricas ao nível do lóbulo pulmonar secundário, o que assume especial importância no estudo de doenças que afetam o interstício pulmonar, assim com em alterações das vias aéreas e vasculares (Kazerooni et al, 2014).

A partir dos anos noventa do século XX surgiram os tomógrafos helicoidais que realizam uma aquisição volumétrica. Os primeiros *scanners* possuíam apenas uma linha de detetores tendo evoluído ao longo dos anos para várias linhas de detetores justapostas, permitindo a aquisição de múltiplas secções em simultâneo. Uma representação esquemática da aquisição helicoidal volumétrica, que associa a rotação constante do *gantry* ao movimento longitudinal da mesa, encontra-se na Figura 3.1 (Silva, 2005). Atualmente, o Aquilion One™ (Toshiba Medical Systems)[3] possui 320 linhas de detetores e permite a aquisição de 16 mm de espessura em apenas 275 milissegundos. A existência de múltiplas linhas de detetores possibilita a aquisição de um maior volume de dados por intervalo de tempo, reduzindo os artefactos inerentes ao movimento do paciente, nomeadamente ao movimento respiratório. Os atuais *scanners* helicoidais multidetetores ou multicorte possibilitam a aquisição de dados quase isotrópicos de todo o tórax num único suster de respiração.

[3] http://www.toshibamedicalsystems.com , consultado em julho de 2016.



a) b)

Figura 3.1: Aquisição helicoidal volumétrica. a) Com uma linha de detetores. b) Com várias linhas de detetores.

Duas abordagens genéricas estão atualmente disponíveis para a aquisição de imagens de TCAR em pacientes com suspeita ou com diagnóstico confirmado de DIP (Pisco, 2009; Salzer, 2012; Prosch, Schaefer-Prokop, Eisenhuber, Kienzl, & Herold, 2013; Kazerooni, 2014):

- A mais tradicional corresponde à aquisição sequencial de imagens axiais de TCAR longitudinalmente ao longo do tórax espaçadas de 10 mm a 20 mm, designada por TCAR axial, incremental ou descontínua. Esta abordagem teve início nos anos 80 do século XX, quando a tecnologia não permitia a aquisição de forma contínua de toda a extensão do pulmão;

- A segunda abordagem designa-se por TCAR volumétrica ou helicoidal e resulta da possibilidade de aquisição de dados volumétricos nos atuais tomógrafos helicoidais multicorte (Figura 3.1). Associando o movimento contínuo da *gantry* e da mesa, onde o paciente se encontra deitado, os dados são adquiridos ao longo de uma trajetória helicoidal desenhada em torno do paciente. As secções são adquiridas sem compassos de espera, permitindo uma maior qualidade de reconstrução da imagem em múltiplos planos e não apenas no plano axial. Os dados obtidos podem ser reconstruídos em cortes bidimensionais ou recorrendo a outros tipos de algoritmos reconstrutivos 3D como o *Maximum Intensity Projection* (MIP) e *Minimum Intensity Projection* (mIP) que permitem a representação de estruturas com valores do voxel acima (MIP) ou abaixo (mIP) de um limiar selecionado previamente. Os algoritmos de reconstrução mostraram ser vantajosos no estudo de diferentes padrões, por exemplo, o MIP mostrou-se vantajoso no estudo de padrões nodulares e o mIP no estudo de alterações enfisematosas (Salzer, 2012; Prosch et al., 2013).

Em qualquer uma das abordagens descritas, o exame é usualmente realizado com o paciente na posição decúbito dorsal, em inspiração máxima. No entanto, há casos em que as imagens devem ser adquiridas na posição decúbito ventral para avaliação de atelectasias passivas que surgem em regiões posteriores do pulmão, que ocorrem devido à estase provocada pela posição decúbito dorsal. Também se pode recorrer a esta posição, em expiração máxima, para avaliar suspeitas de aprisionamento respiratório (Kazerooni et al, 2014).

As indicações para a realização de TCAR pulmonar incluem, mas não se limitam, às seguintes situações (Pisco, 2009; Kazerooni et al, 2014):

- Na avaliação das alterações pulmonares em pacientes com um exame radiológico do tórax normal, mas que apresentam sintomatologia e testes da função respiratória alterados;
- Na avaliação de suspeita de doenças das pequenas vias respiratórias;
- Avaliação de pacientes com DIP diagnosticada ou com suspeita clínica da doença que são avaliados de forma incompleta por exames de TC *standard* ou radiografia torácica;
- Na quantificação da extensão da DIP de forma a avaliar a eficiência dos tratamentos prescritos;
- Como guia na seleção do local para execução de biópsia pulmonar.

Verifica-se que existem opções diferentes nas técnicas usadas em pacientes com suspeita ou diagnóstico de patologias difusas do interstício. Um estudo realizado a membros da *European Society of Thoracic Imaging* (ESTI) (Prosch et al., 2013) permitiu conhecer as opções tomadas no estudo destas patologias pelos imagiologistas no que diz respeito à abordagem de aquisição de TC que adotam, axial ou volumétrica, no recurso a métodos de redução da dose de radiação, nos algoritmos de reconstrução utilizados, se procedem à aquisição de imagens em expiração ou em posição decúbito ventral, entre outros aspetos. Apesar de muitas das aquisições realizadas com os *scanners* de TC serem automáticas, alguns parâmetros permanecem dependentes do técnico, sendo importante a transmissão do conhecimento entre imagiologistas. Através de inquéritos enviados a 173 membros da ESTI, com uma taxa de resposta de 37%, foi possível perceber quais os protocolos de TCAR atualmente usados na avaliação da DIP. Das respostas obtidas 77% recorrem aos protocolos de rotina. Relativamente ao tipo de aquisição,

15% indicam que recorrem apenas à TCAR axial, 63% das respostas avaliam os pacientes exclusivamente recorrendo à TCAR volumétrica e 22% das respostas indicam a aquisição de TCAR axial e volumétrico usadas de forma complementar. Quando a opção é a conjugação dos dois protocolos, 13% das respostas indicaram que a aquisição volumétrica efetuada é de baixa dose. O número de reconstruções efectuadas nos exames volumétricos, varia de dois a oito, excluíndo as reconstruções MIP e mIP. Todas as respostas indicam que é reconstruída uma série de imagens axiais otimizada para a visualização do parênquima pulmonar, combinada com uma segunda série axial otimizada para o mediastino, à excepção de duas das respostas. A terceira série de reconstrução mais frequente é a coronal numa janela pulmonar, 65% das respostas, seguida da coronal na janela do mediastino, em 41% das respostas. A reconstrução volumétrica MIP é obtida sempre em 33% das respostas e em 49% apenas se houver uma indicação específica. No caso da reconstrução mIP as respostas indicam os valores de 41% se houver indicação nesse sentido e apenas 10% das respostas indicam que a série é obtida sempre. O estudo realçou também a preocupação dos profissionais em adaptar os protocolos à idade e/ou peso do paciente (45% das respostas) e a efetuar os exames com modelação da dose de radiação automática (90% das respostas). A indicação da aquisição em modo expiratório e na posição decúbito ventral, sem uma requisição explícita, são efetuados em 58% e 59%, respetivamente. Como se verifica pelas respostas há preferência pela aquisição volumétrica e preocupação com os fatores de exposição do paciente à radiação.

A utilização plena das potencialidades da TCAR implica a familiarização com as vantagens e desvantagens de cada uma das abordagens. A escolha deve ser função do equipamento disponível, indicações clínicas e doses de radiação aplicadas. Apesar das vantagens do TC volumétrico, na avaliação de pacientes com suspeita ou diagnóstico de DIP, a sua utilização deve ser sempre equacionada devido à dose de radiação envolvida e à qualidade das imagens. O princípio ALARA ("*as low as reasonably achievable*") deve ser sempre norteador das decisões. O recurso à TCAR axial, em detrimento da aquisição volumétrica, permite reduzir a dose de radiação até 80% (Bendaoud et al., 2011, referido em Prosch et al., 2013). É possível recorrer a protocolos de TC volumétricos que permitem reduzir a dose de radiação por secção, mas inevitavelmente, a relação sinal/ruído diminui podendo comprometer a qualidade da imagem. Por outro lado, verifica-se que na rotina clínica a análise dos padrões intersticiais se efetua, tradicionalmente, em cortes axiais e não em cortes sagitais e/ou coronais ou em visualizações 3D possibilitadas pela aquisição volumétrica (Prosch et al., 2013).

Usualmente, os exames de TC são constituídos por séries de imagens armazenadas de acordo com a norma DICOM resultantes da aplicação de algoritmos de reconstrução. DICOM é uma norma para a manipulação, armazenamento, impressão e transmissão digital de imagens médicas. Para além da informação da imagem possui informação relativa ao protocolo de aquisição do exame, informação sobre o equipamento, dados do paciente, entre outros.

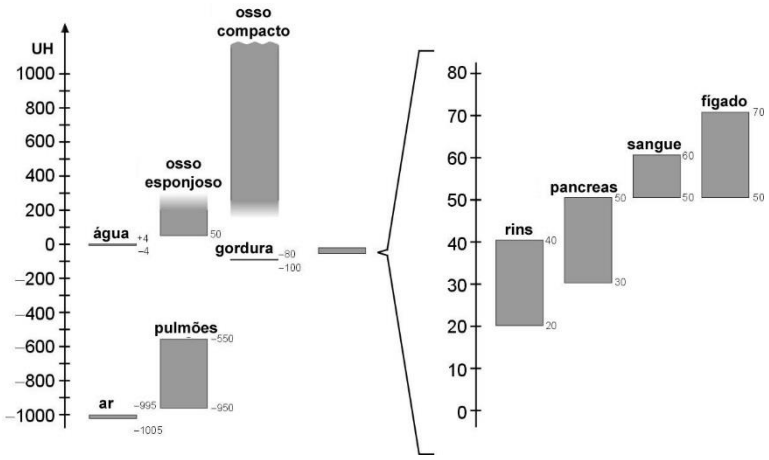


Figura 3.2: Escala de Hounsfield.

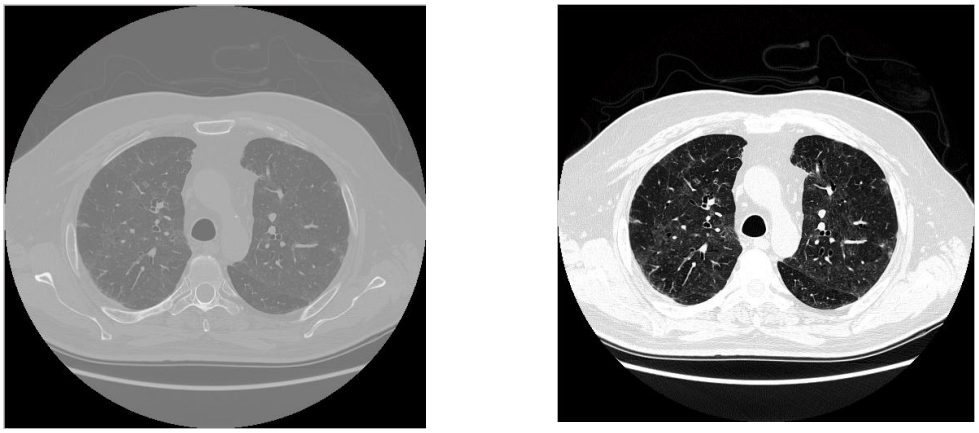
Nas imagens TC, cada pixel é a representação bidimensional de um voxel, cuja terceira dimensão corresponde à espessura do corte especificada na aquisição do exame. A dimensão de cada pixel traduz a razão entre o *field of view* e a matriz de computação. O seu nível de cinzento está relacionado com a atenuação aos raios X, sendo expressa em UH. O valor de Hounsfield de um pixel corresponde ao coeficiente de atenuação linear médio do material biológico que constitui cada voxel relativamente ao coeficiente de atenuação linear da água (Silva, 2005). Um tecido com coeficiente de atenuação linear μ_T corresponde em valores de Hounsfield a:

$$valor\ de\ Hounsfield = \frac{\mu_T - \mu_{água}}{\mu_{água}} \text{ x } 1000 \text{ UH} . \qquad (3.1)$$

Analisando a escala de Hounsfield, verifica-se que apenas os pulmões e a gordura possuem valores de Hounsfield negativos, uma vez que a sua densidade é inferior à da água. Os restantes órgãos e tecidos do corpo humano correspondem a valores positivos, como por exemplo, os ossos e a maioria dos órgãos moles. Nesta escala de níveis de cinzento alargada o valor -1000 UH corresponde ao ar e 0 UH à água, como se pode verificar na Figura 3.2 (Silva, 2005).

As imagens atualmente adquiridas são codificadas com 16 bits por pixel o que que corresponde a $2^{16} =$ 65 536 valores distintos. Uma vez que o sistema de visão humano só permite discriminar entre 20 a 30 níveis de cinzento diferentes, as imagens obtidas são manipuladas de forma a reduzir a escala de Hounsfield e criar uma janela de visualização que permita realçar as estruturas em estudo. Esta janela de visualização caracteriza-se por dois parâmetros: a largura e o centro. Dependendo da região em estudo (mediastino, parênquima pulmonar, osso ou fígado) são aplicados diferentes parâmetros largura/centro.

A largura restringe a gama de valores em uso, centrados no valor médio das estruturas que se pretendem evidenciar, que é dado pelo centro da janela. Todos os valores de Hounsfield acima do valor superior da largura da janela ficam a "branco" e os valores abaixo do valor inferior da largura da janela ficam a "preto" (Pisco, 2009). Esta informação está disponível no cabeçalho dos ficheiros DICOM.



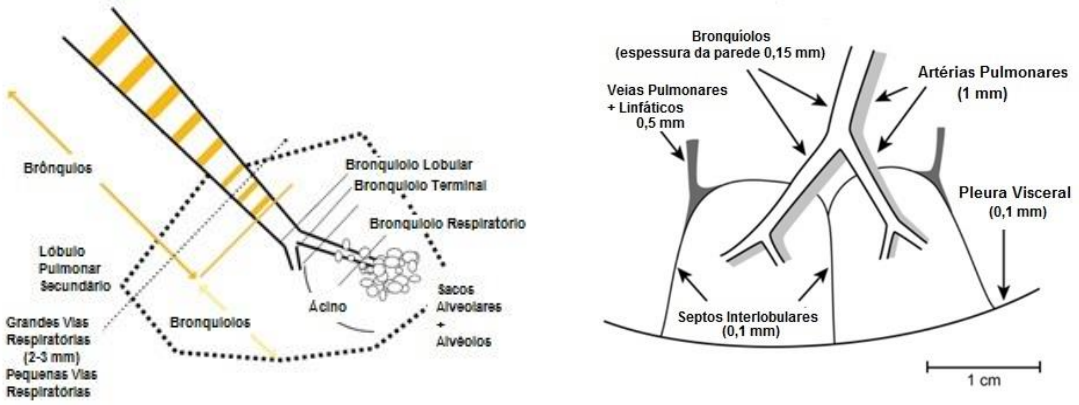
a) b)

Figura 3.3: Imagem de TC do tórax. a) Como se encontra guardada no ficheiro DICOM. b) Após a aplicação da janela pulmonar.

A importância da escolha da janela correta de visualização pode ser observada na Figura 3.3. Do lado esquerdo pode observar-se uma imagem TC visualizada diretamente a partir do ficheiro DICOM e no lado direito a imagem TC após aplicação da janela pulmonar, com os parâmetros: [largura; centro] = [1500; -700]. Pode-se verificar o destaque dado à zona do parênquima pulmonar em detrimento de outras estruturas ou órgãos.

3.2 Estrutura Pulmonar ao Nível do Lóbulo Pulmonar Secundário

O conhecimento da anatomia pulmonar é essencial para perceber as características da TC de patologias pulmonares não só associado à sua aparência, mas também à distribuição das zonas afetadas. Os lóbulos pulmonares secundários ou lóbulos secundários de Miller são as unidades anatómicas e fisiológicas fundamentais do pulmão. Correspondem à periferia da árvore brônquica, como se encontra esquematizado na Figura 3.4 a), adaptada de (Verschakelen et al., 2008). Possuem a forma de um poliedro e medem entre 1 a 2.5 cm, sendo delimitados pelos septos interlobulares onde se encontram as veias pulmonares e vasos linfáticos. No seu centro encontram-se o bronquíolo e a artéria pulmonar acompanhante que têm o diâmetro de cerca de 1 mm. Cada lóbulo pulmonar secundário é constituído por 3-12 ácinos pulmonares, sendo cada um deles composto por: bronquíolo respiratório, alvéolos e sacos alveolares, como se pode ver no esquema da Figura 3.4 a), adaptada de (Verschakelen et al., 2008). Como se pode observar na Figura 3.4 b) a maioria destas estruturas possuem uma espessura inferior a 1 mm.



a) b)

Figura 3.4: Estrutura pulmonar. a) Representação anatómica da periferia da árvore brônquica. b) Anatomia de um lóbulo pulmonar secundário sem patologia.

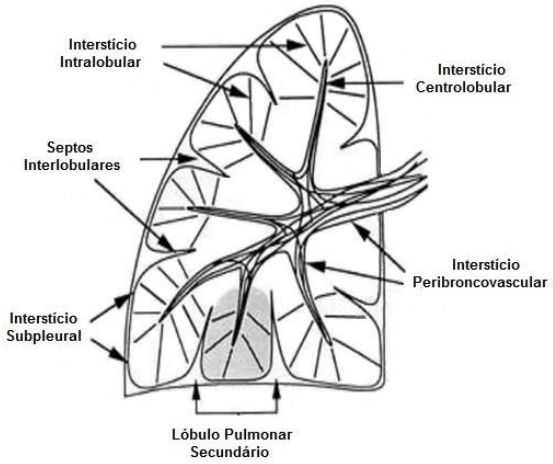


Figura 3.5: Lóbulo pulmonar secundário e identificação dos tipos de tecidos intersticiais e sua localização[4].

Resumidamente, podemos dizer que o pulmão é constituído por um tecido de suporte designado por interstício pulmonar, que é constituído essencialmente por fibras elásticas e de colagénio. Este tecido conetivo serve de base de sustentação às artérias, veias, vasos linfáticos, entre outros. O interstício pulmonar pode ser dividido em três setores que comunicam entre si, todos presentes no lóbulo pulmonar secundário, identificados na Figura 3.5 (Verschakelen et al., 2008; Pisco, 2009):

- Tecido conetivo axial – Corresponde ao conjunto de fibras que envolvem as estruturas broncovasculares, tendo início no hilo e terminando no lóbulo pulmonar secundário. Designa-se por interstício centrolobular e interstício peribroncovascular;

- Tecido conetivo periférico – Nele situam-se as veias e os vasos linfáticos. Localiza-se na periferia e engloba o interstício subpleural, o interstício interlobular e os septos entre ácinos;

- Tecido intralobular – Na região distal ao bronquíolo, os alvéolos e a rede de vasos capilares estão em contacto para permitir a difusão gasosa. Para além deles, estão também presentes fibras e colagénio que fazem parte do parênquima. Estas fibras sustentam alvéolos, vénulas, capilares, entre outros (Pisco, 2009). Este tecido intersticial estabelece a ligação entre o tecido conetivo axial e periférico. Note-se que muitas vezes a designação

[4] https://www.slideshare.net/fernferretie/pulmonary-interstitium , consultado em julho de 2016.

parênquima pulmonar é utilizada de uma forma mais ampla referindo-se ao tecido pulmonar em geral.

Quando ocorrem situações patológicas o tecido intersticial pulmonar reage de forma mais ou menos monótona, sendo possível distinguir quatro tipo de padrões básicos em imagens de TCAR (Pisco, 2009):

- Padrão intersticial;
- Padrão alveolar;
- Padrão broncovascular;
- Padrão cavitário.

Os padrões imagiológicos abordados nesta tese são do tipo intersticial. Em TCAR os padrões intersticiais podem ser agrupados em quatro grandes grupos, de acordo com a sua aparência (Pisco, 2009):

- Opacidades lineares e reticulares - Resultam do espessamento anormal de um ou vários setores do interstício por líquido, fibrose, processos inflamatórios ou tumorais. Podem ocorrer associadas ao edema pulmonar, sarcoidose ou fibrose;
- Opacidades nodulares - Surgem na forma de nódulos intersticiais bem definidos e usualmente periféricos, associados frequentemente à sarcoidose, tuberculose, linfangite e silicose. Outra forma de opacidade nodular caracteriza-se por massas de fibrose, encontradas frequentemente em pacientes com silicose ou sarcoidose em fase terminal;
- Aumento da densidade pulmonar - Estes padrões ocorrem quando há aumento de densidade do parênquima pulmonar. Podem surgir na forma de opacidade em vidro despolido ou consolidação do espaço aéreo;
- Diminuição da densidade pulmonar - Neste tipo de padrões a densidade do parênquima pulmonar encontra-se diminuída. Podem tomar a forma de quistos pulmonares, enfisema ou favo de mel.

3.3 Enfisema Pulmonar

O enfisema pulmonar caracteriza-se por um alargamento anormal e permanente dos espaços aéreos distais do bronquíolo terminal, acompanhado pela destruição das paredes dos espaços aéreos envolvidos (Verschakelen et al., 2007; Silva, Marchiori, Júnior, & Müller, 2010). O principal fator de risco do enfisema é o tabagismo, embora haja outros fatores que contribuem para o seu desenvolvimento, como condições genéticas e a qualidade do ar envolvente. Caso o paciente se mantenha em contacto com o fator de risco a doença torna-se progressiva e irreversível, incapacitando lenta mas inexoravelmente o doente, tornando-o num insuficiente respiratório. Clinicamente, a dispneia é o sintoma principal. O padrão enfisema surge também como achado imagiológico associado à DIP por exemplo na pneumonia intersticial descamativa ou na pneumonite de hipersensibilidade (Depeursinge, 2010).

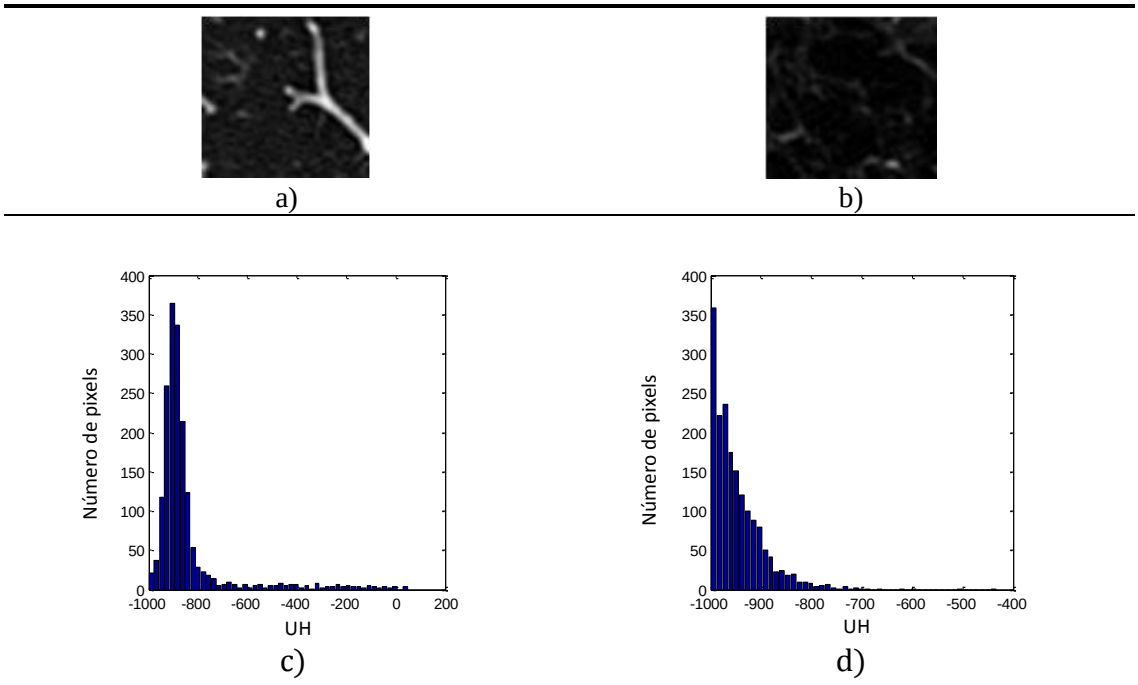


Figura 3.6: Regiões do parênquima pulmonar extraídas de imagens de TCAR e respetivos histogramas. a) e c) Região normal. b) e d) Região com enfisema centrolobular.

O diagnóstico do enfisema pulmonar em imagens de TCAR do tórax é baseado na deteção de regiões focais de atenuação muito baixa, muito próximas do ar, usualmente sem paredes visíveis, que contrastam com o pulmão saudável envolvente. Nessas imagens as regiões de enfisema surgem como zonas muito escuras dentro do pulmão, com áreas e distribuições distintas. Na Figura 3.6 estão representadas duas regiões: normal e com enfisema centrolobular

severo, e os respetivos histogramas. Como se pode verificar o histograma da região enfisematosa encontra-se deslocado para a esquerda, com valores UH próximos do valor do ar (-1000 UH) comparativamente ao histograma da região normal.

A aparência visual das imagens de TCAR de pacientes com enfisema depende do tipo de enfisema e do estádio da doença. Existem essencialmente quatro tipos de enfisemas: enfisema centrolobular, enfisema paraseptal, enfisema panlobular e enfisema bolhoso (Verschakelen et al., 2008; Pisco, 2009). Esta classificação baseia-se na aparência e na distribuição do padrão. O enfisema centrolobular corresponde a áreas centrolobulares com atenuação diminuída, usualmente sem paredes visíveis que ocorrem predominantemente no lobo superior dos pulmões e com uma distribuição não uniforme. É frequente as artérias centrolobulares serem visíveis, como pontos brilhantes, no interior de áreas hipodensas. É o enfisema mais comum em pacientes com hábitos tabágicos (Silva et al., 2010). O enfisema paraseptal caracteriza-se por múltiplas áreas de baixa atenuação nas regiões subpleurais numa única camada, frequentemente envoltas pelos septos interlobulares que surgem visíveis como paredes finas regulares brancas.

Por seu lado, o enfisema panlobular afeta todo o lóbulo pulmonar e manifesta-se como uma diminuição generalizada da atenuação do pulmão, com uma redução do número e diâmetro dos vasos sanguíneos nas áreas afetadas, com ou sem a sua distorção. É o tipo de enfisema tipicamente associado à deficiência da proteína alfa-1 antitripsina no sangue, apesar de poder também ser encontrado em enfisemas severos associados ao tabagismo (Zhang et al., 2013). No caso do enfisema bolhoso ocorrem bolhas circunscritas por paredes finas, que correspondem a zonas de baixa atenuação, cuja dimensão pode variar de um a vários centímetros de diâmetro. Tem predominância subpleural e é assimétrico. Ocorre frequentemente associado ao enfisema centrolobular e paraseptal. A Figura 3.7 mostra imagens de TCAR em corte axial dos diferentes tipos de enfisema pulmonar descritos.

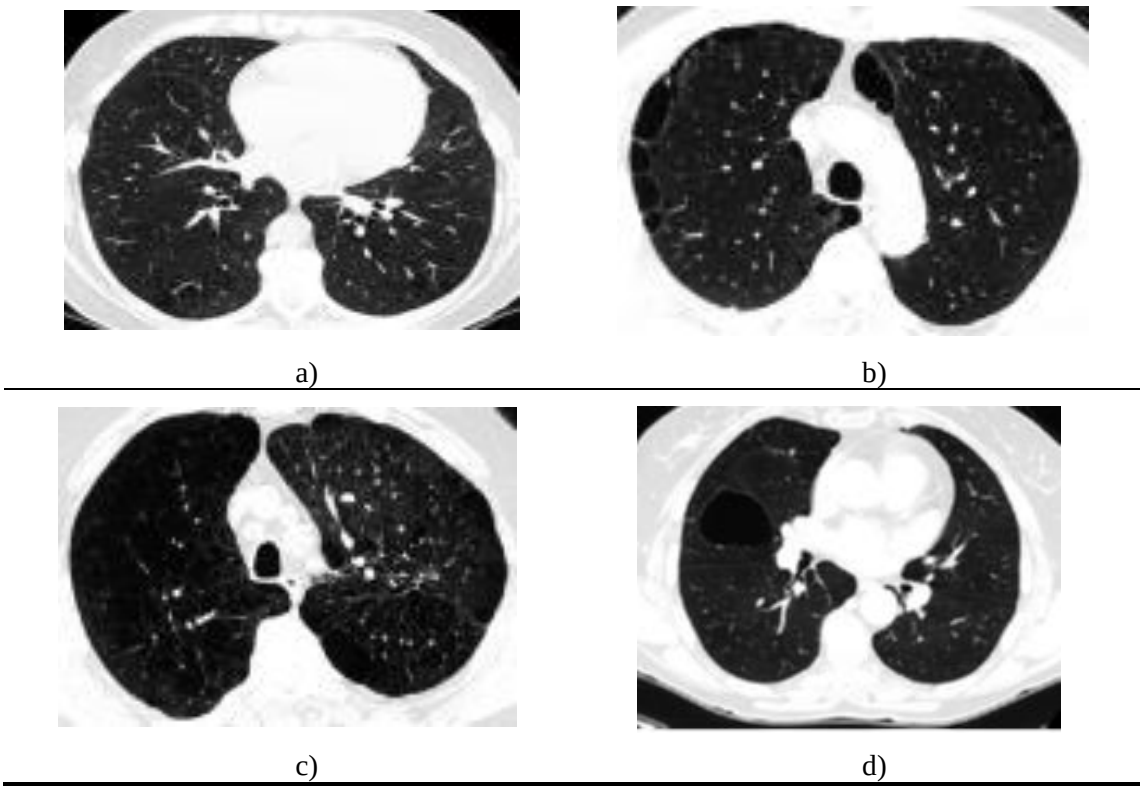


Figura 3.7: Imagens de TCAR do tórax de pacientes com enfisema. a) Enfisema centrolobular. b) Enfisema paraseptal. c) Enfisema panlobular. d) Enfisema bolhoso.

3.4 Doença Intersticial Pulmonar

O padrão intersticial encontra-se associado à DIP. Esta engloba um grupo numeroso e diversificado de alterações agudas e crónicas que afetam essencialmente o interstício pulmonar, correspondendo a 15% das patologias pulmonares (Hamzah et al., 2017). No entanto, frequentemente, os espaços alveolares são também envolvidos o que leva a que a designação doença infiltrativa difusa seja também bastante comum. O grupo de patologias que se enquadram na DIP afetam as mesmas regiões pulmonares e possuem características clínicas, radiológicas e testes funcionais semelhantes, o que justifica uma designação única e revela também a dificuldade que se coloca no seu diagnóstico diferencial.

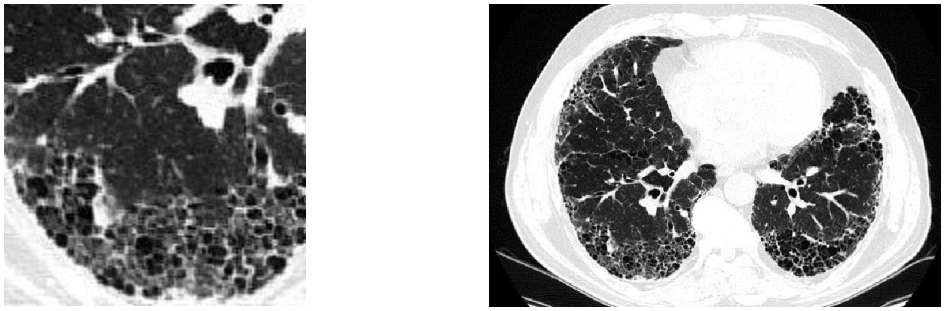
A DIP pode ser uma doença pulmonar primária ou uma manifestação pulmonar de uma doença primária com outra localização. A sua avaliação é multidisciplinar colocando desafios ao clínico, ao imagiologista e ao anatomopatologista. O seu diagnóstico e subsequente orientação terapêutica envolvem exames clínicos, testes laboratoriais, estudos fisiológicos, exames

imagiológicos, broncofibroscopia com lavado, surgindo por vezes a necessidade de realização de biópsia e exame anatomopatológico (Pisco, 2009).

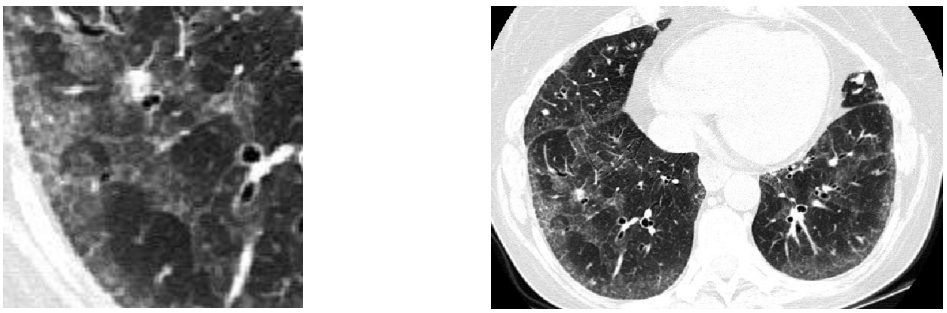
Os sintomas mais comuns dos pacientes com DIP são a dispneia e a tosse seca, apresentando ralas e fervores crepitantes aquando do exame físico. Verifica-se a alteração das trocas gasosas com hipoxemia (pressão parcial anormalmente baixa de oxigénio no sangue arterial) sendo os testes da função pulmonar do tipo restritivo, podendo ocorrer diminuição do volume pulmonar. A DIP reúne patologias de etiologias múltiplas e muitas vezes idiopáticas. Existem mais de 150 causas de DIP, sendo 10% a 15% destas, responsáveis por 90% das doenças. Exemplos de DIP são: sarcoidose, pneumonia intersticial aguda, pneumonia intersticial linfócita, fibrose pulmonar idiopática, pneumoconioses, entre muitas outras (Pisco, 2009; Depeursinge, 2010; Raghu et al., 2011).

O diagnóstico imagiológico da DIP é difícil dada a diversidade dos padrões intersticiais que podem estar presentes, que podem surgir associados, à subtileza das suas manifestações, ao facto do mesmo padrão poder estar associado a várias patologias e poder apresentar características visuais diferentes com a evolução da doença (Verschakelen et al., 2008). Em (Depeursinge, 2010) encontram-se listadas as patologias mais comuns que integram a DIP, assim como, os padrões que lhes podem estar associados em imagens de TCAR. Na Figura 3.8 é mostrado o aspeto visual, em corte axial de imagens TCAR do tórax, dos padrões: normal, vidro despolido e favo de mel.

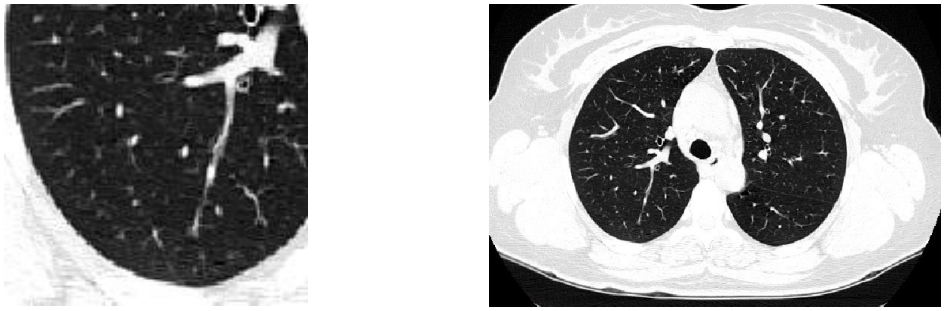
O tecido intersticial não é visível em imagens TCAR, no entanto na presença de DIP este tecido conetivo assume um conjunto de padrões imagiológicos característicos. O padrão normal caracteriza-se por uma densidade pulmonar ligeiramente superior ao ar, determinada pelas estruturas pulmonares como o tecido intersticial, paredes das pequenas vias áreas, as veias e o sangue. O padrão favo de mel caracteriza-se pela presença de múltiplas lesões quísticas dispostas em várias camadas localizadas geralmente na região subpleural. Como os quistos surgem associados em *clusters* têm a aparência de favos de mel. Os quistos possuem diâmetros que podem variar de um a vários centímetros e partilham paredes espessas irregulares bem definidas. Este padrão encontra-se associado a fases avançadas de fibrose pulmonar.



a) Favo de mel



b) Vidro despolido



c) Normal

Figura 3.8: Padrões do tecido pulmonar associados à DIP em imagens TCAR do tórax. Na coluna da esquerda uma vista de pormenor do padrão e à direita a respetiva imagem. a) Favo de mel; b) Vidro despolido; c) Normal.

O padrão opacidade em vidro despolido corresponde a um aumento da densidade do parênquima pulmonar, em que as estruturas broncovasculares continuam visíveis. É causado pelo espessamento das paredes alveolares e colapso do lúmen alveolar, doença alveolar pura ou combinação das anteriores. Este padrão pode surgir difuso ou ocupar de modo homogéneo todo o pulmão, podendo surgir sozinho ou associado a outros padrões pulmonares. É um padrão não específico, que se encontra associado à grande maioria das DIP, no entanto em 80% dos casos corresponde a doença ativa associada a pneumonia intersticial, sarcoidose ou fibrose pulmonar (Pisco, 2009; Depeursinge, 2010).

3.5 Considerações Finais

Ao longo deste capítulo foi descrita a técnica imagiológica baseada em imagem usada no estudo de patologias pulmonares: a TC e o protocolo de Alta Resolução, considerado o *gold standard* na deteção, quantificação e caracterização de patologias que afetam o interstício pulmonar. São descritas as estruturas anatómicas pulmonares afetadas pela DIP, assim como os padrões imagiológicos associados, que foram objeto de estudo ao longo do trabalho descrito na tese.

Capítulo 4
EXTRAÇÃO E SELEÇÃO DE CARACTERÍSTICAS

A escolha das características apropriadas para os requisitos de um problema é uma etapa fundamental em tarefas de aprendizagem automática, como a classificação de padrões. O objetivo é reunir um conjunto de características que sejam uma representação compacta dos padrões que se pretendem classificar.

No contexto do sistema de DAC, e sendo a textura uma componente de grande importância na interpretação de imagens TCAR na presença de patologias pulmonares, os métodos de análise de textura devem permitir definir um conjunto de características que descrevam os diferentes padrões imagiológicos pulmonares. Dada a complexidade da sua textura a análise é efetuada recorrendo a vários métodos, nomeadamente à análise da lacunaridade diferencial. São expostas as principais linhas orientadoras dos métodos de seleção de características que permitem identificar as características do conjunto inicial que são relevantes, e que, portanto, devem ser mantidas, e as características irrelevantes e/ou redundantes que devem ser retiradas.

4.1 Estatística de Níveis de Cinzento de Primeira Ordem

A análise de textura com o recurso a Estatística de Primeira Ordem (FOS - *First-Order Statistics*) baseia-se no histograma de níveis de cinzento. Considerando n_p o número de pixels da imagem, L o número de níveis de cinzento da imagem e # o número de elementos do conjunto, o histograma normalizado de primeira ordem é dado pela seguinte distribuição de probabilidades (Gonzalez, & Woods, 2008):

$$h(i) = \#\{(x,y) \in I(x,y) = i\}/n_p, \qquad 0 \leq i < L \tag{4.1}$$

Nesta análise é apenas tido em consideração o valor de cada pixel individualmente, não existindo qualquer interação com o valor dos pixels vizinhos.

2	2	1	1	0
0	2	4	3	1
2	3	0	3	0
3	4	1	0	2
3	3	0	2	2

$I(x,y)$

0	1	2	3	4
6/25	4/25	7/25	6/25	2/25

$h(i), 0 \leq i \leq 4$

Figura 4.1: Exemplo do cálculo do histograma normalizado da imagem *I*(*x*,*y*).

A partir do histograma normalizado é possível extrair um conjunto de medidas:

- Média (μ): é a média dos níveis de cinzento da imagem. É definida por:

$$\mu = \sum_{i=0}^{L-1} ih(i)\,. \tag{4.2}$$

- Variância (VAR): é uma medida de largura do histograma, mede a dispersão dos níveis de cinzento relativamente à média:

$$VAR = \sum_{i=0}^{L-1} h(i)(i-\mu)^2. \tag{4.3}$$

- Simetria (SIM): medida do grau de simetria da distribuição de probabilidades, dada por:

$$SIM = \frac{1}{\sigma^3}\sum_{i=0}^{L-1} h(i)(i-\mu)^3\,. \tag{4.4}$$

- Curtose (CUR): é uma medida do grau de achatamento da distribuição de probabilidades relativamente à distribuição normal:

$$CUR = \frac{1}{\sigma^4}\sum_{i=0}^{L-1} h(i)(i-\mu)^4 - 3\,. \tag{4.5}$$

- Energia (ENE): é uma medida da variação da intensidade na imagem. No caso de a imagem conter apenas um nível de cinzento a energia é máxima (*ENE* = 1). Por outro lado, se o histograma for perfeitamente uniforme, i.e., todas as probabilidades $h(i)$ são iguais o valor da energia é mínimo *ENE* = 1/*L*:

$$ENE = \sum_{i=0}^{L-1} h^2(i)\,. \tag{4.6}$$

- Entropia (ENT): mede a não uniformidade da imagem, quanto mais uniforme for a imagem menor será a entropia, $0 \leq ENT \leq \log(L)$:

$$ENT = -\sum_{i=0}^{L-1} h(i)\log(\,h(i))\,. \tag{4.7}$$

4.2 Método de Dependência Espacial de Níveis de Cinzento

O método de análise de textura proposto por Haralick (Haralick, 1979) descreve a dependência da distribuição de níveis de cinzento entre pixels vizinhos, que pode ser expressa através de matrizes de coocorrência. No Método de Dependência Espacial de Níveis de Cinzento (SGLDM - *Spatial Gray Level Dependence Method*) a função de densidade de probabilidade condicional de segunda ordem $p(i,j|d,\theta)$ pode ser estimada em várias direções θ e para várias distâncias entre pixels *d*. Assim, $p(i,j|d,\theta)$ é a probabilidade de dois pixels colineares segundo a direção θ e à distância *d* possuírem o nível de cinzento *i* e *j*. As medidas baseadas no SGLDM são também conhecidas como estatística de níveis de cinzento de segunda ordem, pois são analisados dois pixels de cada vez. Podem ser consideradas várias direções, sendo as mais habituais as correspondentes aos ângulos θ = {0°, 45°, 90° e 135°}. As funções $p(i,j|d,\theta)$ definem-se da seguinte forma, para os ângulos considerados:

$$\begin{aligned} p(i,j|d,0°) = \#\{((k,l),(m,n) \in (L_r \times L_c) \times (L_r \times L_c): \\ k = m, |l-n| = d, I(k,l) = i, I(m,n) = j\}/T(d,0°)\,; \end{aligned} \tag{4.8}$$

$$p(i,j|d,45°) = \#\{((k,l),(m,n) \in (L_r \times L_c) \times (L_r \times L_c): \\ (k-m=d, l-n=-d) \text{ ou } (k-m=-d, l-n=d), \\ I(k,l)=i, I(m,n)=j\}/T(d,45°)\ ; \tag{4.9}$$

$$p(i,j|d,90°) = \#\{((k,l),(m,n) \in (L_r \times L_c) \times (L_r \times L_c): \\ |k-m|=d, l=n, I(k,l)=i, I(m,n)=j\}/T(d,90°)\ ; \tag{4.10}$$

$$p(i,j|d,135°) = \#\{((k,l),(m,n) \in (L_r \times L_c) \times (L_r \times L_c): \\ (k-m=d, l-n=d) \text{ ou } (k-m=-d, l-n=-d), \\ I(k,l)=i, I(m,n)=j\}/T(d,135°)\ . \tag{4.11}$$

O # representa número de elementos do conjunto, $(L_r \times L_c)$ as coordenadas espaciais, $I(k, l)$ a intensidade do pixel (k, l) e $T(d, \theta)$ o número total de pares de pixels da imagem à distância d na direção θ. As funções $p(i,j|d,\theta)$ podem ser representadas na forma de matrizes, como descrito pela equação:

$$\Omega(d,\theta) = p(i,j|d,\theta),\ 0 \leq i,j < L. \tag{4.12}$$

As matrizes $\Omega(d,\theta)$ são quadradas $L \times L$, sendo L o nível de cinzento máximo da imagem e simétricas, pois $p(i,j|d,\theta) = p(j,i|d,\theta)$. A sua análise permite compreender a textura da região. Se a textura é "grosseira" e a distância d é pequena, comparativamente ao tamanho dos elementos da textura, os pares de pixels analisados usualmente possuem níveis de cinzento semelhantes. Isto significa que a distribuição de probabilidade na matriz de coocorrência está concentrada na diagonal principal ou perto dela. Por outro lado, no caso de texturas finas os níveis de cinzento dos pixels separados pela distância d devem ser bastante diferentes, portanto a distribuição de probabilidade encontra-se afastada da diagonal. A Figura 4.2 mostra o cálculo das matrizes $\Omega(d,\theta)$ segundo as direções 0º e 45º, para $d = 1$.

2	2	1	1	0
0	2	4	3	1
2	3	0	3	0
3	4	1	0	2
3	3	0	2	2

I(*x*, *y*)

	0	1	2	3	4
0	0	2/40	3/40	4/40	0
1	2/40	2/40	1/40	1/40	1/40
2	3/40	1/40	4/40	1/40	1/40
3	4/40	1/40	1/40	2/40	2/40
4	0	1/40	1/40	2/40	0

$\Omega(1,0°)$

	0	1	2	3	4
0	4/32	0	1/32	2/32	1/32
1	0	0	1/32	3/32	1/32
2	1/32	1/32	4/32	0	0
3	2/32	3/32	0	2/32	2/32
4	1/32	1/32	0	2/32	0

$\Omega(1,45°)$

Figura 4.2: Matrizes Ω(1,0º) e Ω(1,45º), correspondentes à imagem *I*(*x*,*y*).

As matrizes Ω(*d*,θ) constituem a base para o cálculo de diversas medidas estatísticas conhecidas como descritores de Haralick que permitem descrever propriedades das texturas (Haralick, 1979). Para cada par (*d*,θ) é calculada uma matriz Ω(*d*,θ) e um conjunto de seis descritores de textura são extraídos:

- Segundo Momento Angular (SMA) ou Energia: mede o grau de uniformidade. Um valor baixo do SMA é característico de texturas finas, sendo o valor máximo obtido para texturas constantes na janela de análise:

$$SMA = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} p^2(i,j)\,. \tag{4.13}$$

- Entropia (ENT): é uma medida de aleatoriedade:

$$ENT = -\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} p(i,j)\log\big(p(i,j)\big)\,. \tag{4.14}$$

- Momento Inverso da Diferença (MID): fornece uma medida de homogeneidade local:

$$MID = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\frac{p(i,j)}{1+(i-j)^2}. \tag{4.15}$$

- Correlação (COR): mede a dependência linear dos níveis de cinzento. Valores elevados de correlação indicam uma certa ordem local entre os níveis de cinzento:

$$COR = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}p(i,j)\left[\frac{\left(i-\mu_i\right)\left(j-\mu_j\right)}{\sigma_i\sigma_j}\right]. \tag{4.16}$$

Na expressão do cálculo da correlação, a média e desvio padrão segundo a coluna *i* são dados por:

$$\mu_i = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}p(i,j)i\ ; \tag{4.17}$$

$$\sigma_i = \sqrt{\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}p(i,j)\left(i-\mu_i\right)^2}. \tag{4.18}$$

Sendo os cálculos semelhantes para a média e desvio padrão segundo a coluna *j*.

- Variância (VAR): é uma medida de dispersão em torno da média. Na fórmula da Variância μ representa a média:

$$VAR = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}p(i,j)(i-\mu)^2. \tag{4.19}$$

- Contraste (CON): corresponde à variação de intensidade entre um pixel e os seus vizinhos ao longo da imagem. Se existir uma grande variação na região de análise as entradas da matriz $\Omega(d,\theta)$ vão concentrar-se fora da diagonal principal e o contraste será elevado:

$$CON = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}p(i,j)(i-j)^2. \tag{4.20}$$

4.3 Método de Comprimento de Primitivas de Níveis de Cinzento

O Método de Comprimento de Primitivas de Níveis de Cinzento (*Gray Level Run-Length Method* – GLRLM) foi originalmente descrito e aplicado por Galloway (Galloway, 1975). Nele a caracterização da textura baseia-se no cálculo de primitivas *run length*, que correspondem a conjuntos consecutivos de pixels colineares com o mesmo nível de cinzento, numa dada direção. As primitivas *run length* são caracterizadas pelo seu comprimento (número de pixels), direção e nível de cinzento. Numa primeira fase são calculadas as matrizes de primitivas $\Psi(\theta)$ a partir das quais é possível extrair descritores de textura:

$$\Psi(\theta) = M(a, r|\theta), 0 \leq a < L, 0 < r \leq N_r \, . \tag{4.21}$$

Cada elemento da matriz $M(a,r|\theta)$ representa o número de vezes que ocorrem na imagem primitivas com nível de cinzento a e comprimento r segundo a direção θ. A matriz $M(a,r|\theta)$ possui uma dimensão de $L \times Nr$, sendo L o número de níveis de cinzento e N_r o número de primitivas com comprimentos distintos.

2	2	1	1	0
0	2	4	3	1
2	3	0	3	0
3	4	1	0	2
3	3	0	2	2

$I(x, y)$

Comprimento

Níveis de cinzento	1	2	3	4	5
0	6	0	0	0	0
1	2	1	0	0	0
2	3	2	0	0	0
3	4	1	0	0	0
4	2	0	0	0	0

$\Psi(0^o)$

Comprimento

Níveis de cinzento	1	2	3	4	5
0	6	0	0	0	0
1	4	0	0	0	0
2	3	2	0	0	0
3	2	2	0	0	0
4	2	0	0	0	0

$\Psi(90^o)$

Figura 4.3: Matrizes de primitivas *run length* para as direções 0º e 90º, correspondentes à imagem $I(x,y)$.

O cálculo da matriz de primitivas, segundo as direções 0° e 90°, encontra-se ilustrado na Figura 4.3. Após o cálculo dessas matrizes é possível extrair um conjunto de medidas (descritores ou características) que captam as propriedades da textura (Galloway, 1975; Chu, Sehgal, & Greenleaf, 1975; Dasarathy & Holder, 1991). Galloway foi o primeiro autor a propor um conjunto de cinco medidas extraídas a partir das matrizes de primitivas.

- Número total de primitivas existentes na matriz, n_r, é um fator de normalização das restantes medidas:

$$n_r = \sum_{a=1}^{L} \sum_{r=1}^{N_r} M(a,r) \,. \tag{4.22}$$

- Ênfase em Primitivas Curtas (EPC): o valor de cada primitiva é dividido pelo quadrado do seu comprimento e normalizado pelo total de primitivas n_r. Mede a predominância das primitivas curtas:

$$EPC = \frac{1}{n_r} \sum_{a=1}^{L} \sum_{r=1}^{N_r} \frac{M(a,r)}{r^2} \,. \tag{4.23}$$

- Ênfase em Primitivas Longas (EPL): o valor de cada primitiva é multiplicado pelo quadrado do seu comprimento e normalizado pelo total de primitivas n_r, enfatizando as primitivas longas:

$$EPL = \frac{1}{n_r} \sum_{a=1}^{L} \sum_{r=1}^{N_r} M(a,r) r^2 \,. \tag{4.24}$$

- Não Uniformidade dos Níveis de Cinzento (NUNC): o número de primitivas *run length* para cada nível de cinzento *L* é elevado ao quadrado. A soma dos quadrados é posteriormente dividida pelo fator de normalização n_r. Esta característica mede a não uniformidade dos níveis de cinzento na imagem e terá o seu valor mínimo se as primitivas se encontrarem uniformemente distribuídas por todos os níveis de cinzento:

$$NUNC = \frac{1}{n_r} \sum_{a=1}^{L} \left(\sum_{r=1}^{N_r} M(a,r) \right)^2 \,. \tag{4.5}$$

- Não Uniformidade do Comprimento das Primitivas (NUCP): mede a não uniformidade do comprimento das primitivas *run length*. O número de primitivas *run length* para cada comprimento *r* é elevado ao quadrado, posteriormente a soma dos quadrados é dividida pelo fator de normalização n_r. Se as primitivas se encontrarem uniformemente distribuídas pelos diferentes comprimentos o valor da NUCP terá o seu valor mínimo:

$$NUCP = \frac{1}{n_r}\sum_{r=1}^{N_r}\left(\sum_{a=1}^{L} M(a,r)\right)^2 . \tag{4.25}$$

- Percentagem de Primitivas (PP): é a relação entre o número total de primitivas e o número total de primitivas possíveis, se todas as primitivas tivessem o comprimento de um *pixel*. O seu valor máximo será obtido quando todas as primitivas tiverem o comprimento de $r = 1$:

$$PP = \frac{n_r}{n_p} = \frac{1}{n_p}\sum_{a=1}^{L}\sum_{r=1}^{N_r} M(a,r). \tag{4.26}$$

Posteriormente, Chu et al. (Chu et al., 1990) introduziram duas novas características que permitem integrar informação relativamente aos níveis de cinzento das primitivas e distinguir texturas com valores similares de acordo com EPC e EPL, mas que diferem na distribuição dos níveis de cinzento das primitivas:

- Ênfase em Primitivas com Níveis de Cinzento Baixos (EPNCB): o valor de cada primitiva é dividido pelo quadrado do seu nível de cinzento e normalizado pelo total de primitivas n_r. Mede a predominância das primitivas com níveis de cinzento baixo:

$$EPNCB = \frac{1}{n_r}\sum_{a=1}^{L}\sum_{r=1}^{N_r}\frac{M(a,r)}{a^2} . \tag{4.27}$$

- Ênfase em Primitivas com Níveis de Cinzento Altos (EPNCA): o valor de cada primitiva é multiplicado pelo quadrado do seu nível de cinzento e normalizado pelo total de primitivas n_r, enfatizando as primitivas com níveis de cinzento alto:

$$EPNCA = \frac{1}{n_r}\sum_{a=1}^{L}\sum_{r=1}^{N_r} M(a,r)a^2 . \tag{4.28}$$

Em (Dasarathy, & Holder, 1991) são apresentadas quatro características que procuram tirar partido da conjunção da informação dos níveis de cinzento e do comprimento das primitivas:

- Ênfase em Primitivas Curtas com Níveis de Cinzento Baixos (EPCNCB): mede a distribuição conjunta de primitivas curtas e de valores baixos do nível de cinzento. EPCNCB terá valores elevados em imagens com muitas primitivas pequenas e valores de nível de cinzento baixos:

$$EPCNCB = \frac{1}{n_r}\sum_{a=1}^{L}\sum_{r=1}^{N_r}\frac{M(a,r)}{a^2 r^2}. \tag{4.29}$$

- Ênfase em Primitivas Curtas com Níveis de Cinzento Altos (EPCNCA): mede a distribuição conjunta de primitivas curtas e de valores elevados do nível de cinzento. EPCNCA terá valores elevados em imagens com muitas primitivas pequenas e valores de nível de cinzento elevados:

$$EPCNCA = \frac{1}{n_r}\sum_{a=1}^{L}\sum_{r=1}^{N_r}\frac{M(a,r)a^2}{r^2}. \tag{4.30}$$

- Ênfase em Primitivas Longas com Níveis de Cinzento Baixos (EPLNCB): mede a distribuição conjunta de primitivas longas e de valores baixos do nível de cinzento. EPLNCB terá valores elevados em imagens onde predominam primitivas longas e valores de nível de cinzento baixos;

$$EPLNCB = \frac{1}{n_r}\sum_{a=1}^{L}\sum_{r=1}^{N_r}\frac{M(a,r)r^2}{a^2}. \tag{4.31}$$

- Ênfase em Primitivas Longas com Níveis de Cinzento Altos (EPLNCA): mede a distribuição conjunta de primitivas longas e de valores baixos do nível de cinzento. EPLNCA terá valores elevados em imagens onde predominam primitivas longas e valores de nível de cinzento baixos:

$$EPLNCA = \frac{1}{n_r}\sum_{a=1}^{L}\sum_{r=1}^{N_r}M(a,r)r^2\,a^2\,. \tag{4.32}$$

4.4 Método de Diferenças de Níveis de Cinzento

O Método de Diferenças de Níveis de Cinzento (*Gray Level Difference Method* - GLDM) baseia-se na coocorrência de pares de pixels que possuem uma determinada diferença absoluta em termos de nível de cinzento e se encontram a uma distância *d* segundo a direção θ. O histograma normalizado *hd*(*i*|*d*,θ) indica a probabilidade de ocorrência da diferença de níveis de cinzento *i* entre dois pixels à distância e direção (*d*,θ) (Unser, 1986):

$$hd(i|d,0°) = \#\{((k,l),(m,n) \in (L_r \times L_c) \times (L_r \times L_c): k = m, l - n = -d, |I(k,l) - I(m,n)| = i\}/T(d,0°); \quad (4.33)$$

$$hd(i|d,45°) = \#\{((k,l),(m,n) \in (L_r \times L_c) \times (L_r \times L_c): k - m = d, l - n = -d, |I(k,l) - I(m,n)| = i\}/T(d,45°); \quad (4.34)$$

$$hd(i|d,90°) = \#\{((k,l),(m,n) \in (L_r \times L_c) \times (L_r \times L_c): k - m = d, l = n, |I(k,l) - I(m,n)| = i\}/T(d,90°); \quad (4.35)$$

$$hd(i|d,135°) = \#\{((k,l),(m,n) \in (L_r \times L_c) \times (L_r \times L_c): k - m = d, l - n = d, |I(k,l) - I(m,n)| = i\}/T(d,135°)\,. \quad (4.36)$$

Nas equações anteriores o # corresponde ao número de elementos do conjunto, *I*(*k*,*l*) à intensidade do pixel (*k*, *l*), $(L_r \times L_c)$ às coordenadas espaciais e *T*(*d*, θ) ao número total de pixels da imagem a uma distância *d* na direção θ. Para cada par (*d*,θ) é obtido o histograma:

$$HD(d,\theta) = hd(i|d,\theta), 0 \leq i < L. \quad (4.37)$$

Exemplos do cálculo de histogramas *HD*(*d*,*θ*) para as direções 0º e 45º encontram-se ilustrados na Figura 4.4.

2	2	1	1	0
0	2	4	3	1
2	3	0	3	0
3	4	1	0	2
3	3	0	2	2

I(*x*, *y*)

0	1	2	3	4
4/20	6/20	5/20	5/20	0/20

HD(1|0º)

0	1	2	3	4
5/16	3/16	4/16	3/16	1/16

HD(1|45º)

Figura 4.4: Exemplo do cálculo dos histogramas *HD*(1,0º) e *HD*(1,45º) para a imagem *I*(*x*,*y*).

Partindo de cada histograma é possível extrair um conjunto de descritores de textura (Unser, 1986):

- Segundo Momento Angular (SMA_{hd}):

$$SMA_{hd} = \sum_{i=0}^{L-1} hd^2(i). \tag{4.38}$$

- Entropia (ENT_{hd}):

$$ENT_{hd} = -\sum_{i=0}^{L-1} hd(i)\log(hd(i)). \tag{4.39}$$

- Momento Inverso da Diferença (MID_{hd}):

$$MID_{hd} = \sum_{i=0}^{L-1} \frac{hd(i)}{1+i^2}. \tag{4.40}$$

- Variância (VAR_{hd}):

$$VAR_{hd} = \sigma^2 = \sum_{i=0}^{L-1} hd(i)(i-\mu)^2. \tag{4.41}$$

- Contraste (CON_{hd}):

$$CON_{hd} = \sum_{i=0}^{L-1} hd(i)(i)^2. \tag{4.42}$$

- Correlação (COR_{hd}):

$$COR_{hd} = \sum_{i=0}^{L-1} hd(i)\frac{(i-\mu)}{\sigma}. \tag{4.43}$$

4.5 Lacunaridade Diferencial

A terminologia fractal, cuja origem vem da palavra Latina *fractus* que significa fragmentado, foi introduzida pelo matemático Benoit Mandelbrot para descrever superfícies não Euclidianas, que se apresentam autossemelhantes em diferentes escalas, isto é, o conjunto total é constituído por réplicas dele próprio qualquer que seja a escala de visualização (Mandelbrot, 1982). Muitas das estruturas existentes na natureza são complexas, apresentam descontinuidades e fragmentações não podendo ser medidas com precisão à luz da geometria Euclidiana, tendendo a possuir uma dimensão fractal, isto é, "entre dimensões". Enquanto a geometria euclidiana aborda as formas perfeitas (e.g. quadrados, círculos) a dimensão fractal foca-se nas "imperfeições" das formas que existem na natureza. A dimensão fractal possui potencial para uma descrição mais rica das superfícies que a dimensão Euclidiana, assemelhando-se em, analogia, à relação da lógica *fuzzy* com a lógica digital.

Num espaço Euclidiano de dimensão *n* um conjunto fechado *S* em $\Re^n$diz-se autossimilar se *S* resulta da união de N_r cópias distintas de si mesmo, não sobrepostas, cada uma das quais semelhante a *S*, em escalas reduzidas por um fator de *r*. Assim, a dimensão fractal pode ser obtida a partir da seguinte relação (Mandelbrot, 1982; Al-Kadi, & Watson, 2008):

$$DF = \frac{\log(N_r)}{\log(1/r)}. \tag{4.44}$$

A dimensão fractal é uma propriedade de objetos fractais que indica o grau de preenchimento do espaço Euclidiano por parte do objeto. Para um objeto compacto e uniforme, a dimensão fractal é um valor inteiro correspondente à sua dimensão Euclidiana. No caso de se tratar de um fractal a dimensão fractal é um número fracionário. Assim, a dimensão fractal pode ser usada

para caracterizar a complexidade/irregularidade do conjunto (Al-Kadi, & Watson, 2008). A autossimilaridade perfeita raramente se encontra na natureza, sendo frequente nos fractais obtidos através de funções matemáticas. Na verdade, a maioria dos fractais naturais são estocásticos o que significa que existe alguma aleatoriedade envolvida, ocorrendo semelhanças de forma entre escalas diferentes. Um exemplo de fractal estocástico natural é a árvore vascular pulmonar.

Acontece que a dimensão fractal por si só não caracteriza um fractal. Alguns objetos fractais, com aspeto substancialmente diferentes possuem uma dimensão fractal igual. De forma a distinguir estas situações Mandelbrot (Mandelbrot, 1982) propôs uma medida complementar à dimensão fractal, designada por lacunaridade, expressão que tem origem no Latim *Lacuna*. Apesar da lacunaridade ter sido originalmente desenvolvida para descrever propriedades de fractais (Mandelbrot, 1982; Gefen, Meir, Mandelbrot, & Aharony, 1983; Allain, & Cloitre, 1991) não se baseia na autossimilaridade e pode portanto ser aplicada para descrever a distribuição de padrões espaciais fractais ou não fractais, quer sejam binários ou não (Plotnick, Gardner, Hargrove, Prestegaard, & Perlmutter, 1996).

Segundo Mandelbrot esta característica descreve a distribuição das lacunas num fractal. Por seu lado Gefen et al. (Gefen et al., 1983) referem-se à lacunaridade como sendo o desvio de um fractal da invariância translacional, numa determinada escala. Enquanto a dimensão fractal mede o quão preenchido se encontra o espaço Euclidiano, a lacunaridade quantifica a forma como esse espaço está ocupado. Esta propriedade é altamente dependente da escala. Padrões homogéneos observados em escalas pequenas podem apresentar-se bastante heterogéneos se observados em escalas maiores e vice-versa. Portanto, a lacunaridade pode ser considerada uma medida da heterogeneidade em múltiplas escalas (Allain, & Cloitre, 1991; Plotnick et al., 1996). Valores baixos de lacunaridade estão relacionados com padrões espaciais homogéneos, enquanto valores altos de lacunaridade estão associados a padrões heterogéneos. Em geral, quanto maior for a dispersão do tamanho das lacunas na textura mais alta será a lacunaridade calculada, e vice-versa.

Allain e Cloitre (Allain, & Cloitre, 1991) propuseram o método *gliding-box* para o cálculo da lacunaridade em imagens binárias. É considerada uma caixa de largura *r* que se coloca inicialmente no canto superior esquerdo da imagem e que desliza ao longo das linhas e colunas

da imagem, pixel a pixel. Para cada posição da caixa contabiliza-se a massa *M* da caixa, isto é o número de sítios ocupados dentro da caixa. O número de caixas de tamanho *r* que possuem a massa *M* é contabilizado ao longo da imagem sendo dado por $n(M,r)$. Considerando *N*(*r*) o número total de caixas de tamanho *r*, a distribuição de probabilidades *Q*(*M*, *r*), tal que uma caixa de tamanho *r* contenha a massa *M* é dada por:

$$Q(M,r) = \frac{n(M,r)}{N(r)}. \tag{4.45}$$

O primeiro e segundo momentos desta distribuição de probabilidades são dados por:

$$Z^{(1)}(r) = \sum_{M} M\, Q(M,r), \tag{4.46}$$

$$Z^{(2)}(r) = \sum_{M} M^2\, Q(M,r). \tag{4.47}$$

A lacunaridade $\Lambda(r)$ para uma caixa deslizante de tamanho *r* é definida como o quociente do segundo e o quadrado do primeiro momentos da distribuição de probabilidades *Q*(*M*, *r*):

$$\Lambda(r) = \frac{Z^{(2)}(r)}{[Z^{(1)}(r)]^2} = \frac{\sum_M M^2\, Q(M,r)}{[\sum_M M\, Q(M,r)]^2}. \tag{4.48}$$

No caso de conjuntos invariantes à translação o numerador e denominador da Equação (4.48) são iguais sendo o valor da lacunaridade $\Lambda(r) = 1$, independentemente do valor de *r*. Se os conjuntos forem não invariantes à translação então possuem valores de lacunaridade superiores a um. No caso de existirem lacunas de todos os tamanhos é expectável que a lacunaridade seja muito superior a um, enquanto se as lacunas tiverem todas o mesmo tamanho o valor da lacunaridade é um (Allain, & Cloitre, 1991).

É comum a representação da variação da lacunaridade *versus* o tamanho da caixa *r* em gráficos com configuração *log-log* (Dougherty, & Henebry, 2002). O valor de *r* varia de *r* = 1 até ao tamanho da imagem. Esta representação permite entender a estrutura espacial da imagem quando analisada em diferentes escalas. A análise dessas curvas é muitas vezes baseada na interpretação de $Z^{(1)}(r) = \bar{s}(r)$ corresponde à média e $s_s^2(r)$ à variância do número de sítios ocupados por caixa. Assim, o segundo momento pode ser representado por:

$$Z^{(2)}(r) = s_s^2(r) + \bar{s}^2(r). \tag{4.49}$$

Substituindo a Equação (4.49) na Equação (4.48) o valor da $\Lambda(r)$ é determinado por (Plotnick et al., 1996):

$$\Lambda(r) = \frac{s_s^2(r)}{\bar{s}^2(r)} + 1. \tag{4.50}$$

Através da análise da Equação (4.51) e pela observação das curvas de lacunaridade de diferentes padrões (e.g. regulares, aleatórios, fractais, *clusters*) Plotnick et al. concluíram que a lacunaridade de padrões espaciais binários é função da percentagem de sítios ocupados, do tamanho da caixa deslizante e da geometria do padrão da imagem. Com base nestes três aspetos é possível tecer algumas considerações importantes que ajudam na interpretação das curvas de lacunaridade (Plotnick et al., 1996; Reiss, Lemmerer, Hanslmeier, & Ahammer, 2016):

- Se a média dos espaços ocupados tende para zero o valor da lacunaridade tende para infinito, portanto o valor da lacunaridade de padrões densos é menor que a de padrões esparsos, para o mesmo valor da caixa *r*;

- À exceção de padrões espaciais com *clusters* muito marcados, quando o tamanho da caixa *r* aumenta, o valor da lacunaridade $\Lambda(r)$ diminui. A aparência das caixas mais pequenas difere mais do que a aparência das caixas maiores, ou seja há maior invariância à translação nas caixas com maior *r,* o segundo momento $Z^{(2)}(r)$ decresce relativamente ao primeiro. No caso de caixas com a menor largura possível (r = 1) o valor da lacunaridade é função da percentagem de sítios ocupados, sendo independente do tamanho e do arranjo espacial do padrão, não contribuindo portanto para a sua compreensão. De modo semelhante, se o tamanho da caixa corresponder ao tamanho da imagem o valor da lacunaridade será um, uma vez que $s_s^2(r) = 0$;

- No caso de padrões organizados em *clusters* a lacunaridade é alta, uma vez que há uma grande variação no valor da massa da caixa tornando o valor de $s_s^2(r)$ elevado, portanto para a maioria dos tamanhos da caixa a lacunaridade será elevada. Quando *r* se aproxima do tamanho de referência do *cluster* o conteúdo das caixas torna-se semelhante e, portanto, a lacunaridade começa a decair rapidamente. Neste contexto, padrões

autossimiliares são caracterizados por curvas de lacunaridade lineares, dado que as lacunas surgem similares em diferentes escalas.

Os primeiros estudos sobre a análise da lacunaridade diziam respeito apenas a conjuntos binários (Gefen et al., 1983; Allain, & Cloitre, 1991). Posteriormente, Plotnick et al. mostraram que o conceito se aplicava também a conjuntos não binários. Em (Dong, 2000) o autor propõe um método para o cálculo da lacunaridade diferencial que conjuga o método *gliding-box* (Allain, & Cloitre, 1991), proposto para conjuntos binários, com o método *diferential box-counting* proposto por Sarkar e Chaudhuri para o cálculo da dimensão fractal (Sarkar, & Chaudhuri, 1992). O método proposto por Dong permite lidar com imagens de níveis de cinzento. Dong recorreu ao método *diferential box-counting* para o cálculo do valor da massa *M* presente na Equação (4.48). O método considera a imagem como um objeto 3D onde (*x*, *y*) corresponde às coordenadas do pixel e a terceira dimensão corresponde ao seu nível de cinzento. A imagem é dividida em janelas de tamanho $w \times w$ que se deslocam ao longo de toda a imagem e em caixas de tamanho $r \times r$ que percorrem cada uma das janelas. O tamanho da janela, *w*, deve ser ímpar e respeitar a condição $r<w$. Para cada caixa $r \times r$ são sobrepostas colunas de caixas cúbicas de volume $r \times r \times r$, sendo $r>1$. O número de cubos de cada coluna depende do valor máximo dos valores da intensidade dos pixels existentes em cada caixa. A cada caixa cúbica que constitui a coluna é atribuído um número sequencial crescente de baixo para cima (1, 2, 3,…).

Se o valor máximo e mínimo do nível de intensidade dos pixels contidos da caixa pertencerem à caixa cúbica forem *v* e *u*, respetivamente, a altura relativa da coluna, nas coordenadas (*i*, *j*), é dada pela por:

$$n_r(i,j) = v - u - 1. \tag{4.51}$$

O valor da massa M_r representa cada possível valor da altura relativa da coluna n_r e obtém-se fazendo deslizar a caixa ao longo de toda a janela $w \times w$:

$$M_r = \sum_{i,j} n_r\,(i,j)\,. \tag{4.52}$$

A lacunaridade diferencial $\Lambda(r)$ na janela $w \times w$ é obtida substituindo na Equação (4.49) M por M_r. Este valor é atribuído à coordenada central da janela. O processo é repetido para toda a imagem à medida que a janela se desloca.

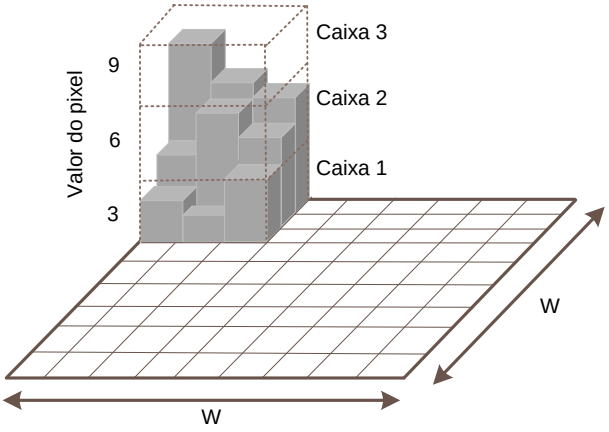


Figura 4.5: Algoritmo *differential box-counting* ($w = 9$, $r = 3$).

Na Figura 4.5 é ilustrada a aplicação do algoritmo *differential box-counting* para uma caixa $r = 3$ e uma janela $w = 9$. Uma coluna de caixas sobrepostas de 3×3×3 é gerada de acordo com os valores dos pixels presentes. O valor mais elevado pertence à caixa $v = 3$ e o menor valor à caixa $u = 1$. De acordo com a Equação (4.51) altura relativa da coluna é $n_r(i, j) = 3 - 1 - 1 = 1$.

O cálculo da lacunaridade diferencial é exemplificado na Figura 4.6 através de uma janela de 4×4 pixels sobre a qual é aplicada uma caixa deslizante sobreposta de 3×3 pixels. As várias fases do cálculo da massa M_r são mostrados na Figura 4.6 a) - e). A Figura 4.6 b) representa a primeira posição da caixa deslizante, na qual os valores dos níveis de cinzento máximo e mínimo são 18 e 7, respetivamente. Considerando cubos sobrepostos de 3×3×3 pixels estes valores encontram-se contidos nas caixas $v = 6$ e $u = 3$, respetivamente. A altura relativa da coluna é $n_r(2, 2) = v - u = 6 - 3 - 1 = 2$. A altura relativa da coluna correspondente à posição da caixa deslizante da Figura 4.6 c) é $n_r(2, 3) = v - u = 7 - 1 - 1 = 5$ correspondente aos valores dos pixels máximo de 19 e mínimo de 3. No caso das posições da caixa deslizante ilustradas na Figura 4.6 d) e e) o valor do pixel máximo é 22 e o mínimo é 3, valores que correspondem às caixas $v = 8$ e $u = 1$ e a colunas de altura relativa 6 (adaptado de (Reiss et al., 2016)).

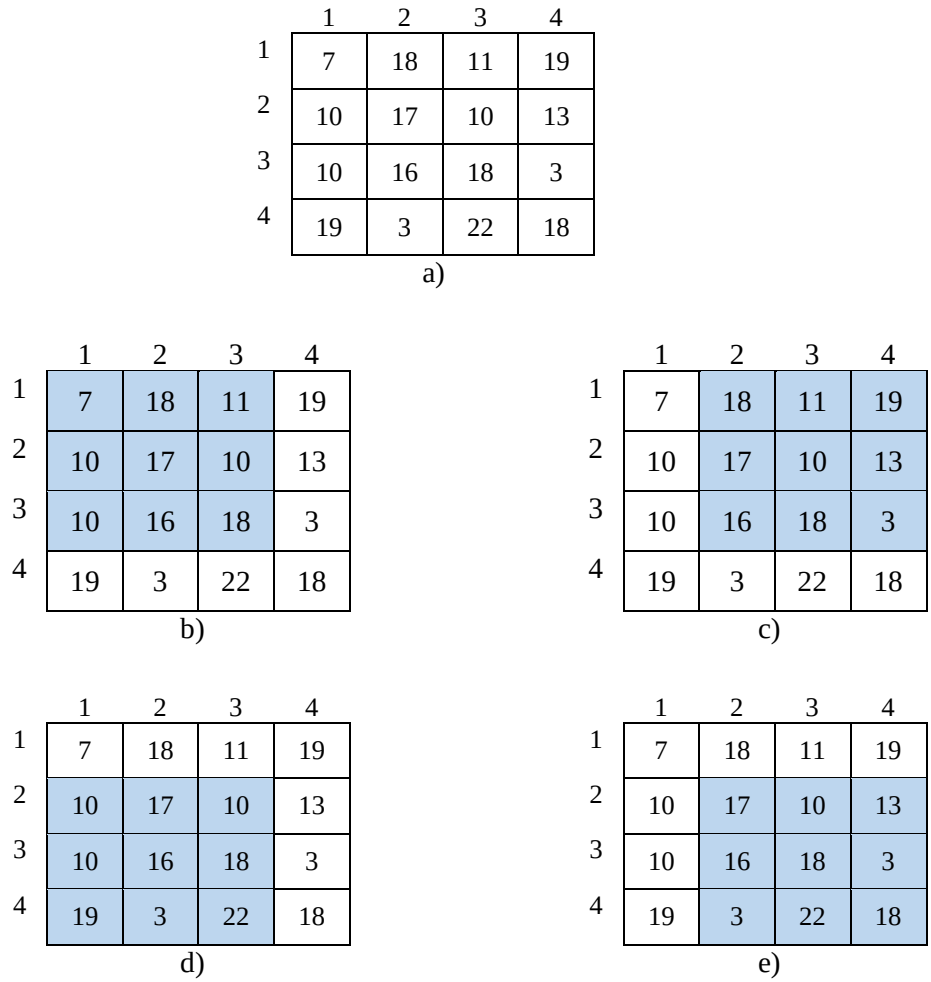


Figura 4.6: Cálculo da lacunaridade diferencial. a) Janela de 4×4 pixels. De b) a e) As quatro posições possíveis da caixa deslizante.

O valor da lacunaridade diferencial para a janela da Figura 4.6 a) é obtido através de:

$$\Lambda(r) = \frac{\sum_{M_r} M_r^2 \, Q(M_r, r)}{\left[\sum_{M_r} M_r \, Q(M_r, r)\right]^2} = \tag{4.53}$$

$$= \frac{(2\text{x}2)\text{x}(1/4) + (5\text{x}5)\text{x}(1/4) + 2\text{x}(6\text{x}6)\text{x}(1/4)}{[2\text{x}(1/4) + 5\text{x}(1/4) + 2\text{x}6\text{x}(1/4)]^2} = 1.12\,.$$

A lacunaridade tem sido utilizada com sucesso na análise da textura em imagens médicas, obtidas a partir de vários tipos de equipamentos, no estudo de patologias que afetam diferentes órgãos e tecidos e na avaliação da resposta a tratamentos. O estudo da osteoporose é objeto de estudo nos trabalhos de (Dougherty, & Henebry, 2002; Zaia, Eleonori, Maponi, Rossi, & Murri, 2006; Zaia, 2015; Slim et al., 2016) usando as técnicas imagiológicas TC, RM e raio X,

respetivamente. Em (Al-Kadi, & Watson, 2008) é proposta a diferenciação entre tumores pulmonares malignos agressivos e não agressivos, em imagens de TC. O trabalho de Smitha et al. mostrou a potencialidade da lacunaridade na diferenciação dos graus de malignidade de gliomas, tumores primários do cérebro, em imagens de RM (Smitha, Gupta, & Jayasree, 2015). Em (Charisis, Hadjileontiadis, Liatsos, Mavrogiannis, & Sergiadis, 2010; Charisis, Hadjileontiadis, Liatsos, Mavrogiannis, & Sergiadis 2012a; Charisis, Hadjileontiadis, Liatsos, Mavrogiannis, & Sergiadis, 2012b) é proposto um trabalho para a diferenciação de regiões normais de regiões ulcerosas em imagens endoscópicas adquiridas através de cápsulas endoscópicas *wireless*. Em (Cordeiro, Backes, Júnior, Gonçalves, & Oliveira, 2016) é proposta a análise dos padrões de osso normal e com displasia fibrosa da região maxilofacial recorrendo à análise da lacunaridade em imagens de TC. Em (Al-Kadi, Chung, Coussios, & Noble, 2016) é avaliada a resposta aos tratamento de quimioterapia a tumores do fígado em imagens de ultrassonografia. No Anexo A, encontra-se uma descrição mais detalhada dos trabalhos referidos neste parágrafo sendo destacado o contributo da análise da lacunaridade em diferentes contextos.

4.6 Seleção de Características

Em problemas de classificação, após a etapa de extração de característica é comum ocorrer a um processo de seleção de características que visa reduzir a dimensão e o "ruído" do conjunto de dados. Para encontrar o subconjunto ótimo de característica é necessário adoptar uma estratégia de pesquisa que permita, a partir da totalidade das característica, escolher apenas algumas e um critério que avalie a qualidade do subconjunto gerado.

O subconjunto de características escolhido deve descrever corretamente o problema, mantendo o mesmo nível de desempenho ou mesmo melhorando-o. O objetivo primordial desta etapa é detetar as características relevantes e descartar as irrelevantes e/ou redundantes (Liu, & Motoda, 2007). As características irrelevantes podem ser removidas sem que o desempenho do algoritmo de aprendizagem seja afetado. As características redundantes implicam a presença de outra característica, sendo que as duas são individualmente relevantes, no entanto, a remoção de uma delas não irá afetar o desempenho da aprendizagem. Assim, a seleção de características é importante uma vez que (Guyon, 2006):

- Reduz o efeito do *Curse of Dimensionality*. O número de características determina a dimensão do espaço das hipóteses que contém todas as hipóteses que podem ser inferidas a partir dos dados, e que permitem a predição das classes. Quanto maior o número de características maior será a dimensão do espaço das hipóteses. A um aumento linear do número de características corresponde um aumento exponencial do espaço das hipóteses. Assim, a redução da dimensionalidade do problemas permite encontrar mais facilmente a hipótese correta (Liu, & Motoda, 2007);

- Identifica as características o que aumenta a compreensão do problema, melhorando, muitas vezes, a visualização dos dados;

- Reduz os custos de futuras análises de dados ou predição.

No contexto da classificação de padrões, existe na literatura um grande número de métodos de seleção de características (Guyon, 2006; Liu, & Motoda, 2007; Hira, & Gillies, 2015). Os métodos diferem uns dos outros na forma como combinam a pesquisa do subconjunto de características com o algoritmo de aprendizagem automático utilizado para inferir o modelo do classificador. Neste sentido, os métodos de seleção de características podem dividir-se em três grupos: os métodos de filtragem, os métodos embebidos e os métodos *wrapper*.

Aos métodos que selecionam as características sem procederem à otimização do desempenho do preditor são genericamente conhecidos como métodos de filtragem. Estes avaliam a relevância das características baseando-se apenas nas propriedades intrínsecas dos dados, como por exemplo coeficientes de correlação ou testes estatísticos. As características que possuem um menor valor de acordo com o índice de relevância são retiradas. Estes métodos são facilmente adaptados a bases de dados de elevada dimensão, são computacionalmente simples e rápidos. Habitualmente, são executados apenas uma vez, podendo ser associados a diferentes algoritmos de aprendizagem. Uma desvantagem comum aos métodos de filtragem é a sua independência relativamente ao algoritmo estimador, não existindo interação entre a escolha das características e a escolha do modelo ótimo do classificador. Por outro lado, muitos dos métodos de filtragem são "míopes" no sentido em que estimam a relevância de cada característica de uma forma independente do contexto das restantes características. Não sendo considerada, a dependência entre características pode piorar o desempenho. Por exemplo, no critério de Fisher a relevância de cada característica é calculada de forma independente das

restantes características, através da correlação das características relativamente às classes das instâncias que constituem o conjunto de treino (Liu, & Motoda, 2007). Na Figura 4.7 encontram-se esquematizados os métodos de seleção de características por filtragem em que o algoritmo de aprendizagem é um classificador.

Figura 4.7: Métodos de filtragem em classificação.

Nos métodos embebidos a interação entre a seleção de características e o algoritmo de aprendizagem é dedicada, ou seja, é específica do algoritmo de aprendizagem utilizado. Estas duas etapas não podem ser separadas. Na Figura 4.8 encontram-se esquematizados os métodos de seleção de características embebidos aplicados num contexto de classificação. Um exemplo de um método embebido são as árvores de decisão em que em cada ramo tem de ser escolhida uma característica.

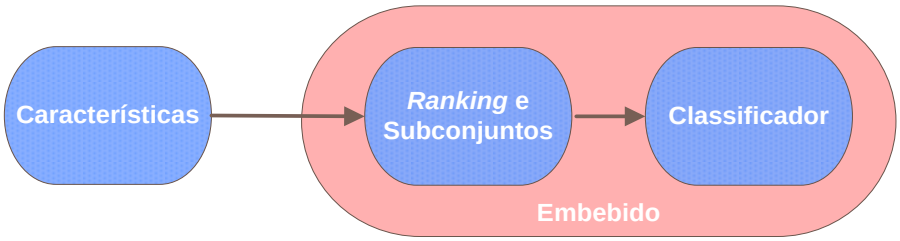


Figura 4.9: Métodos embebidos em classificação.

Os métodos *wrapper* envolvem o algoritmo de aprendizagem como parte do processo de seleção de características. Nestes métodos é estabelecida uma estratégia de pesquisa dentro dos possíveis subconjuntos de características. Vários subconjuntos são gerados e avaliados. Para cada subconjunto é treinado e testado um modelo específico de um algoritmo de aprendizagem, tornando esta abordagem adaptada a um algoritmo de aprendizagem específico. O subconjunto ótimo é aquele que proporciona o melhor desempenho do algoritmo de aprendizagem e que, portanto, possui o maior poder preditivo. Os subconjuntos de características possíveis crescem exponencialmente com a dimensão do espaço das características.

A grande desvantagem dos métodos *wrapper* é o seu custo computacional. Este argumento encontra-se fragilizado pois atualmente existem supercomputadores acessíveis à comunidade científica com uma grande capacidade computacional. Outra desvantagem muitas vezes associada a este tipo de métodos é o risco de ocorrer *overfitting* comparativamente aos métodos de filtragem. A grande vantagem dos métodos *wrapper* consiste na forte interação entre a pesquisa do subconjunto ótimo e a seleção do modelo ótimo do algoritmo de aprendizagem. Na Figura 4.9 encontram-se esquematizados os métodos de seleção de características *wrapper* num contexto de classificação.

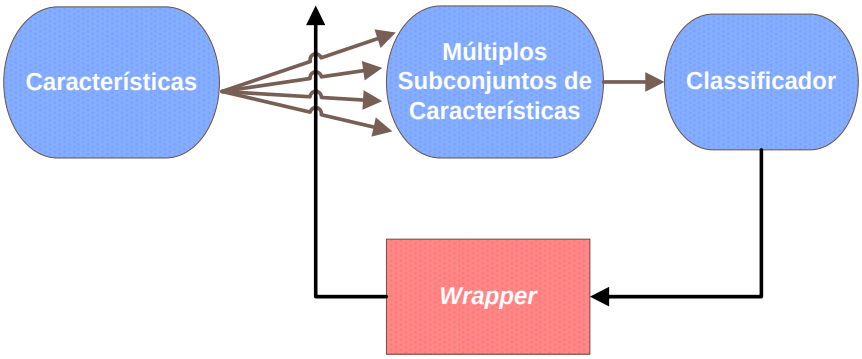


Figura 4.9: Métodos *wrapper* em classificação.

As estratégias de pesquisa existentes na literatura são comuns tanto em situações de classificação supervisionada como não supervisionada. Uma solução ótima implica uma pesquisa exaustiva. Como habitualmente o número de características a selecionar, *d*, das *m* características originais não é conhecido é necessário uma pesquisa exaustiva em $\binom{m}{1} + \binom{m}{2} + \cdots + \binom{m}{d}$ subconjuntos que corresponde a $2^m - 1$ vezes até obter o subconjunto de características, o que mesmo para problemas com moderada dimensão se torna difícil de avaliar em tempo útil devido à complexidade computacional que envolve. No caso de ser conhecida previamente a cardinalidade *d* do subconjunto de características a serem selecionadas a partir do conjunto *m* a pesquisa é restringida a $\binom{m}{d}$ subconjuntos, que, dependendo do valor de *d* e de *m*, pode ser ainda um valor elevado. As dificuldades das estratégias de pesquisa ótimas levaram à adoção de algoritmos de pesquisa heurísticos que não garantem soluções ótimas, mas que apesar de não garantirem a otimização global de todos os subconjuntos, permitem encontrar subconjuntos de características com grande potencial discriminante. Como exemplos podemos referir os algoritmos clássicos SFS e SBS. O algoritmo SBS tem início com um conjunto que inclui todas as características. Sequencialmente é eliminada uma característica de cada vez. Em

cada passo cada uma das características é retirada e o resultado da sua exclusão é comparado com cada uma das outras usando a função de avaliação. Será excluída do conjunto a característica que ao ser retirada leva à obtenção dos melhores resultados. O SFS é semelhante aos SBS mas tem início com um conjunto vazio ao qual são adicionadas características. Em cada passo do algoritmo cada uma das características que ainda não faz parte do conjunto é adicionada, e o conjunto resultante é avaliado.

Ao longo dos anos foram surgindo muitas modificações às estratégias de pesquisa sequencial *forward* e *backward*, como por exemplo o algoritmo *floating search*. Existe ainda um terceiro tipo de algoritmos designados por aleatorizados (*randomized*) que introduzem alguma aleatoriedade na estratégia de pesquisa o que ajuda a evitar mínimos locais, como por exemplo os algoritmos genético.

A seleção de características que será utilizada neste trabalho baseia numa estratégia SBS com recurso a validação cruzada.

É de referir que é comum o recurso à expressão "extração de características" no contexto de redução da dimensionalidade do problema, por exemplo em problemas de classificação. Contrariamente, aos métodos de seleção de características em que o subconjunto de características escolhidas faz parte do conjunto inicial, tendo o mesmo significado físico, no caso da "extração de características" a redução do número de características é conseguida através da construção de novas características obtidas a partir de transformações ou combinações do conjunto de características originais. Um exemplo é o algoritmo análise de componentes principais (Guyon, 2006). Neste documento quando é referido a extração de características é no contexto da aplicação de métodos que permitam caracterizar a textura da imagem.

4.7 Considerações Finais

No contexto da análise da imagem médica a textura é descrita como um dos seus atributos mais importantes, desempenhando um papel preponderante na extração de informação das regiões analisadas. No âmbito de problemas de classificação de padrões, como é o caso dos sistema de DAC baseados em imagem, a análise da textura tem como objetivo caracterizar as variações

espaciais de regiões da imagem possibilitando a sua descrição através de um conjunto de medidas. Este conjunto de características é posteriormente utilizado para discriminar vários padrões.

Ao longo deste capítulo foram descritos os métodos de análise de textura implementados com o objetivo de obter um vetor de características que permita descrever, de uma forma compacta, os padrões pulmonares em imagens de TCAR do tórax. O vetor de características é construído a partir de características extraídas através dos métodos: análise estatística de níveis de cinzento de primeira ordem, método de dependência espacial de níveis de cinzento, método de comprimento de primitivas de níveis de cinzento, método de diferenças de níveis de cinzento e análise da lacunaridade diferencial.

Por fim, é realçada a importância da etapa de seleção de características, nomeadamente na redução da dimensionalidade através da eliminação de características irrelevantes e/ou redundantes sem degradar o desempenho do sistema e apresentadas as linhas orientadoras dos métodos de seleção de características.

Capítulo 5
CLASSIFICAÇÃO DE PADRÕES

As tarefas comuns e aparentemente simples que desempenhamos no dia-a-dia, como reconhecer caracteres ou identificar objetos, entre muitas outras situações, não são inatas ao ser humano, resultam de anos de aprendizagem. À semelhança do que acontece com o ser humano é possível ensinar sistemas computorizados a reconhecer padrões através de técnicas de aprendizagem automática. Neste capítulo são introduzidos conceitos que nos permitem entender os princípios da teoria da aprendizagem estatística, que descreve os aspetos estatísticos da aprendizagem automática na qual o algoritmo SVM se fundamenta. São apresentadas as formulações matemáticas do algoritmo de classificação utilizado – SVM. Inicialmente é introduzida a formulação mais simples correspondente aos dados linearmente separáveis, posteriormente para dados não linearmente separáveis e por fim a formulação mais complexa correspondente ao SVM não linear. São descritos os principais métodos de validação e apresentadas as métricas de avaliação de desempenho utilizadas neste trabalho.

5.1 Metodologia de Aprendizagem

Tal como acontece na aprendizagem humana, os sistemas computorizados de aprendizagem automática têm de ser ensinados a reconhecer padrões. Numa abordagem de aprendizagem supervisionada é necessário um conjunto de padrões (que também podem ser designados por exemplos, dados ou instâncias) de treino e respetivas classes. Consideremos o conjunto de treino *T* constituído por *n* padrões:

$$(\mathbf{x}_1, \cdots, \mathbf{x}_n) \quad \text{com} \quad \mathbf{x}_i \in \Re^m, \qquad \forall\, i = 1, \ldots, n, \tag{5.1}$$

e respetivas classes

$$(y_1, \cdots, y_n) \quad \text{com} \quad y_i \ \epsilon \ \{+1, -1\}, \quad \forall\, i = 1, \ldots, n\,. \tag{5.2}$$

Cada padrão de treino *i* é caracterizado por um vetor $\mathbf{x}_i$ de *m* características ou atributos. No caso concreto, cada vetor é constituído por descritores de textura extraídos das regiões do parênquima pulmonar. Ao padrão *i* encontra-se associada uma classe específica y_i que toma o valor {+1} ou {-1}. A classe {+1} corresponde à ocorrência de patologia e a classe {-1} a pulmão considerado normal. O mapeamento entre o espaço das características em que cada padrão é representado e a respetiva classe pode, normalmente, ser descrito por uma função, que no caso de uma tarefa de classificação se designa por função de decisão. O algoritmo que a partir do conjunto de padrões de treino constrói a função de decisão designa-se por algoritmo de aprendizagem ou máquina de aprendizagem (Borges, 2003).

Numa aprendizagem supervisionada, durante a fase de treino, o algoritmo de aprendizagem deve aprender a reconhecer as principais características associadas à classe {+1} e associadas à classe {-1} e ajustar os seus parâmetros aos padrões que lhe são "mostrados". A função de decisão resultante deve permitir separar corretamente as duas classes no espaço das características $\Re^m$, como ilustrado na Figura 5.1 A sua capacidade de generalização é avaliada no conjunto de teste, constituído por um novo conjunto de padrões cuja classe é conhecida mas que não integraram o conjunto de treino *T*. Este conjunto simula a ocorrência de futuras instâncias cuja classe é desconhecida.

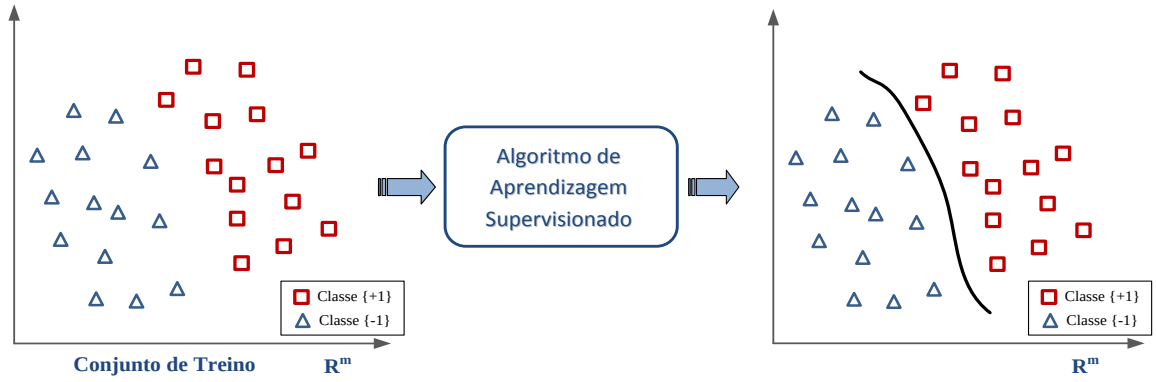


Figura 5.1: Abordagem supervisionada em classificação de padrões.

5.2 Teoria da Aprendizagem Estatística

Consideremos um conjunto *T* constituído por um conjunto de padrões de treino e respetivas classes $(\mathbf{x}_i, y_i),\ i = 1, \ldots, n$ explicitados nas Equações (5.1) e (5.2). Assume-se que os dados são gerados de forma independente e identicamente distribuída (i. i. d.) de acordo com a distribuição de probabilidades $P(\mathbf{x}, y)$. Como referido na Secção 5.1, durante a fase de treino, o algoritmo de aprendizagem estabelece um mapeamento $f: \Re^m \rightarrow \{+1, -1\}$ entre os padrões de entrada e as respectivas classes. Espera-se que *f* seja capaz de classificar novos dados $y = f(\mathbf{x})$ gerados a partir da mesma distribuição de probabilidades, $P(\mathbf{x}, y)$, que os dados de treino (Burges, 1998; Schölkopf, Burges, & Smola, 1999).

O melhor mapeamento *f* é aquele que minimiza o risco esperado (ou erro esperado) para um conjunto de dados de teste (**x**, *y*), quantificado pela equação seguinte:

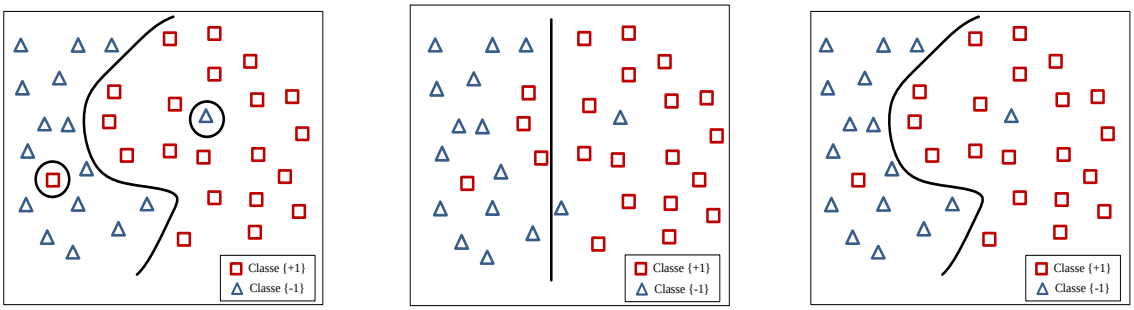
$$R(f) = \int \frac{1}{2} |f(\mathbf{x}) - y| \, dP(\mathbf{x}, y). \tag{5.3}$$

O termo $1/2|f(\mathbf{x}) - y|$ é designado por função de perda e retorna 0 se o padrão **x** é classificado corretamente e 1 no caso contrário. Como a distribuição de probabilidade $P(\mathbf{x}, y)$, a partir da qual os dados são gerados, é geralmente desconhecida, não é possível minimizar diretamente o erro esperado. Recorrendo ao princípio da indução, a minimização do risco esperado é aproximada pela minimização do risco empírico (ou erro de treino) calculado através do erro médio obtido no conjunto de treino, permitindo obter uma estimação da função *f* (Schölkopf et al., 1999):

$$R_{emp}(f) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} \, |f(\mathbf{x}_i) - y_i|. \tag{5.4}$$

É de realçar que um erro pequeno obtido no conjunto de treino não implica necessariamente uma boa capacidade de generalização, nomeadamente um erro pequeno num conjunto de teste independente. Consideremos o conjunto de treino representado na Figura 5.2. O objetivo do processo de aprendizagem é encontrar uma função *f* que discrimine a classe {+1} da classe {-1}. Na Figura 5.2 estão representadas as fronteiras de decisão correspondentes a três funções *f* diferentes. No caso da Figura 5.2 a) a função *f* considerada permite a classificação correta de

todas as instâncias do conjunto de treino, no entanto, encontra-se demasiado ajustada aos exemplos de treino, sendo suscetível de possuir uma baixa capacidade de generalização quando confrontada com novos exemplos (*overfitting*). Por vezes, este ajustamento da função *f* é conseguido através de funções com elevados graus de complexidade, em particular em conjuntos de treino pequenos. No caso da Figura 5.2 b) existem dados pertencentes à mesma classe que o algoritmo de aprendizagem não teve em consideração (*underfitting*). A fronteira de decisão da Figura 5.2 c) representa um compromisso entre as duas situações anteriores, permitindo classificar de forma correta a maioria dos exemplos sem se ajustar demasiado aos dados de treino e permitindo uma maior capacidade de generalização quando confrontada com novos exemplos.



a) b) c)

Figura 5.2: Conjunto de treino e três fronteiras de decisão diferentes.

É possível encontrar várias funções *f* que permitem classificar corretamente todos os exemplos do conjunto de treino, no entanto, cada uma destas funções pode ter desempenhos muito diferentes perante novos dados. Então como escolher a melhor função? A teoria da aprendizagem estatística introduzida por Vapnik estabelece condições matemáticas que auxiliam na escolha da função *f*, a partir de um conjunto de dados de treino. Estas condições têm em consideração o desempenho da função de decisão *f* nos dados de treino e a complexidade da classe de funções a que a função *f* pertence (Vapnik, 1995).

Uma forma concreta de limitar a complexidade das funções *f* é através do princípio da minimização do risco estrutural. Comecemos por introduzir o conceito de dimensão de Vapnik-Chervonenkis (VC), designada por *h*. A dimensão VC é uma propriedade de uma classe de funções que mede a sua complexidade. Quanto maior a dimensão VC, mais complexas são as funções dessa classe. Esta é definida como o maior número de exemplos, em qualquer disposição possível, que pode ser separado por funções da referida classe (Borges, 2003). Por

exemplo, no caso de um problema de classificação binária com três exemplos em $\Re^2$ qualquer que seja a combinação das classes que esses exemplos possam assumir, é sempre possível a sua separação através de uma reta. Nesta situação a dimensão VC da classe das funções é de três. No entanto, se consideramos quatro exemplos em $\Re^2$ existem combinações em que a separação dos exemplos não é possível através de retas, sendo necessário recorrer a funções mais complexas.

Pode-se construir um conjunto de classes de funções encadeadas $F_1 \subset F_2 \subset \cdots \subset F_k$ cuja dimensão de VC satisfaz $h_1 \leq h_2 \leq \cdots \leq h_k$. O princípio da minimização do erro estrutural visa escolher a função $f_i \in F_i$ de modo a que o limite superior do erro esperado (lado direito da Equação (5.5)) seja minimizado:

$$R(f) \leq R_{emp}(f) + \sqrt{\left(\frac{h\left(log\left(\frac{2n}{h}\right)+1\right) - log(\theta/4)}{n}\right)}. \qquad (5.5)$$

O segundo termo do lado direito da Equação (5.5) designa-se por termo de confiança. O limite é garantido com uma probabilidade de $1-\theta$, em que $0 \leq \theta \leq 1$ (Vapnik, 1995). Como ilustra a Figura 5.3, adaptada de (Lorena, & Carvalho, 2007), a minimização do erro esperado $R(f)$ pode ser conseguido através de um compromisso entre a obtenção de um pequeno erro de treino $R_{emp}(f)$ mantendo a dimensão VC a que a função *f* pertence tão pequena quanto possível.

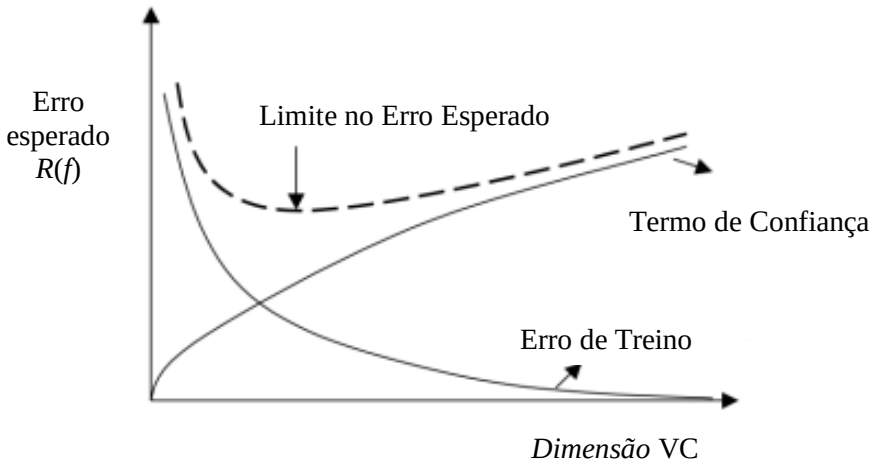


Figura 5.3: Ilustração da minimização do erro estrutural.

Embora o limite apresentado na Equação (5.5) seja importante do ponto de vista teórico, na prática surgem alguns problemas. O cálculo da dimensão VC de uma classe de funções nem sempre é fácil de obter, podendo ser desconhecida ou infinita (Muller, Mika, Ratsch, Tsuda, & Scholkopf, 2001, referenciado por (Lorena, & Carvalho, 2007)).

Para algoritmos de aprendizagem lineares em que as classes são separadas através de hiperplanos no espaço das características, como é o caso do algoritmo SVM, a dimensão VC pode ser limitada por uma quantidade chamada margem – a distância entre os padrões das duas classes mais próximos do hiperplano de separação (Vapnik, 1995). O conceito de margem é ilustrado na Figura 5.4.

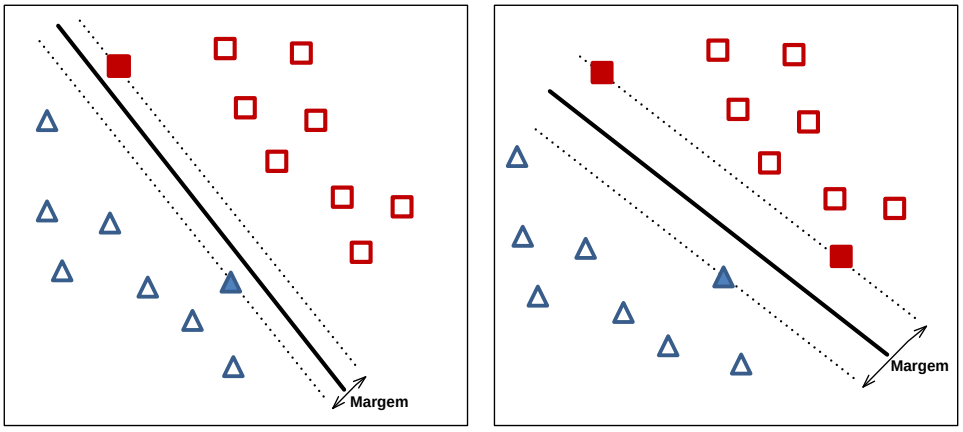


Figura 5.4: Conceito de margem. À esquerda um hiperplano com uma pequena margem de separação entre classes e à direita um hiperplano com uma margem maior.

Quanto maior o valor da margem menor a dimensão VC. Assim, de forma a manter o erro esperado pequeno (Equação (5.5)) os hiperplanos de separação das classes devem ser construídos de forma a maximizar o valor da margem e separar os padrões de treino com o menor erro de classificação possível. Estes resultados são a base do algoritmo SVM apresentado nas secções seguintes.

5.3 Algoritmo SVM

O algoritmo SVM baseia-se no princípio da minimização do erro estrutural que deriva da teoria da aprendizagem estatística, cujos princípios foram resumidamente descritos nas secções anteriores. A sua descrição será efetuada considerando inicialmente o caso de SVM lineares tendo em consideração dados separáveis e não separáveis, por fim será abordado o caso do SVM não linear.

5.3.1 *SVM para Dados Linearmente Separáveis*

A situação mais simples corresponde a um conjunto de treino em que os dados são linearmente separáveis. Considere-se o conjunto *T* de *n* padrões, cada um descrito através de um vetor de *m* características, definido por:

$$\{\mathbf{x}_i, y_i\} \qquad \mathbf{x}_i \in \Re^m, \quad y_i \in \{+1, -1\}, \ \forall i = 1, \ldots, n \ . \tag{5.6}$$

O conjunto *T* é linearmente separável se existir pelo menos um hiperplano que permita separar completamente os padrões correspondentes às duas classes.

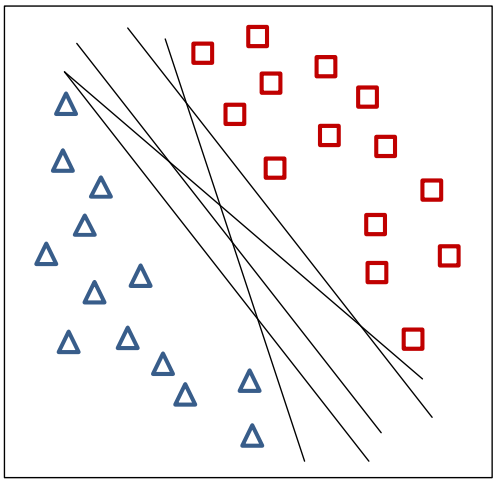


Figura 5.5: Possíveis soluções para um problema de classificação de dados linearmente separáveis.

Como ilustrado na Figura 5.5 existem vários hiperplanos que permitem a separação das duas classes. O objetivo do classificador SVM é encontrar o hiperplano de separação ótimo, que maximize o valor da margem. Esse hiperplano é formado pelos pontos **x** que satisfazem a equação:

$$\mathbf{w}.\mathbf{x} + b = 0\,, \tag{5.7}$$

sendo $\mathbf{w}$ um vetor normal ao hiperplano e $|b|/\|\mathbf{w}\|$ a distância do hiperplano à origem, com b e $\Re$. O espaço das características fica dividido em duas regiões $\mathbf{w}.\mathbf{x} + b > 0$ para a classe {+1} e $\mathbf{w}.\mathbf{x} + b < 0$ para a classe {-1}. Define-se o hiperplano canónico relativamente a T de modo a que os padrões mais próximos do hiperplano $\mathbf{w}.\mathbf{x} + b = 0$ satisfaçam a condição:

$$|\mathbf{w}.\mathbf{x}_i + b| = 1. \tag{5.8}$$

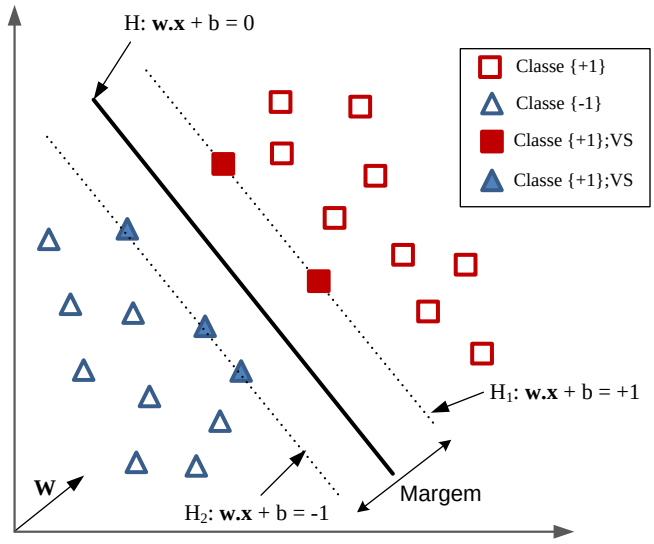


Figura 5.6: Hiperplano separador ótimo.

Seja $\mathbf{x}_1$ um ponto do hiperplano H_1, definido pelos pontos que verificam a igualdade $H_1: \mathbf{w}.\mathbf{x} + b = +1$ e $\mathbf{x}_2$ um ponto do hiperplano H_2, definido pelos pontos que verificam a igualdade $H_2: \mathbf{w}.\mathbf{x} + b = -1$. Sendo $\mathbf{w}.\mathbf{x}_1 = +1 - b$ e $\mathbf{w}.\mathbf{x}_2 = -1 - b$, considerando o comprimento do vetor que une os pontos $\mathbf{x}_1$ e $\mathbf{x}_2$, segundo a direção $\mathbf{w}$, obtemos a distância entre H_1 e H_2, isto é a largura da margem (Schölkopf et al., 1999):

$$\left(\frac{\mathbf{w}}{\|\mathbf{w}\|}.(\mathbf{x}_1 - \mathbf{x}_2)\right) = \left(\frac{\mathbf{w}.\mathbf{x}_1}{\|\mathbf{w}\|} - \frac{\mathbf{w}.\mathbf{x}_2}{\|\mathbf{w}\|}\right) = \frac{2}{\|\mathbf{w}\|}. \tag{5.9}$$

A maximização da margem pode ser obtida através da minimização de $1/2\|\mathbf{w}\|^2$. Assim, o hiperplano de separação ótimo que assegura o menor erro esperado é aquele que maximiza a

margem entre as duas classes ou dito de outra forma, minimiza a dimensão VC (Equação (5.10)) e simultaneamente permite obter um erro empírico zero (Equação (5.11)):

$$\min_{\mathbf{w},\, b} \quad \frac{1}{2}\|\mathbf{w}\|^2, \tag{5.10}$$

$$\text{sujeito a} \quad y_i(\mathbf{w}.\mathbf{x}_i + b) \geq 1, \qquad \forall i = 1, \ldots, n\,. \tag{5.11}$$

Trata-se da minimização de uma função objetivo quadrática sujeita a restrições lineares. Sendo a função objetivo convexa e o conjunto de pontos que satisfazem as restrições também convexo, este problema possui um único mínimo global. O problema de otimização pode ser resolvido através do método dos multiplicadores de Lagrange, que na sua formulação primal possui a seguinte forma:

$$L_p(\mathbf{w}, b, \alpha) = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{n} \alpha_i \left(y_i(\mathbf{w}.\mathbf{x}_i + b) - 1\right), \tag{5.12}$$

sendo $\alpha_i \geq 0,\ i = 1, \ldots, n$ os multiplicadores de Lagrange. A função Lagrangiana L_p deve ser minimizada, o que implica maximizar em relação a α_i e minimizar em relação às variáveis **w** e *b*. A formulação dual obtém-se derivando L_p relativamente a **w** e *b*, e igualando-as a zero:

$$\frac{\partial L_p(\mathbf{w}, b, \alpha)}{\partial \mathbf{w}} = 0 \Leftrightarrow \mathbf{w} - \sum_{i=1}^{n} \alpha_i\, y_i\, \mathbf{x}_i = 0 \Leftrightarrow \mathbf{w} = \sum_{i=1}^{n} \alpha_i\, y_i\, \mathbf{x}_i\,, \tag{5.13}$$

$$\frac{\partial L_p(\mathbf{w}, b, \alpha)}{\partial b} = 0 \Leftrightarrow \sum_{i=1}^{n} \alpha_i\, y_i = 0\ . \tag{5.14}$$

Substituindo as Equações (5.13) e (5.14) na Equação (5.12) obtém-se a formulação dual:

$$\max_{\alpha} \ L_d(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \left(\mathbf{x}_i.\mathbf{x}_j\right), \tag{5.15}$$

$$\text{sujeito a} \quad \begin{cases} \sum_{i=1}^{n} \alpha_i\, y_i = 0 \\ \alpha_i \geq 0, \qquad \forall i = 1, \ldots, n \end{cases}. \tag{5.16}$$

As duas formulações, primal L_p e dual L_d, permitem obter a solução ótima do problema inicial. No entanto, a formulação dual é mais atrativa uma vez que apresenta restrições mais simples e o problema de otimização recorre apenas ao produto interno dos dados de treino e respetivas classes, o que será muito útil na generalização para SVM não lineares. O treino das SVM envolve agora a maximização da formulação dual relativamente a α_i.

A existência de solução para o problema de otimização da Equação (5.15) é garantida pelas condições de Karush - Kuhn - Tucker (KKT), provenientes da teoria de otimização sujeita a restrições. Uma descrição detalhada pode ser encontrada em (Burges, 1998; Schölkopf et al., 1999). Pela "condição de complementaridade" de KKT dada por:

$$\alpha_i(y_i(\mathbf{w}.\mathbf{x}_i + b) - 1) = 0,\ i = 1, \dots, n\ . \tag{5.17}$$

Verifica-se que $\alpha_i \neq 0$ apenas para pontos que se encontram sobre os hiperplanos H_1 ou H_2, representados a cheio na Figura 5.6. Estes pontos designam-se por vetores de suporte (VS), sendo os únicos considerados no cálculo do hiperplano de separação ótimo e na classificação de novos dados.

Após a obtenção da solução do problema dual $\boldsymbol{\alpha}$, o valor de $\mathbf{w}$ é calculado através da Equação (5.13) e o valor de *b* pela Equação (5.17), sendo dado por

$$b = \frac{1}{n_{\text{vs}}} \sum_{\mathbf{x}_j \in \text{vs}} \left(\frac{1}{y_j} - \sum_{\mathbf{x}_i \in \text{vs}} \alpha_i y_i \mathbf{x}_i.\mathbf{x}_j \right), \tag{5.18}$$

onde n_{vs} é o número de VS. Uma vez estimados os parâmetros que definem o hiperplano ótimo obtém-se a função de decisão, que nos permite classificar novos dados:

$$f(\boldsymbol{x}) = \text{sgn}(\boldsymbol{w}.\boldsymbol{x} + b) = \text{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i (\boldsymbol{x}_i.\boldsymbol{x}) + b \right). \tag{5.19}$$

Assim, para um novo exemplo $\mathbf{x}$ a classe que lhe é atribuída será {+1} se o valor de $f(\mathbf{x})$ for positivo e {-1} se for negativo. Como se verifica pela análise da Equação (5.19) apenas os padrões que são VS, para os quais $\alpha_i \neq 0$, influenciam a classificação de novos padrões.

5.3.2 *SVM para Dados Não Linearmente Separáveis*

Na prática é difícil encontrar situações em que os dados sejam totalmente separáveis por fronteiras lineares, seja pela presença de ruído ou pela própria natureza dos dados. Nesta situação não é possível manter o erro empírico igual a zero.

Nesta formulação, designada por SVM de margem suave, permite-se que alguns exemplos sejam mal classificados ou se encontrem entre as margens, desrespeitando as restrições impostas pela Equação (5.11). A estes padrões é associada uma penalização através da introdução de variáveis folga ξ_i (Cortes, & Vapnik, 1995; Burges, 1998). Se o padrão de treino $\mathbf{x}_i$ se encontra mal classificado então ξ_i>1 e se $\mathbf{x}_i$ se encontra entre a margem então $\xi_i \in]0,1]$, como se encontra esquematizado na Figura 5.7.

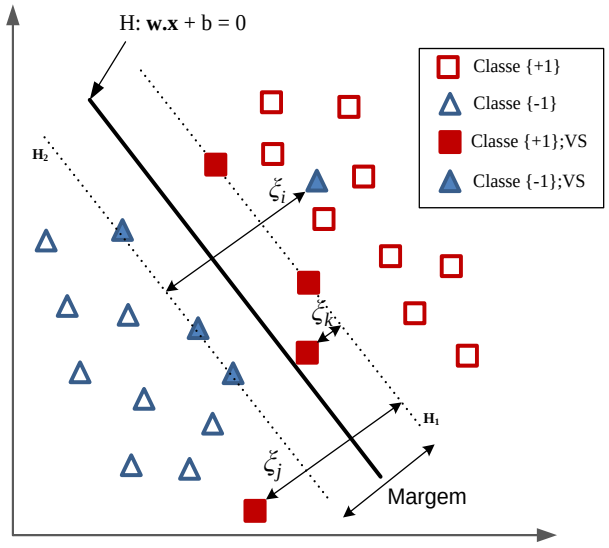


Figura 5.7: Hiperplano separador ótimo para dados não linearmente separáveis.

A função objetivo é reformulada como:

$$\min_{\mathbf{w},\, b,\, \xi} \quad \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\xi_i\,, \tag{5.20}$$

$$\text{sujeito a} \quad \begin{cases} y_i(\mathbf{w}.\mathbf{x}_i + b) \geq 1 - \xi_i \\ \xi_i \geq 0, \qquad\qquad \forall i = 1, \ldots, n \end{cases}. \tag{5.21}$$

De forma análoga à Equação (5.10), a minimização do primeiro termo da Equação (5.20) corresponde a minimizar a dimensão de VC da classe de funções em consideração, isto é à maximização da margem, de modo a garantir uma boa capacidade de generalização. Por outro lado, o termo $C\sum_{i=1}^{n}\xi_i$ coloca um limite no número de dados incorretamente classificados no conjunto de treino, portanto, a sua minimização corresponde à minimização do erro de treino e tende a diminuir a largura da margem. Assim, o parâmetro de regularização *C* é um compromisso entre a largura da margem e a minimização do erro empírico. Trata-se novamente de um problema de otimização quadrático sujeito a restrições lineares semelhante ao descrito para o caso de dados linearmente separáveis sendo os multiplicadores de Lagrange obtidos através de:

$$\max_{\alpha} L_d(\alpha) = \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\,\alpha_j\,y_i y_j\left(\mathbf{x}_i.\mathbf{x}_j\right), \tag{5.22}$$

$$\text{sujeito a} \quad \begin{cases}\sum_{i=1}^{n}\alpha_i\,y_i = 0 \\ 0 \le \alpha_i \le C, \quad \forall i = 1,\ldots,n\end{cases}. \tag{5.23}$$

Relativamente ao SVM linearmente separável os valores dos multiplicadores de Lagrange α_i são limitados pelo parâmetro *C*. Neste caso as "condições de complementaridade" de KKT são:

$$\alpha_i\,(y_i(\mathbf{w}.\mathbf{x}_i + b) - 1 + \xi_i) = 0, \tag{5.24}$$

$$(C - \alpha_i)\xi_i = 0. \tag{5.25}$$

À semelhança dos SVM de margem rígida, os dados que participam na construção do hiperplano separador são os VS, para os quais $\alpha_i > 0$. No entanto, neste caso, existem diferentes tipos de VS, como se encontra ilustrado na Figura 5.7 (Masotti, 2005; Lorena, & Carvalho, 2007). Podem ocorrer as seguintes situações:

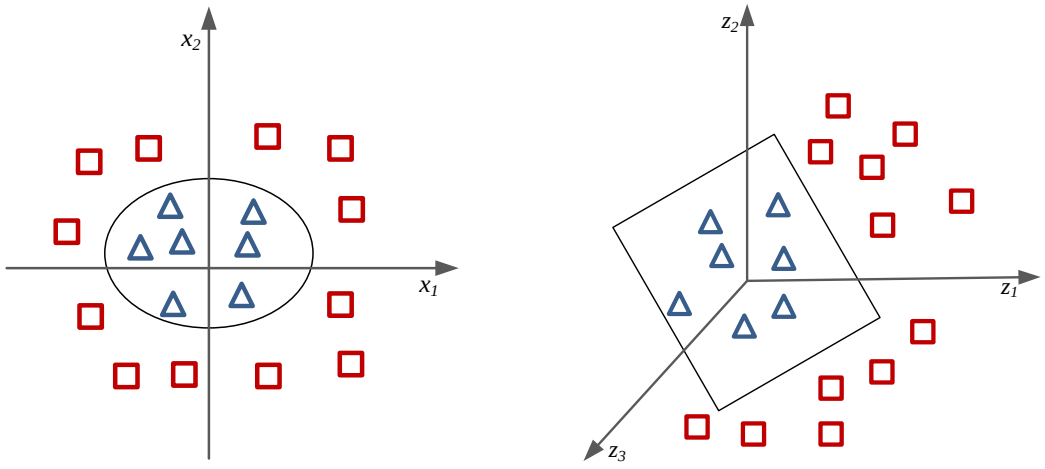
- Se $0 < \alpha_i < C$ então pela Equação (5.25) $\xi_i = 0$, e pela Equação (5.24) verifica-se que $y_i(\mathbf{w}.\mathbf{x}_i + b) = 1$, isto é, o exemplo $\mathbf{x}_i$ está bem classificado e encontra-se sobre um dos hiperplanos que delimitam a margem;

- Se $\alpha_i = C$ pode ter vários significados: erro de classificação se $\xi_i > 1$; pontos corretamente classificados, porém entre a margem, se $0 < \xi_i \leq 1$; ou pontos sobre hiperplanos que delimitam a margem se $\xi_i = 0$.

Os dados para os quais $\alpha_i = 0$ e $\xi_i = 0$ encontram-se fora das margens e corretamente classificados. A função de decisão é a mesma da Equação (5.19) obtida para dados linearmente separáveis, sendo os valores de α_i obtidos através da formulação da Equação (5.22) sujeita às restrições da Equação (5.23). O valor de *b* é calculado através da Equação (5.18) para todos os VS sobre as margens $0 < \alpha_i < C$ ($\xi_i = 0$). Os resultados obtidos mostram que a solução do algoritmo SVM é esparsa, pois apenas um número restrito de dados participam na solução. Este aspeto é muito importante do ponto de vista do tempo de execução do algoritmo, especialmente na fase de treino.

5.3.3 *SVM Não Lineares*

Em muitos casos não é possível separar satisfatoriamente no espaço de entrada os dados de treino através de um hiperplano. A estratégia do algoritmo SVM é mapear os exemplos de treino num novo espaço (espaço das características) de dimensão superior através de $\Phi: \Re^m \rightarrow \Im$. A escolha apropriada de Φ permite que os dados de treino mapeados em $\Im$ sejam separados linearmente.



a) b)

Figura 5.8: a) Padrões não separáveis linearmente em $\Re^2$. b) Mapeamento dos padrões para $\Re^3$, onde é possível a separação linear das classes.

No exemplo ilustrado na Figura 5.8, o conjunto de padrões não é linearmente separável no espaço de entrada. Através do mapeamento Φ: $\Re^2 \rightarrow \Re^3$ é possível encontrar um hiperplano de separação em $\Re^3$:

$$\Phi: \Re^2 \rightarrow \Re^3, \tag{5.26}$$

$$\Phi(\mathbf{x}) = \Phi(x_1, x_2) \rightarrow \left((x_1)^2, \sqrt{2}\, x_2\, x_1, (x_2)^2\right) = (z_1, z_2, z_3), \tag{5.27}$$

$$\mathbf{w}.\Phi(\mathbf{x}) + b = 0. \tag{5.28}$$

Substituindo $\Phi(\mathbf{x})$ na Equação (5.28) obtém-se o hiperplano de separação dado por:

$$w_1 (x_1)^2 + w_2\sqrt{2}\, x_2\, x_1 + w_3 (x_2)^2 + b = 0\,. \tag{5.29}$$

Ao mapear os dados para um espaço de maior dimensão, através de Φ, estamos perante um problema de SVM linear nesse espaço. Partindo da formulação SVM linear com margens suaves (Equações (5.20) e (5.21)), aplica-se o mapeamento Φ aos dados, obtendo-se:

$$\min_{\mathbf{w},\, b,\, \xi} \quad \frac{1}{2}\|\mathbf{w}\|^2 + \mathrm{C}\sum_{i=1}^{n} \xi_i, \tag{5.30}$$

$$\text{sujeito a} \quad \begin{cases} y_i(\mathbf{w}.\Phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \\ \xi_i \geq 0, \qquad\qquad \forall i = 1, \ldots, n \end{cases}. \tag{5.31}$$

O problema de otimização é resolvido de forma semelhante

$$\max_{\alpha} \; L_d(\alpha) = \sum_{i=1}^{n} \alpha_i \; - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\, \alpha_j\, y_i y_j \left(\Phi(\mathbf{x}_i).\Phi(\mathbf{x}_j)\right), \tag{5.32}$$

sujeito às restrições da Equação (5.23).

Uma vez que os vetores são mapeados para espaços de maior dimensão, o cálculo do produto interno $\Phi(\mathbf{x}_i).\Phi(\mathbf{x}_j)$ torna-se computacionalmente pesado. Esta complexidade é ultrapassada através do recurso a funções *kernel* que efetuam este mapeamento implicitamente. No caso do mapeamento da Equação (5.27), para $\mathbf{x}_i = (x_{i1}, x_{i2})$ e $\mathbf{x}_j = (x_{j1}, x_{j2})$ o *kernel* é dado por:

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \,.\, \mathbf{x}_j)^2, \tag{5.33}$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \left((x_{i1})^2 \,, \sqrt{2}\, x_{i2}\, x_{i1} \,, (x_{i2})^2\right).\left(\left(x_{j1}\right)^2 \,, \sqrt{2}\, x_{j2}\, x_{j1} \,, \left(x_{j2}\right)^2\right), \tag{5.34}$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i).\Phi(\mathbf{x}_j). \tag{5.35}$$

As Equações (5.33), (5.34) e (5.35) mostram que através de funções *kernel* não é necessário proceder ao mapeamento direto dos vetores de entrada através de Φ, que muitas vezes não é conhecido. Para que possam garantir a resolução do problema de otimização formulado na Equação (5.32) as funções *kernel* devem respeitar um conjunto de condições estabelecidas pelo teorema de Mercer (Schölkopf et al., 1999). Assim, recorrendo a funções *kernel*, a função de decisão é dada pela Equação (5.36), sendo apenas necessário calcular o produto interno $\Phi(\mathbf{x}_i).\Phi(\mathbf{x}_j)$:

$$f(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i \Phi(\mathbf{x}).\Phi(\mathbf{x}_i) + b\right), \tag{5.36}$$

$$f(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b\right). \tag{5.37}$$

Alguns exemplos de funções *kernel* utilizadas na prática são:

- Linear:

$$K(\mathbf{x}, \mathbf{x}_i) = (\mathbf{x}.\mathbf{x}_i). \tag{5.38}$$

- *Gaussian Radial Basis Function* (RBF):

$$K(\mathbf{x}, \mathbf{x}_i) = \exp\left(\frac{-\|\mathbf{x} - \mathbf{x}_i\|^2}{2\sigma^2}\right). \tag{5.39}$$

- Polinomial:

$$K(\mathbf{x}, \mathbf{x}_i) = (\gamma(\mathbf{x} \,.\, \mathbf{x}_i) + \delta)^d. \tag{5.40}$$

- Sigmoidal:

$$K(\mathbf{x}, \mathbf{x}_i) = \tanh(\gamma(\mathbf{x} \,.\, \mathbf{x}_i) + \delta). \tag{5.41}$$

Sendo γ, δ e d parâmetros usualmente otimizados durante a fase de treino. No caso do *kernel* sigmoidal apenas alguns valores de γ e d respeitam as condições do teorema de Mercer.

5.3.4 *Exemplos de Fronteiras de Decisão*

Existem vários *sites* [5, 6] que disponibilizam *applets* que permite simular o funcionamento de um classificador através da escolha de um conjunto de exemplos, respetivas classes, e da função *kernel*, permitindo visualizar a fronteira de decisão encontrada. A Figura 5.9 ilustra algumas das situações simuladas. Na Figura 5.9 a) e b) a função de *kernel* escolhida foi a linear, em c) e d) a função de *kernel* escolhida foi a RBF. Junto à fronteira de decisão é possível ver os exemplos (obtidos através de uma *applet* disponibilizada em [5] que estão na base da construção da fronteira de decisão – os VS, que surgem circunscritos.

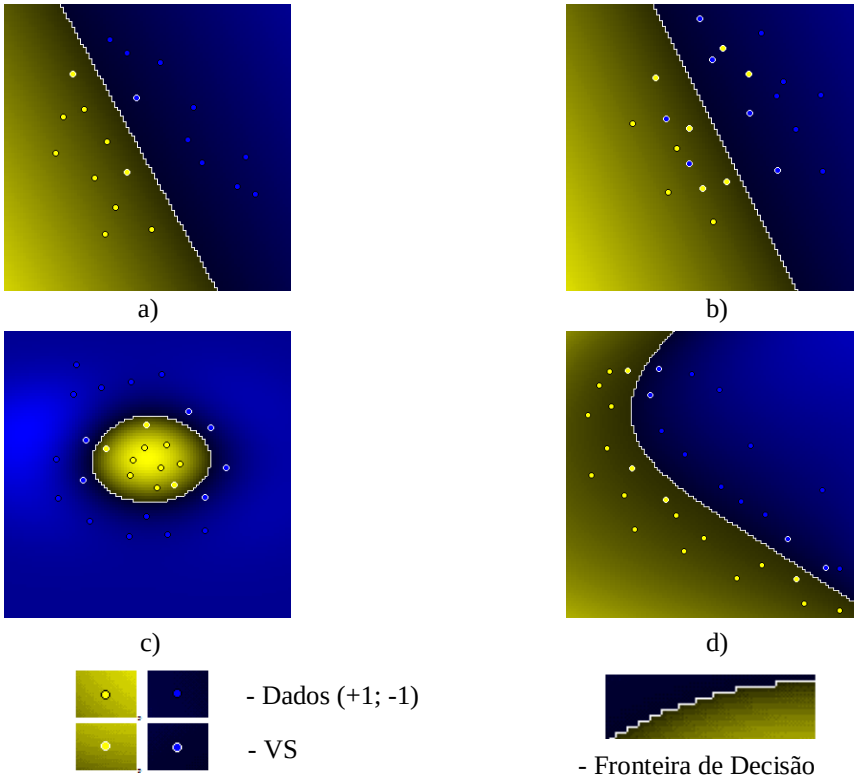


Figura 5.9: Exemplos de fronteiras de decisão. SVM lineares a) e b) e não lineares c) e d).

[5] http://svm.dcs.rhbnc.ac.uk/pagesnew/GPat.shtml, consultado em maio de 2017.
[6] http://www.csie.ntu.edu.tw/~cjlin/libsvm/, consultado em julho de 2017.

5.3.5 *SVM Multiclasse*

O algoritmo de aprendizagem SVM foi desenvolvido para cenários binários em que a função de decisão deve separar duas classes – classificador binário. No entanto, em muitas situações somos confrontados com a necessidade de classificar *c* classes. Várias estratégias têm sido propostas para estender o SVM para um cenário multiclasse. Tipicamente, estes métodos são baseados em combinações de classificadores binários para construir um classificador SVM para *c* classes (Duan, Keerthi, Chu, Shevade, & Poo, 2003). As estratégias mais populares são: "um - *versus* - um" e "um - *versus* - todos".

No método "um – *versus* – um", para *c* classes, são treinados $c(c - 1)/2$ classificadores binários, correspondentes a todos os pares de classes diferentes. No caso do classificador c_{ij} são treinados os exemplos da classe *i* e da classe *j*, sendo considerados os exemplos da classe *i* como positivos (+1) e os exemplos da classe *j* como negativos (-1). A classificação de um novo exemplo **x** é efetuada através de uma estratégia conhecida como *max-wins-voting*. Se o classificador c_{ij} indica que **x** pertence à classe *i* então os votos para esta classe são incrementados de um, caso contrário, são incrementados de um os votos atribuídos à classe *j*. Após o treino de todos os classificadores binários é atribuída ao exemplo **x** a classe com maior número de votos.

No método "um – *versus* – todos" são construídos *c* classificadores binários. No caso do treino do classificador C_i todos os exemplos da classe *i* são considerados positivos e os exemplos das restantes classes são considerados negativos. Para classificar um novo exemplo **x** todos os *c* classificadores são executados. A classe atribuída a **x** é a que corresponde ao classificador cuja função de decisão retorna o maior valor, numa estratégia conhecida como *winner-takes-all*.

A questão da classificação multiclasse tem sido amplamente analisada. Para além dos métodos anteriormente referidos e "um–*versus*–todos", suportados por múltiplos classificadores binários, têm sido propostas outras metodologias mais pesadas computacionalmente e conceptualmente mais complexas que se podem dividir em duas aproximações distintas: "*single-machine"* e "correção de erro" (Rifkin, & Klautau, 2004). A primeira aproximação constrói um classificador multiclasse através da resolução de apenas um problema de otimização. O segundo grupo baseia-se na conjugação de estratégias da teoria de codificação de correção de erros para escolher um conjunto de classificadores binários e um método para

combinar os classificadores binários. Em (Rifkin, & Klautau, 2004) é efetuada uma revisão bibliográfica de trabalhos que se inserem nas duas aproximações de classificação multiclasse atrás enumeradas e avaliado o desempenho dos classificadores, comparativamente aos métodos "um-*versus*-um" e "um–*versus*–todos". Os resultados mostraram que quando os classificadores binários são corretamente treinados, através de um ajustamento fino dos seus parâmetros, o desempenho dos métodos é semelhante. Sendo assim, segundo Rifkin e Klautau, não existe vantagens em recorrer a métodos mais complexos e computacionalmente mais exigentes.

5.4 Avaliação de Desempenho

A maioria dos sistemas de DAC apresenta uma estimativa do seu desempenho para um conjunto de dados específico. No entanto, os métodos de validação propostos na literatura utilizados para avaliar o desempenho dos sistemas variam muito, estando as métricas de desempenho fortemente dependentes da técnica de validação efetuada. Nas subsecções seguintes descrevem-se as técnicas de validação mais comuns.

5.4.1 *Técnicas de Validação*

As técnicas de validação são principalmente motivadas por dois problemas centrais em classificação de padrões: a seleção do modelo do algoritmo de aprendizagem e a estimação do desempenho. O modelo de um algoritmo de aprendizagem corresponde a uma configuração específica dos seus parâmetros para a técnica de validação usada na solução de um problema específico de classificação de padrões. No caso do algoritmo SVM, a seleção do modelo consiste em ajustar os parâmetros associados à função de *kernel* escolhida e o parâmetro de regularização *C* (Guyon, 2006).

Numa situação ideal em que o conjunto de dados é composto por um número ilimitado de instâncias, seria escolhido o modelo do algoritmo de aprendizagem que permitisse obter menor erro na totalidade dos *n* dados, e este seria o erro verdadeiro (risco verdadeiro). Claro que em aplicações reais o conjunto de dados é geralmente bastante limitado, portanto caso o treino do algoritmo de aprendizagem e a estimação do erro sejam efetuadas na totalidade do conjunto de dados leva, habitualmente, a um ajustamento do modelo aos dados (*overfitting*), sendo comum obter um erro de treino igual a zero. Como consequência deste ajustamento entre o modelo do

classificador e os dados de treino resulta uma capacidade de generalização pobre. Para ultrapassar esta situação deve-se recorrer a técnicas de validação que permitam obter uma estimativa o mais imparcial possível do erro verdadeiro através do cálculo do erro no conjunto de treino (risco empírico) (Vapnik, 1995).

Sempre que possível a estratégia de validação usada deve simular as condições para as quais o sistema foi desenvolvido. No caso de um sistema de DAC o objetivo central é a sua integração na rotina de um serviço de saúde. No caso concreto do sistema proposto neste trabalho, a predição da classe associada a novos padrões será baseada no treino e teste do algoritmo de aprendizagem no conjunto de dados recolhidos, que constitui a referência (*ground-truth*).

As técnicas de validação baseadas em amostragem são as que permitem obter resultados mais fidedignos (Duda, Hart, & Stork, 2012). O princípio subjacente a esta técnica consiste em recorrer a conjuntos de dados disjuntos para treinar e testar o algoritmo de aprendizagem. Algumas das técnicas de validação mais comuns são a validação cruzada e o *holdout*. A validação cruzada pode subdividir-se em: validação cruzada *k-fold* e validação cruzada *leave-one-out* (LOO).

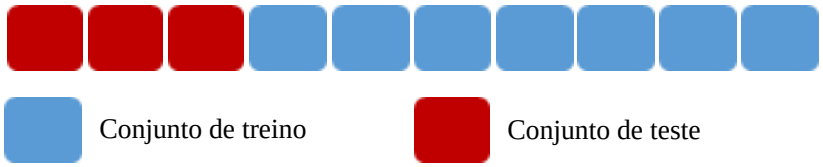


Figura 5.10: A técnica *holdout* com $p = 0.7$.

A técnica *holdout* consiste na divisão da totalidade do conjunto de dados em dois conjuntos mutuamente exclusivos, o conjunto de treino p e o conjunto de teste $(1-p)$. São comuns as divisões do conjunto total dos dados em partes iguais ($p = 0.5$) ou em 70% para o conjunto de treino e 30% para o conjunto de teste ($p = 0.7$), como se encontra esquematizado na Figura 5.10.

A técnica *holdout* dá origem a uma estimativa pessimista do desempenho do classificador, uma vez que apenas uma parte dos dados é utilizado no treino do algoritmo de aprendizagem. Por outro lado, o facto de este método dividir os dados apenas uma vez, o desempenho do classificador pode variar muito entre diferentes partições dos dados. Pode acontecer que as instâncias mais difíceis de classificar corretamente se agrupem no conjunto de teste dando

origem a um erro demasiado pessimista ou, pelo contrário, as instâncias mais simples de classificar constituam o conjunto de teste dando origem a um erro demasiado otimista.

Nas técnicas de validação cruzada todos os dados são utilizados para treino e para teste de uma forma rotativa. Por exemplo, a técnica de validação cruzada *k-fold* consiste em dividir aleatoriamente o conjunto de dados em *k* partições mutuamente disjuntas de tamanho aproximadamente igual, conhecidas como *folds*. No caso particular da validação cruzada *k-fold* estratificada a proporção das instâncias de cada classe no conjunto de treino e teste devem ser aproximadamente as mesmas que na base de dados original.

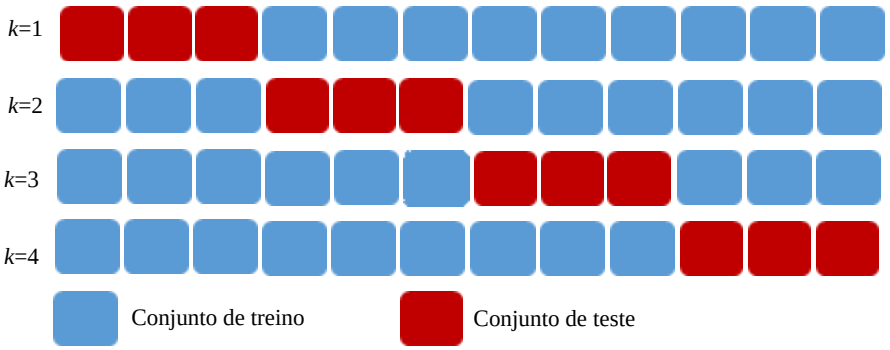


Figura 5.11: A técnica de validação cruzada *k-fold* (*k* = 4).

De modo a avaliar a capacidade de generalização do modelo do classificador, uma das partições é separada para teste (conjunto de teste) e as restantes *k*-1 partições são concatenadas e usadas para treino do algoritmo de aprendizagem (conjunto de treino). O conjunto de teste simula instâncias novas, cujas classes são desconhecidas. De modo a reduzir a variância das métricas de desempenho, o processo é repetido *k* vezes, permutando de forma circular as partições (ver Figura 5.11) (Masotti, 2005; Depeursinge, 2010). O erro verdadeiro (*Erro*) é estimado através da média dos erros obtidos nos conjuntos de teste ao longo das *k* iterações (Equação (5.42). A divisão da base de dados em *k* partições permite evitar a estimação do erro verdadeiro em partições pouco representativas da totalidade dos dados:

$$Erro = \frac{1}{k}\sum_{i=1}^{k} Erro_i \,. \tag{5.42}$$

Um caso particular da técnica de validação cruzada *k-fold* é a validação cruzada LOO. Nesta técnica o tamanho das partições é de 1, isto é, o número de partições *k* corresponde ao número de instâncias *n*. O treino é efetuado em *k*-1 instâncias e o teste na instância sobrante. São efetuadas *n* iterações com diferentes conjuntos de treino de tamanho *n*-1. A estimação do erro verdadeiro é efetuado recorrendo à Equação (5.42), substituindo *k* por *n*. Apesar da técnica de validação cruzada LOO permitir uma estimativa imparcial (*unbiased*) do erro verdadeiro, a sua variância é elevada. Por outro lado, uma vez que são treinados e testados *k* modelos diferentes, esta técnica torna-se computacionalmente pesada sendo apropriada para bases de dados de dimensão relativamente pequena.

O valor escolhido para *k* depende essencialmente tamanho da base de dados. Caso o valor de *k* seja elevado a estimativa do erro verdadeiro será bastante precisa, mas com uma variância e custo computacional elevados, dado o número elevado de iterações *k*. Com um valor de *k* pequeno a variância e o custo computacional serão reduzidos, mas a estimativa do erro verdadeiro será conservativa ou maior que o erro verdadeiro. Caso o conjunto de dados seja pequeno mesmo um valor de $k = 3$ permite obter estimativas corretas. Por outro lado, caso existam poucos dados de cada classe o recurso à técnica validação cruzada LOO será uma opção a considerar, para que o conjunto de treino integre o maior número de exemplos possíveis (Mather, & Tso, 2016). Em aplicações práticas o valor de $k = 10$ é bastante usual.

No contexto de um sistema de DAC baseado em imagem as instâncias correspondem habitualmente a ROIs. Estas podem assumir várias formas e dimensões, serem áreas delineadas manualmente ou selecionadas de forma automática com um tamanho e forma uniforme, ou mesmo corresponder à totalidade da imagem. Um aspeto para o qual se deve estar atento na escolha da estratégia de validação e no modo como ela é implementada, é a forma como se efetua a divisão das ROIs entre o conjunto de treino e o conjunto de teste. Geralmente, de cada paciente, e por vezes da mesma imagem, são retiradas várias áreas de interesse que posteriormente são subdivididas em ROIs. Habitualmente, o número de ROIs é superior, sendo frequentemente muito superior, ao número de pacientes. A existência simultaneamente nos conjuntos de treino e teste de instâncias pertencentes à mesma área de interesse, à mesma imagem ou ao mesmo paciente pode influenciar positivamente o desempenho do classificador. O recurso à validação cruzada baseada no paciente, designada por *Leave-One-Pacient-Out* (LOPO) permite minimizar esta questão (Dundar et al., 2004; Depeursinge, 2010). Na validação

cruzada LOPO o número de partições é igual ao número de pacientes $k = N_p$, portanto o número de classificadores a treinar é N_p. O conjunto de teste é constituído pelas ROIs de um paciente e o conjunto de treino reúne todas as ROIs correspondentes a N_p-1. Na LOPO é assegurado que todas as instâncias que pertencem ao mesmo paciente se encontram no mesmo conjunto, ou de treino ou de teste.

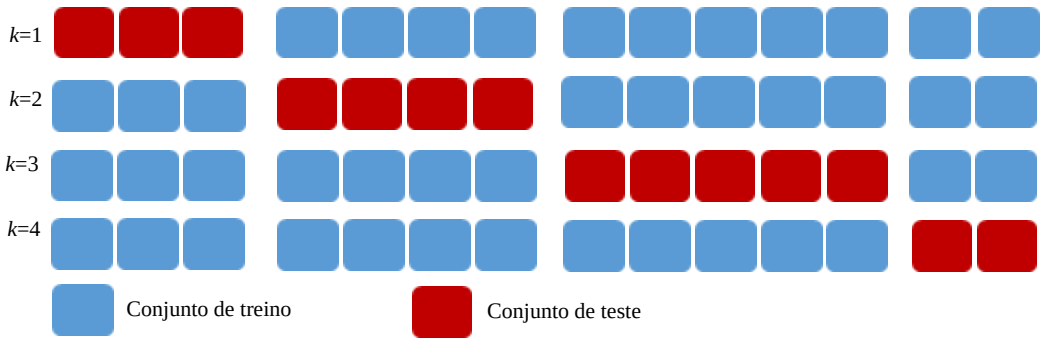


Figura 5.12: Validação cruzada LOPO para $k = N_p = 4$ pacientes.

Esta técnica é ilustrada na Figura 5.12 para $N_p = 4$. Relativamente à validação cruzada, onde os conjuntos são gerados aleatoriamente, a validação cruzada LOPO tem a vantagem de eliminar a variância das medidas de desempenho, uma vez que os conjuntos de treino e teste são sempre constituídos da mesma forma. Esta situação é desejável na pesquisa dos parâmetros ótimos, que não é influenciada pela aleatoriedade da divisão das instâncias.

5.4.2 *Métricas*

Um aspeto essencial no projeto de um classificador e, em particular, num sistema de DAC é a avaliação do seu desempenho. Para o cálculo das métricas de desempenho do classificador é construída a matriz de confusão, também conhecida como matriz de classificação, onde são categorizadas as predições corretas e incorretas do classificador, para todas as classes. O cálculo das métricas que permitem avaliar o desempenho do classificador é efetuado com base na tabela de contingência, também conhecido pela matriz de confusão, como exemplificado na Tabela 5.1 para um cenário multiclasse. Cada elemento de matriz tem dois índices, o primeiro que corresponde à classe real, enquanto o segundo à classe predita pelo classificador.

Tabela 5.1: Matriz de confusão para um cenário de *c* classes.

	Classe Predita		
Classe Real	Classe 1	⋯	Classe *c*
Classe 1	a_{11}	⋯	a_{1c}
⋮	⋮	⋱	⋮
Classe *c*	a_{c1}	⋯	a_{cc}

Na matriz de confusão, os valores da diagonal principal correspondem a decisões acertadas, isto é, a instância foi classificada na sua classe correta. Os valores fora da diagonal principal correspondem a erros de classificação. Os elementos da diagonal principal possuem índices iguais (a_{11}, a_{22} ,…, a_{cc}) que representam as classificações corretas para todas as classes. Todos os outros elementos da matriz correspondem às classificações incorretas. Por exemplo, a_{31} indica o número de instâncias da classe 3 que foram erroneamente classificadas como sendo da classe 1. Após o processo de classificação, a matriz de confusão é preenchida com as classificações de cada uma das classes e a partir dela é possível extrair quatro medidas específicas para cada uma classe:

- Verdadeiros Positivos da classe *i* (VP_i): número de instâncias da classe *i* (instâncias positivas) que são corretamente classificadas como sendo da classe *i*.

- Verdadeiros Negativos da classe *i* (VN_i): número de instâncias de todas as classes, à exceção da classe *i* (instâncias negativas), que não são classificadas como classe *i*.

- Falsos Positivos da classe *i* (FP_i): número de instâncias que não sendo da classe *i* são classificadas como classe *i*.

- Falsos Negativos da classe *i* (FN_i): número de instâncias da classe *i* que são classificadas como outras classes.

Estas medidas são posteriormente utilizadas para o cálculo das métricas que permitem avaliar o desempenho do classificador. Uma métrica comum na avaliação de desempenho global do classificador é a Exatidão Global, que mede a fração de instâncias corretamente classificadas de todas as classes e Exatidão(*i*) que corresponde à exatidão da classe *i*. A Sensibilidade(*i*) da classe *i* mede a fração das instâncias dessa classe que estão corretamente classificadas, enquanto

a Precisão(i) da classe i mede a fração das instâncias corretamente classificadas da classe i, relativamente a todas as instâncias classificadas como classe i. Por outro lado, a Especificidade(i) da classe i mede a fração de instâncias negativas da classe que são classificadas corretamente. Estas métricas podem ser calculadas através das equações:

$$Sensibilidade(i) = \frac{VP_i}{VP_i + FN_i} = \frac{a_{ii}}{\sum_j a_{ij}}. \tag{5.43}$$

$$Precisão(i) = \frac{VP_i}{VP_i + FP_i} = \frac{a_{ii}}{\sum_j a_{ji}}. \tag{5.44}$$

$$Especificidade(i) = \frac{VN_i}{VN_i + FP_i} = \frac{\sum_i \sum_j a_{ij} - \sum_j a_{ij-} \sum_{j \neq i} a_{ji}}{\sum_i \sum_j a_{ij} - \sum_j a_{ij}}. \tag{5.45}$$

$$Exatidão(i) = \frac{VP_i + VN_i}{VP_i + VN_i + FP_i + FN_i} = \frac{\sum_i \sum_j a_{ij} - \sum_{j \neq i} a_{ij-} \sum_{j \neq i} a_{ji}}{\sum_i \sum_j a_{ij}}. \tag{5.46}$$

$$Exatidão\ Global = \frac{\sum_i VP_i}{\sum_i \sum_j (VP_i + VN_i + FP_i + FN_i)} = \frac{\sum_{i=j} a_{ij}}{\sum_i \sum_j a_{ij}}. \tag{5.47}$$

Muitos sistemas de DAC visam a diferenciação de instâncias patológicas, isto é, que correspondem à presença de doença, de instâncias normais ou ausência de doença. Neste cenário uma instância positiva corresponde a uma ROI onde está presente um padrão patológico. Uma instância negativa é uma ROI correspondente a um padrão normal, isto é, um tecido pulmonar considerado saudável. A matriz de confusão surge simplificada com a dimensão 2×2 sendo uma simplificação da classificação multiclasse descritas nas subsecções anteriores.

5.5 Considerações Finais

Neste capítulo foi introduzida a teoria da aprendizagem estatística na qual o algoritmo de aprendizagem SVM se baseia. Foram descritos os princípios matemáticos subjacentes ao algoritmo SVM nas suas diferentes formulações. Este algoritmo apresenta características que o tornam num classificador popular, tais como, a ausência de mínimos locais, possuir uma solução esparsa que depende apenas dos vetores de suporte, depender de poucos parâmetros e apresentar uma boa capacidade de generalização na presença de novos exemplos.

O treino do algoritmo de aprendizagem deve ser efetuado recorrendo a técnicas de validação que permitam uma correta estimação do erro verdadeiro e que não se ajustem ao conjunto de dados de treino. Neste capítulo foram descritas técnicas de validação que evitam a situação descrita. Foram também apresentadas métricas que permitem avaliar o desempenho do modelo do classificador.

Capítulo 6
SISTEMA DE DAC PROPOSTO

O sistema de DAC proposto neste trabalho foi inspirado no sistema de visão do imagiologista. Numa primeira fase são analisadas as áreas de interesse e extraídas as informações relevantes e, posteriormente, ocorre a tomada de decisão, sustentada por uma aprendizagem anterior resultante da sua experiência. No sistema proposto, a informação das regiões pulmonares é concentrada num conjunto de medidas descritivas da sua textura, obtidas através de diferentes métodos de análise de textura. Numa segunda fase, o sistema é induzido a aprender a reconhecer as características discriminatórias de cada classe de padrões em estudo.

Ao longo do capítulo é apresentado o sistema de DAC proposto que visa a classificação de padrões pulmonares em imagens de TCAR. Foi desenvolvida uma aplicação designada Lung Study Tool cujas funcionalidades são descritas. Através desta aplicação foram recolhidas e anotadas amostras de padrões pulmonares a partir de imagens de TCAR de pacientes com pulmão normal, com diagnóstico confirmado de enfisema e *DIP, que permitiram a construção de duas bases de dados. As amostras recolhidas são determinantes para o melhor conhecimento do aspeto visual dos padrões pulmonares e da sua distribuição, assim como para o treino e teste do sistema.*

6.1 Aplicação Lung Study Tool

Com o objetivo de construir bases de dados de padrões pulmonares obtidos a partir de sequências de imagens de TCAR do tórax foi desenvolvida uma aplicação designada por *Lung Study Tool* (Vasconcelos, Silva, & Barroso, 2009). Esta aplicação visa agilizar o processo de visualização das sequências de imagens e facilitar a seleção das áreas de interesse e respetiva

caracterização. Foi desenvolvida uma aplicação intuitiva, *user-friendly* e que minimiza a digitação de dados pelo utilizador, tanto no ato da recolha e anotação das áreas de interesse, como no módulo de visualização.

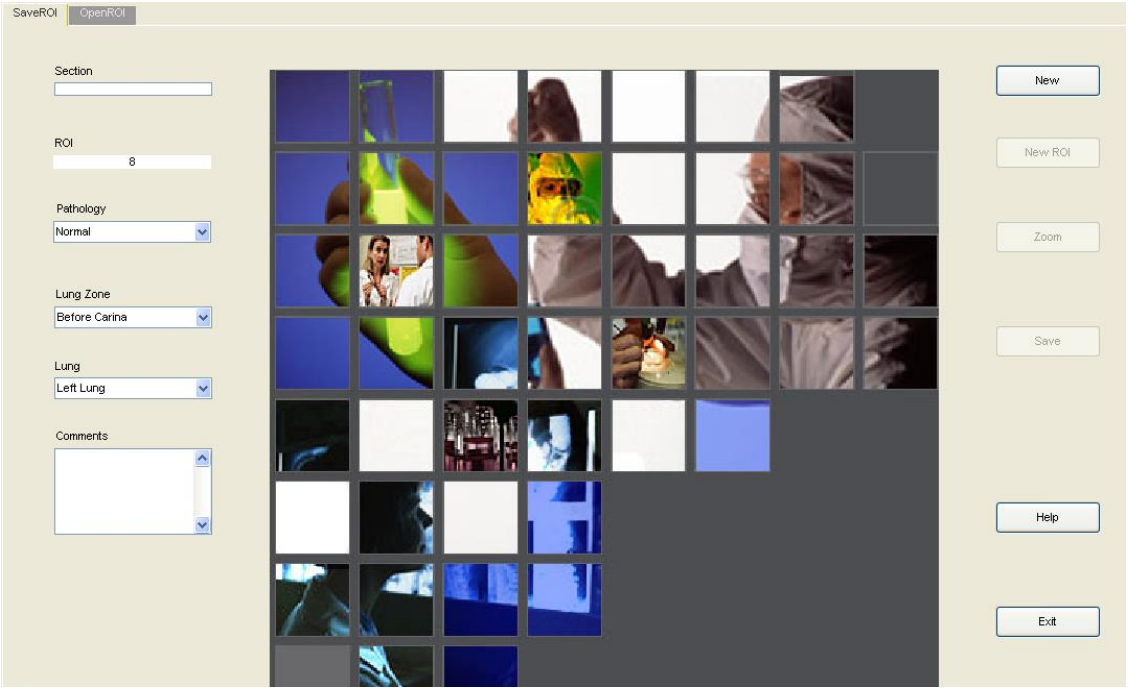


Figura 6.1: Ecrã inicial da aplicação *Lung Study Tool*.

Esta aplicação foi projetada com dois objetivos principais, aos quais correspondem dois separadores distintos na aplicação: "SaveROI" e "OpenROI". Os dados guardados pela aplicação *Lung Study Tool* encontram-se num formato propício ao desenvolvimento de aplicações que integrem uma classificação supervisionada. Nas subsecções seguintes são descritas as suas potencialidades e modo de funcionamento. De forma a reduzir os erros de digitação recorreu-se, sempre que possível, ao preenchimento automático dos campos ou ao recurso a menus *pop-up*. O ecrã de entrada da aplicação pode-se visualizar na Figura 6.1.

A aplicação foi desenvolvida em MATLAB (MATLAB 2010b) a partir de uma *Graphical User Interface* (GUI) e através de um "TabPanel" com os dois separadores referidos. De modo a que a aplicação possa ser instalada sem que para isso seja necessária a instalação da totalidade do *software* MATLAB foi gerado um pacote de instalação que ao ser descompactado executa a *Matlab Component Runtime* e instala a aplicação *Lung Study Tool*. Em futuras utilizações a aplicação comporta-se como uma aplicação *stand-alone* tradicional.

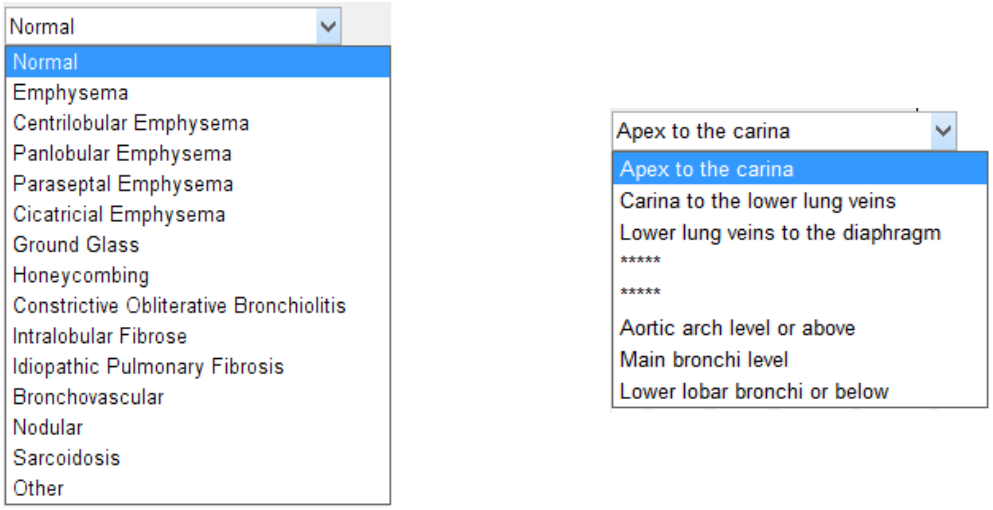
As funcionalidades da aplicação foram sendo ajustadas de acordo com as necessidades do imagiologista, sempre com o objetivo de otimizar e simplificar as tarefas monótonas e consumidoras de tempo da recolha e caracterização dos dados. Os conteúdos de carácter médico como as patologias, padrões pulmonares de interesse e a região anatómica da seleção das áreas de interesse foram escolhidas de acordo com a informação presente em trabalhos afins, a opinião dos imagiologistas e os exames disponíveis à data.

6.1.1 *Modo de Funcionamento do Separador "SaveROI"*

O modo de funcionamento correspondente ao separador "SaveROI" permite a navegação ao longo da série de imagens que constituem cada exame, o que possibilita ao utilizador ter uma visão geral da distribuição das patologias/padrões ao longo dos pulmões.

Figura 6.2: Ecrã da aplicação *Lung Study Tool* em modo de funcionamento "SaveROI".

Na Figura 6.2 é mostrado o ecrã da aplicação no modo de visualização do exame, seleção e caracterização de áreas de interesse que corresponde ao separador "SaveROI". É possível a navegação em modo manual, visualizando imagens consecutivas, a visualização de uma imagem específica através da indicação da sua posição relativa no conjunto das imagens que constituem o exame ou em modo automático (*slideshow*). Nesta opção é possível configurar o intervalo temporal entre a visualização de imagens consecutivas.



a) b)

Figura 6.3: a) Opções disponíveis para caracterizar o padrão/patologia pulmonar. b) Nível anatómico pulmonar da área selecionada.

A seleção da área de interesse é iniciada clicando no botão "NewROI". O número da FH-ROI (do Inglês *Free Hand* – ROI) é atribuído de modo sequencial, relativamente ao número da última FH-ROI guardada, na atual sessão de marcações ou em sessões anteriores. A marcação é efetuada através da digitação de um conjunto de pontos que contornem a região a selecionar, com um mínimo de três pontos. O fecho do contorno da FH-ROI é conseguido através de duplo clique do rato. Dada a diversidade da forma das regiões a marcar e a sua presença em zonas de fronteira do pulmão (região subpleural, por exemplo), considerou-se que a flexibilidade dada por esta metodologia de contornar a FH-ROI permite uma maior precisão, comparativamente à seleção de áreas de interesse através de formas geométricas predefinidas. A seleção da região pode ser efetuada após a ampliação da zona da imagem recorrendo ao botão "Zoom". Após a marcação da FH-ROI é necessário proceder à sua caraterização de acordo com:

- Secção a que corresponde a imagem;
- Padrão ou patologia presente. As opções disponíveis são as apresentadas na Figura 6.3 a) resultante da extensão do menu de *pop-up* "Pathology";
- Pulmão esquerdo ou direito;
- Nível anatómico pulmonar a partir da qual a FH-ROI do pulmão é selecionada. As duas opções presentes na aplicação foram as escolhidas pelo imagiologista, após sugestão de

várias alternativas resultantes de uma pesquisa bibliográfica e podem ser observadas na Figura 6.3 b). Opção 1 (Kauczor et al., 2000): "Entre o vértice e a carina", "Entre a carina e as veias pulmonares inferiores" e "Entre as veias pulmonares inferiores e o diafragma". Na opção 2 (Uchiyama et al., 2003) os limites anatómicos são: "Nível do arco aórtico ou acima", "Nível dos brônquios principais" e "Nível dos brônquios lobulares inferiores ou abaixo";

- Observações consideradas relevantes para a caracterização da FH-ROI podem ser adicionadas ao campo de texto "Comments".

Após a seleção e caracterização da FH-ROI é necessária a confirmação por parte do utilizador para que os dados da FH-ROI e respetiva caracterização sejam adicionados à base de dados. Para além da informação descrita anteriormente é ainda guardada automaticamente na base de dados a seguinte informação de cada FH-ROI:

- Data e hora da marcação;
- Nome do ficheiro DICOM;
- As coordenadas e respetivos níveis de cinzento, na escala de Hounsfield de todos os pontos da FH-ROI, isto é, que se encontram dentro do contorno efetuado;
- Os dados da *spline* utilizados para desenhar o contorno da FH-ROI: as coordenadas dos pontos, as coordenadas do seu centro e a cor do contorno. Estes dados permitem a reconstrução fidedigna da FH-ROI no modo "OpenROI".

6.1.2 *Modo de Funcionamento do Separador "OpenROI"*

No modo de funcionamento "OpenROI" a base de dados é carregada e todos os ficheiros a partir dos quais foram selecionadas FH-ROIs são listados. A seleção de um destes ficheiros permite a visualização das FH-ROIs marcadas e respetivas anotações, permitindo a confirmação do conjunto de dados selecionados. Na Figura 6.4 podemos observar o ecrã no modo "OpenROI" que permite a visualização das marcações e anotações existentes.

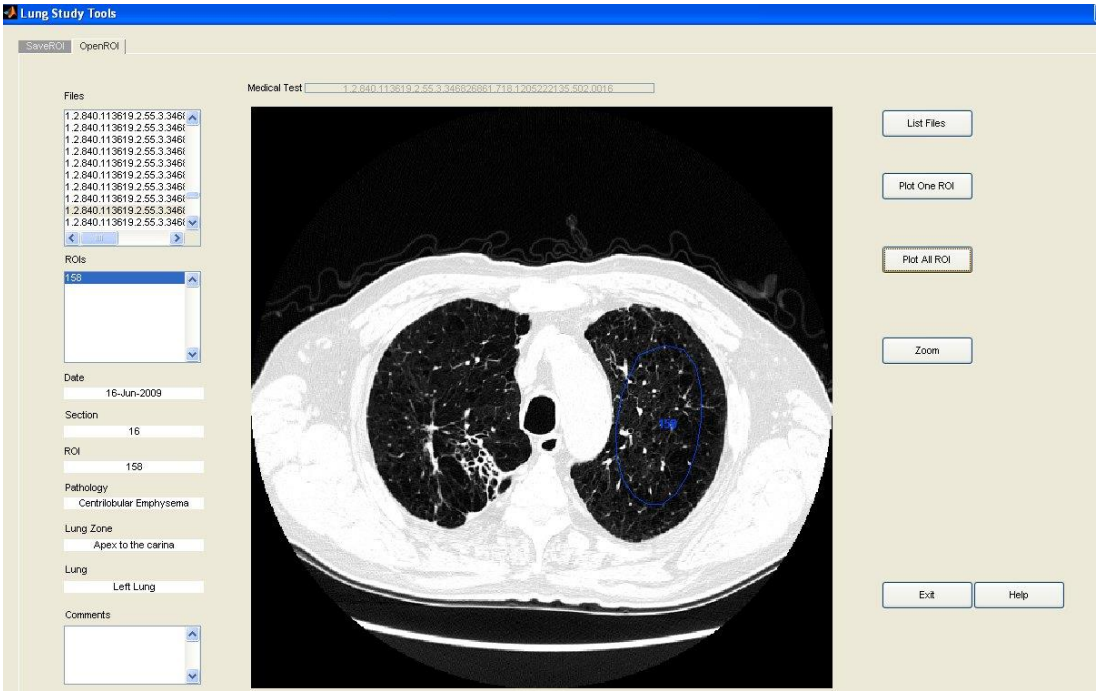


Figura 6.4: Ecrã da aplicação *Lung Study Tool* em modo "OpenROI".

Esta componente da aplicação *Lung Study Tool* pode ser utilizada como apoio ao estudo de patologias e padrões pulmonares em imagens TC, possibilitando aos seus utilizadores:

- Compreender a anatomia pulmonar;
- Compreender o aspeto visual das patologias pulmonares em imagens de TCAR e os padrões pulmonares associados;
- Reconhecer a aparência e distribuição dos padrões pulmonares em diferentes fases da doença;
- Aprender a utilizar o conhecimento sobre os padrões pulmonares e sua distribuição para estabelecer o diagnóstico ou para restringir as possibilidades de diagnóstico.

6.2 Caracterização das Bases de Dados

Os dados usados no âmbito deste trabalho foram recolhidos com a colaboração de imagiologistas do Departamento de Imagiologia do CHUC. O acesso a dados de pacientes é um assunto delicado devido aos problemas éticos que coloca. As equipas de investigação que manipulam dados médicos têm, por isso, o compromisso de proteger a identidade dos pacientes

e médicos envolvidos. Para que a recolha de dados fosse possível foi necessário proceder a um pedido de colaboração junto do Diretor do Serviço de Imagiologia, à data, o Professor Doutor Filipe Caseiro Alves.

Na Figura 6.5 a) pode-se observar uma das salas do Departamento de Imagiologia do CHUC onde se efetuam exames de TC e na Figura 6.5 b) o tomógrafo LightSpeed VCT 64 da General Electric Healthcare, que corresponde ao modelo utilizado na aquisição dos dados. A seleção e anotação de exames considerados relevantes, no ato imediato à sua aquisição, não faz parte do *modus operandi* do Departamento de Imagiologia do CHUC. Assim, para a construção das bases de dados foi necessário, *a posteriori*, localizar e aceder a exames e relatórios de pacientes diagnosticados com a patologia pulmonar de interesse. O imagiologista teve de visualizar, novamente, o exame e o respetivo relatório de modo a efetuar as marcações e respetivas anotações, tornando a recolha de dados um processo bastante moroso.

a)

b)

Figura 6.5: a) Visão geral de uma sala do Departamento de Imagiologia do CHUC. b) *Scanner* da General Electric Healthcare.

A análise dos padrões do tecido pulmonar em exames TCAR é uma tarefa complexa e demorada mesmo para imagiologistas experientes. A baixa concordância intra e interobservador verificada na marcação de exames de padrões associados à DIP são reveladores da dificuldade em interpretar algumas imagens. De forma a garantir a fiabilidade dos dados, alguns estudos recorrem à opinião de vários imagiologistas de modo a encontrar um consenso e apenas as regiões onde existe concordância são consideradas (Uppaluri et al., 1999b; Uchiyama et al., 2003; Sluimer et al., 2003; Kim et al., 2009; Lee et al., 2009; Depeursinge, 2012a; Song et al.,

2013). Neste trabalho, a aquisição dos dados foi efetuada por fases de acordo com a disponibilidade dos médicos para proceder à pesquisa e marcação dos exames. As bases de dados foram reunidas maioritariamente entre novembro de 2008 e dezembro de 2012 tendo sido elaboradas com a colaboração, em regra, de dois imagiologistas.

Foram apenas considerados exames de pacientes aos quais não houve administração de agentes contrastantes. Os exames de TC são constituídos por conjuntos de imagens armazenadas de acordo com a norma DICOM em ficheiros de 512×512 pixels, com 16 bits por pixel e com um tamanho de pixel médio de 0.69 mm, variando de 0.56×0.56 mm^2 a 0.79×0.79 mm^2. A espessura média das secções é de 1.25 mm, variando de 0.74 mm a 1.3 mm, a tensão no tubo é de 140 kV e a corrente no tubo varia de 300 mA a 580 mA. A seleção das regiões foi sempre efetuada em cortes axiais. A área e a forma de cada área delimitada dependem do tamanho e localização dos padrões pulmonares que o imagiologista pretende selecionar.

Figura 6.6: Exemplo de uma região selecionada pelo imagiologista (contorno a azul) e a grelha que permite obter ROIs de 40×40 pixels.

De modo a obter conjuntos de ROIs de forma e tamanho uniforme cada FH-ROI é coberta por uma grelha que permite a seleção de ROIs de dimensão configurável. A grelha pode ser gerada de forma a dar origem a ROIs contínuas, não sobrepostas ou sobrepostas, de forma quadrangular ou retangular. Apenas as ROIs que se encontram 100% dentro do contorno das FH-ROIs são consideradas para análise. Cada ROI guarda a referência da FH-ROI a partir da qual foi extraída. Por exemplo, a FH-ROI*x*S*y* corresponde à ROI *y* extraída da FH-ROI *x*. Na Figura 6.6 é ilustrada a metodologia de obtenção de ROIs de 40×40 pixels, sem sobreposição, a partir do contorno (linha azul) da região selecionada pelo imagiologista. Na situação exemplificada só as ROIs que se encontram completamente dentro dos limites do contorno são adicionadas à base de dados, neste caso as ROIs 8, 13, 14 e 18.

6.2.1 *Base de Dados I - Enfisema versus Normal*

A recolha de dados foi efetuada progressivamente consoante a disponibilidade dos imagiologistas associados ao projeto e ao aparecimento de pacientes diagnosticados com as patologias pretendidas. Numa primeira fase, de forma a ser possível dar início aos primeiros testes foram recolhidos essencialmente exames de pacientes com enfisema e de pacientes com pulmão saudável. A escolha do enfisema pulmonar surgiu naturalmente por ser uma patologia comum e frequente na prática clínica do CHUC, sendo possível obter marcações com alguma rapidez. O número de FH-ROIs selecionados por exame/paciente foi deixado ao critério do imagiologista, com a restrição de selecionar apenas uma FH-ROIs de cada lado do pulmão, para evitar redundâncias.

A primeira base de dados constituída foi assim criada a partir da seleção de 290 regiões de 82 indivíduos diferentes, 55 do sexo masculino e 27 do sexo feminino, com uma idade média de 65 ± 15 anos. As FH-ROIs encontram-se distribuídas da seguinte forma: 185 FH-ROIs com o padrão enfisema, que englobam vários tipos e graus de severidade e 105 FH-ROIs com pulmão normal. Estas FH-ROIs foram subdivididas em ROIs de diferentes tamanhos dando origem a uma base de dados de 1 311 ROIs de 40×40 pixels, 3 334 ROIs de 30×30 pixels e 9 655 ROIs de 20×20 pixels, considerando a obtenção das ROIs com uma sobreposição de 50%. Para todas as dimensões escolhidas o rácio entre o número de ROIs com a presença de enfisema e ROIs consideradas normais é mantido sensivelmente o mesmo: 70% de ROIs com a presença de enfisema e 30% normal.

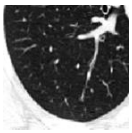
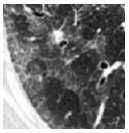
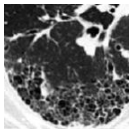
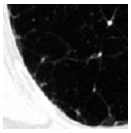
Numa segunda fase foram analisadas cuidadosamente as FH-ROIs que integram a Base de Dados I e os resultados obtidos a partir dela. Verificaram-se situações que nos levaram a remover algumas FH-ROIs marcadas. Destacamos as seguintes:

- Algumas FH-ROIs ao serem marcadas englobaram ligeiramente parte do abdómen. Apesar da região exterior acabar por ser maioritariamente descartada no processo de divisão em ROIs, verificou-se a existência de algumas ROIs incorretas;
- A dimensão de algumas FH-ROIs era demasiado pequena, o que após a sua subdivisão em ROIS levava a um número pequeno de ROIs por paciente.

6.2.2 *Base de Dados II - Padrões Intersticiais*

De forma a incluir na base de dados outros padrões imagiológicos característicos da DIP, a recolha de exames e subsequente marcação foi continuada e adicionada à Base de Dados I. Foram reunidos exames de TCAR de 57 sujeitos (29 mulheres e 28 homens) com uma idade média de 61 ± 16 anos, os imagiologistas selecionaram 417 FH-ROIs. Como descrito previamente, cada região correspondente a cada FH-ROI é amostrada. Na Tabela 6.1 está resumido o conteúdo da Base de Dados II - Padrões Intersticiais considerando ROIs de 40×40 pixels obtidas de forma não sobreposta, num total de 1 261.

Tabela 6.1: Resumo do conteúdo da Base de Dados II – Padrões Intersticiais.

Classe	Normal	Vidro Despolido	Favo de Mel	Enfisema
Aspeto Visual				
Nº de Pacientes	16	20	7	14
Nº de FH-ROIs	87	166	72	92
Nº de ROIs	253	396	217	395

6.3 Testes Experimentais

Uma das dificuldades que existe na implementação de um sistema de DAC é a grande quantidade de opções que se colocam. No caso do sistema proposto foi necessário tomar várias decisões e considerar múltiplas hipóteses, que influenciam, umas mais que outras, o desempenho global do sistema, nomeadamente:

- Que patologias/padrões devem ser integradas na base de dados e, portanto, ser objeto de estudo. Este aspeto foi muitíssimo condicionado pela disponibilidade dos exames à data da recolha dos dados;

- Que tipo de ROIs usar: tamanho, forma, níveis de quantização, zona pulmonar e secções de onde são extraídas;

- Como caracterizar cada ROI. Que características devem integrar o vetor que descreve cada uma. Que tipo de pré-processamento efetuar, por exemplo, proceder ou não à segmentação das estruturas broncovasculares;

- Que algoritmo de aprendizagem automático utilizar, como proceder ao seu treino e validação. No caso específico do classificador SVM, que função de *kernel* escolher.

Alguns dos aspetos descritos levaram à realização de um conjunto de testes, a partir da Base de Dados I - Enfisema *versus* Normal, descrita na Subsecção 6.2.1, que consiste em FH-ROIs de pacientes com enfisema e de pacientes com pulmão saudável. Através da consideração de várias opções e de vários vetores de características foi possível analisar o comportamento e desempenho do classificador SVM na diferenciação de regiões enfisematosas e saudáveis. Apesar de se tratar de um cenário binário, os resultados obtidos permitiram fixar um conjunto de opções, descartar outras, e um maior entendimento do problema.

6.3.1 *Características das ROIs e Vetores de Características*

Num primeiro estudo as 290 FH-ROIs da Base de Dados I- Enfisema *versus* Normal foram subdivididas em ROIs de 40×40 pixels, 30×30 pixels e 20×20 pixels, com uma sobreposição de 50%. Para cada dimensão das ROIs foram considerados os níveis de cinzento 2^q, com q = 4, 5, 6, 7, 8.

Recorrendo aos métodos SGLDM, GLRLM e GLDM de análise estatística da textura, descritos no Capítulo 4, cada ROI é caracterizada através de um vetor de características correspondente a cada método. Para todos os métodos implementados foram consideradas as quatro direções $\theta = \{0º, 45º, 90º, 135º\}$. Recorrendo ao método SGLDM, são extraídas seis características, descritas na Secção 4.2. Considerando que para cada direção é calculada uma matriz de coocorrência e tendo sido consideradas as distâncias entre pixels de d = 1 e d = 2, os vetores de características obtidos possuem uma dimensão de 6×4 = 24 e 6×8 = 48, respetivamente. No método GRLM são considerados três grupos de características, descritas na Secção 4.3. O primeiro grupo é constituído pelas primeiras cinco características EPC, EPL, NUNC, NUCP e PP correspondentes às Equação (4.23) até (4.26), num total de 5×4 = 20 características. O segundo grupo acresce ao primeiro as características EPNCB e EPNCA, Equações (4.27) e (4.28), sendo cada ROI descrita por 7×4 = 28 características. Por fim, o terceiro grupo

considerado reúne o primeiro e segundo, sendo ainda acrescentadas as características EPCNCB, EPCNCA, EPLNCB e EPLNCA, correspondentes às Equações (4.29) até (4.32) num total de 11×4 = 44 características. O método GLDM é testado considerando distâncias entre pixels de $d = 1$ a $d = 4$, o que significa que são considerados vetores de características com dimensões que variam de 6×4 = 24 a 6×16 = 96.

Todas as características extraídas são normalizadas de forma independente com média zero e desvio padrão um.

Tendo em consideração as várias opções para as ROIs e os diferentes vetores de características obtidos através dos três métodos de análise da textura do parênquima pulmonar foram efetuadas várias experiências que permitissem avaliar qual o método, tamanho da ROI e nível de quantização que melhor permite a discriminação das ROIs nas duas classes: Enfisema *versus* Normal. Por se tratar de uma classificação binária recorreu-se ao classificador SVM disponível na *Bioinformatics Toolbox* do MATLAB (MATLAB 2010b). A classificação foi realizada recorrendo ao algoritmo SVM com a função *kernel* RBF. O treino e teste do classificador foi efectuado através de validação cruzada, com $k = 10$, com a totalidade dos dados usados para treino e teste. Os valores da exatidão apresentados correspondem ao valor médio das k iterações calculados nos dez conjuntos de teste. Os parâmetros (C, σ) foram obtidos através da metodologia que será descrita na Subsecção 6.3.2, sem a fase de refinamento. Nas Tabelas 6.2, 6.3 e 6.4 são apresentados os valores obtidos para os três métodos em análise, considerando as diferentes hipóteses (Vasconcelos, Silva, Marques, & Barroso, 2010). Para cada método e para cada tamanho da ROI são destacados os melhores resultados. Fazendo uma análise conjunta dos valores apresentados pode-se concluir que os métodos considerados se adaptam à análise de textura do tecido pulmonar. Qualquer dos métodos permitiu uma diferenciação entre as duas classes superior a 70%, independentemente do tamanho da ROI e número de níveis de cinzento.

As características extraídas através do método GLDM são as que possuem o menor poder discriminatório, para qualquer das hipóteses consideradas. A melhor exatidão alcançada foi 91.1%, com uma sensibilidade de 91.2% e uma especificidade de 91.0%, valores obtidos para uma distância entre pixels de $d = 4$, ROI de dimensão 40×40 pixels, com 64 níveis de cinzento. Verificou-se que no método GLDM o valor da distância entre pixels é o fator que mais influencia o desempenho obtido. O aumento de d de 1 para 4 pixels permite reduzir

substancialmente os erros de classificação, chegando essa diminuição a atingir 12.5%, para o tamanho de ROI de 30×30 e 128 níveis de cinzento.

Tabela 6.2: Valores médios da exatidão para as várias combinações de ROIs e vetores de características extraídas através do SGLDM.

SGLDM			
		Exatidão (%)	
ROIs	**Número de Níveis**	**Número de Características**	
Tamanho / Número		**# 24**	**# 48**
40 / 1311	256	97.1	97.4
	128	97.1	97.5
	64	97.3	**97.7**
	32	97.3	**97.7**
	16	96.6	96.8
30 / 3334	256	96.4	96.8
	128	96.7	97.1
	64	97.0	**97.3**
	32	96.7	97.0
	16	96.0	96.3
20 / 9655	256	93.1	94.5
	128	94.6	95.0
	64	94.7	95.1
	32	94.8	**95.2**
	16	93.8	94.1

Tabela 6.3: Valores médios da exatidão para as várias combinações de ROIs e vetores de características extraídas através do GLRLM.

GLRLM				
		Exatidão (%)		
ROIs	**Número de Níveis**	**Número de Características**		
Tamanho / Número		**# 20**	**# 28**	**# 44**
40 / 1311	256	95.8	95.9	95.9
	128	97.2	97.2	96.9
	64	97.4	97.6	97.6
	32	97.6	**97.7**	**97.7**
	16	97.3	97.6	97.3
30 / 3334	256	95.1	95.3	95.3
	128	96.4	96.4	96.5
	64	96.7	96.8	96.8
	32	**97.3**	97.2	97.2
	16	97.0	96.8	97.0
20 / 9655	256	91.3	91.3	91.4
	128	93.3	93.2	93.5
	64	94.4	94.5	94.4
	32	**95.1**	**95.1**	95.0
	16	94.8	94.7	94.8

O recurso às características obtidas através dos métodos SGLDM e GLRLM permitem um bom desempenho do classificador, com resultados sempre superiores a 91%, para qualquer das hipóteses. Os melhores resultados foram uma exatidão de 97.7%, obtidos com o GLRLM, recorrendo a vetores de características com a dimensão 28 e 44, uma ROI de 40×40 pixels e 32 níveis de cinzento. Resultados similares são obtidos recorrendo ao método SGLDM com $d = 2$, ROI de dimensão 40×40 pixels, com 32 ou 64 níveis de cinzento. No último caso foi obtida uma sensibilidade de 97.5% e uma especificidade de 97.9%. Como se pode verificar pela análise dos resultados o aumento da dimensão do vetor de características no método GLRLM não representa um ganho significativo no poder discriminatório do sistema.

Tabela 6.4: Valores médios da exatidão para as várias combinações de ROIs e vetores de características extraídas através do GLDM.

GLDM					
		Exatidão (%)			
ROIs	**Número de Níveis**	**Número de Características**			
Tamanho / Número		**# 24**	**# 48**	**# 72**	**# 96**
40 / 1311	256	79.4	87.5	90.2	90.5
	128	79.9	87.4	90.4	90.7
	64	80.0	88.2	90.5	**91.1**
	32	80.5	87.6	89.6	90.4
	16	80.1	86.1	88.1	89.2
30 / 3334	256	75.8	83.8	85.6	87.4
	128	75.4	84.4	86.8	87.9
	64	75.9	85.3	87.1	**88.1**
	32	75.9	84.1	86.3	87.3
	16	76.2	82.3	83.6	84.4
20 / 9655	256	71.2	78.1	80.1	80.7
	128	72.0	79.3	81.9	82.3
	64	72.5	78.3	81.9	**82.5**
	32	72.0	78.4	81.2	82.2
	16	72.5	79.4	82.3	**82.5**

Como referido, dos três métodos de análise de textura utilizados o que apresenta menor capacidade de discriminação é o GLDM. Muitos métodos baseiam-se na aplicação de um *threshold* para a distinção dos padrões Enfisema e Normal, o que se se mostra falível em fases iniciais da presença de enfisema, assim como devido à sobreposição em termos de valores UH que ocorre entre estes dois padrões (Xu et al., 2005; Silva, 2005). No entanto, a informação dos níveis de cinzento é essencial. Em imagens de TC a intensidade de cada pixel tem um significado físico que permite a identificação de tecidos e órgãos, podendo também ser associado a padrões patológicos (Depeursinge et al., 2012a). As características extraídas através

do GLDM baseiam-se em histogramas de diferenças absolutas entre níveis de cinzento de pixels vizinhos em que a informação dos níveis de cinzento "se perde". Por exemplo, uma zona homogénea, com entradas no histograma semelhantes, pode corresponder a uma região de baixa atenuação, típica de uma região de enfisema, como de alta atenuação correspondente a um corte da árvore vascular. A análise da vizinhança obtida através das características extraídas para as diferentes distâncias mostrou-se uma mais-valia pois permite distinguir muitas destas situações, como se verifica pelo aumento de ROIs bem classificadas.

Relativamente às características das ROIs, podemos concluir, genericamente, que os melhores resultados se obtiveram para ROIs de dimensão 40×40 pixels com 32 ou 64 níveis de cinzento para qualquer um dos métodos considerados. A quantização em 128 e 256 níveis de cinzento deu origem a matrizes esparsas, portanto menos informativas.

6.3.2 *Parâmetros Ótimos do Modelo de Classificador*

Num segundo conjunto de experiências, foram testadas as funções de *kernel* linear, RBF e polinomial (Subsecção 5.3.3) e avaliada a sua capacidade de diferenciação no cenário binário Enfisema *versus* Normal. É dada especial ênfase à escolha dos parâmetros ótimos para cada modelo do algoritmo de aprendizagem. São consideradas as opções que permitiram obter o melhor desempenho do sistema no conjunto de experiências descritas na secção anterior:

- As ROIs com dimensão de 40×40 pixels foram quantizadas para 32 níveis de cinzento, por uma questão de uniformização, apesar do melhor desempenho obtido com o GLDM ser para uma quantização de 64 níveis de cinzento (0.7% melhor). Esta decisão é consentânea com os resultados obtidos em. (Xu et al., 2005);
- Foram consideradas as quatro direções principais para todos os métodos de extração de características;
- No método GLDM uma distância entre pixels de $d = 1$ a $d = 4$ num total de 96 características;
- No método SGLDM a distância entre pixels é de $d = 1$ e $d = 2$, resultando num total de 48 características;
- No método GLRLM são consideradas a totalidade das 44 características.

O conjunto de dados constituído por 290 FH-ROIs foi dividido no conjunto de treino e teste numa percentagem, 70% e 30%, respetivamente. No conjunto de treino procedeu-se à pesquisa dos parâmetros ótimos do algoritmo de aprendizagem que permitem obter o melhor desempenho e no conjunto de teste procedeu-se ao cálculo das métricas de desempenho. A pesquisa dos parâmetros deve ser efetuada de uma forma sistemática. Numa fase inicial foram realizados alguns testes empíricos com diferentes funções de *kernel* e correspondentes parâmetros, de forma a ter alguma sensibilidade sobre a função e os parâmetros mais apropriados. Posteriormente, foi escolhido um espaço de pesquisa e, para cada uma das coordenadas do espaço, o classificador é treinado no conjunto de treino recorrendo à técnica de validação cruzada *k-folds* (*k* = 10). Inicialmente, o espaço de pesquisa é abrangente o que permite compreender a influência dos parâmetros nos resultados do classificador, posteriormente, este é reduzido em torno das coordenadas do espaço de pesquisa que permitem obter o melhor desempenho do classificador, com vista ao refinamento dos parâmetros.

No caso do *kernel* polinomial, após alguma experimentação, foram mantidos constantes os parâmetros $\gamma = 1$, $\delta = 1$ e $d = 3$. A metodologia de pesquisa dos parâmetros é realizada para o parâmetro de regularização *C* no caso do *kernel* linear e polinomial e para o par (C, σ) no caso do *kernel* RBF.

Inicialmente, foi considerado o espaço de pesquisa $C = 2^{-5}, 2^{-4.5}, \ldots, 2^{15}$ e $\sigma = 2^{-2}, 2^{-1.5}, \ldots, 2^{7}$. O classificador é treinado para cada par de coordenadas $(2^{C}, 2^{\sigma})$ sendo calculada a exatidão média para as *k* iterações da validação cruzada. Numa segunda fase os parâmetros que permitiram obter a melhor exatidão média são considerados o centro de um novo espaço de pesquisa e em torno deles é efetuada uma nova pesquisa em que a variação dos parâmetros se faz agora com um passo mais pequeno, de forma a encontrar os parâmetros ótimos. Para obter o modelo final do classificador este é treinado em todos os dados do conjunto de treino, considerando os parâmetros ótimos obtidos. De forma a avaliar a capacidade de generalização do classificador as métricas de desempenho exatidão, sensibilidade e especificidade são obtidas no conjunto de teste (Vasconcelos, Marques, Barroso, & Silva, 2011).

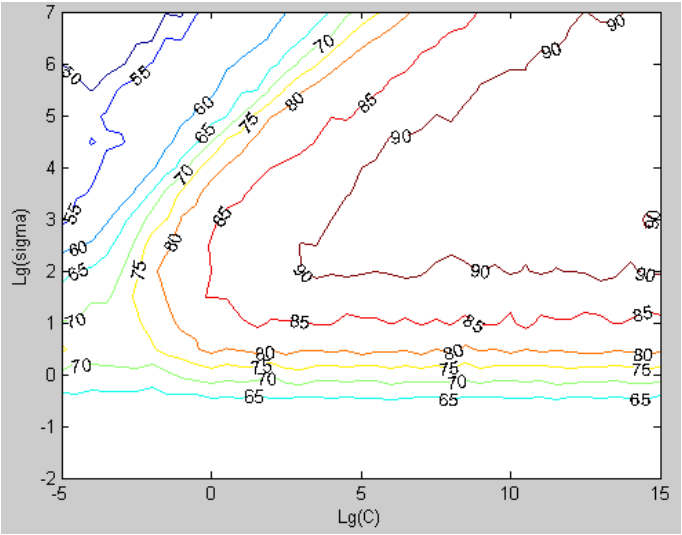


Figura 6.7: Gráfico de contornos correspondente à exatidão média (%) considerando o conjunto de características GLDM e o kernel RBF.

A Figura 6.7 mostra o gráfico de contornos relativo à exatidão média obtida durante a validação cruzada, considerando as características extraídas do método GLDM, para d = 4, ROIs de 40×40 pixels, 32 níveis de cinzento e *kernel* RBF. Uma análise heurística do gráfico permite verificar a importância do valor dos parâmetros escolhidos para o modelo do classificador nos resultados da exatidão. Permite também encontrar a região do espaço de pesquisa onde são obtidos os melhores resultados e consequentemente proceder ao refinamento dos parâmetros. No caso dos valores apresentados na Figura 6.7, ao longo do espaço de pesquisa a variação da exatidão é da ordem dos 45%. O pior resultado é de 41.5% obtido para os parâmetros ($C = 2^{-4}$, $\sigma = 2^{7}$) e o melhor de 92.8% obtido para ($C = 2^{5}$,$\sigma = 2^{2.5}$). Estes valores enfatizam a importância da escolha do modelo do classificador.

A Tabela 6.5 resume os resultados obtidos tendo em consideração os três grupos de características e os três *kernels* em análise. As métricas exatidão, sensibilidade e especificidade mais elevadas foram alcançadas com o *kernel* RBF, para qualquer um dos métodos. De acordo com as experiências realizadas, o *kernel* polinomial originou os modelos SVM que alcançaram as piores métricas nos conjuntos de treino e teste. O SVM baseia-se no princípio da minimização do erro estrutural que engloba a minimização do erro nos dados de treino e a maximização da margem de separação das duas classes. No caso de valores elevados do parâmetro de regularização C a penalização dos erros de classificação nos dados de treino ($C\sum_{i=1}^{n}\xi_i$) leva a situações de *overfitting*, comprometendo a capacidade de generalização do modelo do classificador encontrado. De facto, uma análise comparativa entre a exatidão obtida

no conjunto de treino e a exatidão obtida no conjunto de teste, permite verificar que o decréscimo destas métricas é maior quando utilizado o *kernel* polinomial, chegando a atingir uma variação de 9.9%, no caso do método GLDM. Estes resultados podem ser justificado pelos elevados valores do parâmetro de regularização *C* e pelo facto de não ter sido realizada a otimização dos parâmetros γ, δ e *d*, que ao longo das experiências referidas se mantiveram fixos. O ajuste destes parâmetros poderá conduzir a melhores resultados que os descritos.

Tabela 6.5: Resultados obtidos para os três vetores de características SGLDM, GLRLM e GLDM e os *kernels* linear, RBF e polinomial.

Método	***Kernel***	**Lg(*C*); Lg(σ)**	**Exatidão Treino (%)**	**Exatidão Teste (%)**	**Sen (%)**	**Esp (%)**
SGLDM	Linear	4.25	95.3	89.8	91.3	88.1
	RBF	5.25; 2.5	99.0	**94.3**	96.5	91.8
	Polinomial	11	96.3	87.7	89.6	85.5
GLRLM	Linear	3.50	98.4	96.1	99.4	92.5
	RBF	9; 4.5	98.9	**96.4**	99.4	93.1
	Polinomial	13	98.5	92.5	93.1	91.8
GLDM	Linear	3	87.0	81.9	76.9	87.4
	RBF	5; 2.75	93.1	**88.6**	85.0	92.5
	Polinomial	12	90.1	80.2	75.4	85.3

Sen – Sensibilidade; Esp – Especificidade

Os melhores resultados dizem respeito à conjugação das características GLRLM com o *kernel* RBF, com exatidão de 96.4%, sensibilidade de 99.4% e especificidade de 93.1%. Por outro lado, é o que apresenta melhor capacidade de generalização, quando confrontado com os dados de teste. É de realçar também que para as mesmas características o *kernel* linear permite obter resultados semelhantes ao *kernel* RBF. O método GLRLM é sensível não só às variações dos níveis de cinzento mas também à forma dos padrões.

6.4 Categorização de Padrões Intersticiais em Imagens TC de Alta Resolução

A categorização dos padrões do tecido pulmonar é efetuada apenas com base na informação da textura, uma fonte inquestionável de informação tanto na presença de patologias como no caso de pulmão saudável. Em imagens de TCAR os padrões imagiológicos associados à DIP são complexos, podem surgir combinados e estar presentes em várias patologias. Apesar de ser possível reconhecer texturas típicas, o aspeto visual do parênquima pulmonar depende de um elevado número de fatores como a idade, sexo, especificidade de cada ser humano, estádio da doença, entre outros. A Figura 6.8 ilustra o conjunto de etapas necessárias à implementação do sistema de DAC proposto, desde o processo de análise visual das imagens por parte do imagiologista até aos resultados obtidos pelo sistema. As opções tomadas para cada uma das etapas serão descritas nas subsecções seguintes.

6.4.1 *Seleção de FH-ROIs*

Sendo o sistema de DAC proposto baseado num algoritmo de aprendizagem automático cujo treino é supervisionado é necessário extrair dos exames de TCAR imagens onde sejam encontradas regiões com padrões associados à DIP. Este processo foi efetuado recorrendo ao *software Lung Study Tool* desenvolvido especificamente para esse objetivo e que se encontra descrito detalhadamente na Secção 6.1. As regiões delineadas e anotadas por especialistas, designadas por FH-ROIs, dizem respeito a quatro tipos de padrões e foram reunidas nas bases de dados descrita nas Subsecções 6.2.1 e 6.2.2.

6.4.2 *Processamento das ROIs*

Cada FH-ROI selecionada é amostrada em ROIs de 40×40 pixels contíguas, não sobrepostas, de acordo com a metodologia descrita na Secção 6.2. As ROIs consideradas devem estar totalmente contidas dentro da FH-ROI. Uma vez que o imagiologista tem liberdade na forma e tamanho das FH-ROIs que seleciona, o número de ROIs resultante é muito distinto. Para que seja possível saber a origem de cada ROI, estas guardam a informação da FH-ROIs a partir da qual cada ROI foi obtida.

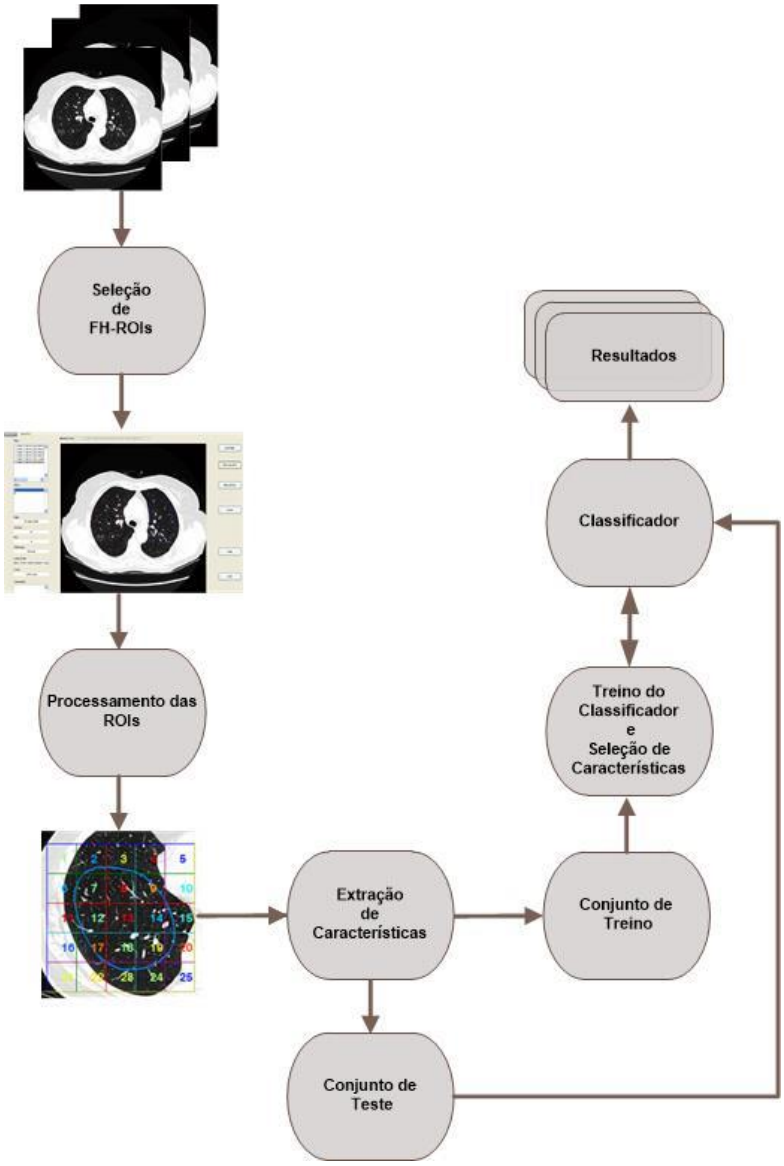


Figura 6.8: Principais etapas do sistema de DAC proposto.

Originalmente a escala das regiões selecionadas é a extensa escala de Hounsfield [-1000 UH; +1000 UH]. A utilização direta desta escala quando combinada com os métodos de extração de características FO, SGLDM, GLRLM e GLDM dá origem, tendencialmente, a matrizes esparsas, com muitos elementos a zero ou com valores pouco significativos. Assim, na fase de pré-processamento foi realizada a quantização das ROIs para 32 níveis de cinzento, tendo em consideração os resultados obtidos na Subsecção 6.3.1.

6.4.3 *Extração de Características*

A escolha das características que permitam uma representação compacta, mas assertiva, da textura dos padrões do tecido pulmonar que se pretendem classificar é um desafio e o cerne de qualquer sistema de DAC baseado em imagens. Dada a complexidade da textura associada a padrões DIP em imagens TCAR as medidas usadas para a descrever devem ser reunidas através de diferentes métodos que sejam sensíveis a vários aspetos da textura. Nas marcações e caracterização de regiões do parênquima pulmonar é possível observar que as FH-ROIs possuem padrões visualmente diferentes e com diferentes gamas de variação em termos de UH, para o mesmo padrão pulmonar e para o mesmo indivíduo. De forma a descrever corretamente a textura das ROIs com estas características foi utilizada uma técnica multiescala sensível às variações dos níveis de cinzento, a lacunaridade diferencial.

Para o cálculo da lacunaridade diferencial de cada ROI, esta é dividida em janelas sobrepostas *w* sobre as quais deslizam caixas de lado *r*. Para cada janela é calculada a massa M_r. A distribuição de probabilidade $Q(M_r, r, w)$ obtêm-se dividindo o número de janelas com massa M_r, dado por $n(M_r, r, w)$, pelo número total de janelas. A análise da textura multiescala faz-se considerando diferentes valores para o tamanho da caixa *r* e da janela *w*, (Vasconcelos, Barroso, Marques, & Silva, 2015):

$$\Lambda(r, w) = \frac{\sum_{M_r} M_r^2 \, Q(M_r, r, w)}{\left[\sum_{M_r} M_r \, Q(M_r, r, w)\right]^2}. \qquad (6.1)$$

O tamanho da janela *w* define a área para o cálculo da massa. Dada a diversidade dos padrões a lacunaridade é calculada para diferentes tamanhos de janelas. Por outro lado, o valor da caixa *r* deve ser pequeno para que seja possível identificar pequenas variações de vizinhança. Mantendo o tamanho da caixa *r* constante e fazendo variar o tamanho da janela *w* podemos obter curvas da lacunaridade $\Lambda(r = constante, w)$ para os diferentes padrões pulmonares.

Foram calculadas curvas para várias combinações (*r*, *w*), considerando *r*<*w*, de forma a aferir quais as características que melhor diferenciam os padrões pulmonares em análise. Para garantir um referencial comum e obter uma maior diferenciação entre as curvas da lacunaridade, estas são normalizadas em relação ao valor da lacunaridade correspondente ao valor da menor janela *w*: $\Lambda(r = constante, w_{min})$. A Figura 6.9 mostra as curvas da lacunaridade média normalizadas

calculadas a partir de todas as ROIs que integram a Base de Dados II - Padrões Intersticiais, considerando $r = 2$ e $w = \{3,\ldots,30\}$.

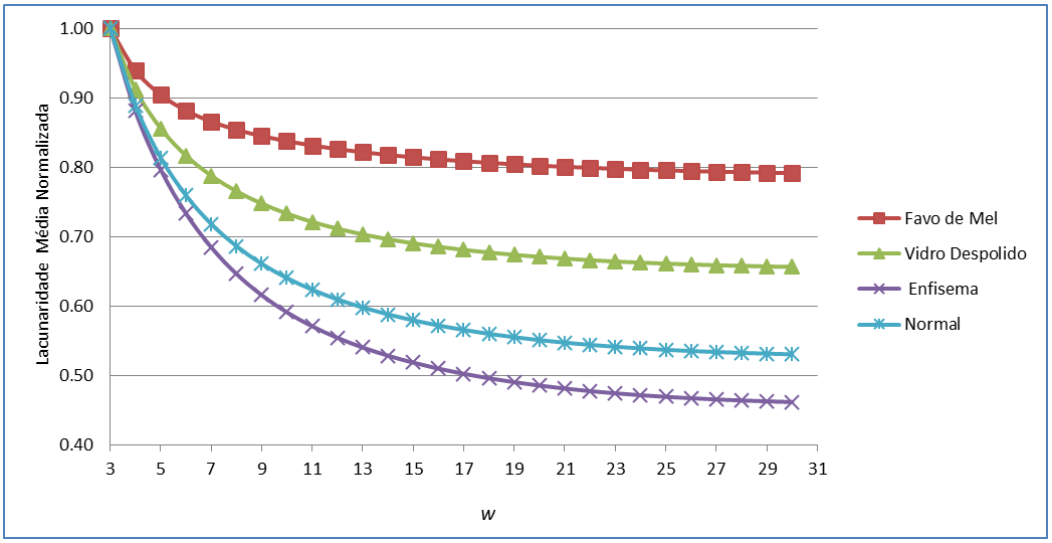


Figura 6.9: Curvas da lacunaridade média normalizadas.

Apesar da diversidade da textura das ROIs existentes na base de dados, mesmo catalogadas com o mesmo padrão, as curvas da lacunaridade média normalizadas descrevem hipérboles diferentes (Vasconcelos, Marques, Silva, & Barroso, 2013; Vasconcelos et al., 2015). Foram consideradas as combinações (r, w) para $r = \{2, \ldots ,5\}$ e $w = \{3, \ldots ,35\}$, sendo as curvas apresentadas na Figura 6.9 as mais discriminatórias, obtidas por experimentação. O conjunto de experiências realizadas com a Base de Dados I - Enfisema *versus* Normal permitiu obter linhas orientadoras para as opções a tomar em cada um dos métodos de análise de textura aquando da sua aplicação ao sistema de DAC proposto. Cada ROI é descrita por um vetor que integra as características extraídas dos métodos FOS, SGLDM, GLRLM e as características da lacunaridade diferencial normalizada obtidas para $r = 2$ e $w = \{3, \ldots ,30\}$. O método GLDM não foi considerado uma vez que apresentou baixo poder discriminatório em todas as experiências realizadas.

6.4.4 *Conjunto de Treino e Teste*

Para proceder ao treino e teste do algoritmo de aprendizagem, a Base de Dados II – Padrões Intersticiais foi dividida em dois conjuntos - treino e teste - numa proporção de 70% e 30%, respetivamente. As ROIs que integram os dois conjuntos foram escolhidas aleatoriamente

através da técnica de *holdout* com estratificação, o que assegura partições mutuamente exclusivas nas quais a proporção das classes em cada conjunto é aproximadamente a mesma que na base de dados integral. O procedimento de *holdout* baseia-se nas FH-ROIs assegurando que ROIs extraídas da mesma FH-ROI são colocadas em apenas um dos conjuntos, treino ou teste, minimizando o *bias* do sistema.

6.4.5 *Treino do Classificador e Seleção de Características*

Durante o treino do algoritmo SVM é realizada a pesquisa dos parâmetros ótimos no conjunto de treino, considerando todas as características, o que permite obter o modelo inicial do classificador. A pesquisa dos hiperparâmetros ótimos foi realizada recorrendo a uma grelha de pesquisa no espaço dos hiperparâmetros (C, σ), uma vez que foi considerada a função *kernel* RBF, como descrito na Subsecção 6.3.2.

O método de seleção de características considerado foi o SBS conjugado com validação cruzada durante o treino do algoritmo SVM. É um método *wrapper* que envolve o classificador na avaliação da qualidade do subconjunto de características ótimo. A ideia base do SBS é considerar todas as características como candidatas e iterativamente remover as características cuja eliminação provoca a menor degradação do desempenho do classificador, neste caso medida através do número de erros de classificação. O processo está descrito no algoritmo seguinte:

Algoritmo: SBS com validação cruzada usando SVM

Candidatas = {todas}
REPETE
    PARA cada característica de *Candidatas* **FAZ**
        Retirar característica de *Candidatas*
        Calcula *Erro* usando validação cruzada para treino do SVM
        Retirar de *Candidatas* a característica cuja ausência provocou menor *Erro*
    FIM_PARA
ENQUANTO o menor *Erro* obtido na iteração atual é superior ao *Erro* obtido na iteração anterior

A inicialização passa por considerar um conjunto que engloba todas as características disponíveis - *Candidatas*. Posteriormente, é retirada cada uma das características de *Candidatas* e para cada um dos subconjuntos obtidos o algoritmo SVM é treinado através de validação cruzada com $k = 10$. Os hiperparâmetros (C, σ) foram selecionados na etapa anterior. Para cada uma das características removidas o *Erro* avaliado é a média dos erros obtidos nos dez conjuntos de teste ao longo da validação cruzada. A característica que ao ser removida deu origem ao menor *Erro* é eliminada do conjunto *Candidatas.* O processo é repetido iterativamente até que o menor *Erro* obtido, considerando todas as possíveis remoções de características numa iteração, seja superior ao *Erro* que levou à eliminação da característica na iteração anterior, isto é, não há nenhuma característica que ao ser retirada do conjunto *Candidatas* diminua o erro do classificador.

6.4.6 *Resultados*

Após a escolha dos hiperparâmetros ótimos e do subconjunto ótimo de características foi obtido o modelo final do classificador. Este modelo é aplicado ao conjunto de teste independente. É sobre este conjunto que será avaliado o desempenho do sistema.

A divisão da base de dados nos dois conjuntos de teste e treino pode dar origem a resultados não realistas. Otimistas no caso dos dados fáceis de classificar serem maioritariamente colocados no conjunto de teste ou pessimistas caso os dados mais difíceis sejam encaminhados para o conjunto de teste. De modo a evitar estes enviesamentos dos resultados, aumentar a reprodutibilidade e fiabilidade dos resultados, a divisão da base de dados entre conjuntos de teste e treino foi realizada cinquenta vezes. Assim, as etapas de treino, seleção do subconjunto ótimo de características e cálculo das métricas de desempenho são realizadas em cinquenta conjuntos de dados diferentes. As métricas de avaliação de desempenho do sistema proposto correspondem à média dos resultados obtidos ao longo de cinquenta divisões da base de dados.

De forma a avaliar o potencial do sistema de DAC proposto e a importância das características de lacunaridade na melhoria do desempenho de DAC foram considerados dois cenários, um binário e um multiclasse. Numa primeira abordagem considerou-se a diferenciação entre o padrão normal e padrões DIP. A segunda situação considerada consiste na diferenciação das quatro classes de padrões. Em ambos os cenários, o vetor de característica foi obtido usando

dois conjuntos diferentes. O *Set 1* inclui as características: FOS + SGLDM + GLRLM. O *Set 2* acrescenta ao vetor anterior as características da lacunaridade.

6.4.6.1 Normal *versus* Patológico

Quando é considerado o cenário padrão Normal *versus* padrões DIP estamos perante uma classificação binária onde as ROIs (#253) com padrão normal são consideradas como negativas e as restantes (#1008) como positivas, uma vez que há a presença de padrões associados à DIP (designados por patológicos).

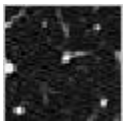
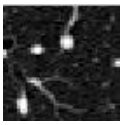
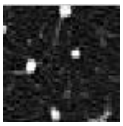
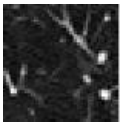
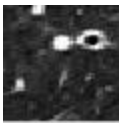
Na Tabela 6.6 são apresentados a média e desvio padrão da exatidão, sensibilidade, precisão e especificidade obtidos para as cinquenta divisões da base de dados, considerando os conjuntos de características *Set 1* e *Set 2*. Recorrendo ao *Set 2* os resultados melhoraram globalmente e o desvio padrão diminui, tendo a exatidão aumentado de 94.6 ± 1.3% para 96.4 ± 1.3%. Os valores da sensibilidade são elevados tanto com o *Set 1* como com o *Set* 2 tendo sido obtidos os valores 96.9 ± 1.1% e 97.6 ± 1.0%, respetivamente.

Tabela 6.6: Resultados obtidos para o cenário normal *versus* patológico, considerando os dois conjuntos de características *Set* 1 e *Set* 2. Valores apresentados em percentagem.

	Set 1	***Set 2***
Exatidão	94.6 ± 1.3	96.4 ± 1.3
Sensibilidade	96.9 ± 1.1	97.6 ± 1.0
Precisão	96.3 ± 1.6	97.9 ± 1.3
Especificidade	85.9 ± 6.6	91.9 ± 5.2

Os valores da sensibilidade e precisão obtidos são semelhantes, com desvios padrões também aproximados, pelo que podemos inferir um número similar de ocorrências de falsos positivos e falsos negativos, considerando tanto o *Set* 1 como o *Set* 2. Dado que o número de ROIs normais é cerca de 25% das ROIs com patologia podemos afirmar que o sistema de DAC proposto tem uma grande capacidade de sinalizar a presença de padrões patológicos. O valor da especificidade é comprometido pelo desequilíbrio da base de dados, sendo também a métrica que apresenta maior desvio padrão.

Tabela 6.7: Exemplos de ROIs mal classificadas entre as duas classes Normal e Patológica.

Normal classificada como Patologia		Patologia classificada como Normal	
		Vidro Despolido	Enfisema
		 Vidro Despolido	Enfisema

Na Tabela 6.7 encontram-se exemplos de ROIs mal classificadas escolhidas aleatoriamente. Como se pode observar as diferenças são subtis e as semelhanças muitas. A classificação correta de instâncias da classe normal é difícil devido à alta variabilidade do tecido pulmonar saudável entre sujeitos e à sua semelhança com os padrões patológicos em estudo.

6.4.6.2 Categorização de Padrões Intersticiais

O segundo cenário considerado consiste na diferenciação dos quatro padrões presentes na Base de Dados II – Padrões Intersticiais. Trata-se de uma classificação multiclasse implementada através do método um-*versus*-um recorrendo à biblioteca LIBSVM (Chang & Lin, 2011). Os valores médios e desvio padrão das métricas específicas de cada classe: exatidão, sensibilidade, precisão e especificidade, ao longo das cinquenta divisões da base de dados são apresentados nas Tabelas 6.8 e 6.9.

Tabela 6.8: Resultados obtidos para o cenário multiclasse com o conjunto de características *Set 1*. Valores apresentados em percentagem.

	Classes			
	Normal	Vidro Despolido	Favo de Mel	Enfisema
Exatidão	96.0 ± 0.9	94.8 ± 1.1	97.9 ± 0.9	98.6 ± 0.7
Sensibilidade	92.3 ± 3.3	92.2 ± 2.5	93.7 ± 3.5	96.0 ± 2.4
Precisão	89.3 ± 2.6	89.8 ± 3.1	94.9 ± 2.7	99.9 ± 0.4
Especificidade	97.0 ± 0.8	95.9 ± 1.4	98.9 ± 3.5	99.9 ± 0.2

Fazendo uma análise global, considerando os dois conjuntos de vetores de características utilizados para a diferenciação das classes de padrões, verificamos que o desempenho global do sistema de DAC melhorou quando foi usado o *Set 2* de características. A exatidão global do modelo obtido com as características do *Set 1* foi de 93.7±1.4% e recorrendo ao *Set 2* foi de 95.6±1.2%.

A primeira linha das tabelas corresponde à exatidão de cada classe, que como se verifica melhorou para os quatro padrões. Estes resultados mostram que a análise da lacunaridade permitiu aumentar a diferenciação entre padrões reduzindo o número de falsos positivos e falsos negativos. Como consequência, os valores das métricas específicas aumentaram em maior ou menor percentagem. A exceção ocorreu na classe Enfisema que manteve os mesmos valores em três das quatro métricas consideradas, duas delas com praticamente 100%. Outro aspeto importante é a diminuição do desvio padrão nos resultados obtidos. Este facto mostra que a introdução das características de lacunaridade tornaram o sistema mais estável quando confrontado com dados de treino e teste diferentes, ao longo de cinquenta iterações.

Tabela 6.9: Resultados obtidos para o cenário multiclasse com o conjunto de características *Set 2*. Valores apresentados em percentagem.

	Classes			
	Normal	Vidro Despolido	Favo de Mel	Enfisema
Exatidão	97.3 ± 0.8	96.6 ± 0.9	98.7 ± 0.4	98.6 ± 0.6
Sensibilidade	95.2 ± 1.6	96.6 ± 1.1	94.6 ± 1.5	96.5 ± 2.1
Precisão	92.5 ± 2.3	91.9 ± 2.2	98.5 ± 0.8	99.9 ± 0.4
Especificidade	97.9 ± 0.7	96.7 ± 0.9	99.6 ± 1.5	99.9 ± 0.2

Para a classe Enfisema as métricas mantiveram os mesmos valores elevados de exatidão, precisão e especificidade nos dois cenários, verificando-se um aumento do valor da sensibilidade de 96.0±2.4% para 96.5±2.1%, o número de falsos negativos diminui. A métrica cuja média mais aumentou, acompanhada da diminuição do desvio padrão, foi a sensibilidade da classe Vidro Despolido, aumentando 4.4%. Tendo em consideração que é um padrão que corresponde a 31% da Base de Dados este aumento é significativo, mais regiões com este padrão são classificadas corretamente, o número de falsos negativos diminui substancialmente.

O erro de classificação mais frequente ocorre entre o padrão Vidro Despolido e o padrão Normal. O segundo maior aumento registou-se na precisão da classe Favo de Mel aumentando 3.6%, o que significa que menos instâncias de outras classes foram classificadas erradamente como Favo de Mel. No entanto, o padrão Favo de Mel corresponde apenas a 17% da Base de Dados.

A sensibilidade média da classe Normal aumentou 2.9%, havendo um incremento de 92.3±3.3% para 95.2±1.6%. Num cenário multiclasse estes resultados significam que o número de falsos negativos específicos da classe Normal diminuiu, i.e., o número de instâncias da classe Normal que é classificada como patológica diminui. Em ambiente clínico isto significa que o número de pacientes sujeitos ao stresse de realizar exames médicos adicionais desnecessários é menor. A precisão da classe Normal aumentou de 89.3±2.6% para 92.5 ± 2.3% (+3.2%) e a especificidade de 97.0 ± 0.8 % para 97.9 ± 0.7 % (+0.9). Estes resultados significam que o número de falsos positivos relativos à classe Normal, ou seja, instâncias com padrões patológicos classificadas como Normal, diminuíram com o *Set 2*. Este tipo de erros de classificação tem um significado grave ou seja: o sistema de DAC não sinaliza a presença de um padrão patológico.

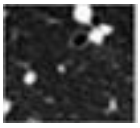
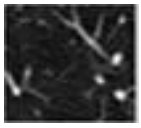
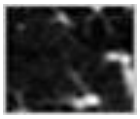
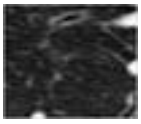
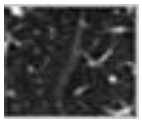
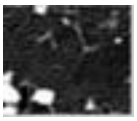
No contexto do sistema de DAC os resultados obtidos com o *Set 2* contribuem para aumentar o diagnóstico diferencial da DIP, uma vez que há padrões que se encontram associados a certas patologias DIP. A análise dos resultados permite concluir que as características de lacunaridade para além de melhorarem globalmente o desempenho do sistema de DAC, permitindo uma maior diferenciação entre classes, também melhoraram a estabilidade do classificador, uma vez que os valores do desvio padrão das métricas diminuíram na grande generalidade dos casos.

As patologias em estudo possuem aspetos localizados em pequenas áreas, cuja análise de vizinhança é essencial, associada a uma grande heterogeneidade visível a maiores escalas. O facto do recurso à lacunaridade ter melhorado o desempenho do sistema resulta da conjugação destes dois aspetos. Por um lado a sensibilidade aos pormenores de vizinhança através do cálculo da massa recorrendo a um valor pequeno da caixa r ($r = 2$), por outro a análise multiescala considerando vários tamanhos da janela w ($w = \{3, \ldots, 30\}$) que permitem a análise da heterogeneidade dos padrões ao longo da ROI de 40×40 pixels.

Considerando o *Set 2*, ao longo de cinquenta divisões da base de dados integraram os cinquenta conjuntos de teste 18 915 ROIs das quais 786 ROIs foram mal classificadas. Apesar da distribuição do número de ROIs por classe na Base de Dados II não ser equitativa, após a análise dos erros verificou-se uma discrepância entre o número de classificações erradas que ocorrem entre as diferentes classes. Nas Tabelas 6.10, 6.11 e 6.12 são apresentados exemplos de ROIs mal classificadas entre as classes cujos erros foram maiores. Os exemplos das ROIs apresentados foram escolhidos de forma aleatória.

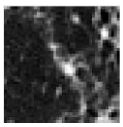
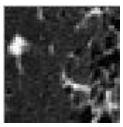
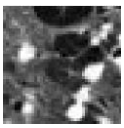
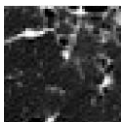
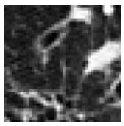
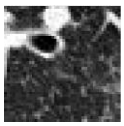
Das seis combinações possíveis de classificações erradas entre classes, a maior percentagem de erros ocorreu entre a classe Normal e a classe Vidro Despolido. Entre estas duas classes ocorreram 36.8% de todos os erros de classificação nos cinquenta conjuntos de teste. Deste valor, 20.3% corresponde a ROIs da classe Normal que foram classificadas como Vidro Despolido.

Tabela 6.10: Exemplos de ROIs mal classificadas entre as classes Vidro Despolido e Normal.

Normal classificada como Vidro Despolido		Vidro Despolido classificada como Normal	
			
			

A Tabela 6.10 ilustra exemplos de ROIs da classe Normal que foram classificados como Vidro Despolido, na coluna da esquerda, e exemplos de ROIs com padrão Vidro Despolido que foram classificados como Normal, na coluna da direita. Uma das razões destes erros deve-se ao facto da opacidade Vidro Despolido se caracterizar por áreas de maior atenuação, mas que por vezes não são suficientemente densas para "esconder" a presença da árvore broncovascular, especialmente nas fases iniciais das patologias DIP, associadas à presença do padrão Vidro Despolido, tornando os dois padrões semelhantes.

Tabela 6.11: Exemplos de ROIs mal classificadas entre as classes Favo de Mel e Vidro Despolido.

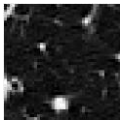
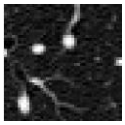
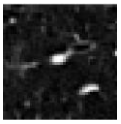
Favo de Mel classificada como Vidro Despolido		Vidro Despolido classificada como Favo de Mel	
			
			

A segunda maior percentagem de erros, 27.8%, ocorreu entre as classes Vidro Despolido e Favo de Mel, sendo que 21.6 % corresponde à classificação de ROIs com o padrão Favo de Mel como sendo o padrão Vidro Despolido, elevado valor de falsos negativos da classe Favo de Mel. Na Tabela 6.11 podemos analisar exemplos de ROIs mal classificadas entre as duas classes. Visualmente o padrão Favo de Mel surge nas imagens de TCAR como regiões escuras grosseiramente circulares envoltas por paredes espessas com valores de Hounsfield altos. Estas zonas claras elevam os valores de nível de cinzento da região podendo levar a uma classificação errónea como Vidro Despolido, especialmente em ROIs com poucos quistos, fazendo com que a densidade média local seja similar entre os dois padrões. Por outro lado, os quistos presentes no padrão Favo de Mel, que correspondem a zonas de baixa atenuação, podem tomar um aspeto semelhante à anatomia normal da árvore brônquica que em corte axial surge como regiões aproximadamente circulares de baixa atenuação (vias áreas) adjacentes a áreas de alta atenuação (vias vasculares). Há ainda a acrescentar o facto do padrão Favo de Mel ser muitas vezes acompanhado por pequenas áreas de Vidro Despolido (Depeursinge et al., 2012a). A classificação do padrão Favo de Mel continua a ser um dos padrões mais desafiantes da DIP (Vo, & Sowmya, 2010).

A terceira maior percentagem de erro, 25.2%, verificou-se entre as classes Normal e Enfisema, sendo que a grande maioria dos erros, 24.6 %, corresponde à classificação de ROIs com o padrão Enfisema como sendo o padrão Normal. A Tabela 6.12 ilustra exemplos de ROIs mal classificadas destas duas classes. Na fase inicial do Enfisema, especialmente no Enfisema

Centrolobular a semelhança com o padrão Normal é grande, uma vez que o Enfisema surge como pequenas zonas de baixa atenuação com uma distribuição não uniforme. Ocorrem também situações em que as zonas de baixa atenuação do Enfisema são confundidas com as zonas de baixa atenuação correspondentes a cortes na árvore brônquica.

Tabela 6.12: Exemplos de ROIs mal classificadas entre as classes Normal e Enfisema.

Normal classificada como Enfisema		Enfisema classificada como Normal	
			

Analisando os erros de classificação mais frequentes, referidos nos parágrafos anteriores, verificamos que são coerentes com as curvas de lacunaridade média normalizada da Figura 6.9. As curvas de lacunaridade mais próximas correspondem a padrões entre os quais é maior o número de classificações erradas: Normal *versus* Vidro Despolido, Favo de Mel *versus* Vidro Despolido e Normal *versus* Enfisema.

Uma das maiores dificuldades na classificação dos padrões pulmonares é a elevada variabilidade dentro da mesma classe, associada, muitas vezes, a pequenas diferenças entre classes que em muitas situações visualmente se tornam difíceis de distinguir (Song et al., 2013). Os exemplos de ROIs apresentados nas tabelas anteriores permitem constatar esta realidade, que é especialmente verdade quando a classificação se efetua ao nível de pequenas áreas 2D da imagem sem informação clínica adicional. É de referir ainda que todos os padrões estudados podem surgir visualmente em imagens de TC com aspetos muito diferentes e que estes padrões surgem associados/misturados em muitas patologias o que dá origem a uma classificação errada e, portanto, nem sempre é possível justificar algumas classificações erradas a partir apenas da ROI.

6.5 Considerações Finais

Ao longo deste capítulo foi descrito o sistema de DAC proposto neste trabalho. O sistema visa a classificação de padrões pulmonares normais e associados à DIP em imagens de TCAR do tórax.

A necessidade de lidar com um elevado número de imagens, no sentido de as visualizar, selecionar áreas de interesse, de as catalogar e organizar num formato suscetível de ser utilizado por algoritmos de aprendizagem automática levou ao desenvolvimento da aplicação *Lung Study Tool*, cujas potencialidades e funcionalidades foram apresentadas. Posteriormente, as duas bases de dados construídas com a colaboração de imagiologistas do Departamento de Imagiologia do CHUC foram caracterizadas. Estas bases de dados reúnem ROIs de padrões pulmonares extraídos de imagens de TCAR de pacientes com DIP e de pacientes com pulmão saudável.

Os padrões recolhidos foram considerados o *ground truth* para o treino e teste do sistema de DAC. Cada ROI foi descrita de forma compacta através de um vetor de características recorrendo a métodos baseados no histograma de níveis de cinzento, na matriz de coocorrência, na matriz de primitivas de comprimento e na análise da lacunaridade diferencial. Numa primeira abordagem, foram realizados testes tendo em consideração várias hipóteses, relativamente às ROIs, às características e ao treino do algoritmo de aprendizagem, tendo em consideração regiões normais e enfisematosas da Base de Dados I. Os resultados obtidos permitiram uma maior compreensão do problema e constituíram o ponto de partida para o sistema proposto na Secção 6.4 aplicado aos padrões DIP que integram a Base de Dados II.

Os resultados obtidos mostraram que a integração de informação multiescala através de características de lacunaridade aumentaram o desempenho do sistema de DAC tanto na diferenciação do padrão normal dos padrões patológicos, como na classificação de todos os padrões, contribuindo ainda para uma maior estabilidade dos resultados.

Capítulo 7
CONCLUSÕES E TRABALHO FUTURO

A interação entre sistemas automáticos de auxílio ao diagnóstico e os profissionais de saúde é uma realidade. A exigência de muitos serviços hospitalares torna essencial o recurso a software de apoio que permita ao médico uma segunda opinião. O trabalho descrito nesta tese enquadrou-se na área de apoio ao diagnóstico da DIP em imagens de TCAR do tórax. Neste último capítulo são apresentadas as conclusões do trabalho desenvolvido e apontadas direções para trabalhos futuros.

7.1 Conclusões

A informação fornecida pelas imagens médicas tornou-se atualmente uma parte fundamental nos cuidados de saúde prestados aos pacientes. As técnicas imagiológicas complementares de diagnóstico são prática comum em todos os centros hospitalares das sociedades modernas. Esta realidade deu origem a um aumento enorme das imagens adquiridas diariamente que devem ser processadas e analisadas cuidadosamente. É cada vez maior a necessidade de uma estreita colaboração entre profissionais de saúde e equipas de investigação multidisciplinares no sentido de adaptar e desenvolver técnicas que respondam às exigentes necessidades dos ambientes clínicos. Foi neste âmbito que o trabalho desenvolvido nesta tese se enquadrou.

No estudo das doenças pulmonares, a TCAR é considerada a técnica de referência. A interpretação das imagens obtidas é um trabalho intenso e desafiante. As imagens de TCAR na presença de patologias pulmonares como a DIP caracterizam-se por alterações da textura do parênquima pulmonar, relativamente ao padrão pulmonar considerado normal. A análise visual destas alterações, associadas aos conhecimentos e à experiência clínica adquirida pelo imagiologista permitem detetar e reconhecer padrões associados às diferentes patologias

pulmonares. Nesta tese procurou-se dar um passo em frente na caracterização e diferenciação automática de padrões do tecido pulmonar associados à DIP em imagens de TCAR do tórax através do desenvolvimento de um sistema de DAC.

O primeiro requisito para o estudo de qualquer patologia é reunir um conjunto de casos representativos. A primeira fase deste projeto consistiu no desenvolvimento da aplicação *Lung Study Tool* que permitiu a recolha de padrões imagiológicos em imagens de TCAR. As funcionalidades da aplicação foram desenvolvidas de modo a facilitar a escolha e visualização das imagens, a seleção de áreas de interesse, de diferentes tamanhos e formas, e a sua anotação por parte do imagiologista. Uma das vertentes da aplicação permite a análise visual das regiões selecionadas e anotadas no contexto do exame, o que poderá ser relevante do ponto de vista pedagógico, por exemplo para estudantes.

A recolha de padrões extraídos das imagens de TCAR foi realizada em colaboração com o Departamento de Imagiologia do CHUC. Este processo foi condicionado pela ocorrência no serviço de exames de TC com interesse para este trabalho, assim como, pela disponibilidade dos imagiologistas que ao longo do projeto estiveram a ele associados. Uma vez que exames de TC com a presença de enfisema pulmonar são bastante comuns na rotina diária do serviço, o primeiro conjunto de marcações inclui o padrão Enfisema e o padrão Normal, associado a pacientes sem patologias pulmonares. Posteriormente, foram recolhidas imagens de pacientes com DIP onde os padrões Vidro Despolido e Favo de Mel estavam presentes. Os dados recolhidos foram organizados em duas bases de dados: uma com regiões de padrão Normal e padrão Enfisema e uma outra que reúne quatro padrões: Normal, Enfisema, Vidro Despolido e Favo de Mel. A informação recolhida é de grande valor para o treino e teste de técnicas de reconhecimento de padrões, partes integrantes da maioria dos sistemas de apoio ao diagnóstico.

Para estabelecer um diagnóstico, avaliar a progressão da doença ou planear o tratamento da DIP é determinante a análise do aspeto visual dos padrões em imagens de TCAR. Uma das componentes determinantes dessa análise é a perceção da textura dos padrões imagiológicos. A textura de cada região é descrita por um vetor de características, reunidas através de diversos métodos de modo a captar aspetos diferentes. Foram consideradas características estatísticas de primeira ordem, segunda e de ordem superior extraídas do histograma de níveis de cinzento, matriz de coocorrência, histograma de diferenças e matriz de primitivas de comprimento. Dada

a diversidade do aspeto visual dos padrões intersticiais e normal, foi ainda considerada uma análise multiescala através da análise da lacunaridade diferencial.

Numa primeira fase foram consideradas várias hipóteses no que diz respeito às regiões de interesse, escolhas de parâmetros nos métodos de análise de textura considerados e no algoritmo de aprendizagem adotado: SVM. As análises comparativas efetuadas permitiram uma maior compreensão do problema. Por uma questão temporal, a base de dados utilizada durante esta fase foi a que integra os padrões Enfisema e Normal. Verificou-se que as características extraídas através do SGLDM e GLRLM permitem a diferenciação dos dois padrões com resultados sempre superiores a 91%, para qualquer das hipóteses consideradas. Os melhores resultados obtidos foram uma exatidão de 97.7%, obtidos com o GLRLM, para uma ROI de dimensão 40×40 pixels, quantizada com 32 níveis de cinzento. Resultados similares foram obtidos recorrendo ao método SGLDM com uma distância entre pixels $d = 2$, ROI de dimensão 40×40 pixels, com 32 ou 64 níveis de cinzento. As características extraídas através do método GLDM foram as que apresentaram menor poder discriminatório, para qualquer das hipóteses consideradas. Considerando os *kernels* linear, polinomial e RBF verificou-se que este permite obter os melhores resultados para os métodos de extração de características considerados.

Posteriormente, foi implementado e testado um sistema de DAC aplicado a imagens de TCAR para a diferenciação de padrões do tecido pulmonar presentes na DIP. Foram considerados dois cenários, o binário que classifica cada padrão como Normal ou Patológico (padrão intersticial) e o multiclasse que considera as quatro classes de padrões: Normal, Vidro Despolido, Favo de Mel e Enfisema. Para cada um, foi avaliado o potencial diferenciador das características da lacunaridade diferencial. A metodologia multiescala adotada permitiu melhorar o desempenho do sistema de DAC nos dois cenários considerados. A lacunaridade diferencial foi calculada nas imagens em níveis de cinzento, tendo em consideração por um lado, o aspeto local da textura, e por outro, a sua heterogeneidade em diferentes escalas. No cenário binário, a presença das características de lacunaridade permitiu melhorar todas as métricas e diminuir o seu desvio padrão. A exatidão aumentou de 94.6 ± 1.3% para 96.4 ± 1.3%, a sensibilidade de 96.9 ± 1.1% para 97.6 ± 1.0, a precisão de 96.3 ± 1.6% para 97.9 ± 1.3% e a especificidade de 85.9 ± 6.6% para 91.9 ± 5.2%. O sistema mostrou-se sensível à presença do padrão intersticial, requisito essencial de um sistema de DAC. No cenário multiclasse o sistema é capaz de diferenciar os padrões presentes na Bases de Dados II - Padrões Intersticiais com uma exatidão global de

95.6±1.2%, tendo aumentado 1.9% com a presença das características de lacunaridade. Estas permitiram aumentar as métricas específicas de cada classe, especialmente nos padrões Normal, Vidro Despolido e Favo de Mel. A sensibilidade do padrão Vidro Despolido aumentou em 4.4% e a precisão 2.1%, já a sensibilidade do padrão Favo de Mel aumentou 0.9% e a precisão 3.6%. A classe Normal requer especial atenção. Um falso positivo da classe Normal é grave porque o sistema de DAC não assinala um padrão patológico e um falso negativo induz o médico da presença de uma patologia que pode não existir. Os valores destes dois tipos de erros diminuíram com a presença das características da lacunaridade. A sensibilidade e a precisão aumentaram 2.9% e 3.2%, respetivamente.

A lacunaridade nunca tinha sido utilizada na análise da textura de padrões associados à DIP. A sua aplicação em imagens de níveis de cinzento através da metodologia multiescala proposta melhorou o desempenho global do sistema mostrando que tem potencial na categorização de padrões intersticiais em imagens de TCAR.

7.2 Trabalho Futuro

Após o términus de um trabalho há sempre caminhos que não foram trilhados e novas ideias que surgem já no decorrer/final do trabalho. Num projeto que abrange várias áreas do saber, como medicina, imagiologia, informática, processamento e análise de imagem, *machine learning*, só para destacar as principais, as hipóteses de trabalho futuro são muitas. Assim, destacam-se algumas que são consideradas exequíveis a médio prazo:

- Aumentar o número de amostras das bases de dados, assim como, o tipo de padrões considerados. Estes aspetos permitiriam consolidar os resultados obtidos e tornar mais robusto o sistema;

- No sistema de DAC proposto a seleção das áreas de interesse e a sua classificação encontram-se em módulos independentes. O sistema estaria completo através da integração de todo o sistema a funcionar em tempo real. Podem ser equacionados dois caminhos. O primeiro, mais na continuidade do trabalho existente, em que o imagiologista selecionava uma região pulmonar e o sistema procederia à sua classificação

em tempo real. Uma segunda proposta seria a segmentação automática do pulmão e a classificação das regiões com padrões patológicos presentes;

- Avaliar o interesse de diferentes técnicas de *machine learning*, nomeadamente de algoritmos de *deep learning*. Outro aspeto interessante seria a classificação ser baseada no conhecimento da classe da imagem e não da ROI que integra essa imagem (*weakly supervised classifier*);

- Testar o sistema proposto em bases de dados referidas em outros artigos da área, por exemplo nas bases de dados usadas em (Depeursinge et al., 2012b; Anthimopoulos et al., 2016). Este aspeto implica a utilização de dados com origem em vários domínios, como equipamentos e protocolos.

REFERÊNCIAS BIBLIOGRÁFICAS

Ackerman, L. V., & Gose, E. E. (1972). Breast lesion classification by computer and xeroradiograph. Cancer, 30(4), 1025-1035.

Al-Kadi, O. S., & Watson, D. (2008). Texture analysis of aggressive and nonaggressive lung tumor CE CT images. *IEEE transactions on biomedical engineering*, *55*(7), 1822-1830.

Al-Kadi, O. S., Chung, D. Y., Coussios, C. C., & Noble, J. A. (2016). Heterogeneous tissue characterization using ultrasound: a comparison of fractal analysis backscatter models on liver tumors. *Ultrasound in Medicine and Biology*, *42*(7), 1612-1626.

Allain, C., & Cloitre, M. (1991). Characterizing the lacunarity of random and deterministic fractal sets. *Physical review A*, *44*(6), 3552.

Anthimopoulos, M., Christodoulidis, S., Christe, A., & Mougiakakou, S. (2014). Classification of interstitial lung disease patterns using local DCT features and random forest. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE* (pp. 6040-6043).

Anthimopoulos, M., Christodoulidis, S., Christe, A., & Mougiakakou, S. (2015). Computer-aided diagnosis of interstitial lung diseases based on computed tomography image analysis. *Computational Optical Biomedical Spectroscopy and Imaging*, 175-200.

Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., & Mougiakakou, S. (2016). Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE transactions on medical imaging*, *35*(5), 1207-1216.

Bendaoud, S., Remy-Jardin, M., Wallaert, B., Deken, V., Duhamel, A., Faivre, J. B., & Remy, J. (2011). Sequential versus volumetric computed tomography in the follow-up of chronic bronchopulmonary diseases: comparison of diagnostic information and radiation dose in 63 adults. *Journal of thoracic imaging*, *26*(3), 190-195.

Borges, J. D. S. (2003). *Classificadores e máquinas de suporte vectorial em imagens multi-espectrais.* (Tese de Doutoramento, Universidade do Porto).

Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, *2*(2), 121-167.

Chabat, F., Yang, G. Z., & Hansell, D. M. (2003). Obstructive lung diseases: texture classification for differentiation at CT. *Radiology*, *228*(3), 871-877.

Charisis, V., Hadjileontiadis, L. J., Liatsos, C. N., Mavrogiannis, C. C., & Sergiadis, G. D. (2010). Abnormal pattern detection in wireless capsule endoscopy images using nonlinear analysis in RGB color space. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE* (pp. 3674-3677).

Charisis, V., Hadjileontiadis, L. J., Liatsos, C. N., Mavrogiannis, C. C., & Sergiadis, G. D. (2012a). Capsule endoscopy image analysis using texture information from various colour models. *Computer methods and programs in biomedicine*, *107*(1), 61-74.

Charisis, V., Hadjileontiadis, L., & Sergiadis, G. (2012b). Enhanced ulcer recognition from capsule endoscopic images using texture analysis. In *New Advances in the Basic and Clinical gastroenterology*.

Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, *2*(3), 27.

Chu, A., Sehgal, C. M., & Greenleaf, J. F. (1990). Use of gray value distribution of run lengths for texture analysis. *Pattern Recognition Letters*, *11*(6), 415-419.

Cordeiro, M. S., Backes, A. R., Júnior, A. F. D., Gonçalves, E. H., & Oliveira, J. X. (2016). Fibrous Dysplasia Characterization Using Lacunarity Analysis. *Journal of digital imaging*, *29*(1), 134-140.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

Dasarathy, B. V., & Holder, E. B. (1991). Image characterizations based on joint gray level—run length distributions. *Pattern Recognition Letters*, *12*(8), 497-502.

Delorme, S., Keller-Reichenbecher, M. A., Zuna, I., Schlegel, W., & Van Kaick, G. (1997). Usual interstitial pneumonia: quantitative assessment of high-resolution computed tomography findings by computer-assisted texture-based image analysis. *Investigative radiology*, 32(9), 566-574.

Depeursinge, A. (2010). *Affine–invariant texture analysis and retrieval of 3D medical images with clinical context integration* (Doctoral dissertation, Geneva University).

Depeursinge, A., Van de Ville, D., Platon, A., Geissbuhler, A., Poletti, P. A., & Muller, H. (2012a). Near-affine-invariant texture learning for lung tissue analysis using isotropic wavelet frames. *IEEE Transactions on Information Technology in Biomedicine*, *16*(4), 665-675.

Depeursinge, A., Vargas, A., Platon, A., Geissbuhler, A., Poletti, P. A., & Müller, H. (2012b). Building a reference multimedia database for interstitial lung diseases. *Computerized medical imaging and graphics*, *36*(3), 227-238.

Doi, K. (2005). Current status and future potential of computer-aided diagnosis in medical imaging. *The British journal of radiology*, *78*(suppl_1), s3-s19.

Doi, K. (2007). Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized medical imaging and graphics*, *31*(4), 198-211.

Dolecek, T. A., Propp, J. M., Stroup, N. E., & Kruchko, C. (2012). CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005–2009. *Neuro-oncology*, *14*(suppl_5), v1-v49.

Dong, P. (2000). Test of a new lacunarity estimation method for image texture analysis. *International Journal of Remote Sensing*, *21*(17), 3369-3373.

Dougherty, G., & Henebry, G. M. (2002). Lacunarity analysis of spatial pattern in CT images of vertebral trabecular bone for assessing osteoporosis. *Medical engineering & physics*, *24*(2), 129-138.

Du, G., & Yeo, T. S. (2002). A novel lacunarity estimation method applied to SAR image segmentation. *IEEE Transactions on Geoscience and remote sensing*, *40*(12), 2687-2691.

Duan, K., Keerthi, S. S., Chu, W., Shevade, S. K., & Poo, A. N. (2003). Multi-category classification by soft-max combination of binary classifiers. *Multiple Classifier Systems*, *2709*, 125-134.

Duda, R. O., Hart, P. E., & Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

Dundar, M., Fung, G., Bogoni, L., Macari, M., Megibow, A., & Rao, B. (2004). A methodology for training and validating a CAD system and potential pitfalls. In *International Congress Series* (Vol. 1268, pp. 1010-1014). Elsevier.

Fujita, H., Uchiyama, Y., Nakagawa, T., Fukuoka, D., Hatanaka, Y., Hara, T., ... & Zhou, X. (2008). Computer-aided diagnosis: The emerging of three CAD systems induced by Japanese health care needs. *Computer methods and programs in biomedicine*, *92*(3), 238-248.

Fujita, H., Zhang, X., Kido, S., Hara, T., Zhou, X., Hatanaka, Y., & Xu, R. (2010). An introduction and survey of computer-aided detection/diagnosis (CAD). In *2010 international conference on future computer, control and communication* (*FCCC2010*) (pp. 200-205).

Galloway, M. M. (1975). Texture analysis using gray level run lengths. *Computer graphics and image processing*, *4*(2), 172-179.

Gangeh, M. J., Sørensen, L., Shaker, S. B., Kamel, M. S., De Bruijne, M., & Loog, M. (2010). A texton-based approach for the classification of lung parenchyma in CT images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 595-602). Springer, Berlin, Heidelberg.

Gao, M., Bagci, U., Lu, L., Wu, A., Buty, M., Shin, H. C., ... & Xu, Z. (2016). Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 1-6.

Gefen, Y., Meir, Y., Mandelbrot, B.B. & Aharony, A. (1983). Geometric implementation of hypercubic lattices with noninteger dimensionality by use of low lacunarity fractal lattices, *Physical Review Letters*, Vol.50, No.3, pp. 145–148

Gilman, M. J., Laurens Jr, R. G., Somogyi, J. W., & Honig, E. G. (1983). CT attenuation values of lung density in sarcoidosis. *Journal of computer assisted tomography*, *7*(3), 407-410.

Gonzalez, W., & Woods, R. E. (2008). Eddins, Digital Image Processing. *3rd Edition New Jersey: Prentice Hall*.

Guyon, I. (2006). Feature Extraction Foundations and Applications. *Pattern Recognition*.

Hamzah, M. F. M., Kasim, R. M., Yunus, A., Rijal, O. M., & Noor, N. M. (2017). Detection of Interstitial Lung Disease using correlation and regression methods on texture measure. In *Imaging, Vision & Pattern Recognition (icIVPR), 2017 IEEE International Conference on* (pp. 1-4).

Hanen, A., Imen, B., Asma, B. A., Patrick, D., & Hédi, B. M. (2009). Multifractal modelling and 3D lacunarity analysis. *Physics Letters A*, 373(40), 3604-3609.

Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), 786-804.

Hira, Z. M., & Gillies, D. F. (2015). A review of feature selection and feature extraction methods applied on microarray data. *Advances in bioinformatics*, *2015*.

Hu, S., Hoffman, E. A., & Reinhardt, J. M. (2001). Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. *IEEE transactions on medical imaging*, *20*(6), 490-498.

Jennane, R., Touvier, J., Bergounioux, M., & Lespessailles, E. (2014). A variational model for trabecular bone radiograph characterization. In *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on* (pp. 1283-1286).

Kauczor, H. U., Heitmann, K., Heussel, C. P., Marwede, D., Uthmann, T., & Thelen, M. (2000). Automatic detection and quantification of ground-glass opacities on high-resolution CT using multiple neural networks: comparison with a density mask. *American Journal of Roentgenology*, *175*(5), 1329-1334.

Kazerooni, E. A., Austin, J. H., Black, W. C., Dyer, D. S., Hazelton, T. R., Leung, A. N., ... & Pipavath, S. (2014). ACR–STR practice parameter for the performance and reporting of lung cancer screening thoracic computed tomography (CT): 2015 (Resolution 17). *Journal of thoracic imaging*, *29*(5), 310-316.

Kim, N., Seo, J. B., Lee, Y., Lee, J. G., Kim, S. S., & Kang, S. H. (2009). Development of an automatic classification system for differentiation of obstructive lung disease using HRCT. *Journal of digital imaging*, *22*(2), 136-148.

Korfiatis, P. D., Karahaliou, A. N., Kazantzi, A. D., Kalogeropoulou, C., & Costaridou, L. I. (2010). Texture-based identification and characterization of interstitial pneumonia patterns in lung multidetector CT. *IEEE transactions on information technology in biomedicine*, 14(3), 675-680.

Lee, Y., Seo, J. B., Lee, J. G., Kim, S. S., Kim, N., & Kang, S. H. (2009). Performance testing of several classifiers for differentiating obstructive lung diseases based on texture analysis at high-resolution computerized tomography (HRCT). *Computer methods and programs in biomedicine*, *93*(2), 206-215.

Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D. D., & Chen, M. (2014). Medical image classification with convolutional neural network. In *Control Automation Robotics & Vision (ICARCV), 2014 13th International Conference on* (pp. 844-848).

Lin, T. W., Huang, P. Y., & Cheng, C. W. C. (2016). Computer-aided diagnosis in medical imaging: Review of legal barriers to entry for the commercial systems. In *e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on* (pp. 1-5). IEEE.

Liu, H., & Motoda, H. (Eds.). (2007). *Computational methods of feature selection*. CRC Press.

Lodwick, G. S., Haun, C. L., Smith, W. E., Keller, R. F., & Robertson, E. D. (1963). Computer diagnosis of primary bone tumors: A preliminary report. *Radiology*, *80*(2), 273-275.

Long, L. R., Antani, S., Deserno, T. M., & Thoma, G. R. (2009). Content-based image retrieval in medicine: Retrospective assessment, state of the art, and future directions. *International journal of*

healthcare information systems and informatics: official publication of the Information Resources Management Association, *4*(1), 1.

Lorena, A. C., & Carvalho, A. C. (2007). Uma introdução às support vector machines. *Revista de Informática Teórica e Aplicada*, *14*(2), 43-67.

Mandelbrot, B.B. (1982). *The Fractal geometry of nature*, W. H. Freeman, New York, USA.

Mariolis, I., Korfiatis, P., Costaridou, L., Kalogeropoulou, C., Daoussis, D., & Petsas, T. (2010). Investigation of 3D textural features' discriminating ability in diffuse lung disease quantification in MDCT. In *Imaging Systems and Techniques (IST), 2010 IEEE International Conference on* (pp. 135-138).

Martinez, J. A. B. (1998). Doenças intersticiais pulmonares. *Medicina (Ribeirao Preto. Online*, *31*(2), 247-256.

Masotti, M. (2005). *Optimal image representations for mass detection in digital mammography* (Doctoral dissertation, Bologna University).

Mather, P., & Tso, B. (2016). *Classification methods for remotely sensed data*. CRC press.

MATLAB 2010b. The MathWorks, Natick, 2010b Mathworks Inc., Natick, USA.

Meyers, P. H., Nice Jr, C. M., Becker, H. C., Nettleton Jr, W. J., Sweeney, J. W., & Meckstroth, G. R. (1964). Automated computer analysis of radiographic images. *Radiology*, *83*(6), 1029-1034.

Müller, N. L., Staples, C. A., Miller, R. R., & Abboud, R. T. (1988). "Density mask": an objective method to quantitate emphysema using computed tomography. *Chest*, *94*(4), 782-787.

Muller, K. R., Mika, S., Ratsch, G., Tsuda, K., & Scholkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE transactions on neural networks*, *12*(2), 181-201.

Müller, H., Michoux, N., Bandon, D., & Geissbuhler, A. (2004). A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. *International journal of medical informatics*, *73*(1), 1-23.

Pisco, J. B. M. (Ed.). (2009). *Imagiologia básica: texto e atlas*. LIDEL, 2ª ed.

Plotnick, R. E., Gardner, R. H., Hargrove, W. W., Prestegaard, K., & Perlmutter, M. (1996). Lacunarity analysis: a general technique for the analysis of spatial patterns. *Physical review E*, *53*(5), 5461.

Prosch, H., Schaefer-Prokop, C. M., Eisenhuber, E., Kienzl, D., & Herold, C. J. (2013). CT protocols in interstitial lung diseases—a survey among members of the European Society of Thoracic Imaging and a review of the literature. *European radiology*, *23*(6), 1553-1563.

Raghu, G., Collard, H. R., Egan, J. J., Martinez, F. J., Behr, J., Brown, K. K., ... & Lynch, D. A. (2011). An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *American journal of respiratory and critical care medicine*, *183*(6), 788-824.

Reiss, M. A., Lemmerer, B., Hanslmeier, A., & Ahammer, H. (2016). Tug-of-war lacunarity—A novel approach for estimating lacunarity. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *26*(11), 113102.

Rifkin, R., & Klautau, A. (2004). In defense of one-vs-all classification. *Journal of machine learning research*, *5*, 101-141.

Salzer, R. (2012). *Biomedical Imaging: Principles and Applications*, Wiley.

Schölkopf, B., Burges, C., & Smola A. (1999). *Advances in Kernel Methods—Support Vetor Learning*. Cambridge: MIT Press, 1999.

Shikata, H., Hoffman, E. A., & Sonka, M. (2004, April). Automated segmentation of pulmonary vascular tree from 3D CT images. In *Proc. SPIE*(Vol. 5369, No. 107-116, pp. 0277-786X).

Silva. J. (2005). *Segmentação Pulmonar em Estudos de Tomografia Axial Computorizada*. (Tese de Doutoramento, Universidade de Aveiro).

Silva, C. I. S., Marchiori, E., Júnior, A. S. S., & Müller, N. L. (2010). Consenso brasileiro ilustrado sobre a terminologia dos descritores e padrões fundamentais da TC de tórax. *Jornal Brasileiro de Pneumologia*, *36*(1), 99-123.

Slim, I., Akkari, H., Abdallah, A. B., Bhouri, I., Lespessailles, E., Jennane, R., & Bedoui, M. H. (2016). Trabecular Bone Radiograph Characterization Using Lacunarity Measurement. In *Computer Graphics, Imaging and Visualization (CGiV), 2016 13th International Conference on*(pp. 196-199).

Sluimer, I. C., van Waes, P. F., Viergever, M. A., & van Ginneken, B. (2003). Computer-aided diagnosis in high resolution CT of the lungs. *Medical physics*, *30*(12), 3081-3090.

Sluimer, I., Schilham, A., Prokop, M., & van Ginneken, B. (2006). Computer analysis of computed tomography scans of the lung: a survey. *IEEE transactions on medical imaging*, *25*(4), 385-405.

Smitha, K. A., Gupta, A. K., & Jayasree, R. S. (2015). Fractal analysis: fractal dimension and lacunarity from MR images for differentiating the grades of glioma. *Physics in medicine and biology*, *60*(17), 6937.

Song, Y., Cai, W., Zhou, Y., & Feng, D. D. (2013). Feature-based image patch approximation for lung tissue classification. *IEEE transactions on medical imaging*, *32*(4), 797-808.

Sonka, M., Tschirren, J., Ukil, S., Zhang, X., Xu, Y., Reinhardt, J. M., ... & Hoffman, E. A. (2007, April). Pulmonary CT image analysis and computer aided detection. In *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on* (pp. 500-503).

Sørensen, L., Shaker, S. B., & de Bruijne, M. (2010). Quantitative analysis of pulmonary emphysema using local binary patterns. *IEEE transactions on medical imaging*, *29*(2), 559-569.

Stark, P. (2008). High resolution computed tomography of the lungs. *UpToDate*. Waltham, MA, Denise S. Basow edition.

Tschirren, J., Hoffman, E. A., McLennan, G., & Sonka, M. (2005). Intrathoracic airway trees: segmentation and airway morphology analysis from low-dose CT scans. *IEEE transactions on medical imaging*, *24*(12), 1529-1539.

Tulder, G., & de Bruijne, M. (2014). Learning features for tissue classification with the classification restricted Boltzmann machine. In *International MICCAI Workshop on Medical Computer Vision* (pp. 47-58). Springer, Cham.

Tulder, G., & de Bruijne, M. (2016). Combining generative and discriminative representation learning for lung CT analysis with convolutional restricted boltzmann machines. *IEEE transactions on medical imaging*, *35*(5), 1262-1272.

Uchiyama, Y., Katsuragawa, S., Abe, H., Shiraishi, J., Li, F., Li, Q., ... & Suzuki, K. (2003). Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography. *Medical Physics*, 30(9), 2440-2454.

Unser, M. (1986). Sum and difference histograms for texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1), 118-125.

Uppaluri, R., Mitsa, T., Sonka, M., Hoffman, E. A., & McLennan, G. (1997). Quantification of pulmonary emphysema from lung computed tomography images. *American journal of respiratory and critical care medicine*, 156(1), 248-254.

Uppaluri, R., Hoffman, E. A., Sonka, M., Hunninghake, G. W., & McLennan, G. (1999a). Interstitial lung disease: a quantitative study using the adaptive multiple feature method. *American journal of respiratory and critical care medicine*, *159*(2), 519-525.

Uppaluri, R., Hoffman, E. A., Sonka, M., Hartley, P. G., Hunninghake, G. W., & McLennan, G. (1999b). Computer recognition of regional lung disease patterns. *American journal of respiratory and critical care medicine*, *160*(2), 648-654.

Vapnik, V. (1995). *The nature of statistical learning theory*. Springer-Verlag, New York.

Vasconcelos, V., Silva, J. S., & Barroso, J. (2009). CAD lung system: texture based classifier of pulmonary pathologies. In *Proceedings of the 8th Iberian Conference on Information Systems and Technologies* (pp. 383-386).

Vasconcelos, V., Silva, J. S., Marques, L., & Barroso, J. (2010). Statistical textural features for classification of lung emphysema in CT images: a comparative study. In *Information Systems and Technologies (CISTI), 2010 5th Iberian Conference on* (pp. 1-5).

Vasconcelos, V., Marques, L., Barroso, J., & Silva, J. S. (2011). Comparative Performance Analysis of Support Vector Machines Classification Applied to Lung Emphysema in HRCT Images. In *IMAGAPP/IVAPP* (pp. 134-139).

Vasconcelos, V., Marques, L., Silva, J. S., & Barroso, J. (2013, June). Lacunarity analysis of pulmonary emphysema in high-resolution CT images. In *Information Systems and Technologies (CISTI), 2013 8th Iberian Conference on* (pp. 1-5).

Vasconcelos, V., Barroso, J., Marques, L., & Silva, J. S. (2015). Enhanced classification of interstitial lung disease patterns in HRCT images using differential lacunarity. *BioMed research international, 2015*.

Verschakelen, J. A., De Wever, W., & Matamoros, A. (2008). Computed tomography of the lung: a pattern approach. *Journal of Nuclear Medicine*, *49*(1), 164-164.

Vo, K. T., & Sowmya, A. (2010, August). Multiple kernel learning for classification of diffuse lung disease using HRCT lung images. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE* (pp. 3085-3088).

Winsberg, F., Elkin, M., Macy Jr, J., Bordaz, V., & Weymouth, W. (1967). Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis. *Radiology*, *89*(2), 211-215.

Wong, A. K., & Sahoo, P. K. (1989). A gray-level threshold selection method based on maximum entropy principle. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(4), 866-871.

Xu, Y., Sonka, M., McLennand, G., Guo, J., & Hoffmanc, E. (2005). Sensitivity and specificity of 3-D texture analysis of lung parenchyma is better than 2-D for discrimination of lung pathology in stage 0 COPD. In *Proc. of SPIE Vol* (Vol. 5746, p. 475).

Xu, Y., Sonka, M., McLennan, G., Guo, J., & Hoffman, E. A. (2006a). MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. *IEEE transactions on medical imaging*, *25*(4), 464-475.

Xu, Y., van Beek, E. J., Hwanjo, Y., Guo, J., McLennan, G., & Hoffman, E. A. (2006b). Computer-aided classification of interstitial lung diseases via MDCT: 3D adaptive multiple feature method (3D AMFM). *Academic radiology*, *13*(8), 969-978.

Xu, R., Hirano, Y., Tachibana, R., & Kido, S. (2013). A bag-of-features approach to classify six types of pulmonary textures on high-resolution computed tomography. *IEICE TRANSACTIONS on Information and Systems*, *96*(4), 845-855.

Zaia, A., Eleonori, R., Maponi, P., Rossi, R., & Murri, R. (2006). MR imaging and osteoporosis: fractal lacunarity analysis of trabecular bone. *IEEE transactions on information technology in biomedicine*, *10*(3), 484-489.

Zaia, A. (2015). Fractal lacunarity of trabecular bone and magnetic resonance imaging: new perspectives for osteoporotic fracture risk assessment. *World journal of orthopedics*, *6*(2), 221.

Zavaletta, V. A., Bartholmai, B. J., & Robb, R. A. (2007). High resolution multidetector CT-aided tissue analysis and quantification of lung fibrosis. *Academic radiology*, *14*(7), 772-787.

Zhang, M., Zhoua, X., Goshimab, S., Chenc, H., Muramatsua, C., Haraa, T., ... & Fujitaa, H. (2013). Pulmonary emphysema classification based on an improved texton learning model by sparse representation. In *Proc. of SPIE Vol* (Vol. 8670, pp. 86700F-1).

Zhao, W., Xu, R., Hirano, Y., Tachibana, R., & Kido, S. (2013). Classification of diffuse lung diseases patterns by a sparse representation based method on HRCT images. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE* (pp. 5457-5460).

Zhao, W., Xu, R., Hirano, Y., Tachibana, R., & Kido, S. (2015). A sparse representation based method to classify pulmonary patterns of diffuse lung diseases. *Computational and mathematical methods in medicine*, 2015.

Zhou, W., Li, H., & Tian, Q. (2017). Recent Advance in Content-based Image Retrieval: A Literature Survey. arXiv preprint arXiv:1706.06064.

Zrimec, T., & Wong, J. S. (2007). Improving computer aided disease detection using knowledge of disease appearance. In *Medinfo 2007: Proceedings of the 12th World Congress on Health (Medical) Informatics; Building Sustainable Health Systems* (p. 1324). IOS Press.

Anexo A
APLICAÇÃO DA LACUNARIDADE NA ANÁLISE DE IMAGENS MÉDICAS

Neste anexo resumem-se trabalhos que recorrem à análise da lacunaridade para o estudo de diferentes órgãos ou tecidos em imagens adquiridas a partir de diferentes tipos de equipamentos.

A osteoporose é uma doença óssea metabólica que fragiliza o osso trabecular, ou esponjoso, tornando-o suscetível a fraturas. A principal componente do osso trabecular designa-se por trabécula e corresponde a uma rede de suporte constituída por finas colunas de osso entrelaçadas de forma irregular. Com a idade e em pacientes com osteoporose, a arquitetura trabecular começa a deteriorar-se perdendo progressivamente conetividade e aumentando os espaços intertrabeculares. A presença de osteoporose caracteriza-se por uma baixa densidade mineral óssea e por uma deterioração da microarquitectura do tecido ósseo. São vários os autores que recorrem à lacunaridade para analisar a microarquitectura óssea e correlacioná-la com presença de osteoporose. Os estudos recorrem a imagens obtidas através de diferentes técnicas imagiológicas como TC, RM e raio-X (Dougherty, & Henebry, 2002; Zaia et al., 2006; Zaia, 2015; Slim et al., 2016).

Em (Dougherty, & Henebry, 2002) é investigada a identificação de parâmetros de textura *in vivo*, que possuam informação estrutural do osso trabecular das vértebras e que permitam a distinção entre pacientes com e sem osteoporose. Estes parâmetros são baseados na análise da lacunaridade calculada a partir de imagens de TC, e visam aumentar a sensibilidade dos valores obtidos através de densitometria mineral óssea (DMO) constituindo um instrumento complementar de diagnóstico da osteoporose. O cálculo da lacunaridade é efetuado em imagens binarizadas. A partir de imagens de teste de TC axiais do osso trabecular das vértebras de quatro dadoras recolhidas em autópsia com secções de espessura 1, 3, 5 e 10 mm (i.e. 16 imagens de

teste) são investigados diferentes métodos de segmentação com base em *thresholds* globais (média, mediana e baseado na análise discriminante) e *threshold* local adaptativo (aplicado a janelas de 7×7 pixels). A segmentação mais coerente com as curvas de lacunaridade expectáveis foi obtida considerando o *threshold* local adaptativo. Num conjunto de imagens de TC de dez mulheres voluntárias, com espessura de 5 mm, foram medidos os valores de DMO. A partir das curvas de lacunaridade obtidas para diferentes tamanhos de vizinhança em imagens binarizadas, os autores extraíram parâmetros como o desvio máximo da curva de lacunaridade relativamente à curva linear (característica de padrões autossimilares em diferentes escalas) e a derivada máxima desse desvio que se mostraram sensíveis aos vários valores de DMO previamente medidos. Apesar dos resultados obtidos, o trabalho descrito é considerado pelos autores como sendo preliminar, devido ao pequeno número de pacientes e à falta de resolução espacial das imagens. O projeto acabou por não ter continuidade (informação fornecida pelo primeiro autor).

Em (Zaia et al., 2006; Zaia, 2015) a análise da lacunaridade foi aplicada em imagens de RM de vértebras da região lombar para a extração de características relevantes que permitam a diferenciação entre três tipos de estrutura óssea trabecular de pacientes do sexo feminino com idades e estados fisiopatológicos diferentes (jovens saudáveis, saudáveis em perimenopausa e pacientes com osteoporose). O cálculo da lacunaridade baseia-se no algoritmo *gliding-box* (Allain & Cloitre, 1991). Os autores propõem um método que aproxima as curvas de lacunaridade, calculada a partir de imagens binarizadas, através do modelo de uma hipérbole dependente de três parâmetros (α, β e γ) que caracterizam de forma unívoca a estrutura trabecular dos padrões estudados. O parâmetro α representa a convergência da curva, β a concavidade da hipérbole e γ corresponde ao termo translacional. O parâmetro β encontra-se correlacionado com o estado fisiopatológico possuindo valores sucessivamente decrescentes do grupo de jovens saudáveis, para pacientes em perimenopausa e pacientes com osteoporose. Este parâmetro encontra-se também correlacionado com a idade da paciente. Segundo os autores os resultados do estudo permitem considerar o parâmetro β como um índice que permite caracterizar a arquitetura trabecular óssea, com um potencial preditivo no risco de fratura.

O trabalho desenvolvido por Slim et al. (Slim et al., 2016) recorre à lacunaridade para a analisar a arquitetura óssea em imagens radiográficas com vista à discriminação entre pacientes com osteoporose, que apresentam diferentes tipos de fraturas, e pacientes de controlo. O conjunto de imagens utilizado é composto por 174 radiografias do osso do calcanhar, de mulheres na

pós-menopausa, sendo 87 de controlo e 87 de mulheres com fraturas osteoporóticas de diferentes tipos (punho, quadril, vertebras, entre outras). Uma vez que a idade é um dos fatores que influenciam a textura do osso trabecular e a densidade óssea, a média de idades dos dois grupos é semelhante, cerca de 71.3 ± 10.5 anos. São excluídas as pacientes sujeitas a medicação e portadoras de doenças que possam interferir no metabolismo ósseo. O método utilizado para o cálculo da lacunaridade corresponde à versão 2D do método *Relative Differential Box-Counting* 3D (RDBC3) proposto por (Hanen, Imen, Asma, Patrick, & Hédi, 2009). O estudo analisa o poder discriminatório da lacunaridade entre o grupo de controlo e cada um dos tipos de fratura obtendo um valor de AUC que varia de 90% (fratura do quadril) a 72% (fratura das vértebras). No caso de grupo de controlo *versus* pacientes com fraturas o valor de AUC é de 78%. A metodologia proposta por (Slim et al., 2016) produziu melhores resultados que os obtidos por Jennane et al. (Jennane, Touvier, Bergounioux, & Lespessailles, 2014) usando a mesma base de dados de imagens. Neste trabalho a imagem é sujeita a um pré-processamento que a decompõe em três componentes através do método da decomposição variacional, a macrotextura, a microtextura, o contraste e a luminância. A imagem original e as imagens correspondentes às três componentes foram analisadas recorrendo aos métodos SGLM e GLRLM a partir dos quais foram extraídas as características da energia e NUNC, respetivamente, nas quatro direções principais. Em (Slim et al., 2016) sem proceder ao pré-processamento das imagens, foi possível aumentar a classificação de 85%, obtida por Jennane et al., para 90% relativamente à discriminação entre fratura da anca e o de grupo de controlo, sendo também possível proceder à discriminação entre todos os tipos de fratura e o grupo de controlo, o que não aconteceu no trabalho de Jennane et al..

Em (Al-Kadi, & Watson, 2008) os autores propõem a diferenciação de tumores pulmonares malignos agressivos de tumores pulmonares malignos não agressivos. Um tumor não agressivo encontra-se em estádios iniciais e possui um baixo metabolismo enquanto tumores agressivos correspondem a estádios avançados aos quais corresponde um elevado metabolismo, acompanhado de maior vascularização. O estudo foi efetuado em séries temporais de imagens de TC de pacientes aos quais foi administrado um agente contrastante. Como a ramificação dos vasos sanguíneos assemelham-se a um fractal, as regiões dos tumores mais vascularizadas podem ser analisadas com base em propriedades fractais. As imagens originais são melhoradas, transformando a intensidade de cada pixel no correspondente valor da sua dimensão fractal. Neste domínio a segmentação das regiões correspondentes aos tumores é mais fácil, uma vez

que as arestas correspondentes a diferentes tecidos e a ramificação dos grandes vasos sanguíneos tornam-se mais explícitas. Na dimensão fractal, o valor médio da dimensão fractal e da lacunaridade da área do tumor é calculado para cada imagem, sendo selecionado o maior valor. Estes valores são correlacionados com informação médica que se sabe capaz de predizer a agressividade do tumor, como o grau do tumor e o valor de FDG (*fluorodeoxyglucose*) determinados pela tomografia por emissão de positrões (PET - *Positron Emission Tomography*). Os resultados mostram que valores elevados do valor médio da dimensão fractal encontram-se associados a tumores em estádios avançados e, portanto, mais agressivos. Em 83.3% dos casos as características fractais permitiram a diferenciação entre tumores agressivos e não agressivos.

Os gliomas são tumores primários do cérebro que apresentam vários subtipos que na sua totalidade são responsáveis por 30% dos tumores e 80% dos tumores malignos (Dolecek, Propp, Stroup, & Kruchko, 2012). A determinação do grau ou estádio do tumor (I, II, III ou IV) é determinante para o prognóstico e planeamento do tratamento a adotar. Para a diferenciação do grau do tumor é necessário recorrer à integração da informação obtida através de várias técnicas avançadas de RM, o que implica exames extras e um custo elevadíssimo. O trabalho de Smitha et al. (Smitha et al., 2015) centra-se na procura de técnicas de processamento de imagem que possam ser incorporadas na rotina de aquisição de imagem de RM de forma a melhorar a eficiência na determinação do grau de malignidade do tumor. Com este objetivo os autores focaram-se na análise fractal – dimensão fractal e lacunaridade. As imagens são sujeitas a um conjunto de etapas de pré-processamento. Começam por ser binarizadas recorrendo a um *threshold* determinado através do princípio da entropia máxima, complementada com informação espacial e distribuição dos níveis de cinzento proposta por (Wong, & Sahoo, 1989). Às imagens binarizadas é aplicado o operador morfológico dilatação e posteriormente são detetados os contornos do tumor. Por fim, as regiões de interesse são selecionadas por um neurologista experiente. Nas regiões obtidas é aplicado o algoritmo *box-counting*. Este método consiste em sistematicamente sobrepor sobre a imagem grelhas de tamanhos sucessivamente menores e contabilizar o número de caixas ocupadas com um ou mais pixels. A dimensão fractal obtém-se a partir da reta de regressão num gráfico *log-log* entre o número de caixas contabilizadas $N(r)$ e a largura da caixa r necessária para cobrir a região. Neste estudo a variação do comprimento da caixa r foi de dois pixels até 45% da ROI considerando 12 orientações diferentes para cada valor de r. À medida que a dimensão fractal aumenta a complexidade e

irregularidade do objeto fractal também aumenta. O valor da lacunaridade é calculado através da fórmula:

$$L = \left(\sigma/\mu\right)^2_{r,g} \quad \text{(A.1)}$$

sendo μ o valor médio da densidade dos pixels na caixa de tamanho *r*, σ o desvio padrão e o valor *g* a orientação da grelha. A densidade dos pixels é definida como o rácio entre o número de pixels em cada caixa e o número total de pixels da ROI. Para cada ROI é calculada a dimensão fractal e a lacunaridade tendo em consideração os vários valores de *r* e *g*, sendo posteriormente calculada a média dos valores individuais. O estudo foi aplicado a uma população de sessenta e seis pacientes com uma média de idades de 38 anos com gliomas não tratáveis. De cada paciente foram adquiridas sequências de imagens de RM T1, T2, FLAIR (*Fluid-Attenuated Inversion Recovery*) e T1 com contraste. A análise fractal é efetuada a partir das sequências FLAIR 2D. Os resultados mostraram que valores mais altos da dimensão fractal e de lacunaridade estão associados a tumores com maior malignidade, permitindo a diferenciação entre tumores benignos ou malignos. A lacunaridade surge com maior potencial na diferenciação entre vários graus de malignidade enquanto a dimensão fractal apenas diferencia o grau II do grau IV. Segundo os autores a complementaridade entre a análise fractal e as imagens FLAIR tornam mais preciso o diagnóstico do que apenas as correspondentes sequências de imagens.

Charisis et al. (Charisis et al., 2012b) propuseram um trabalho para a diferenciação entre regiões normais e regiões ulcerosas em imagens endoscópicas adquiridas através de cápsulas endoscópicas *wireless*. A metodologia proposta baseia-se na informação da cor e da textura, inspirada pela prática dos gastroentrologistas na sua prática clínica. Trabalhos anteriores (Charisis et al., 2010; Charisis et al., 2012a) permitiram concluir que o espaço de cor mais apropriado para a análise das imagens em consideração é o RGB, comparativamente aos espaços HSV e o CIE Lab. Mais concretamente a maioria da informação de textura encontra-se no canal *Verde*, sendo nele que se foca a análise. A técnica de obtenção de imagens utilizada é muito ruidosa devido ao *hardware* usado na cápsula e às condições de aquisição das imagens. O pré-processamento das imagens é efetuado recorrendo à técnica *Bidimensional Ensemble Empirical Mode Decomposition* (BEEMD) que decompõe a imagem em oito 2D *intrinsic mode functions* (IMFs) mais resíduo. As duas primeiras IMFs correspondem às componentes de alta

frequência da imagem. De modo a reduzir o ruído estas IMFs são descartadas. Curvas de lacunaridade são calculadas para cada IMF. As curvas com declives dentro de uma determinada gama de valores, escolhidos de acordo com as curvas de imagens com úlceras, correspondem a IMFs com maior informação de textura e são portanto as selecionadas. De modo semelhante a Zaia et al. (Zaia, 2006), as curvas de lacunaridade são modeladas através de uma função hiperbólica com base nos parâmetros (α, β e γ) que descreve o comportamento global da curva de lacunaridade. O vetor de características consiste nos três parâmetros (α, β e γ) e num conjunto de valores da lacunaridade escolhidos por experimentação. A classificação das imagens em normal *versus* úlcera é efetuada recorrendo a classificadores baseados na análise discriminante, com as funções discriminantes linear, quadrática e Mahalanobis e SVM. Os melhores resultados foram obtidos quando o vetor de características é extraído a partir das imagens reconstruídas com as IMFs ótimas mais resíduo e a classificação é efetuada com SVM, sendo obtido uma diferenciação correta em 96,7% dos casos.

A displasia fibrosa é uma anomalia benigna do desenvolvimento que afeta a região maxilofacial de muitos jovens, podendo causar deformidades severas e assimetrias faciais. Esta anomalia caracteriza-se pelo preenchimento do espaço medular do osso normal por tecido fibroso. A TC é um dos meios de diagnóstico apropriado para a deteção precoce e correto diagnóstico das lesões, assim como no suporte à cirurgia maxilofacial. Em (Cordeiro et al., 2016) é proposta a análise dos padrões de osso normal e com displasia fibrosa recorrendo à análise da lacunaridade. Foram considerados dez pacientes que não apresentavam doenças sistémicas capazes de alterar a estrutura óssea como a osteoporose, assim como pacientes que usassem drogas que pudessem interferir na atividade óssea visível nas imagens. Após a seleção dos pacientes são realizados novos exames TC, sendo considerados cortes axiais e coronais. Para cada paciente são descritos e anotados os padrões tomográficos predominantes, que podem ser: escleróticos, císticos ou mistos. Na amostra considerada, seis pacientes apresentam a predominância do padrão misto e os restantes do padrão esclerótico. De cada paciente, são selecionadas três ROIs de osso com displasia e três ROIs contralaterais de osso normal, com gamas de densidade diferentes (alta, intermédia e baixa) de acordo com a escala de Hounsfield. Dada a discrepância entre o tamanho das imagens de TC obtidas e as áreas de interesse os autores aumentaram o tamanho original da imagem (512×512 pixels) para duas e três vezes através de interpolação bicúbica. É calculada a média e desvio padrão da lacunaridade para os dois cortes, axial e coronal, e analisada a soma das interseções entre as curvas médias da lacunaridade para os dois cortes. As

curvas do corte coronal apresentam uma menor interseção entre os valores da lacunaridade das amostras de osso normal ou na presença de displasia, especialmente nas imagens de 1536×1536 pixels. O cálculo da lacunaridade baseia-se no método *relative diferential box-counting* na abordagem proposta por Du e Yeo (Du, & Yeo, 2002). É considerada uma janela de dimensão 33×33 que corresponde ao tamanho da área de osso normal considerada. Para cada paciente e para cada valor de $r = \{2, 3, \ldots, 10\}$, é calculado o valor da lacunaridade média das três amostras do osso normal e do osso com displasia, considerando o corte coronal. Os valores obtidos evidenciaram que a lacunaridade média é menor no caso das amostras correspondentes à displasia fibrosa, mostrando que a sua textura é mais homogénea que no caso do osso normal. Através de árvores de decisão a correta classificação das amostras foi de 84.2 % usando apenas a informação de $\Lambda(2)$ e $\Lambda(3)$. Após a aplicação do algoritmo análise de componentes principais, considerando também $r = 2$ e $r = 3$ a exatidão aumentou para 89.5 %.

A ultrassonografia é uma técnica não invasiva que permite a deteção precoce da resposta dos tumores quando sujeitos a tratamentos de quimioterapia, nomeadamente em tumores do fígado. Os padrões de dispersão (*scatterers*) tendem a ser diferentes em tecidos normais e em tecidos com patologia em forma de lesão, assim como as características acústicas asssociadas podem caracterizar diferentes concentrações de *scatterers* e microestruturas, o que permite uma associação a padrões diferentes. Assim, a interação de uma onda sonora com regiões de tecido diferentes pode ser descrita pelo envelope (amplitude) do sinal de rádio-frequência *backscattered*. Este sinal pode ser considerado um processo aleatório e modelado por várias distribuições estatísticas. Em (Al-Kadi et al., 2016) são consideradas cinco distribuições: Rayleigh, Rician, *k*, Nakagami e Nakagami-*generalized inverse of Gaussian* (NIG). Para cada voxel são estimados os parâmetros dos modelos estatísticos referidos a partir do envelope do sinal de rádio-frequência dando origem a imagens paramétricas cujo número depende da quantidade de parâmetros da distribuição estatística. O arranjo local das concentrações de *scatterers* e a sua distribuição espacial ocorridas no tecido do tumor são analisadas através de uma aproximação multifractal. A dimensão fractal é utilizada para caracterizar a distribuição espacial dos *scatterers* e a lacunaridade é aplicada para determinar o número de densidade. A dimensão fractal é estimada através do fractal *Brownian motion* que é um modelo não estacionário conhecido pela capacidade de descrever processos aleatórios. A lacunaridade é obtida através do algoritmo *gliding box-counting* aplicado à imagem fractal. As duas assinaturas são estimadas em diferentes escalas de resolução após a aplicação de *wavelets* de Daubechies.

Segundo os autores, valores elevados da dimensão fractal indicam uma textura heterógenea, associada a uma distribuição dos *scatterers* aleatória, enquanto a lacunaridade está associada ao grau de heterogeneidade (densidade) apresentando valores baixos quando os *scatterers* são grandes. O desempenho foi avaliado em imagens de ultrasonografia de rádio-frequência de tumores do fígado que reúnem imagens de tumores que responderam ao tratamento e outros que não. As características fractais obtidas a partir do modelo de Nakagami e da sua extensão NIG foram as que permitiram obter melhores resultados na caracterização do tecido dos tumores. Estes resultados foram obtidos através da validação *leave–one–tumor–out* recorrendo ao classificador de Bayes. Os resultados alcançados através das assinaturas fractais são superiores aos obtidos diretamente a partir das imagens paramétricas, para todos os modelos estatísticos considerados.