

LUIZA DRI BAGESTEIRO

**CLASSIFICAÇÃO DE PADRÕES RADIOLÓGICOS POR  
BLOCOS EM IMAGENS NÃO SEGMENTADAS DE  
TOMOGRAFIA COMPUTADORIZADA**

Dissertação de mestrado apresentada  
ao Programa de Pós-Graduação em  
Informática, Setor de Ciências Exatas,  
Universidade Federal do Paraná.  
Orientador: Daniel Weingaertner  
Co-orientador: Lucas Ferrari de Oliveira

CURITIBA

2015

LUIZA DRI BAGESTEIRO

**CLASSIFICAÇÃO DE PADRÕES RADIOLÓGICOS POR  
BLOCOS EM IMAGENS NÃO SEGMENTADAS DE  
TOMOGRAFIA COMPUTADORIZADA**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, Universidade Federal do Paraná, como requisito à obtenção do grau de Mestre em Informática.

Orientador: Daniel Weingaertner  
Co-orientador: Lucas Ferrari de Oliveira

CURITIBA

2015

---

B144c

Bagesteiro, Luiza Dri

Classificação de padrões radiológicos por blocos em imagens não segmentadas de tomografia computadorizada/ Luiza Dri Bagesteiro. – Curitiba, 2015.

84 f. : il. color. ; 30 cm.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-graduação em Informática, 2015.

Orientador: Daniel Weingaertner – Co-orientador: Lucas Ferrari de Oliveira.

Bibliografia: p. 76-84.

1. Doença pulmonar - Diagnóstico por imagem. 2 Tomografia computadorizada. 3. Reconhecimento de padrões. I. Universidade Federal do Paraná. II. Weingaertner, Daniel. III. Oliveira, Lucas Ferrari de . IV. Título.

CDD: 616.24028564

---



Ministério da Educação  
Universidade Federal do Paraná  
Programa de Pós-Graduação em Informática

## PARECER

Nós, abaixo assinados, membros da Banca Examinadora da defesa de Dissertação de Mestrado em Informática, da aluna Luiza Dri Bagesteiro, avaliamos o trabalho intitulado, “*Classificação de padrões radiológicos por blocos em imagens não segmentadas de tomografia computadorizada*”, cuja defesa foi realizada no dia 29 de junho de 2015, às 14:00 horas, no Departamento de Informática do Setor de Ciências Exatas da Universidade Federal do Paraná. Após a avaliação, decidimos pela:  
 **aprovação** da candidata.  **reprovação** da candidata.

Curitiba, 29 de junho de 2015.



Prof. Dr. Daniel Weingaertner  
PPGInf - Orientador

Prof. Dr. Lucas Ferrari de Oliveira  
PPGInf - Coorientador

Prof. Dr. Marcel Koenigkam Santos  
USP – Membro Externo

Prof. Dr. Luiz Eduardo de Oliveira  
PPGInf – Membro Interno

## RESUMO

As Doenças Pulmonares Difusas (DPDs) são compostas por mais de 180 patologias que causam danos nos pulmões. Os pacientes que possuem essas doenças apresentam alterações no tecido pulmonar, que podem ser visualizadas em imagens de Tomografia Computadorizada (TC). Dessa forma, torna-se importante identificar essas alterações nas imagens, pois juntamente com outras informações, elas podem definir o tipo de doença que o paciente possui. Com base nisso, o objetivo deste trabalho é reconhecer automaticamente padrões radiológicos em imagens não segmentadas de TC do pulmão, e também classificar áreas externas à região pulmonar, além dos cinco padrões pulmonares: normal, enfisema, opacidade em vidro-fosco, fibrose e micronódulos. Para atingir os objetivos, uma metodologia que avalia diversos descritores é proposta. Em busca de uma complementaridade entre os resultados, são realizadas combinações entre as respostas dos classificadores. Utilizando uma base de imagens pública, com 20.540 blocos de imagens de TC, cada descritor foi submetido a um classificador *Support Vector Machine* que gerou *scores* para representar a probabilidade das amostras pertencerem a cada uma das seis classes. Após isso, os classificadores foram combinados empregando algumas regras de fusão. O classificador CLBP invariante à rotação e uniforme obteve o melhor resultado individual, e combinado com a transformação top-hat, atingiram taxas de 81,65% de sensibilidade e 96,06% de especificidade. Com base na relevância dos resultados obtidos, o próximo passo é analisar a possibilidade de desenvolver uma ferramenta de auxílio ao diagnóstico utilizando a metodologia desenvolvida.

**Palavras-chave:** padrões radiológicos, doenças pulmonares difusas, Tomografia Computadorizada, reconhecimento de padrões.

## ABSTRACT

Diffuse lung diseases (DLDs) are comprised of more than 180 diseases that cause damage to the lungs. Patients who have these diseases present changes in lung tissue, which can be seen on Computed Tomography (CT) images. Thus, it becomes important to identify alterations in these images because, along with other information, they can define the type of disorder that the patient has. The aim of this study is to automatically recognize radiological patterns in non-segmented lung CT images, and also to classify the areas outside the lung region, in addition to five lung patterns: normal, emphysema, ground-glass opacity, fibrosis and micronodules. To achieve the objectives, a methodology that evaluates various descriptors is proposed. Searching for a complementarity of the results, combinations are performed among the classifiers. Using a public dataset containing 20.540 CT blocks, each descriptor was subjected to a classifier Support Vector Machine generating scores to represent the likelihood of the samples belonging to each of the six classes. After that, the classifiers were combined using fusion rules. The rotation invariant uniform CLBP classifier achieved the best individual result, and combined with top-hat transform descriptor, reached rates of 81.65% of sensitivity and 96.06% of specificity. Based on the relevance of the results obtained, the next step is to analyze the possibility of developing a support tool for diagnosis using the developed methodology.

**Keywords:** radiological patterns, diffuse lung diseases, Computed Tomography, pattern recognition.

## LISTA DE FIGURAS

1.1	Imagens do tórax dos tipos (a) Radiografia (Arzhaeva et al., 2006) e (b) TCAR (Depeursinge et al., 2011). . . . .	2
1.2	Exemplo de erro de segmentação em pulmão patológico: (a) Imagem original, (b) Segmentação manual e (c) Segmentação convencional (adaptado de Sluimer et al. (2004)). . . . .	6
2.1	Achados radiológicos mais frequentes em imagens de Tomografia Computadorizada: (a) Padrão normal (b) Enfisema, (c) Vidro-fosco, (d) Fibrose, (e) Micronódulos (Depeursinge et al., 2011). . . . .	11
2.2	Cálculo do LBP (adaptado de Mäenpää et al. (2000)). . . . .	15
2.3	Os 58 padrões uniformes que podem ocorrer em uma vizinhança circular (8, $R$ ) (adaptado de Ahonen et al. (2009)). . . . .	16
2.4	Histograma de níveis de cinza e medidas estatísticas extraídas da imagem (Depeursinge, 2010). . . . .	23
2.5	Exemplo de um problema binário linearmente separável em um espaço bidimensional (Adaptado de Cortes and Vapnik (1995)). . . . .	26
2.6	Arquitetura de combinação de classificadores em paralelo. . . . .	28
4.1	Sequência de etapas do método proposto. . . . .	46
4.2	ROIs demarcadas por especialistas, onde o padrão enfisema é representado pelas áreas azuis, e as áreas vermelhas representam a fibrose. . . . .	47
4.3	Exemplo de extração de bloco de ROI pulmonar de tamanho 32x32 pixels (adaptado de Song et al. (2013)). . . . .	49
4.4	Método de extração de blocos de não-pulmão, onde (a) imagem original, (b) máscara pulmonar, (c) máscara pulmonar dilatada e (d) imagem resultante, onde a parte cinza é a ROI não-pulmonar. . . . .	50

4.5	Exemplo dos blocos extraídos de (a) não-pulmão, (b) pulmão normal, (c) enfisema, (d) vidro-fosco, (e) fibrose e (f) micronódulos. . . . .	51
4.6	Processo de classificação para cada característica extraída. . . . .	55
4.7	Processo de fusão de respostas dos classificadores. . . . .	57
5.1	Curva ROC da melhor combinação de classificadores. . . . .	67

## LISTA DE TABELAS

1.1	13 diagnósticos histológicos comuns de DPDs e os achados de TCAR associados. . . . .	3
2.1	Exemplo de matriz de confusão de duas classes. . . . .	31
3.1	Síntese dos trabalhos de análise e classificação de padrões de doenças pulmonares. . . . .	41
4.1	Quantidade de blocos de ROIs extraídos. . . . .	51
4.2	Resumo dos descritores utilizados. . . . .	55
4.3	Número de pacientes e blocos por <i>fold</i> . . . . .	56
4.4	Número de exemplos de cada classe por <i>fold</i> . . . . .	56
5.1	Resultados gerais dos 15 classificadores individuais (em %) . . . . .	60
5.2	Matriz de confusão da melhor configuração do CLBP ( $CLBP\_SMC_{(8,2)}^{riu2}$ ) (em %) . . . . .	61
5.3	Matriz de confusão da transformação top-hat (em %) . . . . .	62
5.4	Matriz de confusão do histograma (em %) . . . . .	63
5.5	Matriz de confusão da melhor configuração do LBP ( $LBP_{(8,2)}^{u2}$ ) (em %) . . . . .	63
5.6	Matriz de confusão da melhor configuração do LPQ ( $LPQ_{(8,1)}^g$ ) (em %) . . . . .	64
5.7	Matriz de confusão da GLCM (em %) . . . . .	64
5.8	Resultados gerais das 10 melhores combinações (em %) . . . . .	65
5.9	Matriz de confusão da melhor combinação de classificadores (em %) . . . . .	66
5.10	Matriz de confusão simplificada da melhor combinação de classificadores (em %). . . . .	67
5.11	Sensibilidade média por <i>fold</i> da melhor combinação de classificadores (em %). . . . .	68
5.12	Comparativo com os resultados obtidos na literatura (em %). . . . .	71

## LISTA DE SIGLAS

CAD	<i>Computer-Aided Diagnosis</i> - Sistema de Auxílio ao Diagnóstico
CLBP	<i>Completed Local Binary Pattern</i>
DICOM	<i>Digital Imaging and Communication in Medicine</i>
DPD	Doença Pulmonar Difusa
FN	Falso Negativo
FP	Falso Positivo
GLCM	<i>Gray-Level Co-occurrence Matrix</i> - Matriz de Co-ocorrência de Níveis de Cinza
HU	<i>Hounsfield Unit</i> - Unidade Hounsfield
LBP	<i>Local Binary Pattern</i>
LPQ	<i>Local Phase Quantization</i>
RBF	<i>Radial Basis Function</i>
ROI	<i>Region of Interest</i> - Região de Interesse
SVM	<i>Support Vector Machine</i> - Máquina de Vetor de Suporte
TC	Tomografia Computadorizada
TCAR	Tomografia Computadorizada de Alta Resolução
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
1.1	Motivação . . . . .	2
1.2	Descrição do problema . . . . .	5
1.3	Objetivos . . . . .	6
1.4	Desafios . . . . .	7
1.5	Contribuições do trabalho . . . . .	8
1.6	Organização do documento . . . . .	8
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>9</b>
2.1	Achados radiológicos . . . . .	9
2.2	Tomografia Computadorizada . . . . .	11
2.3	Diagnóstico Auxiliado por Computador . . . . .	13
2.4	Extração de características . . . . .	14
2.4.1	Local Binary Pattern . . . . .	14
2.4.2	Completed Local Binary Pattern . . . . .	17
2.4.3	Local Phase Quantization . . . . .	19
2.4.4	Matriz de Co-ocorrência de Níveis de Cinza . . . . .	21
2.4.5	Histograma de níveis de cinza . . . . .	22
2.4.6	Transformação top-hat . . . . .	24
2.5	Classificação . . . . .	25
2.5.1	Support Vector Machine . . . . .	26
2.5.2	Combinação de classificadores . . . . .	28
2.5.3	Métricas de avaliação dos classificadores . . . . .	31
<b>3</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>33</b>
3.1	Metodologias para a classificação de padrões pulmonares . . . . .	34
3.2	Resumo dos trabalhos apresentados . . . . .	40

<b>4</b>	<b>MÉTODO PROPOSTO</b>	<b>45</b>
4.1	Base de imagens . . . . .	45
4.2	Extração dos blocos de imagem . . . . .	47
4.3	Extração de características . . . . .	51
4.3.1	LBP . . . . .	51
4.3.2	CLBP . . . . .	52
4.3.3	LPQ . . . . .	52
4.3.4	GLCM . . . . .	52
4.3.5	Histograma e medidas estatísticas . . . . .	53
4.3.6	Transformação <i>top-hat</i> . . . . .	54
4.3.7	Resumo dos descritores utilizados . . . . .	54
4.4	Classificação . . . . .	54
4.5	Fusão dos classificadores . . . . .	56
<b>5</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>59</b>
5.1	Resultados individuais . . . . .	59
5.2	Resultados da combinação de classificadores . . . . .	64
5.3	Conclusão . . . . .	69
<b>6</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>72</b>
6.1	Trabalhos futuros . . . . .	74
	<b>BIBLIOGRAFIA</b>	<b>76</b>

## CAPÍTULO 1

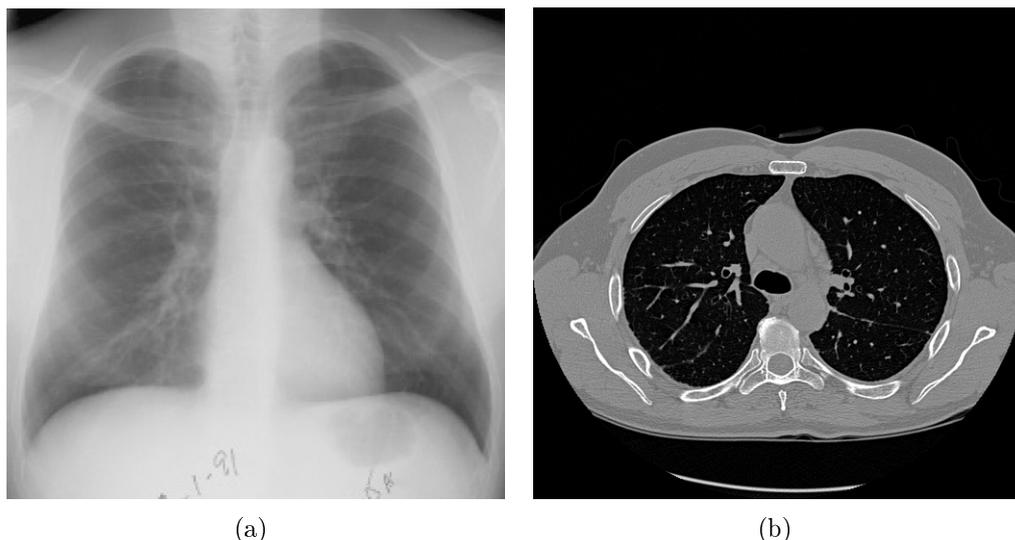
### INTRODUÇÃO

Dentre as diversas doenças que afetam os pulmões, as Doenças Pulmonares Difusas (DPDs) são compostas por mais de 180 patologias heterogêneas que causam danos nos pulmões (Martinez, 1998). De acordo com Depeursinge et al. (2011), as DPDs são caracterizadas pela alteração gradual do parênquima pulmonar, o que acarreta em problemas respiratórios.

Depeursinge et al. (2011) afirmam que, para se diagnosticar uma doença, o especialista leva em consideração o histórico completo do paciente, exame físico, exames laboratoriais, teste de função pulmonar, e também os aspectos visuais da radiografia de tórax. Porém, a radiografia do tórax fornece um diagnóstico confiável em apenas 23% dos casos de doenças pulmonares. Já a Tomografia Computadorizada de Alta Resolução (TCAR) fornece uma avaliação anatômica global dos pulmões, melhorando a sensibilidade e a especificidade do diagnóstico clínico (Elicker et al., 2008). Dessa forma, as imagens de TCAR são necessárias para se ter uma avaliação rápida e correta do tecido pulmonar. A diferença entre uma imagem de radiografia de tórax e uma imagem de TCAR do tórax pode ser visualizada na Figura 1.1.

Pacientes com DPDs apresentam certas alterações no tecido pulmonar, que podem ser visualizadas nas imagens de Tomografia Computadorizada (TC). Essas alterações são chamadas de ‘achados radiológicos’, de acordo com Ponciano-Silva (2009), ou ‘aspectos radiológicos’, conforme Rubin et al. (2012). Ponciano-Silva (2009) ainda define achados radiológicos como: “lesões, anormalidades ou alterações na estrutura normal de um tecido observado nas imagens médicas de exames radiológicos”. Dentre os achados radiológicos mais frequentes em imagens de TC, estão: enfisema, opacidade em vidro-fosco, fibrose e micronódulos (Depeursinge et al., 2011).

Embora um único achado radiológico na TCAR seja inespecífico, a combinação de



**Figura 1.1:** Imagens do tórax dos tipos (a) Radiografia (Arzhaeva et al., 2006) e (b) TCAR (Depeursinge et al., 2011).

vários achados juntamente com sua distribuição anatômica pode fornecer um possível diagnóstico (Elicker et al., 2008). Assim sendo, é importante identificar esses padrões radiológicos anormais pois, juntamente com outras informações, é possível definir que tipo de doença pulmonar o paciente possui. Com base nisso, o problema que este trabalho aborda está relacionado ao reconhecimento dos padrões de doenças pulmonares mais frequentes em imagens de TCAR.

Este capítulo apresenta na seção 1.1 a motivação para o trabalho realizado. A seção 1.2 apresenta a descrição do problema. Na seção 1.3 descreve-se detalhadamente o objetivo geral e os objetivos específicos. Na seção 1.4 são apresentados os desafios que envolvem o tema. A seção 1.5 descreve as contribuições do trabalho. Por fim, na seção 1.6 é apresentada uma descrição sucinta dos demais capítulos deste documento.

## 1.1 Motivação

As DPDs estão associadas a diversos achados radiológicos que ajudam a determinar o tipo de doença que acomete os pulmões. A análise de imagens desempenha um papel importante no diagnóstico dessas doenças pulmonares devido a diversos aspectos. Um importante aspecto está relacionado à detecção de doenças em seu estágio inicial. Martinez (1998) explica que a TCAR fornece informações sobre o grau de inflamação e fibrose

existentes em um paciente e que a presença do achado radiológico ‘opacidade em vidro-fosco’ pode ser o indicador de uma doença inicial, com possibilidade de obter boa resposta ao tratamento. Além disso, por meio da análise da imagem pode-se reduzir as indicações de biópsias pulmonares aos pacientes, pois torna-se possível detectar quadros tomográficos específicos de acordo com a combinação dos achados radiológicos presentes (Martinez, 1998). Essa combinação de achados radiológicos que determinam um diagnóstico pode ser visualizada na Tabela 1.1, que apresenta os 13 diagnósticos histológicos mais comuns, e seus achados radiológicos.

**Tabela 1.1:** 13 diagnósticos histológicos comuns de DPDs e os achados de TCAR associados.

Diagnóstico histológico	Achados radiológicos na TCAR
Pneumonia de hipersensibilidade	vidro-fosco, enfisema, fibrose
Pneumocistose	vidro-fosco, pavimentação em mosaico, cistos, pneumotórax
Pneumonia eosinofílica	vidro-fosco, consolidação, pavimentação em mosaico
Histiocitose de células de Langerhans	cistos, vidro-fosco, micronódulos, reticulação
Sarcoidose	micronódulos, consolidação, macronódulos, vidro-fosco, fibrose (estágio final)
Tuberculose	micronódulos (miliar), árvore em brotamento, consolidação
Bronquiolite respiratória associada à DPD	vidro-fosco, enfisema
Pneumonia Intersticial Usual	fibrose, bronquiectasia, vidro-fosco
Pneumonia intersticial descamativa	vidro-fosco, enfisema, fibrose (incomum)
Pneumonia intersticial aguda	vidro-fosco, consolidação
Pneumonia intersticial não específica	vidro-fosco, consolidação, reticulação, fibrose (incomum)
Pneumonia em organização criptogênica	vidro-fosco, consolidação (irregular), macronódulos, espessamento das paredes brônquicas, pavimentação em mosaico
Pneumonia intersticial linfocítica	vidro-fosco, micronódulos

Fonte: adaptado de Depeursinge et al. (2011).

Como é possível verificar na Tabela 1.1, não existe um padrão radiológico específico para cada diagnóstico. Por exemplo, o padrão ‘vidro-fosco’ está presente na grande maioria das DPDs (Depeursinge et al., 2011). Mesmo assim, a detecção desses padrões mostra-se importante, pois através da quantificação dos padrões presentes é possível obter uma perspectiva sobre o grau de gravidade da patologia (Bagci et al., 2012). Especialis-

tas afirmam que a TCAR pode ser conclusiva no diagnóstico de um paciente quando é aliada aos dados clínicos, permitindo o acerto do diagnóstico em aproximadamente 60% dos casos de DPD (Rubin et al., 2012).

Os vários achados tomográficos podem ser combinados para formar padrões típicos. Esses, em conjunto com a distribuição das anormalidades nos planos axial e coronal e com os dados clínicos fornecem uma abordagem clínica e radiológica ao paciente, podendo estreitar o diagnóstico das doenças pulmonares difusas (Elicker et al., 2008).

De acordo com Zhao et al. (2013), a interpretação de uma imagem de TC pulmonar depende muito da experiência do médico especialista. Também há o problema da grande quantidade de cortes tomográficos, que gera uma carga sobre os radiologistas. Sendo assim, a utilização de uma ferramenta de diagnóstico auxiliado por computador (*Computer-Aided Diagnosis - CAD*) que forneça aos radiologistas uma segunda opinião em relação ao diagnóstico de doenças pulmonares é de extrema utilidade, não só para facilitar o trabalho do especialista como, também, para diminuir o grau de incerteza diagnóstica. Ressalta-se que o computador é utilizado somente como uma ferramenta para obtenção de informação adicional, sendo o diagnóstico final sempre feito pelo radiologista (Azevedo-Marques, 2001).

O desenvolvimento de uma ferramenta de detecção de padrões de doenças pulmonares em imagens de TC pode ser dividido basicamente em três etapas. A primeira etapa consiste em realizar a segmentação pulmonar. Na segunda etapa, é necessário definir métodos de extração de características e de classificação que tornem possível identificar corretamente os diversos padrões radiológicos. Por fim, na terceira etapa ocorre o processo de análise da imagem de TC. Nessa etapa, cada região do pulmão recebe o rótulo de um determinado padrão radiológico, podendo utilizar para isso informações não somente da classificação para aquela área, como também da vizinhança da região, e da localização da anormalidade no pulmão.

Com base nisso, este trabalho está inserido somente no contexto da extração de características e classificação de padrões radiológicos.

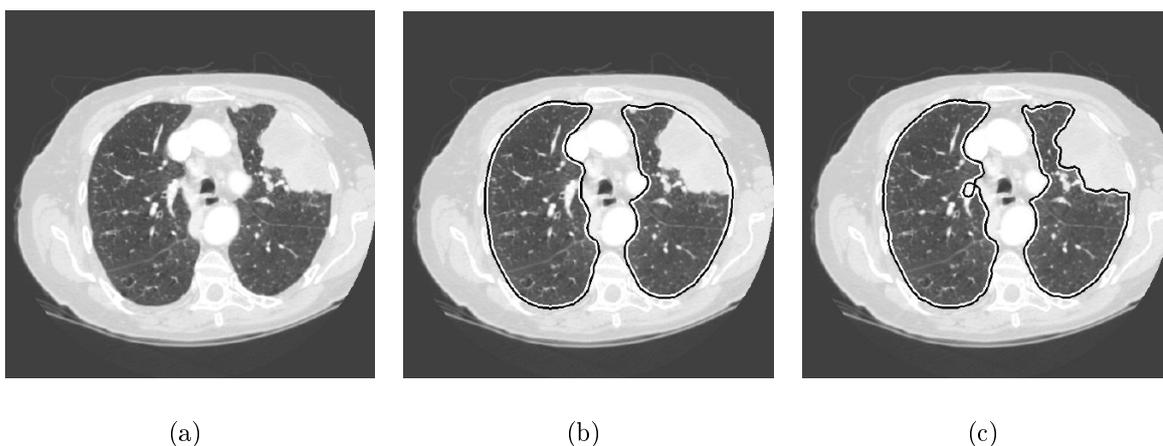
## 1.2 Descrição do problema

Se o diagnóstico de padrões radiológicos não é uma tarefa trivial mesmo para radiologistas, é normal que a classificação desses padrões, quando realizada por um computador, esteja associada a diversos problemas. Sluimer et al. (2006) alegam que as diversas propostas de resolução do problema de classificação de padrões pulmonares encontradas na literatura mostram uma considerável variação, ilustrando que não existe consenso sobre um objetivo comum e definições exatas.

Alguns problemas relacionados ao reconhecimento dos padrões radiológicos são descritos abaixo:

- **Segmentação pulmonar:** A segmentação dos pulmões é uma etapa necessária na classificação do tecido pulmonar, que indica as regiões a serem analisadas pelas rotinas de análise textural (Korfiatis et al., 2007). Bagci et al. (2012) afirmam que a segmentação dos pulmões é um processo desafiante, especialmente quando eles possuem anormalidades. De acordo com Sluimer et al. (2006), isso ocorre devido a falta de contraste entre as regiões patológicas e o tecido ao redor, pois as patologias de alta atenuação ficam conectadas às bordas do pulmão. Uma estratégia bastante utilizada em diversos artigos é a segmentação semi-automática dos pulmões que contém anormalidades mas, como explicado por Sluimer et al. (2006), delimitar manualmente estruturas de interesse nas imagens torna-se um processo exageradamente grande, considerando o número de cortes tomográficos gerados para cada paciente. O problema que envolve realizar a segmentação em pulmões que contém anormalidades pode ser visto na Figura 1.2. Na figura é possível perceber que, ao utilizar técnicas convencionais de segmentação (limiarização, por exemplo), algumas áreas pulmonares anormais acabam por não serem reconhecidas, acarretando no não reconhecimento como parte do pulmão;
- **Variações dos padrões pulmonares:** refere-se à pequena variação interclasse e grande variação intraclasse dos padrões radiológicos. Como explicado por Song et al. (2013), as pequenas variações interclasse são comuns em padrões como enfisema e normal,

vidro-fosco e fibrose, e micronódulos e tecido normal. Por outro lado, a variação intraclasse pode ocorrer em padrões como enfisema e fibrose, que possuem estruturas locais irregulares, apresentando grandes variações entre imagens da mesma classe. No entanto, essas variações dos padrões também geram confusão entre especialistas (Park et al., 2009), mostrando que esse é o principal problema a ser enfrentado neste trabalho.



**Figura 1.2:** Exemplo de erro de segmentação em pulmão patológico: (a) Imagem original, (b) Segmentação manual e (c) Segmentação convencional (adaptado de Sluimer et al. (2004)).

### 1.3 Objetivos

O presente trabalho tem como principal objetivo reconhecer automaticamente padrões radiológicos em blocos de imagens de TCAR não segmentadas, ou seja, imagens que não foram submetidas a uma prévia segmentação pulmonar. Além do objetivo principal, os seguintes objetivos específicos podem ser citados:

- Classificar áreas externas ao pulmão (ex.: gordura, tecido e osso) possibilitando, com isso, obter uma segmentação pulmonar grosseira a fim de reduzir erros decorrentes do processo de segmentação;
- Identificar os seguintes padrões radiológicos: tecido pulmonar normal, enfisema, vidro-fosco, fibrose e micronódulos, além de áreas não pulmonares;

- Avaliar diversos métodos de extração de características com o intuito de identificar, dentre os métodos testados, qual a melhor combinação de características para reconhecer os padrões radiológicos deste trabalho;
- Utilizar blocos de imagens do pulmão que forneçam uma representação adequada dos aspectos visuais de cada padrão, para avaliar os métodos de extração de características;
- Utilizar uma base de imagens pública, tornando possível a replicação do método desenvolvido com as mesmas imagens;
- Comparar os resultados obtidos com os métodos da literatura.

## 1.4 Desafios

Com base nos problemas apresentados na seção 1.2, os seguintes desafios foram identificados:

- Desenvolver um método automático de classificação dos padrões radiológicos que não dependa de uma segmentação prévia dos pulmões, visto que esta é muito sujeita a erros quando os pulmões possuem anormalidades;
- Encontrar características que aumentem a variação interclasse e diminuam a variação intraclasse, a fim de obter melhores resultados de reconhecimento. As variações dos padrões de tecido pulmonar também causam confusão entre os especialistas, e as características devem ser selecionadas de modo que destaquem as diferenças entre cada padrão pulmonar e, ao mesmo tempo, reduzam as variações entre imagens de um mesmo padrão pulmonar.

Em suma, como foi afirmado por Sluimer et al. (2006), o maior desafio provavelmente se encontra em projetar algoritmos que sejam robustos contra as variedades patológicas e anatômicas do pulmão.

## 1.5 Contribuições do trabalho

Como contribuições geradas pela realização deste trabalho, destacam-se as citadas abaixo:

- O desenvolvimento de um método de classificação que não depende de uma imagem pré-segmentada, pois classifica as áreas externas ao pulmão;
- Uma análise de diversos extratores de características e a identificação do conjunto de características que melhor classifica os padrões radiológicos, dentre os métodos estudados;
- Um método automático de classificação de blocos de imagens de TCAR que representa de forma adequada a prática clínica e que pode ser integrado a um sistema de auxílio ao diagnóstico.

Além dessas contribuições, os resultados preliminares obtidos neste trabalho foram publicados no 28th IEEE International Symposium on Computer-Based Medical Systems, com o título “Blockwise Classification of Lung Patterns in Unsegmented CT Images”, tendo como autores Luiza Dri Bagesteiro, Lucas Ferrari de Oliveira e Daniel Weingaertner.

## 1.6 Organização do documento

Esta dissertação de mestrado está estruturada como segue:

Capítulo 2: Descreve os principais conceitos relacionados a padrões radiológicos e reconhecimento de padrões, necessários para o entendimento deste trabalho.

Capítulo 3: Apresenta alguns trabalhos relacionados ao tema de reconhecimento de padrões radiológicos.

Capítulo 4: Descreve, de forma detalhada, o método proposto para realizar a classificação de padrões radiológicos, utilizando combinação de classificadores.

Capítulo 5: Apresenta resultados experimentais e sua discussão.

Capítulo 6: Considerações finais após a realização deste trabalho, abordando também as perspectivas de trabalhos futuros.

## CAPÍTULO 2

### FUNDAMENTAÇÃO TEÓRICA

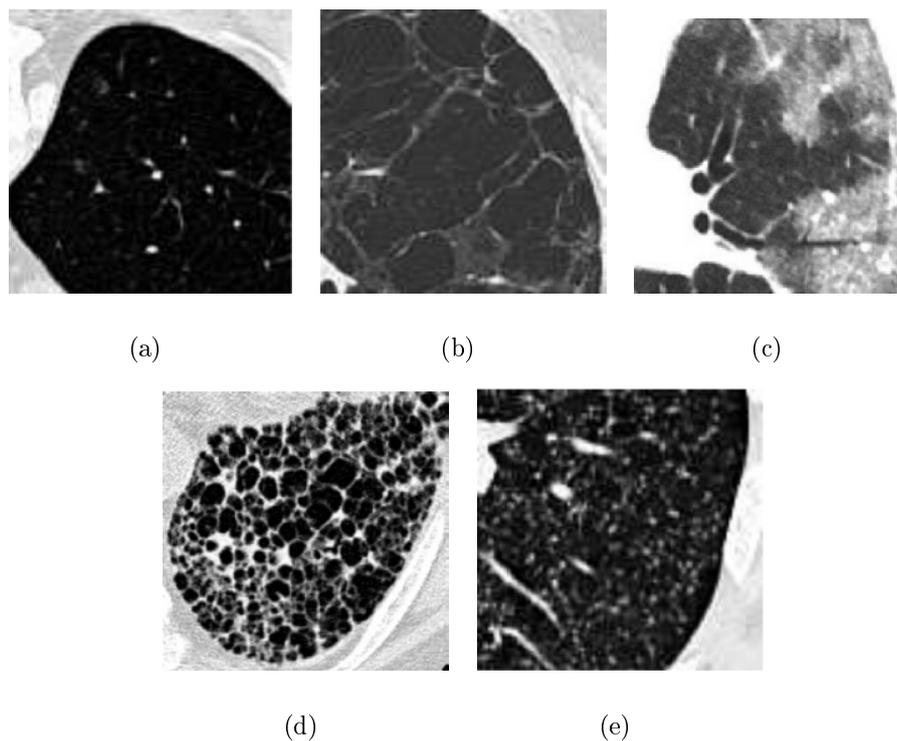
Este capítulo tem por objetivo apresentar os conceitos relacionados à classificação de padrões radiológicos. Isso inclui as definições de achados radiológicos e Tomografia Computadorizada (TC), bem como métodos de extração de características e classificação. Dessa forma, a seção 2.1 apresenta uma descrição da aparência dos achados radiológicos mais frequentes nas imagens de TC. Na seção 2.2, são descritos alguns aspectos das imagens de Tomografia Computadorizada. A seção 2.3 aborda algumas definições referentes ao Diagnóstico Auxiliado por Computador (*Computer-Aided Diagnosis* - CAD). Já na seção 2.4 são explicados os conceitos dos métodos de extração de características utilizados neste trabalho, que são: *Local Binary Pattern* (LBP), *Completed Local Binary Pattern* (CLBP), *Local Phase Quantization* (LPQ), Matriz de Co-ocorrência de Níveis de Cinza (*Gray Level Co-occurrence Matrix* - GLCM), Histograma de Níveis de Cinza e Transformação *Top-Hat*. Por fim, a seção 2.5 apresenta os conceitos de classificação utilizados neste trabalho, como o classificador *Support Vector Machine* (SVM), métodos de fusão de classificadores e as métricas utilizadas para avaliar a performance das características.

#### 2.1 Achados radiológicos

A aparência e avaliação dos tipos de padrões pulmonares possuem informações que são capazes de auxiliar no diagnóstico de uma doença intersticial (Depeursinge et al., 2011). Como mencionado anteriormente, esses padrões são comumente chamados pelos radiologistas de ‘achados radiológicos’ (Ponciano-Silva, 2009).

O padrão de tecido pulmonar normal pode ser visualizado na Figura 2.1a. Já os padrões pulmonares anormais mais encontrados em pacientes com DPD são: enfisema, opacidade em vidro-fosco, fibrose e micronódulos. Esses padrões são descritos de acordo com seus aspectos visuais nas imagens de TC:

- **Enfisema:** Na TC, o enfisema caracteriza-se por regiões de baixa atenuação (Hansell et al., 2008). Regiões que possuem valores abaixo de -950HU na TC geralmente estão associados a este tipo de padrão (ver Figura 2.1b);
- **Opacidade em vidro-fosco:** Na TC, o padrão vidro-fosco aparece como um aumento da opacidade dos pulmões, devido a maior atenuação dos feixes de raio-X, no qual ainda é possível identificar os contornos dos vasos e estruturas brônquicas no interior dos pulmões (Hansell et al., 2008) (ver Figura 2.1c);
- **Fibrose:** A base de imagens de Depeursinge et al. (2011), utilizada neste trabalho, empregou o termo ‘fibrose’ para descrever todos os achados radiológicos que possuem o diagnóstico ‘fibrose pulmonar’, incluindo reticulação, bronquiectasia de tração, distorção arquitetural pulmonar e faveolamento. As características do faveolamento em imagens de TC incluem cistos pulmonares de destruição fibrosados, que possuem espaços císticos agrupados, geralmente com diâmetro variando de 0,3 mm a 1 cm (Souza Jr. et al., 2002). Ponciano-Silva (2009) descreve as paredes dos cistos como sendo bem definidas (ver Figura 2.1d);
- **Micronódulos:** O micronódulo apresenta-se na TC como uma opacidade discreta, arredondada, focal e pequena, que possui diâmetro menor que 7 mm, 5 mm ou até mesmo 3 mm, de acordo com alguns autores (Hansell et al., 2008) (ver Figura 2.1e).



**Figura 2.1:** Achados radiológicos mais frequentes em imagens de Tomografia Computadorizada: (a) Padrão normal (b) Enfisema, (c) Vidro-fosco, (d) Fibrose, (e) Micronódulos (Depeursinge et al., 2011).

## 2.2 Tomografia Computadorizada

A TC é um exame que produz uma coleção de imagens em tons de cinza de determinada parte do corpo ou de um órgão, em que cada imagem mostra uma ‘fatia’, ou seja, uma parte dessa região corporal. Uma Tomografia Computadorizada de Alta Resolução (TCAR), por exemplo, possui a espessura da fatia de 1 mm a 2 mm aproximadamente, um tamanho de  $512 \times 512$  pixels, e o tamanho de um pixel pode variar de 0.4 mm a 1 mm (Depeursinge, 2010).

As imagens geradas representam as propriedades de atenuação de raios-X do corpo (Flokstra et al., 2014). Assim, o valor numérico de cada pixel é obtido através do cálculo da diferença na atenuação dos raios-X entre o material observado (osso, tecido pulmonar, ar) e da água, proporcionando uma medida da atenuação dos órgãos (Depeursinge, 2010). Os valores dos pixels são medidos em Unidades Hounsfield (*Hounsfield Units* - HU), uma escala de níveis de cinza criada especialmente para imagens de TC. Os valores das

Unidades Hounsfield para vários materiais orgânicos que aparecem na TC são descritos abaixo, sendo o intervalo de números de  $[-1000, 3000]$  em geral:

- Ar: -1000
- Pulmão normal: -750 a -850
- Gordura: -40 a -100
- Água: 0
- Partes moles: 20 a 80
- Calcificação: acima de 200
- Osso: acima de 250

Como é possível observar, os valores que correspondem ao tecido pulmonar normal variam de -750 HU a -850 HU, já a gordura e os tecidos moles estão entre -40 HU a -100 HU e 20 HU a 80 HU, respectivamente. Por fim, o osso compreende os valores acima de 250 HU.

Um aspecto importante a ser observado é o tipo de imagem gerada pelos aparelhos modernos de radiologia, como a TC. As imagens são criadas no formato DICOM (*Digital Imaging and Communication in Medicine*), que permite o armazenamento das informações do paciente juntamente com a imagem, e de forma estruturada (Caritá et al., 2004). Esse conjunto de informações fica gravado no início do arquivo de imagem, que é chamado de ‘cabeçalho DICOM’, e os pixels da imagem são armazenados após esse cabeçalho. São utilizados 12 bits para guardar o valor de intensidade do pixel, gerando imagens com 4096 níveis de cinza, porém, são alocados na memória 16 bits por pixel (Pianykh, 2012).

O protocolo de imagens de TCAR será utilizado neste trabalho, visto que é o protocolo padrão no diagnóstico de DPIs (Depeursinge et al., 2011).

## 2.3 Diagnóstico Auxiliado por Computador

É possível definir o termo ‘diagnóstico auxiliado por computador’ como um diagnóstico feito por um médico especialista que apenas utiliza os resultados da análise computadorizada como uma ‘segunda opinião’, ajudando na formação do diagnóstico. Nesse caso, cabe observar que o computador apenas fornece informações adicionais, sendo o diagnóstico final sempre feito por um especialista (Doi, 1999). Isso difere os termos ‘diagnóstico auxiliado por computador’ e ‘diagnóstico automatizado’. Com o diagnóstico automatizado, a performance do computador deve ser maior ou igual que a do especialista. Já com o CAD, a performance do computador não precisa ser comparável ou melhor que a dos especialistas, e sim, complementar (Doi, 2007).

O objetivo do CAD é utilizar a resposta do computador apenas como referência, aprimorando tanto a precisão do diagnóstico quanto a consistência da interpretação da imagem. Como podem existir diferenças entre os radiologistas em relação ao diagnóstico, ou até mesmo certas lesões passarem despercebidas na análise, a resposta do computador mostra-se útil nesses casos (Doi, 1999). Como afirmado por Doi (2007), com CAD é possível combinar o conhecimento do radiologista com a capacidade do computador.

Usualmente, dois tipos gerais de aplicações são utilizadas em sistemas CAD. Uma delas tem como propósito localizar lesões em imagens como, por exemplo, nódulos pulmonares em imagens do tórax. O outro tipo consiste no auxílio ao diagnóstico, quantificando as características da imagem em padrão normal ou anormal como, por exemplo, texturas pulmonares associadas a doenças intersticiais (Doi, 1999). Azevedo-Marques (2001) afirma que um dos objetivos dos sistemas de auxílio ao diagnóstico é reduzir o número de casos benignos que são encaminhados para biópsias, porém, não identificar uma lesão maligna é mais incorreto que classificar erroneamente um caso benigno.

Em suma, a importância de uma ferramenta CAD na prática clínica é compreensível, visto que resultados de estudos mostram um progresso no desempenho quando especialistas empregam sistemas de auxílio computadorizado, se comparado às situações onde esses sistemas não são utilizados (Azevedo-Marques, 2001).

## 2.4 Extração de características

Para se resolver um problema de reconhecimento de padrões, uma das etapas necessárias envolve a extração de características das imagens, que serão utilizadas no treinamento dos classificadores, bem como na classificação dos objetos. Dessa forma, é necessário escolher adequadamente o conjunto de características que melhor representa as diferenças entre as classes dos objetos. Theodoridis and Koutroumbas (2008) afirmam que essa escolha está diretamente relacionada ao domínio do problema e às características disponíveis.

A textura é uma das características que mais se assemelha aos aspectos visuais dos achados radiológicos, sendo que a taxonomia utilizada para descrevê-los muitas vezes referem-se à propriedades texturais (Depeursinge et al., 2011). Todos os trabalhos da literatura estudados recorrem à utilização de descritores de textura para resolver o problema de classificação de padrões de doenças pulmonares. Alguns trabalhos também extraem características de contorno, como um complemento.

As subseções seguintes apresentam os seis descritores de textura utilizados neste trabalho. O descritor LBP já foi utilizado em alguns trabalhos para identificar certos tipos de enfisema (Sørensen et al., 2008, 2010; Gangeh et al., 2010). Os descritores CLBP e LPQ são relativamente novos (Guo et al., 2010; Ojansivu and Heikkilä, 2008), e foram utilizados nesses experimentos visto que nenhum dos trabalhos da literatura explorou esses métodos. Já a GLCM, Histograma de Níveis de Cinza e Transformação *top-hat* compõem as características de diversos trabalhos apresentados.

### 2.4.1 Local Binary Pattern

O LBP é um descritor de textura baseado em um modelo proposto por Wang and He (1990), no qual o conceito básico é de que a textura de uma imagem pode ser decomposta em um conjunto de pequenas unidades, chamadas de unidades texturais. Essas pequenas unidades são representadas por oito elementos, cada um podendo assumir um dos três valores possíveis (0, 1, 2) obtidos em uma vizinhança de  $3 \times 3$  pixels. Isso gera  $3^8 = 6561$  unidades de textura possíveis. Posteriormente, Ojala et al. (1996) apresentaram uma nova

versão com duas possibilidades de valores (0, 1), gerando apenas  $2^8 = 256$  unidades de textura possíveis.

O LBP de um pixel é calculado de acordo com os níveis de cinza dos seus pixels vizinhos. Essa vizinhança é definida de acordo com um conjunto de amostras  $P$  igualmente espaçadas e distribuídas sobre uma circunferência de raio  $R$  com centro sobre o pixel que está sendo analisado. Quando um vizinho não corresponde a uma posição de um pixel, seu valor é obtido por interpolação.

Para cada pixel da imagem, assume-se o pixel em questão como sendo o ponto central, que é então comparado com cada um dos seus vizinhos. Cada vizinho assume o valor de 0 ou 1, de acordo com a seguinte condição: se o nível de cinza do pixel vizinho  $(x_i, y_i)$  for igual ou superior ao nível de cinza do pixel central  $(x_c, y_c)$ , o valor do pixel vizinho é definido como 1, caso contrário define-se 0. Essa comparação gera uma sequência de zeros e uns que é organizada realizando uma leitura sequencial a partir de um vizinho qualquer do ponto central. Cada elemento binário será então multiplicado por um peso binomial  $2^p$  e, ao somar todos os elementos, obtém-se o padrão que caracteriza a região. O cálculo do LBP é descrito na Equação 2.1, e ilustrado pela Figura 2.2.

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} (x_i, y_i) 2^p \quad (2.1)$$

$$(x_i, y_i) \begin{cases} 1, & \text{se } (x_i, y_i) \geq (x_c, y_c) \\ 0, & \text{caso contrário} \end{cases}$$

exemplo	limiarizado	pesos
6 5 2	1 0 0	1 2 4
7 6 1	1 0 0	128 8
9 8 7	1 1 1	64 32 16

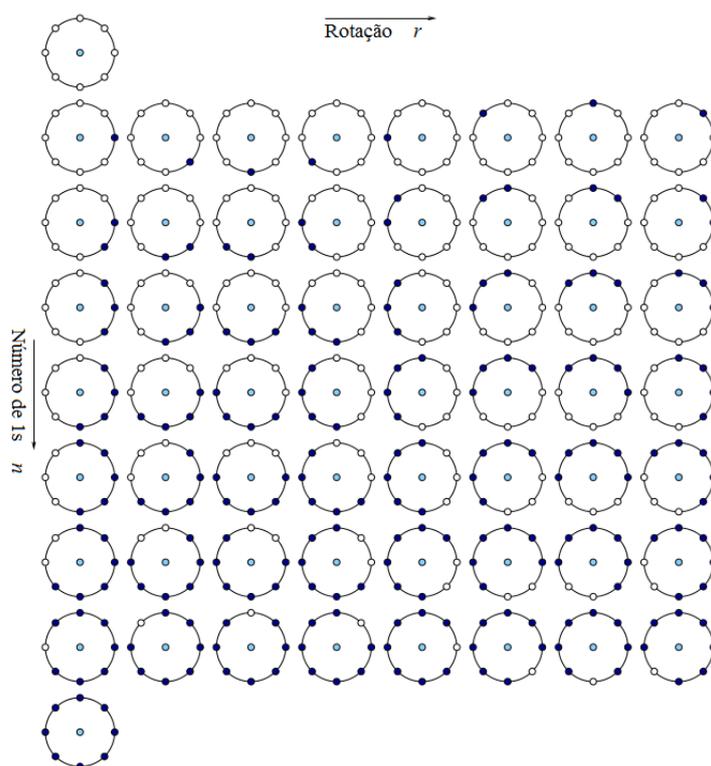
Padrão = **11110001**

**LBP** =  $1 + 16 + 32 + 64 + 128 = 241$

**Figura 2.2:** Cálculo do LBP (adaptado de Mäenpää et al. (2000)).

Mais tarde, Ojala et al. (2002) apresentaram o conceito de uniformidade da sequência obtida no LBP, visto que certos padrões binários local, denominados ‘uniformes’, são propriedades fundamentais da textura, e seu histograma de ocorrência é um poderoso

descritor. Para medir a uniformidade da sequência, basta calcular o número de transições entre zeros e uns presentes. Se o número de transições for menor ou igual a 2, o padrão é considerado uniforme, levando em consideração também o fato de que o código é tratado como uma lista circular. Assim, ao invés de utilizar todos os padrões gerados pelo LBP, apenas os padrões uniformes são considerados. A Figura 2.3 ilustra esses 58 padrões binários uniformes. Ojala et al. (2002) afirmam que, além dos 58 padrões uniformes possíveis, todos os demais padrões não uniformes devem ser armazenados em uma coluna adicional do histograma, totalizando 59 características. Essa variação do LBP que calcula apenas os 59 valores possui a seguinte referência:  $LBP_{P,R}^{u2}$ , sendo  $P$  o número de vizinhos, e  $R$  o valor do raio. A referência ‘u2’ reflete o descritor uniforme com  $U$  valendo, no máximo, 2.



**Figura 2.3:** Os 58 padrões uniformes que podem ocorrer em uma vizinhança circular (8,  $R$ ) (adaptado de Ahonen et al. (2009)).

A forma circular de vizinhança do LBP facilita uma definição de invariância à rotação, no qual cada pixel é considerado o centro de rotação. Para tornar o LBP invariante à rotação, deve-se gerar todos os valores de textura possíveis para o pixel, e então selecionar

o menor valor para representar o padrão binário local. Ou seja, rotacionar o código binário  $P$  vezes, e identificar o menor valor decimal gerado. Assim, um mesmo elemento do descritor acumula todos os padrões binários que possuem o mesmo valor decimal mínimo quando ocorrem  $P$  rotações dos bits. Da mesma forma, a versão uniforme e invariante à rotação *LBP Rotation Invariant and Uniform* ( $LBP_{P,R}^{riu2}$ ) usa apenas os padrões binários e acumula, em um mesmo elemento do descritor, todos os padrões binários que possuem o mesmo valor decimal mínimo quando ocorrem as  $P$  rotações dos bits (Ojala et al., 2002).

A família de descritores LBP já foi utilizada para classificar subtipos de enfisema (Sørensen et al., 2008, 2010; Gangeh et al., 2010), e uma variação do LBP foi proposta por Song et al. (2013) para reconhecer padrões de alterações pulmonares.

## 2.4.2 Completed Local Binary Pattern

O descritor LBP, anteriormente apresentado, possui diversas variações apresentadas na literatura. Algumas delas são: *center-symmetric* LBP (CSLBP) (Heikkilä et al., 2009), e *completed* LBP (CLBP) (Guo et al., 2010).

O método do CLBP foi proposto por Guo et al. (2010), e provê uma modelagem completa do LBP, representando uma região local pelo seu pixel central e também por uma transformada local da diferença sinal-magnitude (*Local Difference Sign-Magnitude Transform* - LDSMT).

No CLBP, a imagem original é representada pelo nível de cinza central (C) e pela diferença local, decomposta pelo sinal (S) e magnitude (M). Para cada pixel central  $g_c$ , o método calcula a diferença entre  $g_c$  e seus vizinhos  $P$  circulares e espaçados uniformemente, definido por  $d_p = g_p - g_c$ . Essa diferença é decomposta nos componentes magnitude ( $m_p$ ) e sinal ( $s_p$ ), como mostra a Equação 2.2.

$$m_p = |d_p|$$

$$s_p = \begin{cases} 1, & \text{se } d_p \geq 0 \\ -1, & \text{caso contrário} \end{cases} \quad (2.2)$$

O operador CLBP\_S corresponde ao operador original LBP, enquanto o operador

CLBP\_M é representado pela Equação 2.3, tornando-o consistente com o operador CLBP\_S.

$$CLBP\_M_{P,R} = \sum_{p=0}^{P-1} t(m_p - c)2^p$$

$$t(x, c) = \begin{cases} 1, & \text{se } x \geq c \\ 0, & \text{caso contrário} \end{cases} \quad (2.3)$$

$R$  é o raio da vizinhança e  $c$  é um limiar a ser determinado de forma adaptativa, que é definido originalmente como o valor médio de  $m_p$  de toda a imagem. Então, codifica-se como CLBP\_C o nível de cinza central, de acordo com a Equação 2.4.

$$CLBP\_C_{P,R} = t(g_c - c_I)$$

$$t(x, c_I) = \begin{cases} 1, & \text{se } x \geq c_I \\ 0, & \text{caso contrário} \end{cases} \quad (2.4)$$

Aqui,  $c_I$  é definido como a média de níveis de cinza de toda a imagem.

O componente do sinal preserva a informação da diferença local, enquanto os componentes da magnitude e o valor de intensidade do pixel podem contribuir com informações discriminantes adicionais. Assim, os operadores CLBP\_C, CLBP\_S e CLBP\_M são propostos para codificar as características C, S e M, respectivamente. A combinação desses operadores forma o mapa de características CLBP da imagem original.

Algumas observações devem ser feitas sobre o CLBP. A primeira refere-se ao fato de que o LBP é um caso especial de CLBP, que utiliza somente o CLBP\_S. Outra consideração a fazer é que o componente do sinal preserva mais informações sobre a estrutura da imagem que o componente da magnitude, e isso explica o fato de o LBP (CLBP\_S) funcionar melhor que o CLBP\_M para reconhecimento por textura.

De forma similar às versões uniforme e invariante à rotação e uniforme do LBP, aqui também pode-se definir as mesmas versões para o CLBP:  $CLBP_{P,R}^{u2}$  e  $CLBP_{P,R}^{riu2}$ , respectivamente.

Todos os três mapas estão em forma binária, de forma que podem ser facilmente combinados para formar o histograma final. Os três operadores podem ser combinados de

duas maneiras, em conjunto, ou de forma híbrida. Na primeira forma, constrói-se um histograma 3D conjunto, denotado CLBP\_S/M/C. Na segunda forma, um histograma 2D conjunto é construído, sendo ele CLBP\_S/C ou CLBP\_M/C, sendo após convertido para um histograma 1D, que é concatenado com CLBP\_M ou CLBP\_S para gerar o histograma conjunto, denotado por CLBP\_S\_M/C. Neste trabalho, a combinação utilizada para gerar o histograma foi a CLBP\_S/M/C, que é o esquema que apresentou melhores resultados no trabalho de Guo et al. (2010).

### 2.4.3 Local Phase Quantization

Ojansivu and Heikkilä (2008) propuseram um método de extração de características texturais denominado *Local Phase Quantization* (LPQ), que é caracterizado por ser insensível ao borramento simétrico de uma imagem, podendo ser utilizado em imagens com ou sem perda de foco. Existem métodos computacionais que permitem retirar o borrão da imagem, porém eles podem introduzir novos artefatos, sendo desejável poder analisar a textura da imagem de uma forma insensível a esse tipo de interferência.

O descritor se baseia na fase quantizada da Transformada Discreta de Fourier (*Discrete Fourier Transform* - DFT), calculado localmente em uma janela para cada posição da imagem e apresentando os códigos resultantes na forma de histograma. Como apenas a informação de fase é utilizada, o método também é invariante a mudanças de iluminação uniforme.

O cálculo do LPQ é feito utilizando informações extraídas da transformada de Fourier de curto termo (*Short-Term Fourier Transform* - STFT) sobre uma vizinhança retangular  $N_x$  de tamanho  $M \times M$ , em cada pixel  $x$  da imagem  $f(x)$  definido pela Equação 2.5.

$$F(u, x) = \sum_{y \in N_x} f(x - y) e^{-j2\pi u^T y} \quad (2.5)$$

sendo  $j = \sqrt{-1}$  e  $f(x)$  a função de janela que define a vizinhança  $N_x$ .

Alguns parâmetros a serem selecionados são o tamanho da vizinhança local e o tipo de função da janela. Pequenos valores para a vizinhança fazem com que os componentes

de baixa frequência capturem mais detalhes do bloco da imagem. Rahtu et al. (2012) apresentaram algumas funções de janela que podem ser utilizadas no LPQ. Dentre elas, estão as janelas uniformes ( $LPQ^u$ , correspondem à versão básica do LPQ), as gaussianas ( $LPQ^g$ ), e com filtros de banda de quadraturas gaussianas derivativas  $LPQ^{gd}$ . Então, na prática, um filtro discreto de tamanho  $M \times M$  percorre a imagem de entrada.

Apenas quatro coeficientes complexos são considerados no LPQ, que correspondem às frequências bidimensionais  $u_1 = [a, 0]^T$ ,  $u_2 = [0, a]^T$ ,  $u_3 = [a, a]^T$  e  $u_4 = [a, -a]^T$ , sendo  $a$  uma frequência escalar.

Considerando que para cada pixel  $x$  tem-se o vetor representado por  $F_x^c = [F(u_1, x), F(u_2, x), F(u_3, x), F(u_4, x)]$ , e seja  $F_x = [Re\{F_x^c\}, Im\{F_x^c\}]^T$  as partes real e imaginária, respectivamente, de um número complexo, a matriz de transformação  $8 \times M^2$  correspondente é definida pela Equação 2.6.

$$F_x = [Re\{f_x\}, Im\{f_x\}] \quad (2.6)$$

Então, o vetor resultante é quantizado utilizando um quantizador escalar simples, apresentado na Equação 2.7:

$$q_j = \begin{cases} 1, & \text{se } g_j \geq 0 \\ 0, & \text{caso contrário} \end{cases} \quad (2.7)$$

sendo que  $q_j$  é a  $i$ -ésima posição do vetor, e  $g_j$  representa o  $i$ -ésimo componente do vetor  $F$ . Esses elementos são transformados da forma binária para decimal através de uma leitura linear sobre o vetor. Dessa forma, eles passam a representar valores inteiros de 0 a 255, em que suas ocorrências na imagem são contabilizadas para formar o vetor final de 256 elementos do LPQ.

Apesar de o foco desse descritor ser para imagens borradas, Ojansivu and Heikkilä (2008) afirmam que o LPQ também pode ser utilizado em texturas não desfocadas, obtendo bons resultados.

### 2.4.4 Matriz de Co-ocorrência de Níveis de Cinza

A matriz de co-ocorrência de níveis de cinza (*Gray Level Co-occurrence Matrix* - GLCM) foi proposta por Haralick et al. (1973) e consiste em um método estatístico de textura que calcula a frequência com que pares de pixels com valores específicos e com determinada relação espacial (horizontal, vertical ou diagonal) ocorrem em uma imagem.

Haralick et al. (1973) definem GLCM da seguinte forma: dada uma imagem  $I$  de tamanho  $N_l \times N_c$ , sendo  $N_l = \{1, 2, \dots, N_l\}$  o domínio espacial vertical,  $N_c = \{1, 2, \dots, N_c\}$  o domínio espacial horizontal, e  $G = \{1, 2, \dots, N_g\}$  o conjunto de níveis de cinza quantizados  $N_g$ . A imagem  $I$  pode ser representada como uma função que atribui determinado nível de cinza de  $G$  para cada célula ou par de coordenadas de  $N_l \times N_c$ . Com base nisso, uma matriz  $N_g \times N_g$  contabiliza a co-ocorrência de pixels com valores de cinza  $i$  e  $j$  em uma dada distância  $d$  e um ângulo  $\theta$ . Essa distância pode ser definida em ângulos com intervalos de 45 graus ( $0^\circ, 45^\circ, 90^\circ, 135^\circ$ ), porém esses ângulos podem ser redefinidos. A possibilidade de variação da distância e dos ângulos permite obter diversas matrizes para extrair informações. Uma vez criada a matriz, diversas medidas estatísticas podem ser extraídas da mesma.

No total, 14 medidas estatísticas foram propostas por Haralick et al. (1973) e Haralick (1979) para se extrair da GLCM. Porém, alguns trabalhos da literatura afirmam que existem redundâncias entre as informações dessas 14 medidas, e acabam utilizando apenas um subconjunto desse total de estatísticas. Das 14 medidas estatísticas existentes, cinco foram utilizadas neste trabalho: Segundo Momento Angular (*Angular Second Moment* - ASM), Contraste, Energia, Homogeneidade e Correlação.

Segundo Momento Angular: é uma medida de uniformidade da distribuição de níveis de cinza local (Equação 2.8).

$$ASM = \sum_{i,j=0}^{N_g-1} P_{i,j}^2 \quad (2.8)$$

Contraste: é a medida do contraste entre as intensidades de um pixel analisado e do pixel vizinho (Equação 2.9).

$$\text{contraste} = \sum_{i,j=0}^{N_g-1} P_{i,j} (i - j)^2 \quad (2.9)$$

Energia: é soma dos elementos elevados ao quadrado dentro da matriz (Equação 2.10).

$$\text{energia} = \sum_{i,j=0}^{N_g-1} P(i, j)^2 \quad (2.10)$$

Homogeneidade: um valor que representa a proximidade da distribuição dos elementos em relação à diagonal da matriz (Equação 2.11).

$$\text{homogeneidade} = \sum_{i,j=0}^{N_g-1} \frac{P_{i,j}}{1 + (i - j)^2} \quad (2.11)$$

Correlação: é a medida de quão correlacionado está um pixel com o seu vizinho (Equação 2.12).

$$\text{correlação} = \sum_{i,j=0}^{N_g-1} P_{i,j} \left[ \frac{(i - \mu_i)(j - \mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right] \quad (2.12)$$

### 2.4.5 Histograma de níveis de cinza

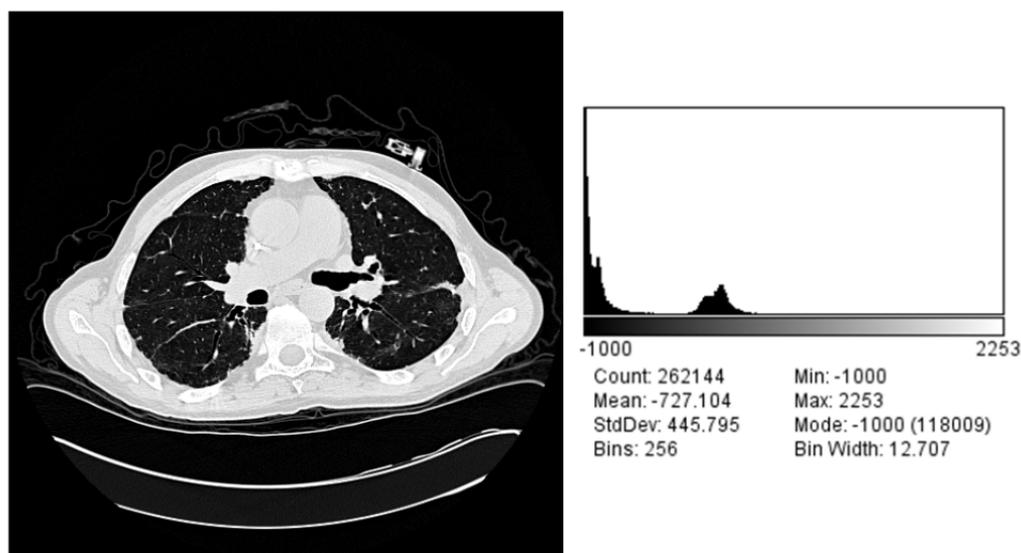
O histograma de níveis de cinza de uma imagem é uma função de distribuição que mostra a frequência com que os tons de cinza aparecem na imagem. O histograma de uma imagem tons de cinza no intervalo  $[0, t - 1]$  pode ser definido como uma função discreta  $f(r_k) = n_k$ , onde  $r_k$  é o  $k$ -ésimo nível de cinza e  $n_k$  é o número de pixels na imagem que possuem o tom de cinza  $r_k$  (Gonzalez and Woods, 2002). É importante observar que, como o histograma simplesmente calcula a frequência de tons da imagem, a mudança de distribuição dos objetos em uma imagem não terá efeito sobre o histograma.

Tradicionalmente, para calcular o histograma de uma imagem com  $t$  tons de cinza, utiliza-se um vetor de tamanho  $t$  para armazenar a contagem de pixels que cada nível de cinza possui. Então, percorre-se a imagem pixel a pixel, incrementando a posição do vetor cujo índice corresponde ao nível de cinza do pixel. Após toda a imagem ser percorrida, cada posição do vetor armazenará a soma de pixels que possuem o tom de

cinza equivalente ao índice do vetor (Marques Filho and Vieira Neto, 1999).

O primeiro passo para se construir um histograma é determinar o número de células (*bins*) a ser utilizado, isto é, como será a divisão do intervalo de valores. Em geral, o número de *bins* corresponde à quantidade de tons de cinza  $t$  da imagem, porém também pode-se dividir o intervalo de valores em vários *bins*, e então calcular a quantidade de valores que está dentro de cada intervalo.

Várias medidas estatísticas podem ser extraídas do histograma de uma imagem, sendo as mais utilizadas o valor máximo, valor mínimo, média, desvio padrão, assimetria e curtose. Um exemplo em que medidas estatísticas são extraídas do histograma de uma imagem de TC pode ser visualizado na Figura 2.4.



**Figura 2.4:** Histograma de níveis de cinza e medidas estatísticas extraídas da imagem (Depeursinge, 2010).

No que se refere às imagens de TC, o histograma de níveis de cinza pode fornecer informações complementares sobre a atenuação das estruturas presentes no pulmão, utilizando a Escala *Hounsfield* (Depeursinge, 2010), sendo a justificativa para sua utilização neste trabalho.

### 2.4.6 Transformação top-hat

As características texturais muitas vezes podem não ser suficientes para capturar certos aspectos de um tipo específico de forma na imagem de TC (Kim et al., 2009). A área da Morfologia Matemática consiste basicamente em extrair de uma imagem desconhecida informações sobre sua geometria e topologia (Marques Filho and Vieira Neto, 1999). O grupo de operações realizadas sobre imagens que se baseiam em sua geometria são chamados de operadores morfológicos. A transformação *top-hat* é um operador morfológico utilizado em várias tarefas de processamento de imagem, inclusive na extração de características. Ele extrai pequenos detalhes e objetos de uma imagem, utilizando para isso um elemento estruturante na operação de abertura ou fechamento (Gonzalez and Woods, 2002). Todos esses cálculos são inicialmente formados por dois operadores morfológicos básicos, a erosão e a dilatação.

Existem dois tipos de transformação *top-hat*: *white top-hat*, conhecida também como *top-hat* por abertura, e *black top-hat*, também chamada de *top-hat* por fechamento. Isso porque o *white top-hat* utiliza a combinação da imagem original e sua correspondente obtida pela abertura. Em contrapartida, o *black top-hat* faz uso da imagem original e sua equivalente obtida pelo fechamento.

A transformação *white top-hat* retorna uma imagem cujos elementos são menores que o elemento estruturante definido, e pontos mais claros que seus vizinhos, ou seja, as regiões de mais alta intensidade. Já a transformação *black top-hat* devolve uma imagem no qual seus elementos são mais escuros que sua vizinhança, isto é, elementos de baixa intensidade, além de serem menores que o elemento estruturante (Dougherty and Lotufo, 2003).

O *white top-hat* de uma imagem  $i$  é definido como a diferença entre a imagem original  $i$  e sua abertura  $\gamma$ , como mostra a Equação 2.13 (Soille, 2003).

$$WTH(i) = i - \gamma(i) \quad (2.13)$$

Por outro lado, o *black top-hat* de uma imagem  $i$  é descrito como a diferença entre o fechamento da imagem  $i$  e a própria imagem original, sendo definido pela Equação 2.14

(Soille, 2003).

$$BTH(i) = \phi(i) - i \quad (2.14)$$

A forma e o tamanho do elemento estruturante geralmente é escolhida de acordo com algum conhecimento sobre a geometria das estruturas da imagem (Soille, 2003). Facon (1996) explica que as operações entre a imagem de entrada e o elemento estruturante são o princípio básico da morfologia matemática, sendo necessário escolher adequadamente o elemento estruturante para obter bons resultados. As duas transformações produzem imagens que só possuem pixels não negativos (Soille, 2003).

A abertura de uma imagem, citada anteriormente, basicamente é constituída pela operação de erosão de uma imagem  $u$  pelo seu elemento estruturante, seguido pela dilatação com o elemento estruturante refletido. De forma contrária, o fechamento é composto pela dilatação de  $u$  usando um elemento estruturante, seguido pelo operador de erosão, usando o elemento estruturante refletido (Soille, 2003).

A transformação *top-hat* já é utilizada em alguns trabalhos da literatura com o intuito de extrair características geométricas (Park et al., 2009; Lim et al., 2011; Chang et al., 2012).

## 2.5 Classificação

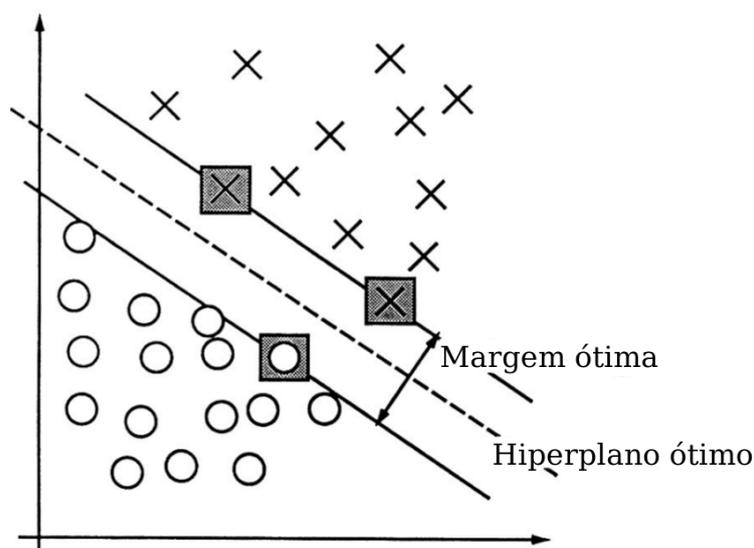
Esta seção apresenta as técnicas de classificação empregadas neste trabalho. A subseção 2.5.1 apresenta as definições básicas relacionadas ao classificador SVM, utilizado para classificar os padrões radiológicos. Na subseção 2.5.2 são descritos os métodos de combinação de classificadores. Por fim, a subseção 2.5.3 descreve as métricas de avaliação dos classificadores.

A escolha pela utilização do SVM como classificador justifica-se por este ter mostrado ser o mais eficiente na identificação de padrões de alterações pulmonares (Depeursinge et al., 2010).

## 2.5.1 Support Vector Machine

*Support Vector Machine* (SVM), ou Máquina de Vetor de Suporte, é um classificador proposto por Vapnik (1982), que tem por objetivo separar as classes de um problema da melhor forma possível, por meio de hiperplanos que maximizam a margem entre essas classes. Muitos classificadores não encontram a melhor fronteira para o problema, utilizando apenas a primeira fronteira encontrada que separe corretamente os dados. O SVM, por sua vez, irá encontrar o hiperplano que apresenta a maior distância possível dos pontos mais próximos do limite entre as classes, definindo assim a maior margem de separação (Theodoridis and Koutroumbas, 2008).

A Figura 2.5 ilustra o conceito de maximização da margem através de pontos próximos do limite entre duas classes. Nela, é possível visualizar que os vetores de suporte, destacados pelos quadrados de cor cinza, definem a margem da separação máxima entre as duas classes. A uma distância de cada uma das retas encontra-se o hiperplano, ilustrado pela linha pontilhada, que é o separador de margem máxima dos dados.



**Figura 2.5:** Exemplo de um problema binário linearmente separável em um espaço bidimensional (Adaptado de Cortes and Vapnik (1995)).

O exemplo ilustrado pela Figura 2.5 apresenta a separação de dados em um espaço bidimensional. Porém, na grande maioria dos problemas, os dados não são linearmente separáveis. Para isso é necessário projetar os dados em dimensões mais altas. Dessa forma, os dados tendem a ser linearmente separáveis por um hiperplano. As funções

de *kernel* permitem realizar esse mapeamento dos dados, utilizando vetores  $x_i$  e  $x_j$  que calculam o produto escalar em um determinado espaço de características (Theodoridis and Koutroumbas, 2008). As funções mais comuns são: Polinomial de grau  $d$  (Equação 2.15), *Radial Basis Function* (RBF), em que  $\gamma$  define o comportamento da exponencial (Equação 2.16), e Tangente Hiperbólica, com os parâmetros ajustáveis  $\beta$  e  $\alpha$  (Equação 2.17).

$$K(x_i, x_j) = (1 + x_i \cdot x_j)^d \quad (2.15)$$

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (2.16)$$

$$K(x_i, x_j) = \tanh(\beta x_i \cdot x_j + \alpha) \quad (2.17)$$

Inicialmente o SVM foi desenvolvido para problemas com duas classes, isto é, problemas binários. Porém, existem dois algoritmos que permitem construir um classificador para problemas com  $N$  classes. O primeiro método utiliza  $N$  classificadores, um para cada classe. Treina-se o  $i$ -ésimo classificador enquanto rotula-se todas as amostras da  $i$ -ésima classe como sendo positivas, e as restantes como negativas (Vapnik, 1998). Esse algoritmo é chamado de SVM *one-versus-all*. O segundo método foi proposto por Knerr et al. (1990), que consiste em construir  $N \times (N - 1)/2$  classificadores, realizando todas as combinações de pares das  $N$  classes. Esse algoritmo é conhecido por SVM *one-versus-one*. Neste trabalho, emprega-se a técnica *one-versus-one* de separação de classes.

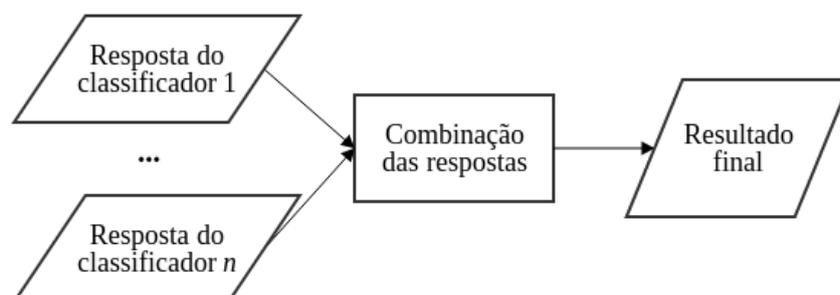
Os resultados obtidos após a execução de classificadores SVM são descalibrados por padrão, não gerando probabilidades *a posteriori*. Platt (1999) propôs um método para converter as saídas dos classificadores em *scores*, pois afirma que construir um classificador que produza probabilidades  $P(\text{classe}|\text{entrada})$  é muito útil em diversas situações de reconhecimento. Por exemplo, probabilidades *a posteriori* são requeridas quando as saídas de diversos classificadores devem ser combinadas para uma decisão final. O método consiste em estimar uma função sigmóide e utilizá-la na conversão das saídas dos clas-

sificadores. Esses *scores* representam a probabilidade de cada amostra pertencer a cada uma das classes definidas no problema. Posteriormente, Wu et al. (2004) apresentaram técnicas que estimam probabilidades em problemas multiclasse.

## 2.5.2 Combinação de classificadores

Muitas vezes, os resultados obtidos por um determinado classificador e/ou conjunto de características não são satisfatórios. Porém, esse mesmo classificador que, de forma geral, não obteve alta taxa de reconhecimento, pode conseguir um desempenho elevado para uma determinada classe apenas. Da mesma forma, outro classificador utilizado pode ter alcançado resultados igualmente razoáveis, porém complementares ao primeiro. O objetivo da combinação de classificadores é tentar obter um resultado melhor de classificação, aproveitando essa complementaridade que pode existir entre classificadores. Mesmo que um dos classificadores seja especialista, a fusão das respostas de diferentes classificadores geralmente resulta em uma decisão final melhor.

De acordo com Jain et al. (2000), existem diversas formas de combinar múltiplos classificadores, que podem ser divididas em categorias de acordo com sua arquitetura: 1) em paralelo, 2) em série e 3) de forma hierárquica. A maioria dos esquemas de combinação da literatura pertence à categoria em paralelo, que consiste em empregar os classificadores de forma independente, apenas combinando seus resultados de acordo com alguma regra. Após os classificadores independentes serem selecionados, aplica-se uma função de combinação (fusão) das saídas geradas por cada classificador individual. Essa arquitetura é ilustrada pela Figura 2.6, sendo a arquitetura de combinação empregada neste trabalho.



**Figura 2.6:** Arquitetura de combinação de classificadores em paralelo.

Diferentes saídas podem ser produzidas pelos classificadores individuais. O resultado pode ser abstrato, baseado em *ranking*, ou baseado em probabilidades. Na saída abstrata, o classificador informa apenas qual a classe definida como correta. Já no resultado por *ranking*, o classificador ordena as classes, da mais provável à menos provável, porém os resultados do *ranking* não podem ser utilizados isoladamente, pois a classe mais provável não significa necessariamente que possui alto nível de confiança na classificação. Por fim, o resultado em probabilidades contém maiores informações, pois o classificador gera um valor numérico para cada classe, indicando a probabilidade de determinada amostra pertencer a cada uma dessas classes (Jain et al., 2000). Neste trabalho, as saídas dos classificadores foram geradas pelo resultado em probabilidades.

As regras mais comuns para realizar a fusão de classificadores são: voto majoritário, produto, soma, média, máximo e mínimo (Kittler et al., 1998; Costa, 2013), sendo descritas a seguir.

## Regra do voto majoritário

É a regra mais simples de combinação de respostas de classificadores. A regra é basicamente realizar a contagem de votos entre as saídas de todos os classificadores combinados e atribuir, para determinada amostra, a classe que recebeu mais votos. Nesse caso, essa regra não requer as saídas *a posteriori* de cada classe. A Equação 2.18 apresenta o cálculo de votação para uma amostra  $x$ , sendo  $n$  o número total de classificadores,  $y_i$  a classe de saída do  $i$ -ésimo classificador para os possíveis rótulos de classe  $\Omega = \omega_1, \omega_2 \dots \omega_c$ . Se houver empate no número de votos, a escolha entre as classes empatadas pode ser aleatória.

$$classe(x) = \arg \max_{k=1}^c \sum_{i=1}^n y_{i,k} \quad (2.18)$$

## Regra do produto

Essa regra consiste em multiplicar as probabilidades associadas às saídas dos classificadores  $c_i$ . Isso torna a regra bastante severa, bastando que o resultado de apenas um dos classificadores seja próximo a zero para que o resultado final seja penalizado. Portanto

a regra do produto é mais indicada em situações críticas, onde erros não são aceitos. Na equação 2.19,  $x$  é o padrão a ser classificado,  $n$  é o número de classificadores,  $y_i$  é a classe de saída do  $i$ -ésimo classificador, para os possíveis rótulos de classe  $\Omega = \omega_1, \omega_2, \dots, \omega_c$  e  $P(\omega_k|y_i(x))$  é a probabilidade de  $x$  pertencer à classe  $\omega_k$  do  $i$ -ésimo classificador.

$$classe(x) = \arg \max_{k=1}^c \prod_{i=1}^n P(\omega_k|y_i(x)) \quad (2.19)$$

## Regra da soma

A regra da soma realiza a combinação calculando o somatório entre as probabilidades associadas às saídas dos classificadores  $c_i$ . Ela não penaliza tanto o resultado final se algum classificador errar, sendo mais indicada para situações menos críticas. O cálculo dessa regra é representado pela Equação 2.20.

$$classe(x) = \arg \max_{k=1}^c \sum_{i=1}^n P(\omega_k|y_i(x)) \quad (2.20)$$

## Regra da média

Essa regra calcula a média entre as probabilidades associadas às saídas dos classificadores, produzindo resultados parecidos com os da soma. É definida pela Equação 2.21.

$$classe(x) = \frac{1}{n} \arg \max_{k=1}^c \sum_{i=1}^n P(\omega_k|y_i(x)) \quad (2.21)$$

## Regra do máximo

Essa regra considera como vencedora a classe que apresentar o maior *score* dentre todos os classificadores, mesmo que isso signifique que a classe apresenta bom desempenho somente para um dos classificadores. Isso torna essa regra de baixa severidade, pois se uma classe obtiver um bom desempenho em qualquer classificador, ela tem uma grande chance de ser escolhida. A equação 2.22 representa a regra do máximo.

$$classe(x) = \arg \max_{k=1}^c \max_{i=1}^n P(\omega_k | y_i(x)) \quad (2.22)$$

## Regra do mínimo

De forma inversa à regra do máximo, a regra do mínimo considera como resultado o menor valor de *score* atribuído às classes pelos classificadores. A regra do mínimo, assim como a do produto, é de alta severidade, pois é sensível a baixos *scores*. Sua definição é apresentada na Equação 2.23.

$$classe(x) = \arg \max_{k=1}^c \min_{i=1}^n P(\omega_k | y_i(x)) \quad (2.23)$$

### 2.5.3 Métricas de avaliação dos classificadores

A matriz de confusão de um método de reconhecimento de padrões (também chamada de tabela de contingência), é uma tabela que mostra as predições de um classificador, facilitando a visualização de ocorrências de predições corretas, bem como das confusões ocorridas entre as  $n$  classes possíveis. A diagonal principal da matriz, em que tem-se *linha = coluna*, armazenam os acertos do classificador. As demais células apresentam as confusões geradas. A Tabela 2.1 exemplifica uma matriz de confusão de duas classes, sendo que as linhas representam a classe correta, e as colunas, a classe predita.

**Tabela 2.1:** Exemplo de matriz de confusão de duas classes.

		Classe predita	
		P	N
Classe correta	P	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	N	Falso Positivo (FP)	Verdadeiro Negativo (VN)

A partir da matriz de confusão, quatro indicadores são extraídos em relação a cada classe:

- Verdadeiro Positivo (VP) - número de exemplos positivos classificados como positivos;
- Falso Positivo (FP) - número de exemplos negativos classificados como positivos;
- Verdadeiro Negativo (VN) - número de exemplos negativos classificados como negativos;
- Falso Negativo (FN) - número de exemplos positivos classificados como negativos.

Com base nesses valores, as métricas calculadas são definidas conforme segue abaixo:

Sensibilidade - taxa verdadeiro positivo (Equação 2.24): identifica a proporção de amostras positivas que foram corretamente identificadas como positivas.

$$\text{sensibilidade} = \frac{VP}{VP + FN} \quad (2.24)$$

Especificidade - taxa de verdadeiro negativo (Equação 2.25): identifica a proporção de amostras negativas que foram corretamente identificadas como negativas.

$$\text{especificidade} = \frac{VN}{VN + FP} \quad (2.25)$$

Precisão (Equação 2.26): calcula a precisão das predições de uma classe (Depeursinge, 2010).

$$\text{precisão} = \frac{VP}{VP + FP} \quad (2.26)$$

Acurácia (Equação 2.27): calcula a proporção de resultados verdadeiros (VP e VN) entre o total de classes.

$$\text{acurácia} = \frac{VP + VN}{VP + FP + VN + FN} \quad (2.27)$$

## CAPÍTULO 3

### REVISÃO BIBLIOGRÁFICA

Este capítulo apresenta alguns trabalhos existentes na literatura que visam resolver o problema de classificação de padrões radiológicos. Sendo assim, o objetivo deste estudo é investigar os diversos métodos utilizados pelos autores, especialmente o processo de extração de características, bem como verificar aspectos semelhantes entre os métodos de cada trabalho.

Como existe um número considerável de trabalhos relacionados, foi necessário definir alguns critérios de inclusão e exclusão de artigos. Dessa forma, optou-se por incluir artigos científicos publicados a partir do ano de 2009. Em adição, foram incluídos nesta pesquisa trabalhos que realizam a classificação dos mesmos padrões pulmonares, ou de grande parte, que este trabalho pretende reconhecer.

Para avaliar e quantificar os métodos desenvolvidos, os autores utilizam imagens de TC de domínio público ou privado. Imagens privadas são todas as imagens obtidas que não são de livre acesso. Dessa forma, sua disponibilidade está limitada apenas ao grupo de pesquisa que desenvolveu o trabalho. Por sua vez, as imagens públicas são aquelas que estão disponíveis livremente para uso. A base de imagens pública utilizada pelos trabalhos aqui apresentados é disponibilizada por Depeursinge et al. (2011), sendo a escolhida para este trabalho.

A seção 3.1 apresenta alguns dos trabalhos mais recentes da literatura sobre classificação de padrões pulmonares. A seção 3.2 mostra um comparativo entre os trabalhos apresentados, fazendo algumas considerações em relação ao estudo dos trabalhos relacionados.

### 3.1 Metodologias para a classificação de padrões pulmonares

O sistema automático de Park et al. (2009) realiza a detecção dos seguintes padrões de doenças pulmonares difusas em imagens de TCAR: enfisema, consolidação, faveolamento, opacidade em vidro-fosco e opacidade reticular. Em adição, o padrão pulmonar normal também foi analisado. O método desenvolvido conta com um conjunto de 37 características de textura e forma. Como descritores de textura, foram empregados histograma, gradiente, matriz de comprimento, GLCM e Insight ToolKit GLCM (ITK-GLCM). Já os descritores de forma utilizados foram os seguintes: transformação *top-hat* e informações das áreas de baixa atenuação. Utilizando 600 regiões de interesse (*Region of Interest - ROI*) circulares com 10 pixels de diâmetro, o sistema foi treinado com o classificador SVM. A performance do SVM foi testada com o método *5-fold* de validação-cruzada, primeiramente classificando os dados randomicamente em cinco grupos, e então usando quatro grupos para a construção do modelo e um grupo para testar a performance do modelo. O sistema de quantificação desenvolvido primeiramente realiza a segmentação dos pulmões por meio de uma função manual de desenho, e a análise do parênquima pulmonar é realizada utilizando uma função de ROI móvel, onde cada pixel é classificado em uma das seis categorias. Nesse experimento, 92 imagens foram utilizadas. O desempenho geral da classificação alcançou 89% e o resultado do sistema de quantificação, por sua vez, atingiu uma média de 61% de concordância entre a detecção realizada pelo sistema e a detecção manual realizada por um dos especialistas.

O trabalho proposto por Vo and Sowmya (2010) apresenta um novo algoritmo para classificação dos seguintes padrões pulmonares: tecido normal, enfisema, faveolamento e opacidade em vidro-fosco. O algoritmo é baseado na análise textural das imagens de TCAR e utiliza características espaço-escalares com base em filtros derivativos gaussianos, e características baseadas nas transformadas *wavelet* e *contourlet* das imagens originais. Esses filtros foram escolhidos para representar a textura pois, de acordo com os autores, possuem vantagens em caracterizar aspectos direcionais e escalares na imagem. Para avaliar o desempenho do sistema de extração de características, utilizou-se o classificador *multiclass Multiple Kernel Learning* (m-MKL), uma combinação ponderada de diferentes

*kernels* que otimiza a combinação de características. Experimentos foram realizados com 73.000 ROIs, de tamanho  $32 \times 32$  pixels, extraídas de 89 imagens de TCAR. A comparação de desempenho entre cada conjunto de características mostrou que, no geral, as características da transformada *contourlet* superam as demais utilizadas. O método desenvolvido atingiu uma sensibilidade média de 94,16% combinada com uma especificidade média de 98,68%.

Lim et al. (2011) desenvolveram um novo método de classificação sensível ao contexto, o *context-sensitive SVM* (csSVM), que classifica seis tipos de padrões pulmonares: normal, enfisema, opacidade em vidro-fosco, opacidade reticular, faveolamento e consolidação. A ideia principal do classificador sensível ao contexto é poder utilizar informações da vizinhança da região analisada. De acordo com os autores, existe uma correlação entre ROIs vizinhas, visto que os padrões de doenças pulmonares usualmente estão presentes em mais de uma ROI. Um total de 21 características texturais e de forma foram extraídas das ROIs. As características de textura empregadas foram: histograma, gradiente, matriz de comprimento e GLCM. Características de forma foram extraídas de diferentes tipos de transformações *top-hat* e análise de agrupamentos. Após extrair as características, o classificador csSVM é empregado, utilizando informações das regiões vizinhas, como as características e a classe das mesmas. Para validar os resultados, um total de 900 ROIs foram utilizadas, empregando a técnica *5-fold* de validação-cruzada. Posteriormente, 84 imagens do pulmão foram quantificadas. A acurácia do método proposto foi de 89,88% a nível de ROI, e de 60,30% para a classificação de toda a imagem pulmonar.

Depeursinge et al. (2012) propõem um conjunto de características texturais com base na transformada *wavelet* e nos níveis de cinza da imagem para identificar cinco tipos de padrões pulmonares em imagens de TCAR: normal, enfisema, vidro-fosco, fibrose e micronódulos. A principal contribuição do trabalho é o desenvolvimento de características invariantes à translação e rotação e covariantes à escala, empregando a transformada *wavelet*. O processo de classificação inicia com a segmentação semiautomática do parênquima pulmonar utilizando crescimento de regiões e morfologia matemática. Após esse processo, cada região segmentada da imagem é dividida em blocos sobrepostos de

tamanho  $32 \times 32$  pixels. Em cada bloco, calcula-se o vetor de características, que é composto pelos valores do histograma de níveis de cinza, e pelos valores dos coeficientes *wavelet*. Classifica-se o parênquima pulmonar utilizando blocos com quatro pixels de distância entre os centros dos blocos. O classificador SVM é utilizado para identificar os padrões pulmonares, empregando a técnica *one-versus-one*. Para estimar o desempenho do método de classificação, a técnica *leave-one-patient-out* (LOPO) foi utilizada, impedindo realizar o treino e teste com imagens de um mesmo paciente. As imagens utilizadas fazem parte de um banco de imagens de TCAR público construído pelo próprio autor e seus colaboradores em um trabalho anterior (Depeursinge et al., 2011). Utilizando um total de 17.848 blocos de ROIs, o método proposto atingiu uma sensibilidade geral de 75,81%.

O método proposto por Chang et al. (2012) tem como objetivo classificar padrões de doenças pulmonares de forma rápida e eficiente utilizando um esquema de classificação hierárquica. Os padrões analisados na pesquisa foram: pulmão normal, enfisema, faveolamento e opacidade reticular, consolidação e opacidade em vidro-fosco. Os autores partem da ideia de que algumas classes necessitam de apenas um subconjunto de características para serem classificadas, mesmo que outras classes precisem de todo o conjunto. Isso gera um custo computacional diferente para cada padrão. Para representar as características de cada ROI, 13 características texturais (descritores: histograma, gradiente, matriz de comprimento e GLCM) e 11 características de forma (descritores: transformações *top-hat* e análise de agrupamentos) foram utilizadas. O classificador SVM é introduzido com uma arquitetura de árvore binária, onde cada nodo distingue dois conjuntos de classes. Treina-se o classificador binário em cada nodo da hierarquia, e em cada classificador, utiliza-se um conjunto de características específico para cada classe, bem como um método de seleção de características sensível ao custo. Para poupar custo computacional, a classificação utiliza primeiramente características simples para, posteriormente, adicionar características mais complexas. O SVM hierárquico empregou o método *one-against-all* e *5-fold* de validação-cruzada para calcular a precisão do sistema. Utilizando 600 ROIs circulares com 32 pixels de diâmetro nos experimentos, o método desenvolvido atingiu uma acurácia

média de 92,63%.

Song et al. (2013) propõem um novo método de classificação de diversos padrões pulmonares, como: padrão normal, enfisema, opacidade em vidro-fosco, fibrose e micronódulos. A classificação é realizada com base na aproximação das características a partir dos blocos de imagem de referência. Primeiramente, características de textura, intensidade e gradiente são extraídas dos blocos de imagem, e dois descritores de características, nomeados *rotation-invariant Gabor-LBP* (RGLBP) e *multi-coordinate histogram of oriented gradients* (MCHOG), são propostos. Em seguida, cada bloco de imagem é classificado com base em dicionários de referência, utilizando um novo algoritmo de aproximação esparsa. Após, a classe de cada ROI é obtida com base nas estimativas probabilísticas da classificação. Utilizando imagens públicas, os experimentos foram realizados com 1.458 ROIs, e 23.731 blocos de imagem de  $31 \times 31$  pixels extraídos dessas regiões. O método proposto atingiu uma sensibilidade média de 82,64%.

Zhao et al. (2013) apresentam um método baseado em representação esparsa para classificar padrões de pulmão normal e cinco tipos de padrões de alterações pulmonares, que são: enfisema, opacidade em vidro-fosco, faveolamento, consolidação e nodular. A ideia principal da representação esparsa é obter uma aproximação das amostras (exemplos) por meio da combinação linear de um número pequeno de características-chave (átomos), selecionados a partir de um dicionário redundante. Primeiramente, as características locais foram extraídas dos volumes de interesse (*Volume of Interest* - VOI) utilizando os autovalores da matriz Hessiana juntamente com valores da TC. Isso porque alguns padrões pulmonares, como consolidação e enfisema, caracterizam-se pelos valores da TC, e outros padrões, como o faveolamento e o nodular, são caracterizados pela forma. Também foram calculadas quatro medidas estatísticas: média, desvio padrão, assimetria e curtose. Após extrair as características locais, os descritores foram gerados utilizando do modelo de representação esparsa. O algoritmo K-SVD (uma generalização do algoritmo *k-means*) foi empregado no aprendizado do dicionário, o qual utiliza os vetores de características da base de treinamento. Dessa forma, os descritores gerados são utilizados como entrada do classificador, de acordo com o dicionário e com os vetores de características originais.

Finalmente, o classificador SVM é utilizado para classificar o modelo de representação esparsa, empregando a técnica *one-against-one*. Os experimentos foram realizados utilizando 2.360 VOIs de 117 indivíduos, obtidas em um serviço hospitalar. Os VOIs foram divididos em dois conjuntos: o primeiro conjunto compõe a base de treinamento, e o segundo compõe a base de teste. Resultados mostraram que a técnica desenvolvida atingiu uma acurácia global de 95,4%, mantendo taxas de reconhecimento de todos os padrões acima de 90% cada.

Li et al. (2013) apresentam uma classificação de blocos de imagens pulmonares utilizando a técnica de aprendizado automático de características. Os padrões pulmonares utilizados são: normal, enfisema, opacidade em vidro fosco, fibrose e micronódulos. Ao invés de definir o conjunto de características manualmente, desenvolveu-se uma técnica de aprendizagem não-supervisionada para construir descritores de características, que são convoluídos com os blocos de imagem para obter o vetor de características. Primeiramente, extratores de características com diferentes escalas são aprendidos a partir de blocos de imagem. Para isso, uma rede neural multi-escala denominada Gaussiana RBM (GRBM) foi empregada para aprender os padrões da base baseado nos exemplos de treinamento. Três redes foram treinadas utilizando blocos de tamanhos variados, para obter características de diferentes tamanhos. Após, os vetores de características são obtidos por meio da convolução dos descritores com os blocos da imagem. O vetor é então classificado em uma das cinco categorias empregando o classificador SVM. Utilizando 16.220 blocos de tamanho  $32 \times 32$  pixels, extraídos de 92 imagens de TCAR de uma base de imagens pública, o método desenvolvido atingiu uma sensibilidade média de 74,2%, mostrando-se superior quando comparado a outros extratores utilizados na literatura, como LBP.

Anthimopoulos et al. (2014) desenvolveram um método para classificar padrões de doenças pulmonares intersticiais baseado na transformada discreta de cosseno (*Discrete Cosine Transform - DCT*). As classes consideradas são: tecido normal, vidro-fosco, consolidação, reticulação, faveolamento, e a combinação de reticulação com vidro-fosco. A extração de características se baseia na análise espectral local usando um banco de filtros baseados em DCT. O histograma de níveis de cinza da imagem original também é uti-

lizado, formando o vector final de características. A classificação das amostras emprega *random forest*, uma combinação de árvores de decisão onde cada uma “vota” em uma das classes existentes. Para determinar o rótulo de determinado exemplo, a floresta escolhe aquele que possui a maioria dos votos entre todas as árvores. Os experimentos empregaram 2.503 blocos de ROIs de tamanho  $21 \times 21$  pixels, extraídos com uma sobreposição de 90% da ROI original, provenientes de uma base de imagens pública. A avaliação do método de classificação é baseado em um esquema *5-fold* de validação cruzada, e seu desempenho foi avaliado em termos da média de *F-score*, definida por:  $F_{avg} = \frac{1}{M} \sum_{c=1}^M F_c$ , onde  $M$  é o número de classes e  $F_c$  é o *F-score* para a classe  $c$ , sendo  $F_c = 2 \cdot \frac{precision_c \cdot sensitivity_c}{precision_c + sensitivity_c}$ . Dessa forma, a média *F-score* obtida pelo método foi de 89%.

Dash et al. (2014) apresentam uma nova técnica para identificar os seguintes padrões pulmonares em imagens de TCAR: tecido normal, enfisema, vidro-fosco, nodular e consolidação. A proposta utiliza características texturais obtidas da Transformada Discreta de Wavelet (*Discrete Wavelet Transform* - DWT), juntamente com os classificadores Naive Bayes e Redes Neurais. A DWT é utilizada para decompor a imagem de entrada em quatro sub-bandas e, após isso, medidas estatísticas são extraídas dos coeficientes *wavelet*. As decisões geradas pelos dois classificadores são combinadas utilizando uma função de decisão, a fim de obter uma resposta final para cada amostra. Os experimentos utilizaram 100 ROIs com tamanhos retangulares variáveis, sendo 20 ROIs de cada classe. O método mostrou resultados melhores do que os obtidos utilizando apenas um classificador, atingindo uma acurácia média de 95%.

O método proposto por Pereyra et al. (2014) tem como objetivo utilizar análise de textura e fractais em sistemas de auxílio ao diagnóstico de doenças pulmonares difusas. Os padrões a serem reconhecidos são: pulmão normal, enfisema, vidro-fosco, faveolamento, consolidação e espessamento dos septos. Inicialmente, as imagens DICOM foram pré-processadas, sendo convertidas para 8 bits por pixel. Também aplicou-se equalização de histograma, juntamente com fechamento morfológico, limiarização binária e seleção de objetos, antes de obter as ROIs. Para extrair características, diversas técnicas são propostas. Os atributos utilizados incluem estatísticas de primeira ordem da distribuição

de níveis de cinza, características texturais de Haralick e Law, informações estatísticas extraídas da Transformadas Discretas de Fourier, e o valor de dimensão fractal. O vetor de características para cada ROI possui 28 elementos, porém uma seleção de atributos também foi empregada. Dessa forma, os resultados da classificação utilizando o vetor de características completo é comparado com os vetores de atributos selecionados. O classificador *k-Nearest Neighbors* (com  $k = 5$ ) foi utilizado para classificar as ROIs, juntamente com um método de validação cruzada de *10-folds*. Os experimentos utilizaram uma base de dados contendo 3.252 ROIs de  $64 \times 64$  pixels, em que a maior taxa de exemplos corretamente classificados atingiu 82,6%, utilizando o vetor obtido por *linear forward selection*.

## 3.2 Resumo dos trabalhos apresentados

O resumo dos trabalhos apresentados na seção anterior pode ser vistos na Tabela 3.1, que mostra os tipos de padrões pulmonares reconhecidos em cada trabalho, os diferentes descritores de características e classificadores empregados, as imagens utilizadas e os resultados gerais alcançados. As taxas reportadas pelos trabalhos são baseadas nos valores de VP, FP, VN e FN. Com exceção da métrica de *F-score*, utilizada no trabalho de Anthimopoulos et al. (2014) e definida anteriormente, as equações de sensibilidade, especificidade, precisão e acurácia foram definidas na Seção 2.5.3, visto que foram utilizadas neste trabalho.

**Tabela 3.1:** Síntese dos trabalhos de análise e classificação de padrões de doenças pulmonares.

Trabalho	Padrões analisados	Características	Classificador	Imagens	Resultado
Park et al. (2009)	normal, enfisema, opacidade em vidro-fosco, opacidade reticular, faveolamento e consolidação	histograma, gradiente, matriz de comprimento, GLCM, ITK-GLCM, transformação top-hat e informações das áreas de baixa atenuação	SVM	Privadas - 600 ROIs	89% de acurácia
Vo and Sowmya (2010)	normal, enfisema, opacidade em vidro-fosco e faveolamento	características espaço-escalares com base em filtros derivativos gaussianos, e características baseadas nas transformadas wavelet e contourlet	multiclass Multiple Kernel Learning	Privadas - 73.000 ROIs	94,16% de sensibilidade
Lim et al. (2011)	normal, enfisema, opacidade em vidro-fosco, opacidade reticular, faveolamento e consolidação	histograma, gradiente, matriz de comprimento, GLCM, transformação top-hat e análise de agrupamentos	context-sensitive SVM	Privadas - 900 ROIs	89,88% de acurácia
Depeursinge et al. (2012)	normal, enfisema, opacidade em vidro-fosco, fibrose e micronódulos	valores do histograma de níveis de cinza, e valores dos coeficientes wavelet	SVM	Públicas - 17.848 ROIs	75,81% de sensibilidade

continua na próxima página

Trabalho	Padrões analisados	Características	Classificador	Imagens	Resultado
Chang et al. (2012)	normal, enfisema, opacidade em vidro-fosco, faveolamento e opacidade reticular, e consolidação	histograma, gradiente, matriz de comprimento, GLCM, transformação top-hat e análise de agrupamentos	SVM com arquitetura de árvore binária	Privadas - 600 ROIs	92,63% de acurácia
Song et al. (2013)	normal, enfisema, opacidade em vidro-fosco, fibrose e micronódulos	rotation-invariant Gabor-LBP (RGLBP) e multi-coordinate histogram of oriented gradients (MCHOG)	estimações probabilísticas	Públicas - 23.731 ROIs	82,64% de sensibilidade
Zhao et al. (2013)	normal, enfisema, nodular, opacidade em vidro-fosco, faveolamento e consolidação	autovalores da matriz Hessiana e valores da TC	SVM	Privadas - 2.360 VOIs	95,4% de acurácia
Li et al. (2013)	normal, enfisema, opacidade em vidro-fosco, fibrose e micronódulos	aprendizagem não-supervisionada de características utilizando a rede neural multi-escala Gaussiana RBM (GRBM)	SVM	Públicas - 16.220 ROIs	74,2% de sensibilidade
Anthimopoulos et al. (2014)	normal, opacidade em vidro-fosco, faveolamento, consolidação, reticulação, e reticulação com vidro-fosco	análise espectral usando banco de filtros baseado em DCT	Random Forest	Públicas - 2.503 ROIs	89% de <i>F-score</i>

continua na próxima página

Trabalho	Padrões analisados	Características	Classificador	Imagens	Resultado
Dash et al. (2014)	normal, enfisema, opacidade em vidro-fosco, nodular e consolidação	características texturais obtidas da DWT	Naive Bayes e Redes Neurais	Privadas - 100 ROIs	95% de acurácia
Pereyra et al. (2014)	normal, enfisema, opacidade em vidro-fosco, faveolamento, consolidação e espessamento dos septos	características de Haralick e Law, valor de dimensão fractal, informações estatísticas da DFT e estatísticas de primeira ordem da distribuição de níveis de cinza	SVM	Privadas - 3.252 ROIs	82,6% de sensibilidade

É possível perceber que, de forma geral, a literatura já provê diversos métodos computacionais que visam classificar imagens de padrões de doenças pulmonares. No entanto, é difícil realizar uma análise comparativa correta entre os resultados obtidos em cada método, pois as abordagens descritas utilizam bases de imagens diferentes. Como algumas bases de imagem não são públicas, isso impede a replicação dos métodos da literatura utilizando as mesmas imagens que os autores. Além disso, não é possível verificar certas características da imagem, como a quantidade e o tamanho das áreas anormais. Pode-se observar que os trabalhos que utilizaram a mesma base de imagens pública atingiram resultados semelhantes entre si, porém resultados inferiores em relação aos demais trabalhos. No entanto, em todos os trabalhos analisados as regiões de interesse foram demarcadas por médicos especialistas, gerando assim o *ground-truth* dos padrões.

Recentemente, muitos estudos sobre classificação de padrões de doenças pulmonares foram desenvolvidos, porém as diferentes técnicas de resolução mostram que ainda não existe um consenso sobre quais características utilizar no processo de classificação. Normalmente, padrões de tecidos pulmonares estão associados à propriedades texturais (Depeursinge et al., 2011). Os trabalhos mostram dificuldade em diferenciar determinados padrões, que também causam confusão entre especialistas (Park et al., 2009) devido à sua similaridade, mostrando que esse é o principal problema a ser enfrentado. Além disso, a segmentação dos pulmões afetados por patologias de alta densidade ainda é um trabalho em andamento (Sluimer et al., 2006), e a maioria dos estudos sobre classificação de padrões de doenças pulmonares usam imagens dos pulmões já segmentados (Anthimopoulos et al., 2014; Dash et al., 2014; Zhao et al., 2013; Song et al., 2013; Li et al., 2013). Em outros casos, uma segmentação semiautomática é aplicada (Depeursinge et al., 2012).

## CAPÍTULO 4

### MÉTODO PROPOSTO

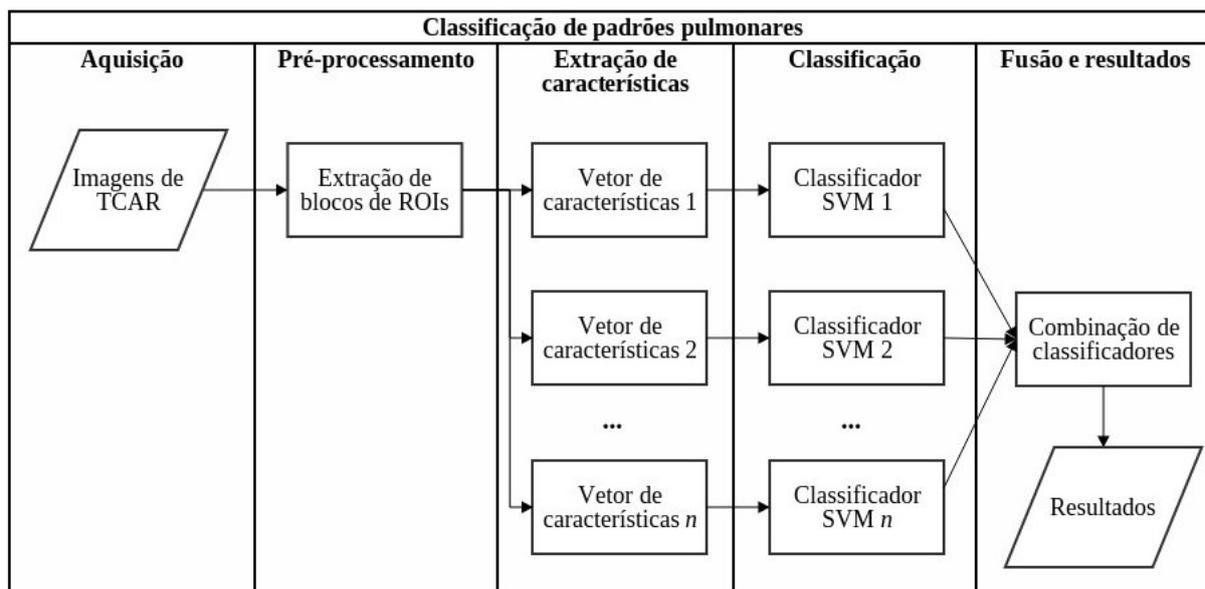
Este capítulo apresenta a metodologia utilizada para o desenvolvimento do trabalho. O principal objetivo, como já citado na Seção 1.3, é realizar o reconhecimento automático de padrões radiológicos em blocos de imagens de TC não segmentadas. Para isso, propõe-se classificar também as áreas externas ao pulmão, a fim de possibilitar uma pré-segmentação pulmonar. Seis padrões radiológicos foram utilizados neste trabalho, sendo eles: não-pulmão, tecido pulmonar normal, enfisema, vidro-fosco, fibrose e micronódulos.

A classificação de doenças pulmonares é um problema de reconhecimento de padrões. Vetores de características foram extraídos das amostras, e então submetidos a um classificador. Uma variedade de descritores foi analisada neste trabalho, visto que certas características podem se comportar melhor para um determinado padrão, e pior para outro. Para buscar uma forma de complementaridade entre os resultados, realizou-se uma combinação entre classificadores.

Primeiramente, este capítulo descreve a base de imagens utilizada, bem como o processo de extração dos blocos de imagens que foram empregados para avaliar a metodologia. Após, os descritores de características utilizados são apresentados. Por fim, o método de classificação e de combinações entre classificadores são explicados. A Figura 4.1 apresenta uma visão geral dos procedimentos realizados, em que cada etapa é descrita detalhadamente nas próximas seções.

#### 4.1 Base de imagens

Neste trabalho, foi utilizada a base de imagens disponibilizada por Depeursinge et al. (2011), que possui casos de Doenças Pulmonares Difusas (DPDs) dos Hospitais Universitários de Genebra. A base contém conjuntos de TCAR de 113 pacientes, que variam de 11 a 61 cortes tomográficos. As imagens de TCAR possuem resolução de  $512 \times 512$



**Figura 4.1:** Sequência de etapas do método proposto.

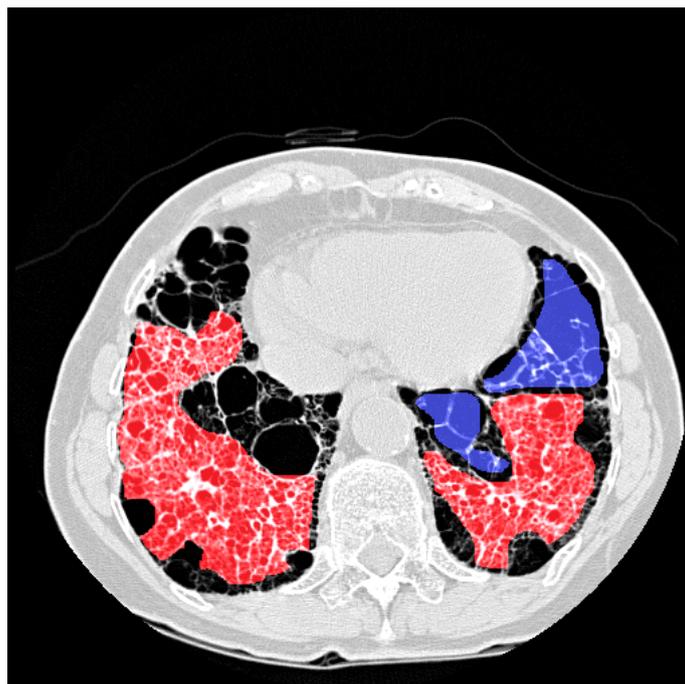
pixels, uma espessura de corte de 1 mm, e tamanho do pixel variando entre 0.4 mm e 1 mm. Além disso, são utilizados 12 bits para armazenar as informações dos pixels, ou seja, 4096 níveis de cinza no intervalo  $[-1024, 3071]$ .

A base provê as máscaras dos pulmões segmentados, bem como o *ground truth* das imagens, isto é, a marcação das regiões pulmonares anormais, realizada por dois radiologistas com 15 e 20 anos de experiência. Ademais, ambos os radiologistas devem concordar com a região delineada. Algumas regiões de tecido pulmonar normal também foram marcadas, para fornecer uma quantidade de exemplos dos aspectos do pulmão normal. No entanto, existem limitações quanto aos exemplos anotados de tecido normal pois as imagens de TCAR pertencem a pacientes que já possuem um histórico médico de DPD.

As regiões delimitadas são chamadas de regiões de interesse (*Regions of Interest - ROIs*). A Figura 4.2 mostra uma imagem de TCAR com quatro ROIs demarcadas por especialistas, onde as ROIs de cor vermelha indicam a presença de fibrose, e as ROIs de cor azul representam o padrão enfisema.

No total, a base possui 17 padrões pulmonares demarcados, porém neste trabalho optou-se por utilizar os cinco padrões mais comuns em casos de DPDs, que são: tecido pulmonar normal, enfisema, vidro-fosco, fibrose e micronódulos (Depeursinge et al., 2011).

Um importante aspecto observado na base de imagens é a presença de seis estudos de



**Figura 4.2:** ROIs demarcadas por especialistas, onde o padrão enfisema é representado pelas áreas azuis, e as áreas vermelhas representam a fibrose.

TCAR que possuem imagens realizadas com o paciente em decúbito ventral. Usualmente, o tecido pulmonar normal apresenta um maior gradiente de densidade do tecido com o paciente em decúbito dorsal, do que em decúbito ventral (Tawhai et al., 2009).

A descrição completa da base de imagens pode ser encontrada no artigo de Depeursinge et al. (2011). A base é disponibilizada por meio de um pedido e da assinatura de um contrato de licença<sup>1</sup>.

## 4.2 Extração dos blocos de imagem

Na etapa de extração de características, os experimentos utilizam pequenos blocos de imagem das ROIs. Esta seção descreve o processo de extração dos blocos de imagem de tecido pulmonar e do corpo (gordura, músculo, osso, etc.).

Os blocos de ROIs de tecido pulmonar foram extraídos como mostra o Algoritmo 1. Para cada ROI pulmonar delimitada pelos radiologistas, que pertença a um dos cinco padrões a serem estudados neste trabalho, uma janela deslizante extrai blocos de tamanho

<sup>1</sup>Disponível em: <http://medgift.hevs.ch/silverstripe/index.php/team/adrien-depeursinge/multimedia-database-of-interstitial-lung-diseases/>

32×32 pixels, com meia sobreposição sobre o eixo  $x$  e o eixo  $y$ , que contenham um mínimo de 75% dos pixels pertencentes ao padrão da ROI. Esse tamanho de bloco foi adotado com base em um estudo anterior de Depeursinge et al. (2012) que afirma que blocos de 32×32 pixels possuem o melhor equilíbrio entre o desempenho da classificação e localização. Em adição, Kim et al. (2008) afirmam que, para preservar as formas dos padrões pulmonares, deve-se utilizar blocos de ROIs com, no mínimo, 20×20 pixels. A Figura 4.3 ilustra esse processo de extração, no qual os contornos amarelos delimitam as ROIs, e o quadrado vermelho representa um bloco de tamanho 32×32 pixels a ser extraído.

---

**Algorithm 1** Pseudocódigo para extração de blocos de tecido pulmonar.

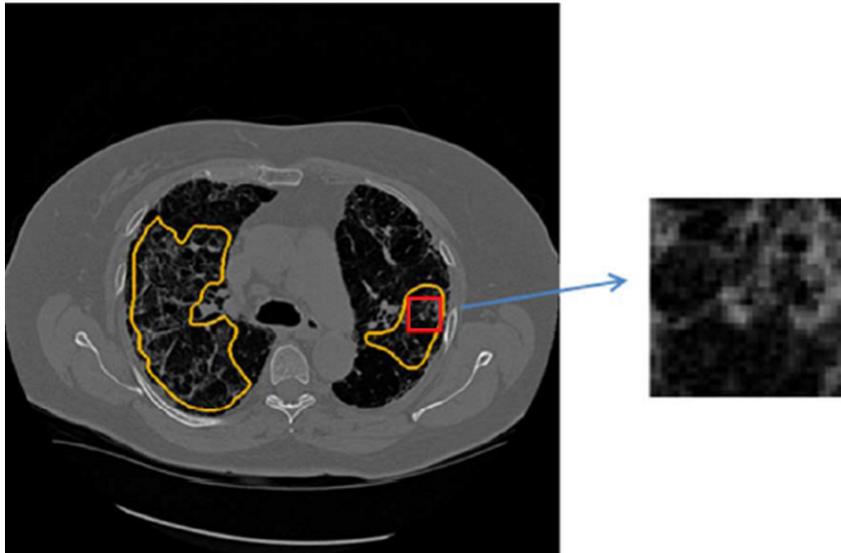
---

**Entrada:** imagem de TCAR original ( $I_{Orig}$ ) e máscara das marcações ( $I_{Marc}$ )

**Saída:** blocos de tecido pulmonar ( $I_{Final}$ )

- 1: encontre todos os contornos de  $I_{Marc}$ , utilizando a função *findContours* da biblioteca OpenCV
  - 2: **para** cada contorno encontrado em  $I_{Marc}$  **faça**
  - 3:     **se** região do contorno pertence a um dos cinco padrões pulmonares **então**
  - 4:         cria bloco de tamanho 32×32 pixels (janela deslizante)
  - 5:         **para** cada meia sobreposição da janela deslizante nos eixos  $x$  e  $y$  **faça**
  - 6:              $num\_px \leftarrow n^o$  de pixels que pertencem ao padrão do contorno
  - 7:             **se**  $num\_px \geq 768$  **então**  $\triangleright$  se possui no mínimo 75% de pixels do padrão
  - 8:                  $I_{Final} \leftarrow bloco$
  - 9:         **fim se**
  - 10:     **fim para**
  - 11:    **fim se**
  - 12: **fim para**
- 

Para extrair blocos externos ao tecido pulmonar, a técnica descrita pelo Algoritmo 2 foi adotada. Utilizando a imagem máscara dos pulmões, fornecida pela base de imagens, aplicou-se uma dilatação morfológica, definida por  $\delta^{(B)}(X)$  (Soille, 2003), com 15 iterações, onde  $X$  é a imagem máscara do pulmão e  $B$  é o elemento estruturante utilizado para aplicar a dilatação (neste caso, o elemento estruturante utilizado é uma elipse circular de tamanho 5). Assim, é obtida a ROI não-pulmonar de cada imagem, definida como a área dilatada que não pertence à máscara pulmonar ( $\delta^{(B)}(X) - (X)$ ). Essa área foi escolhida para representar a ROI de não-pulmão (NP) pois ela contém mais variações por estar conectada às bordas dos pulmões. Após esse procedimento, são extraídos blocos não-sobrepostos de 32×32 pixels que possuam um mínimo de 75% dos seus pixels dentro



**Figura 4.3:** Exemplo de extração de bloco de ROI pulmonar de tamanho 32x32 pixels (adaptado de Song et al. (2013)).

da ROI não-pulmonar. A Figura 4.4 ilustra o método de extração de blocos de não-pulmão. No total, 70.000 blocos foram obtidos, porém foram selecionados randomicamente apenas 3.000 blocos não-pulmonares, um número escolhido com base na média de blocos pulmonares extraídos por classe (ver Tabela 4.1).

---

**Algorithm 2** Pseudocódigo para extração de blocos externos ao tecido pulmonar (não-pulmão).

---

**Entrada:** imagem de TCAR original ( $I_{Orig}$ ) e máscara dos pulmões ( $I_{Mask}$ )

**Saída:** blocos de tecido não-pulmonar ( $I_{Final}$ )

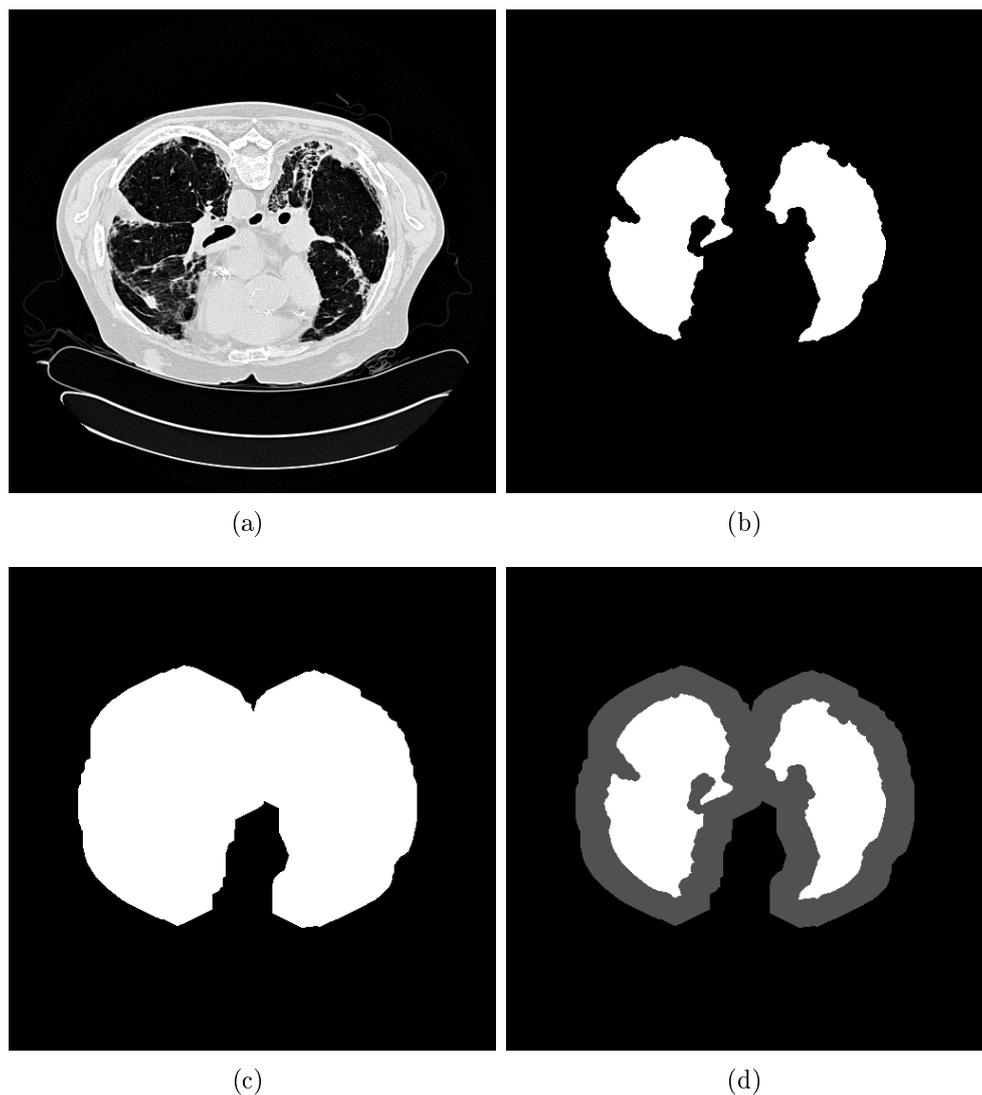
```

1:  $B \leftarrow 5$                                 ▷ elemento estruturante para a dilatação morfológica
2:  $I_{Dil} \leftarrow \delta^{(B)}(I_{Mask})$           ▷ dilatação morfológica
3: para cada contorno encontrado em  $I_{Dil}$  faça
4:   cria bloco de tamanho 32×32 pixels (janela deslizante)
5:   para deslocamento não sobreposto do tamanho da janela nos eixos  $x$  e  $y$  faça
6:      $num\_px \leftarrow n^o$  de pixels que pertencem à região dilatada não pulmonar
7:     se  $num\_px \geq 768$  então                ▷ se possui no mínimo 75% do total de pixels
8:        $I_{Final} \leftarrow bloco$ 
9:     fim se
10:  fim para
11: fim para

```

---

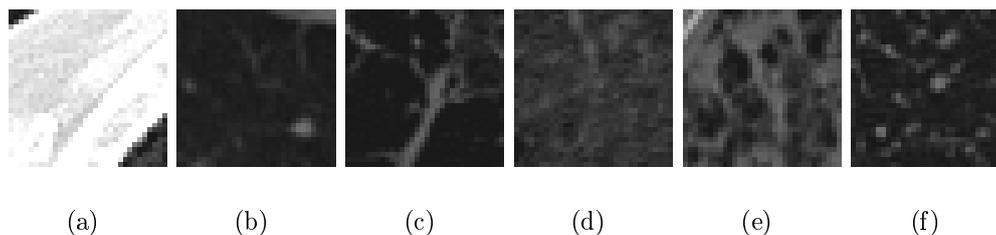
No total, imagens de TCAR de 91 pacientes foram utilizadas para avaliar o desempenho do método proposto. Exemplos dos blocos de imagens extraídos são apresentados na Figura 4.5. Os blocos de imagens são gravados no formato DICOM, para evitar qualquer tipo de conversão dos níveis de cinza. No total, 20.540 blocos são utilizados neste trabalho.



**Figura 4.4:** Método de extração de blocos de não-pulmão, onde (a) imagem original, (b) máscara pulmonar, (c) máscara pulmonar dilatada e (d) imagem resultante, onde a parte cinza é a ROI não-pulmonar.

A Tabela 4.1 mostra a quantidade de blocos obtidos para cada padrão, utilizando os métodos de extração descritos anteriormente.

É importante observar que o objetivo de utilizar blocos com pelo menos 75% dos pixels pertencentes à ROI é obter bons exemplos dos padrões radiológicos. Esse valor de 75% foi adotado com base no trabalho de Depeursinge et al. (2012). Porém, na análise das imagens de TC do pulmão, uma região de  $32 \times 32$  pixels pode conter mais de um padrão de tecido pulmonar e, por consequência, não possuir um mínimo de 75% de determinado padrão.



**Figura 4.5:** Exemplo dos blocos extraídos de (a) não-pulmão, (b) pulmão normal, (c) enfisema, (d) vidro-fosco, (e) fibrose e (f) micronódulos.

**Tabela 4.1:** Quantidade de blocos de ROIs extraídos.

Padrão	Quantidade de blocos
Não-pulmão	3.000
Pulmão normal	5.733
Enfisema	1.017
Vidro-fosco	1.942
Fibrose	2.736
Micronódulos	6.112

### 4.3 Extração de características

Para identificar quais métodos de extração de características melhor se adaptam ao problema de reconhecimento de padrões radiológicos, seis descritores são avaliados neste trabalho, e apresentados nas seções a seguir. O objetivo foi analisar a variedade de descritores para extrair características distintas das imagens, buscando uma forma de complementaridade entre os extratores, visto que algumas características podem se comportar melhor para classificar determinado padrão, e pior para outro.

#### 4.3.1 LBP

O descritor LBP foi avaliado utilizando duas variações. A primeira, denominada  $LBP^{u2}$  refere-se ao padrão uniforme. A segunda consiste no padrão invariante à rotação e uniforme, chamada  $LBP^{riu2}$ . Ambas as variações foram testadas com os valores (8, 1) e (8, 2) para os parâmetros ( $P, R$ ), onde  $P$  é o número de vizinhos e  $R$  é o valor do raio.

### 4.3.2 CLBP

Da mesma forma que o LBP, duas variações do CLBP foram avaliadas, utilizando a combinação CLBP\_S/M/C, descrita anteriormente na Seção 2.4.2. Assim, o  $CLBP\_SMC^{u2}$  utiliza padrões uniformes, e  $CLBP\_SMC^{riu2}$  emprega padrões uniformes invariantes à rotação. Os valores (8, 1) e (8, 2) para os parâmetros  $(P, R)$  também foram adotados para avaliar o descritor. Para extrair as características, utilizou-se o código do CLBP desenvolvido pelo Departamento de Computação da Universidade Politécnica de Hong Kong<sup>2</sup>.

### 4.3.3 LPQ

Para o descritor LPQ, foram extraídas características utilizando funções de janela do tipo Uniforme ( $LPQ^u$ ), Gaussiana ( $LPQ^g$ ), e com filtros de banda de quadraturas Gaussianas Derivativas ( $LPQ^{gd}$ ). Os valores empregados para  $(P, R)$  foram (8, 1), que representam uma 8-vizinhança de raio 1. A aplicação do LPQ utilizada para extração das características foi desenvolvida pelo Departamento de Ciência da Computação e Engenharia da Universidade de Oulu<sup>3</sup>.

### 4.3.4 GLCM

Nesse método de extração de características a imagem de entrada foi submetida a uma conversão de 12 bits para 8 bits. De acordo com Hall-Beyer (2015), quanto mais níveis de cinza forem utilizados no cálculo da GLCM, mais células da matriz são preenchidas com zeros, porque não ocorrem combinações dos níveis de cinza. Como a GLCM aproxima a distribuição de probabilidade conjunta dos pixels, ter muitos zeros na matriz torna essa aproximação muito ruim. Sendo assim, as estatísticas calculadas a partir da GLCM não ajudariam na classificação, se uma grande quantidade de tons de cinza fosse utilizada.

Para transformar os blocos de ROIs em 256 níveis de cinza, optou-se por fazer uma conversão manual, pois as ferramentas de conversão de imagem não trabalham com ima-

<sup>2</sup>Disponível em: <http://www4.comp.polyu.edu.hk/~cslzhang/code/CLBP.rar>

<sup>3</sup>Disponível em: <http://www.cse.oulu.fi/CMV/Downloads/LPQMatlab>

gens de 12 bits, apenas com 16 bits. Se fosse considerada uma conversão de 16 bits para 8 bits, muita informação dos pixels seria perdida. Dessa forma, considerando que a imagem de entrada possui 4096 níveis de cinza, cada célula da imagem foi convertida de acordo com a Equação 4.1:

$$img\_conv_{(l,c)} = \frac{img\_orig_{(l,c)} \times 255}{4096} \quad (4.1)$$

onde  $img\_conv$  é a imagem resultante da conversão,  $l$  e  $c$ , respectivamente, são linha e coluna da matriz, e  $img\_orig$  é a imagem de entrada original, ou seja, o bloco de ROI.

As propriedades de GLCM empregadas nos experimentos foram: contraste, energia, homogeneidade, correlação e segundo momento angular, que também foram utilizadas em trabalhos anteriores (Park et al., 2009; Lim et al., 2011; Chang et al., 2012). Essas cinco características foram extraídas com distância  $d = 1$  em quatro ângulos diferentes ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ ). Assim, obteve-se um total de 20 elementos no vetor de características de GLCM.

### 4.3.5 Histograma e medidas estatísticas

Na Seção 2.2, foi mostrado que os valores da Escala Hounsfield correspondem a medidas de atenuação dos órgãos anatômicos, permitindo identificar componentes do tecido pulmonar, bem como outras substâncias corporais. O histograma de níveis de cinza é utilizado como um método de extração de características para obter informações referentes aos componentes presentes na imagem de TC, transformando o valor de cada célula em um componente do vetor de características. Utilizando a Escala Hounsfield, é possível definir a faixa de dados do histograma em função das estruturas da TC devem ser analisadas.

Para construir o histograma, a faixa de dados definida para análise foi  $[-1000; 700]$ , que compreende as seguintes substâncias: pulmão, gordura, água, músculo, tecidos moles e osso. Esses são os valores de interesse para o método proposto, pois este trabalho objetiva reconhecer não somente tecidos pulmonares, como também tecidos externos ao pulmão. A faixa de dados foi dividida em 25 células (*bins*), de forma que cada célula calcule o número

de pixels de 68 níveis de cinza. Sendo assim, o vetor de características do histograma é composto por 25 elementos.

A partir do histograma construído, um novo conjunto de características é criado. Esse conjunto é composto por medidas estatísticas de distribuição derivadas do histograma, que são: média, desvio-padrão, assimetria e curtose. Dessa forma, esse vetor de características possui quatro elementos.

### 4.3.6 Transformação *top-hat*

As características de forma foram extraídas utilizando as transformações *white top-hat* e *black top-hat*, seguindo a ideia de alguns trabalhos da literatura que realizaram o mesmo (Park et al., 2009; Lim et al., 2011; Chang et al., 2012). A operação *white top-hat* retorna os pontos claros (picos) da imagem, que são menores que um elemento estruturante definido, enquanto o *black top-hat* retorna os pontos escuros (vales) da imagem, menores que o elemento estruturante.

Dessa forma, essas duas operações foram aplicadas empregando um elemento estruturante em forma de disco, com  $R = 5$  para o valor do raio. A partir disso, a média e o desvio-padrão das duas imagens resultantes foram extraídas, de forma a compor o vetor de características da transformação *top-hat*, que possui quatro elementos.

### 4.3.7 Resumo dos descritores utilizados

A Tabela 4.2 sumariza os descritores testados e o tamanho do vetor de características resultante. Os dados dos vetores foram normalizados utilizando o aplicativo *svm-scale* da biblioteca LIBSVM (Chang and Lin, 2011), no intervalo  $[-1, 1]$ .

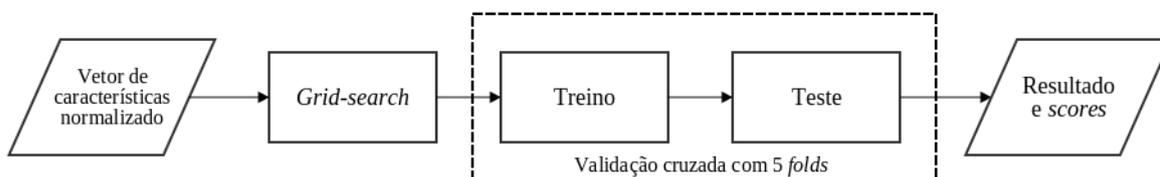
## 4.4 Classificação

O método de classificação consiste em treinar um classificador para cada um dos métodos de extração de características apresentados na Tabela 4.2. Dessa forma, 15 classificadores distintos foram treinados. Posteriormente, foram efetuadas combinações entre

**Tabela 4.2:** Resumo dos descritores utilizados.

Descritor	Tamanho do vetor
$LBP_{(8,1)}^{u2}$	59
$LBP_{(8,2)}^{u2}$	59
$LBP_{(8,1)}^{riu2}$	10
$LBP_{(8,2)}^{riu2}$	10
$CLBP\_S/M/C_{(8,1)}^{u2}$	200
$CLBP\_S/M/C_{(8,2)}^{u2}$	200
$CLBP\_S/M/C_{(8,1)}^{riu2}$	200
$CLBP\_S/M/C_{(8,2)}^{riu2}$	200
$LPQ_{(8,1)}^u$	256
$LPQ_{(8,1)}^g$	256
$LPQ_{(8,1)}^{gd}$	256
GLCM	20
Histograma	25
Medidas estatísticas do histograma	4
Transformação <i>top-hat</i>	4

esses classificadores. A Figura 4.6 apresenta o processo de classificação realizado para cada característica, explicado detalhadamente neste capítulo.

**Figura 4.6:** Processo de classificação para cada característica extraída.

O classificador utilizado foi o *Support Vector Machine* (SVM), empregando a técnica *one-vs-one* na separação das classes. O SVM foi utilizado nessa fase em razão de ser o melhor classificador no reconhecimento de padrões de tecido pulmonar (Depeursinge et al., 2010). Para isso, utilizou-se a biblioteca *scikit-learn*<sup>4</sup> para Python, desenvolvida por Pedregosa et al. (2011). A função de *kernel* utilizada para classificação foi a RBF, e os parâmetros  $C$  (custo) e  $\gamma$  para os classificadores foram determinados através de um método de *grid-search* em cada conjunto de treinamento.

Para executar o treino e teste, os blocos de imagens foram divididos utilizando o método *k-fold* de validação cruzada, com  $k = 5$ . As imagens foram divididas com base nos

<sup>4</sup>Disponível em: <http://scikit-learn.org/>

pacientes, garantindo que todos os blocos de imagem pertencentes a um mesmo paciente estejam em um mesmo grupo, impedindo treino e teste com imagens do mesmo paciente, e evitando o *over-fitting* das imagens de teste. A Tabela 4.3 apresenta a quantidade de pacientes em cada grupo, bem como o número de blocos de imagens presente. Na Tabela 4.4 é possível visualizar a quantidade de exemplos de cada padrão radiológico por *fold*. É importante salientar que a divisão dos *fold*s foi realizada de forma randômica, apenas garantindo a não existência de imagens do mesmo paciente simultaneamente nas etapas de treino e teste.

**Tabela 4.3:** Número de pacientes e blocos por *fold*.

<i>Fold</i>	N <sup>o</sup> de pacientes	N <sup>o</sup> de blocos
1	43	4.044
2	19	4.214
3	21	3.991
4	17	3.962
5	12	4.329

**Tabela 4.4:** Número de exemplos de cada classe por *fold*.

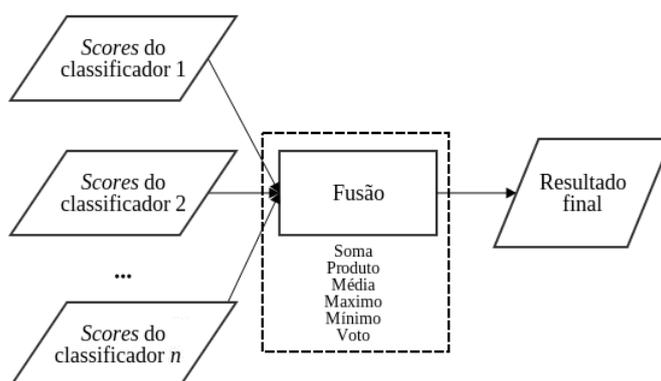
Fold	Não-pulmão	Normal	Enfisema	Vidro-fosco	Fibrose	Micronódulos
1	1160	668	58	559	444	1155
2	620	2305	175	682	352	80
3	546	203	0	390	494	2358
4	389	226	111	311	1195	1730
5	285	2331	673	0	251	789

A fim de realizar posteriores combinações dos classificadores, optou-se por adotar a estratégia proposta por Platt (1999), e adaptada por Wu et al. (2004) para problemas multiclasse, que consiste em estimar probabilidades para as saídas geradas do SVM. Dessa forma, cada classificador gera *scores* que representam a probabilidade das amostras pertencerem a cada uma das seis classes.

## 4.5 Fusão dos classificadores

Com o propósito de combinar as respostas dos classificadores e verificar qual combinação gera os melhores resultados, os classificadores foram treinados para produzir probabili-

dades *a posteriori*, que representam a hipótese de determinado exemplo pertencer a cada uma das classes. Em geral, utiliza-se um esquema de votação para efetuar as combinações, no qual a classe eleita como a correta é aquela que possui mais votos dos classificadores. No entanto, este trabalho avaliou seis diferentes regras de combinação, sendo: Soma, Produto, Média, Máximo, Mínimo e Voto (Kittler et al., 1998). Na Figura 4.7 é apresentado o processo da fusão de classificadores.



**Figura 4.7:** Processo de fusão de respostas dos classificadores.

A fim de identificar quais classificadores formavam a melhor combinação, a fusão consistiu em gerar todas as possíveis combinações de respostas, começando por grupos de 2 classificadores, até chegar na única combinação dos 15 classificadores. A quantidade total de combinações é definida pela Equação 4.2, onde  $n$  é a quantidade total de classificadores (15 nesse caso) e  $r$  representa o número de classificadores combinados em cada processo da fusão.

$$\sum_{r=2}^n \frac{n!}{r! \times (n-r)!} \quad (4.2)$$

No total foram efetuadas 32.752 combinações para os 15 classificadores. Cada uma dessas combinações aplicou as seis regras de fusão descritas anteriormente. O resultado final desse processo retorna a matriz de confusão de cada fusão.

A análise dos resultados é feita por meio da matriz de confusão gerada pelos classificadores, que mostra todas as confusões ocorridas entre as classes. A partir do número de VP, FP, VN e FN obtidos, calculam-se as medidas de sensibilidade, precisão, especificidade

e acurácia. Os resultados de todos os classificadores individuais são analisados e, para a fusão de classificadores, apenas as 10 melhores combinações com base na sensibilidade são apresentadas.

## CAPÍTULO 5

### RESULTADOS E DISCUSSÃO

Este capítulo apresenta os resultados experimentais do método proposto no capítulo anterior, ou seja, os resultados individuais da classificação, com 15 classificadores, bem como os 10 melhores resultados de fusão dos classificadores, com base na sensibilidade.

Nas matrizes de confusão apresentadas, as linhas representam a classe real dos objetos e as colunas representam a classe predita no processo de classificação. Em adição, as classes dos objetos são representadas por siglas, da seguinte forma: não pulmão (NP), pulmão normal (N), enfisema (E), vidro-fosco (VF), fibrose (F) e micronódulos (M).

#### 5.1 Resultados individuais

Cada método de extração de características, com seus parâmetros variados, foi submetido a um classificador SVM, conforme o método descrito na Seção 4.4. Cada classificador gerou como saída a matriz de confusão do resultado, bem como probabilidades *a posteriori* que representam a hipótese de cada amostra pertencer a cada uma das classes.

A Tabela 5.1 mostra os resultados gerais dos 15 classificadores individuais, pela ordem decrescente de sensibilidade. Com base nesses resultados gerais, apresenta-se a seguir apenas a matriz de confusão do melhor classificador de cada família de descritores (ou o único, no caso da transformação *top-hat* e GLCM).

É interessante observar que os melhores resultados foram obtidos pela família CLBP, mostrando a superioridade desse descritor sobre os demais, quando se trata de classificadores individuais. A diferença de taxa de reconhecimento da pior configuração do CLBP para o próximo melhor resultado (transformação *top-hat*) é de 2,43 pontos percentuais (pp), e a diferença do melhor CLBP para a transformação *top-hat* chega a 9,36 pp. A transformação *top-hat*, que extraiu medidas estatísticas dos blocos de imagens, foi o segundo melhor descritor.

**Tabela 5.1:** Resultados gerais dos 15 classificadores individuais (em %)

Classificador	Sensibilidade	Precisão	Especificidade	Acurácia
$CLBP\_S/M/C_{(8,2)}^{riu2}$	78,99	75,95	95,55	93
$CLBP\_S/M/C_{(8,1)}^{riu2}$	75,29	70,84	94,82	91,76
$CLBP\_S/M/C_{(8,2)}^{u2}$	74,18	70,76	94,51	91,39
$CLBP\_S/M/C_{(8,1)}^{u2}$	72,06	67,89	94,12	90,69
Transformação <i>top-hat</i>	69,63	61,86	93,5	89,88
Histograma	67,47	63,44	93,05	89,16
$LBP_{(8,2)}^{u2}$	66,97	64,31	92,88	88,99
$LPQ_{(8,1)}^g$	66,76	61,31	92,95	88,92
$LBP_{(8,1)}^{u2}$	66,58	62,27	92,84	88,86
$LBP_{(8,2)}^{riu2}$	66,01	62,79	92,7	88,67
$LPQ_{(8,1)}^u$	63,53	55,98	92,33	87,84
$LBP_{(8,1)}^{riu2}$	62,57	58,28	91,96	87,52
$LPQ_{(8,1)}^{gd}$	62,1	53,45	92,05	87,37
GLCM	55	53,36	90,51	85
Medidas estatísticas do histograma	41,42	42,82	87,04	80,47

As melhores configurações para o CLBP e LBP utilizaram janelas de 8-vizinhança e raio igual a dois. Porém, a melhor versão do CLBP empregou padrões uniformes e invariantes à rotação, enquanto que o LBP atingiu melhores resultados apenas com o padrão uniforme. O descritor de textura LPQ teve seus melhores resultados aplicando uma janela gaussiana.

O histograma de níveis de cinza atingiu resultados melhores que as medidas estatísticas extraídas do mesmo, que obteve a pior taxa de classificação. Isso condiz com o apresentado por Depeursinge et al. (2012), que afirma que resultados preliminares já mostraram que os valores dos *bins* do histograma permite melhor classificação dos padrões pulmonares, se comparado às medidas estatísticas da distribuição. Embora as medidas estatísticas do histograma sejam utilizadas em diversos trabalhos da literatura, esse descritor obteve os piores resultados individuais. Outro descritor empregado por diversos métodos da literatura é o GLCM, que obteve uma taxa de reconhecimento quase 24 pp mais baixa que o melhor classificador.

Conforme pode-se observar na Tabela 5.1, o melhor resultado de cada família de descritores, por ordem decrescente de sensibilidade foi:

1.  $CLBP\_S/M/C_{(8,2)}^{riu2}$  (Tabela 5.2);
2. Transformação *top-hat* (Tabela 5.3);
3. Histograma (Tabela 5.4);
4.  $LBP_{(8,2)}^{u2}$  (Tabela 5.5);
5.  $LPQ_{(8,1)}^g$  (Tabela 5.6);
6. GLCM (Tabela 5.7).

Analisando a Tabela 5.2, referente à matriz de confusão do  $CLBP\_SMC_{(8,2)}^{riu2}$ , nota-se que o descritor consegue reconhecer as classes normal, fibrose e micronódulos com taxas próximas, por classe, aos trabalhos da literatura que utilizaram a mesma base de imagens (Depeursinge et al., 2012; Li et al., 2013; Song et al., 2013). Além disso, alcançou ótima taxa de reconhecimento para a classe não pulmonar. Porém, o descritor apresenta problemas em identificar as classes enfisema e vidro-fosco. Muitas confusões ocorrem entre os pares de classes enfisema e normal, vidro-fosco e normal, e vidro-fosco com fibrose. Diversas amostras dos padrões enfisema e vidro-fosco foram incorretamente classificadas como pertencentes ao padrão normal, gerando muitos FPs para a classe normal. Mesmo assim, o número de FPs gerados para o padrão normal é o menor dentre todos os obtidos pelos classificadores individuais para a mesma classe.

**Tabela 5.2:** Matriz de confusão da melhor configuração do CLBP ( $CLBP\_SMC_{(8,2)}^{riu2}$ ) (em %)

Classe	Predição					
	NP	N	E	VF	F	M
NP	97,5	0,7	0,17	1,37	0,23	0,03
N	0,68	79,98	0,78	7,01	0,72	10,83
E	2,46	19,96	57,13	5,51	10,91	4,03
VF	3,5	24,46	0,26	39,8	23,43	8,55
F	0,58	1,1	1,86	15,1	77,6	3,76
M	0,02	9,62	0,44	1,42	2,81	85,68

Os resultados obtidos pela transformação *top-hat*, apresentados na Tabela 5.3, mostram que o descritor não conseguiu diferenciar o padrão enfisema do padrão normal, embora o contrário não seja válido. Ao utilizar a transformação, não foi possível extrair da

classe enfisema características que a tornasse distinta da classe normal. É possível ainda afirmar que isso ocorreu em outros pares de classes, porém em menor escala, como em enfisema com micronódulos. Dessa forma, pode-se concluir que esse descritor, aplicando os parâmetros citados na Seção 4.3.6, não conseguiu extrair as características da classe enfisema. Outro ponto negativo do descritor é o alto valor de FP gerado para o padrão normal. Isso indica que muitos exemplos das classes enfisema, vidro-fosco e micronódulos são identificados como tecido pulmonar normal. A classe vidro-fosco obteve, com esse descritor, a maior taxa de reconhecimento dentre todos os classificadores individuais, embora inferior a 50 pp.

**Tabela 5.3:** Matriz de confusão da transformação top-hat (em %)

Classe	Predição					
	NP	N	E	VF	F	M
NP	97,73	0,3	1,2	0,47	0,13	0,17
N	0,47	78,98	3,73	4,2	1,5	11,11
E	1,18	50,05	12,29	0	5,6	30,88
VF	2,06	30,64	0	48,56	14,88	3,86
F	3,03	3,8	1,1	13,16	66,34	12,57
M	10,19	15,38	1,9	1,1	6,66	64,77

A matriz de confusão do histograma, exposta na Tabela 5.4, indica que o descritor obteve a maior taxa de reconhecimento para o padrão não pulmonar dentre todos os classificadores individuais, sendo essa de 99,5%. Como essa classe apresenta pixels de alta intensidade sem nenhum padrão de textura, torna-se mais fácil diferenciá-la das demais classes. O histograma consegue discriminar algumas classes, como enfisema e vidro-fosco, pois a primeira apresenta valores de pixels mais baixos que a segunda. No entanto, devido à semelhança de valores que as demais classes apresentam na HU, torna-se difícil identificá-las corretamente somente com esse descritor. Por exemplo, os padrões normal e enfisema possuem histogramas semelhantes, bem como os padrões vidro-fosco e fibrose (Depeursinge, 2010), considerando seus valores correspondentes na Escala *Hounsfield*. Assim, o histograma apenas modela uma informação que corresponde às atenuações das estruturas na imagem de TC.

Embora o descritor LBP já tenha sido empregado em trabalhos de identificação de

**Tabela 5.4:** Matriz de confusão do histograma (em %)

Classe	Predição					
	NP	N	E	VF	F	M
NP	99,5	0	0,1	0,17	0,23	0
N	0,05	58,21	4,20	14,79	0,02	22,73
E	1,77	23,01	41	0,1	9,44	24,68
VF	0,98	37,8	0	40,94	12,62	7,67
F	1,43	0,37	1,5	9,8	75	11,92
M	0,02	13,27	3,88	2,52	10,42	69,9

subtipos de enfisema, sua variação  $LBP_{(8,2)}^{u2}$  não obteve destaque em nenhuma das classes, como mostra sua correspondente matriz de confusão, na Tabela 5.5. Também apresentou as mesmas confusões já discutidas até aqui, envolvendo as classes enfisema, vidro-fosco e micronódulos com a classe normal, bem como confusões recíprocas entre vidro-fosco e fibrose. Cabe observar que o melhor resultado da família de descritores LBP foi obtido com sua versão uniforme, diferentemente do melhor resultado da família CLBP, que empregou a versão invariante à rotação e uniforme.

**Tabela 5.5:** Matriz de confusão da melhor configuração do LBP ( $LBP_{(8,2)}^{u2}$ ) (em %)

Classe	Predição					
	NP	N	E	VF	F	M
NP	94,2	1,37	0,37	2,3	0,97	0,8
N	1,88	61,47	1,78	9,24	3,47	22,15
E	3,34	27,14	41,99	0,88	10,91	15,73
VF	5,77	36,35	0	34,4	20,03	3,45
F	0,99	7,38	2,81	17,32	62,57	8,92
M	0,02	18,75	1,18	1,01	3,81	75,23

As mesmas observações podem ser feitas para o  $LPQ_{(8,1)}^g$  que, utilizando janela gaussiana, também se deparou com as confusões interclasse citadas anteriormente, como mostra a Tabela 5.6. Em linhas gerais, as confusões ocorridas neste classificador são semelhantes às apresentadas nos outros descritores de textura utilizados.

A matriz de co-ocorrência de níveis de cinza, ou simplesmente GLCM, é um dos métodos mais utilizados pelos trabalhos apresentados no Capítulo 3. No entanto, sua matriz de confusão, apresentada na Tabela 5.7, mostra que além de seu resultado geral ter sido muito inferior aos demais, suas taxas de reconhecimento para as classes não pulmão e pul-

**Tabela 5.6:** Matriz de confusão da melhor configuração do LPQ ( $LPQ_{(8,1)}^g$ ) (em %)

Classe	Predição					
	NP	N	E	VF	F	M
NP	90,6	4,97	0,5	2,6	1,27	0,07
N	7,9	60,47	2,44	7,33	3,37	18,49
E	9,44	26,84	41,2	1,47	5,41	15,63
VF	12,46	24,67	0,41	26,16	26,21	10,09
F	2,23	5,85	1,02	14,51	70,03	6,36
M	0,08	13,81	2,14	2,05	5,28	76,64

mão normal foram as piores dentre as apresentadas por todos os classificadores individuais para essas mesmas classes. Como citado na Seção 4.3.4, as estatísticas calculadas a partir da GLCM podem não ajudar na classificação, se uma imagem com grande número de níveis de cinza é utilizada. Por esse motivo a imagem sofreu uma conversão simples para 8 bits. Pode-se considerar realizar uma conversão diferente nas imagens, em que o intervalo utilizado para conversão seja o mesmo empregado no histograma, que corresponde somente aos valores de interesse na HU.

**Tabela 5.7:** Matriz de confusão da GLCM (em %)

Classe	Predição					
	NP	N	E	VF	F	M
NP	69,2	10,63	0,4	9,3	6,6	3,87
N	16,01	55	1,97	3,24	0,12	23,65
E	9,54	10,23	34,51	2,26	12,49	30,97
VF	19,36	16,84	0	30,79	16,07	16,94
F	7,82	0,95	1,28	13,38	67,36	9,21
M	14,59	17,41	1,69	2,81	9,91	53,58

## 5.2 Resultados da combinação de classificadores

Cada classificador individual gerou matrizes de probabilidades *a posteriori* que permitem realizar a combinação de classificadores utilizando regras de fusão, como explicado na Seção 2.5.2. Aplicou-se as regras de voto majoritário, produto, soma, média, máximo e mínimo sobre cada combinação possível entre classificadores (ver Equação 4.2). As 10 melhores combinações de classificadores são apresentadas na Tabela 5.8 que mostra os resultados calculados por meio das medidas de sensibilidade, precisão, especificidade e

acurácia.

Todos os melhores resultados foram combinados empregando a regra do produto, mostrando portanto ser a melhor técnica de combinação para este problema, dentre as seis avaliadas. A regra do produto é caracterizada por ser bastante severa, pois é baseada na multiplicação dos níveis de confiança calculados pelos classificadores, isto é, as probabilidades *a posteriori*. Basta que uma amostra possua um resultado baixo, isso será refletido ao se multiplicar as probabilidades dos demais classificadores a esta mais baixa.

**Tabela 5.8:** Resultados gerais das 10 melhores combinações (em %)

Combinação	Sensibilidade	Precisão	Especificidade	Acurácia
(a)	81,65	79	96,06	93,88
(b)	81,56	79,82	96,05	93,85
(c)	81,51	79,36	96,06	93,84
(d)	81,51	78,47	96,07	93,84
(e)	81,42	78,94	96,03	93,81
(f)	81,3	79,64	96	93,77
(g)	81,3	79,35	95,99	93,77
(h)	81,23	78,93	95,99	93,74
(i)	81,09	78,99	95,97	93,7
(j)	81,03	79,07	95,94	93,68

(a)  $CLBP_{(8,2)}^{riu2}$ , top-hat

(b)  $CLBP_{(8,2)}^{riu2}$ ,  $CLBP_{(8,2)}^{u2}$ , top-hat

(c)  $CLBP_{(8,1)}^{riu2}$ ,  $CLBP_{(8,2)}^{riu2}$ ,  $CLBP_{(8,2)}^{u2}$ , top-hat

(d)  $CLBP_{(8,1)}^{riu2}$ ,  $CLBP_{(8,2)}^{riu2}$ , top-hat

(e)  $CLBP_{(8,2)}^{riu2}$ ,  $CLBP_{(8,1)}^{u2}$ , top-hat

(f)  $CLBP_{(8,2)}^{riu2}$ ,  $CLBP_{(8,1)}^{u2}$ ,  $CLBP_{(8,2)}^{u2}$ , top-hat

(g)  $CLBP_{(8,2)}^{riu2}$ ,  $CLBP_{(8,2)}^{u2}$ , top-hat,  $LPQ_{(8,1)}^u$

(h)  $CLBP_{(8,1)}^{riu2}$ ,  $CLBP_{(8,2)}^{riu2}$ ,  $CLBP_{(8,2)}^{u2}$ , top-hat,  $LPQ_{(8,1)}^u$

(i)  $CLBP_{(8,1)}^{riu2}$ ,  $CLBP_{(8,2)}^{riu2}$ ,  $CLBP_{(8,1)}^{u2}$ ,  $CLBP_{(8,2)}^{u2}$ , top-hat

(j)  $CLBP_{(8,2)}^{riu2}$ ,  $CLBP_{(8,1)}^{u2}$ ,  $CLBP_{(8,2)}^{u2}$ , top-hat,  $LPQ_{(8,1)}^u$

Os dados apresentados na Tabela 5.8 mostram que os classificadores  $CLBP_{(8,2)}^{riu2}$  e transformação *top-hat* participam das 10 melhores combinações. Além disso, o restante da família de descritores CLBP também se mostra presente entre os 10 resultados. De fato, as quatro variações do CLBP mostraram os melhores resultados individuais, e a transformação *top-hat* apresentou resultados individuais que a colocaram como segundo melhor classificador por família de descritores.

Apesar de não ter obtido bons resultados como classificador individual, inclusive não sendo o melhor classificador da sua família de descritores, o  $LPQ_{(8,1)}^u$  participa de três combinações. Isso indica que ele possui informações que o tornam complementar aos demais classificadores da combinação, e está de acordo com o afirmado por Kittler et al. (1998), que citam que a combinação de classificadores melhora os resultados obtidos no reconhecimento de padrões devido à complementaridade entre eles.

A Tabela 5.9 apresenta a matriz de confusão da melhor combinação de classificadores, isto é  $CLBP_{(8,2)}^{riu2}$  e transformação *top-hat*. Observa-se que, se comparada à matriz de confusão do melhor classificador individual,  $CLBP_{(8,2)}^{riu2}$  (Tabela 5.2), houve um aumento de sensibilidade em todas as classes, exceto na classe enfisema (Não pulmão de 97,5 para 99,2; Normal de 79,98 para 85,92; Vidro-fosco de 39,8 para 51,29; Fibrose de 77,6 para 79,57; Micronódulos de 85,68 para 85,77, em %). A transformação *top-hat* foi o classificador individual que obteve o pior resultado de reconhecimento para o padrão enfisema (12,29%) e sua combinação com o classificador  $CLBP_{(8,2)}^{riu2}$  refletiu esse aspecto negativo do descritor.

**Tabela 5.9:** Matriz de confusão da melhor combinação de classificadores (em %)

Classe	Predição					
	NP	N	E	VF	F	M
NP	99,2	0,03	0,5	0,2	0,03	0,03
N	0,16	85,92	1,53	4,76	0,4	7,22
E	1,47	32,35	44,64	0,79	8,55	12,19
VF	2,11	29,71	0	51,29	13,65	3,24
F	1,46	1,39	1,28	11,15	79,57	5,15
M	0,03	9,21	0,29	0,74	3,96	85,77

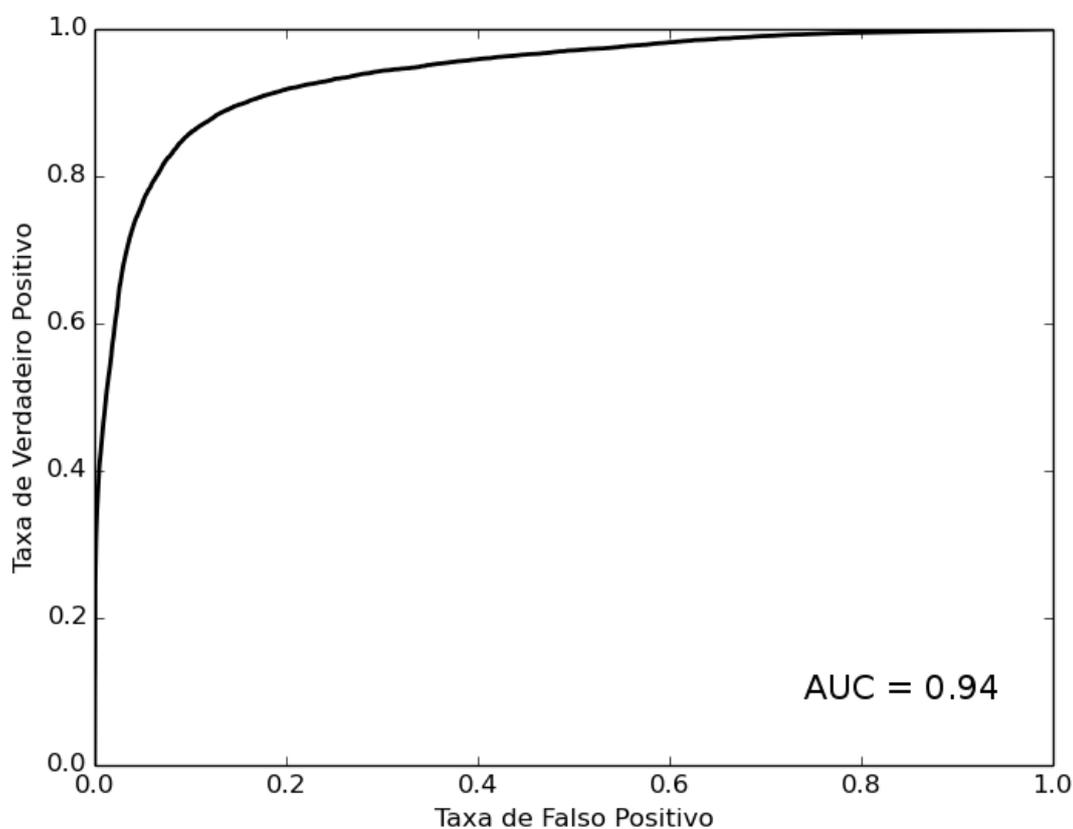
É possível ainda visualizar os dados de classificação de forma mais condensada, considerando apenas três classes: não-pulmão (NP), pulmão normal (N) e pulmão anormal (A), como mostra a Tabela 5.10. Os dados ali apresentados mostram que a combinação  $CLBP_{(8,2)}^{riu2}$  e transformação *top-hat* reconhece os diversos padrões radiológicos anormais com 86,41% de sensibilidade.

Outro dado referente à melhor combinação de classificadores é a curva ROC, ilustrada na Figura 5.1. Para o conjunto  $CLBP_{(8,2)}^{riu2}$  e transformação *top-hat*, unidos pela com-

**Tabela 5.10:** Matriz de confusão simplificada da melhor combinação de classificadores (em %).

Classe	Predição		
	NP	N	A
NP	99,2	0,03	0,77
N	0,16	85,92	13,92
A	0,83	12,76	86,41

binhação do produto, a Área Sob a Curva (*Area Under the Curve* - AUC) calculada é de aproximadamente 94%. A curva ROC mostra a relação de custo-benefício dos classificadores, exibindo a quantidade de falsos positivos pela quantidade de verdadeiros positivos em vários pontos. Isso indica que um determinado ponto localizado mais à esquerda do gráfico fornece mais sensibilidade e menos especificidade, enquanto que um ponto localizado mais acima provê mais especificidade e menos sensibilidade. Com a curva ROC é possível obter o melhor limiar de combinação de valores de especificidade e sensibilidade. A área sob a curva ROC mede o desempenho de um teste, o quão exato ele é. Isso significa que em média 94% das vezes, o teste consegue identificar os VPs e VNs.

**Figura 5.1:** Curva ROC da melhor combinação de classificadores.

Como pode ser observado, a melhor combinação de classificadores, além de todos os classificadores individuais, possui falhas em reconhecimento, especialmente nas classes enfisema e vidro-fosco. Com base nisso, avaliou-se a influência da organização dos *folds* construídos nos resultados apresentados. Isto é, a possibilidade de que a composição dos *folds* estivesse afetando de forma negativa os resultados gerais. Como explicado no capítulo anterior, os *folds* foram construídos de forma randômica, apenas garantindo a não existência de imagens de um mesmo paciente nas etapas de treino e teste simultaneamente. No entanto, como observado na Tabela 4.4, a quantidade de exemplos de cada classe por grupo não é balanceada e, além disso, nem todos os grupos possuem exemplos de todas as classes. Para avaliar a influência dos *folds* sobre o resultado geral da melhor combinação de classificadores, os seguintes dados analisados: taxa de sensibilidade de cada *fold* (Tabela 5.11), taxa de sensibilidade geral e desvio padrão geral.

**Tabela 5.11:** Sensibilidade média por *fold* da melhor combinação de classificadores (em %).

<i>Fold</i>	Sensibilidade
1	77,27
2	84,07
3	82,53
4	89,65
5	75,16

Considerando que a sensibilidade média atingida pela combinação  $CLBP_{S/M/Criu2_{(8,2)}}$  e transformação *top-hat* foi de 81,65% com desvio-padrão de 5,09, ou seja,  $81,65 \pm 5,09$ , e observando também que a diferença da melhor taxa por *fold* para a pior taxa por *fold* é de 14,49 pontos percentuais, pode-se concluir que os resultados de classificação foram afetados pela organização dos *folds*. O *fold* que possui a pior distribuição de amostras por classe é, por consequência, o *fold* que obteve a pior taxa (75,16%). Já o *fold* que possui a melhor distribuição de amostras por classe atingiu o melhor resultado, de 89,65%.

Nesse caso, nota-se a dificuldade em obedecer a restrição da garantia de que imagens de um mesmo paciente permaneçam num mesmo *fold*, e ao mesmo tempo construir *folds* de forma aleatória porém balanceada. Isto é, numa situação ideal, os *folds* deveriam ser produzidos de forma aleatória e ao mesmo tempo possuir uma distribuição de amostras

equilibrada, tanto para a quantidade de amostras por classe, quanto para a quantidade de amostras por *fold*, sem esquecer do requisito de imagens de um paciente no mesmo *fold*.

### 5.3 Conclusão

Este capítulo apresentou os resultados experimentais para os descritores estudados neste trabalho, bem como para os métodos de combinação de resposta dos classificadores individuais.

Por meio dos resultados apresentados, é possível identificar que a maioria das confusões, em todas as matrizes de confusão apresentadas, ocorreu entre as classes enfisema e normal, vidro-fosco e normal, e vidro-fosco e fibrose. Essas confusões interclasse podem ser explicadas pelo fato de que os padrões enfisema e tecido normal podem apresentar baixas intensidades na região pulmonar, sendo difícil diferenciar os detalhes desses dois padrões. Já o vidro-fosco pode mostrar altas atenuações, tornando-se similar à fibrose (Song et al., 2013). Por outro lado, enfisema e fibrose contêm muitas estruturas locais irregulares, mostrando uma ampla variação intraclasse. Ademais, o tecido pulmonar normal não possui uma textura uniforme única, pois depende da idade ou do histórico como fumante do paciente (Depeursinge et al., 2011) e também da aparência dos vasos e bronquíolos (Malone et al., 2004).

Devido às grandes confusões ocorridas nas classes enfisema e vidro-fosco, sentiu-se a necessidade de realizar uma etapa pós-classificação, que consiste em avaliar e resolver as confusões mais relevantes ocorridas e que representam a maior parte do total de erros. A verificação é um procedimento que permite reavaliar determinada amostra, extraindo características específicas para cada confusão que se deseja analisar. Porém, não há um conhecimento sobre quais características melhor representariam os erros ocorridos.

Considerando a melhor combinação de classificadores (Tabela 5.9), afirma-se que os resultados alcançados pelas classes normal, fibrose e micronódulos foram similares aos da literatura, sendo que a categoria dos micronódulos obteve a maior taxa de reconhecimento entre alguns estudos que utilizaram a mesma base de imagens (Depeursinge et al., 2012; Song et al., 2013; Li et al., 2013). O padrão pulmonar normal obteve melhor sensibilidade

que o apresentado por Depeursinge et al. (2012) e Li et al. (2013), porém inferior ao obtido por Song et al. (2013). Entretanto, as classes enfisema e vidro-fosco obtiveram os resultados mais baixos reportados na literatura. Este estudo demonstrou isso de forma clara por meio de uma avaliação completa, comparando as principais medidas descritas na literatura.

Em relação ao padrão não pulmonar, uma sensibilidade de 99,2% foi alcançada na classificação por blocos. Malone et al. (2004) obtiveram 98% de sensibilidade para o padrão não-pulmão, porém é importante observar que eles utilizaram uma base de imagens privada.

Ao se comparar o resultado do melhor classificador com o resultado da melhor combinação de classificadores, tem-se um aumento em todas as medidas gerais calculadas, sendo 2,66 pontos percentuais (pp) na sensibilidade, 3,05 pp na precisão, 0,51 para a especificidade e 0,88 na acurácia.

A principal desvantagem da melhor combinação de classificadores são os diversos erros de classificação entre as classes enfisema e normal, vidro-fosco e normal, e vidro-fosco e fibrose. Isso é uma limitação imposta pela utilização de descritores de textura. Algumas regiões pulmonares não possuem uma textura homogênea. Além disso, na classificação por blocos as imagens perdem propriedades globais de textura, sendo ressaltadas somente as propriedades locais (Paula-Filho, 2012). Porém, cabe ressaltar que a composição final dos *folds* também afetou de forma negativa os resultados atingidos.

É possível realizar uma comparação entre os resultados atingidos e os demais métodos da literatura. Essa comparação é feita considerando as medidas de sensibilidade e acurácia, sendo apresentada na Tabela 5.12. Como pode-se observar, a taxa de sensibilidade do método proposto é maior que as obtidas por Depeursinge et al. (2012) e Li et al. (2013).

**Tabela 5.12:** Comparativo com os resultados obtidos na literatura (em %).

Autor	Padrões analisados	Sensibilidade	Acurácia
Park et al. (2009)	normal, enfisema, opacidade em vidro-fosco, opacidade reticular, faveolamento e consolidação	92,5	89
Vo and Sowmya (2010)	normal, enfisema, opacidade em vidro-fosco e faveolamento	94,16	-
Lim et al. (2011)*	normal, enfisema, opacidade em vidro-fosco, opacidade reticular, faveolamento e consolidação	-	89,88
Depeursinge et al. (2012)*	normal, enfisema, opacidade em vidro-fosco, fibrose e micronódulos	75,81	90,76
Chang et al. (2013)	normal, enfisema, opacidade em vidro-fosco, faveolamento e opacidade reticular, e consolidação	-	92,63
Song et al. (2013)*	normal, enfisema, opacidade em vidro-fosco, fibrose e micronódulos	82,64	-
Zhao et al. (2013)	normal, enfisema, nodular, opacidade em vidro-fosco, faveolamento e consolidação	-	95,4
Li et al. (2013)*	normal, enfisema, opacidade em vidro-fosco, fibrose e micronódulos	74,2	-
Dash et al. (2014)	normal, enfisema, opacidade em vidro-fosco, nodular e consolidação	-	95
Pereyra et al. (2014)	normal, enfisema, opacidade em vidro-fosco, faveolamento, consolidação e espessamento dos septos	82,6	-
Este trabalho*	normal, enfisema, opacidade em vidro-fosco, fibrose e micronódulos	81,65	93,88

\* Utilizaram base de imagens pública de Depeursinge et al. (2011).

## CAPÍTULO 6

### CONSIDERAÇÕES FINAIS

Este trabalho apresentou um método baseado em técnicas de reconhecimento de padrões que tem por objetivo identificar seis tipos de padrões radiológicos em blocos de imagens de TCAR. Esses padrões são das classes: não pulmonar, tecido pulmonar normal, enfisema, vidro-fosco, fibrose e micronódulos.

Na Introdução, foi possível tomar conhecimento sobre a motivação para a realização deste trabalho, bem como os objetivos do mesmo. A Fundamentação Teórica abrangeu os conceitos que envolvem este trabalho, necessários para seu entendimento. Na Revisão Bibliográfica, foram apresentados trabalhos recentes da literatura que visam reconhecer padrões de alterações pulmonares.

O método proposto, descrito no Capítulo 4, avaliou seis extratores de características, com diversas variações, sendo que cada característica foi submetida a um classificador SVM. No total, 15 classificadores SVM foram avaliados individualmente, em que cada um gerou seus resultados na forma de probabilidades *a posteriori*. Em busca de uma forma de complementaridade entre os resultados, as respostas dos classificadores individuais foram combinadas utilizando métodos de fusão.

A base utilizada nos experimentos é composta por 20.540 blocos de imagens que pertencem a um dos seis padrões radiológicos, e representam os aspectos visuais de cada padrão. Essa base de imagens é pública, sendo disponibilizada por Depeursinge et al. (2011), permitindo assim a replicação do método. É importante lembrar que os experimentos não utilizaram imagens de um mesmo paciente nos grupos de treino e teste simultaneamente.

Os resultados experimentais mostraram que o melhor classificador individual é o  $CLBP_{(8,2)}^{riu2}$ , dentre os 15 classificadores avaliados. Já na combinação de classificadores, o conjunto  $CLBP_{(8,2)}^{riu2}$  e transformação *top-hat* atingiu as maiores taxas de reconheci-

mento. Isso indica que os classificadores em questão possuem informação complementares um ao outro.

Os resultados da melhor combinação de classificadores comparam-se aos da literatura para os padrões normal, fibrose e micronódulos. O padrão micronódulos obteve a melhor taxa de reconhecimento dentre os trabalhos da literatura que utilizam a mesma base de imagens. No entanto, classes como enfisema e vidro-fosco obtiveram os piores resultados entre os trabalhos. Ademais, o padrão não-pulmão obteve 99,2% de sensibilidade, mostrando que seria possível classificar as regiões externas ao pulmão em imagens de TCAR, de forma semelhante a uma segmentação. O método apresenta desvantagens, pois realiza muitas classificações incorretas entre os pares enfisema e normal, vidro-fosco e normal, e vidro-fosco e fibrose. Se comparado aos trabalhos da literatura discutidos no Capítulo 3, a combinação de classificadores obteve maior sensibilidade que o apresentado por Li et al. (2013) e Depeursinge et al. (2012), e maior acurácia que a obtida por Park et al. (2009), Lim et al. (2011) e Chang et al. (2013).

Como já explicado na Seção 2.3 deste trabalho, o auxílio ao diagnóstico se mostra importante não apenas para aprimorar a precisão do diagnóstico, como também a consistência da interpretação da imagem. A relevância dos resultados obtidos pode ser percebida observando estudos que tem por objetivo avaliar o desempenho de radiologistas com e sem o uso de sistemas CAD. Em um estudo publicado por Brown et al. (2005), observou-se que o auxílio de sistemas CAD melhoram as taxas de detecção de nódulos pulmonares. Médicos não radiologistas atingem uma média de 69,7% na taxa de detecção, e esse número é aumentado para 80,8% com o uso de sistemas CAD. Da mesma forma, radiologistas residentes têm esse índice melhorado de 77,4% para 85,5%. Esse progresso também é observado para radiologistas do tórax (82,5% sem auxílio, e 88,9% com auxílio de sistemas CAD). O desempenho dos radiologistas pode variar muito, dependendo do grau de especialização e de treinamento. No entanto, em todos os casos, há um progresso na detecção dos nódulos. Com base nisso, e nos resultados obtidos, pode-se considerar inserir o método desenvolvido em uma ferramenta de auxílio ao diagnóstico, bastando integrá-lo aos demais procedimentos que compõem o desenvolvimento de um sistema CAD.

Por fim, os resultados preliminares obtidos por meio deste trabalho foram publicados no 28th IEEE International Symposium on Computer-Based Medical Systems, com o título “Blockwise Classification of Lung Patterns in Unsegmented CT Images”, tendo como autores Luiza Dri Bagesteiro, Lucas Ferrari de Oliveira e Daniel Weingaertner. O artigo em questão apresenta os resultados da família de descritores CLBP para os seis padrões radiológicos apresentados neste trabalho.

Os materiais utilizados no desenvolvimento do método podem ser encontrados em <http://web.inf.ufpr.br/vri/alumni/2015-LuizaDriBagesteiro-Msc>.

## 6.1 Trabalhos futuros

A realização desta pesquisa, bem como seus resultados, podem ser utilizadas em trabalhos futuros como, por exemplo, a criação de uma ferramenta de auxílio ao diagnóstico. Nesta seção, sugerem-se algumas atividades que tornariam o trabalho mais completo:

- **Hierarquia de classificadores:** uma classificação hierárquica pode ser feita ao treinar um classificador binário em cada nodo da hierarquia. A ideia é que a hierarquia comece por identificar os padrões radiológicos mais discrepantes, e termine por distinguir os pares de classes similares;
- **Fusão dos classificadores:** cada classificador individual gerou probabilidades *a posteriori* que podem ser combinadas de outras formas além das apresentadas neste trabalho. O voto majoritário ponderado, por exemplo, é uma técnica de fusão que acrescenta pesos aos classificadores, de acordo com a opinião dos mesmos. Além disso, técnicas de seleção dinâmica de classificadores também podem ser estudadas;
- **Seleção dinâmica de classificadores:** um exemplo de abordagem dinâmica de seleção de classificadores é a utilização do oráculo, que atribui a classe correta para uma amostra se existe pelo menos um classificador do conjunto que acerta a amostra. Dessa forma, é possível determinar os classificadores que melhor classificam cada padrão (Vriesmann, 2012);

- **Pré-segmentação dos pulmões:** a alta taxa de reconhecimento obtida para a classe não pulmonar abre a possibilidade de testar o método como um pré-segmentador, ou seja, identificando na imagem de TCAR todas as regiões que não são de tecido pulmonar;
- **Quantificação da imagem pulmonar:** consiste em analisar as regiões da imagem, identificando os padrões radiológicos presentes e calculando a porcentagem de ocorrência de cada padrão. Nesse âmbito, algumas técnicas podem ser adotadas. Pode-se incluir informações de vizinhança ao investigar o tipo de padrão pulmonar de determinada região, bem como informações da distribuição das anormalidades no pulmão no plano axial.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Ahonen, T., Matas, J., He, C., and Pietikäinen, M. (2009). Rotation Invariant Image Description with Local Binary Pattern Histogram Fourier Features. In *Image Analysis*, volume 5575 of *Lecture Notes in Computer Science*, pages 61–70. Springer Berlin Heidelberg.
- Anthimopoulos, M., Christodoulidis, S., Christe, A., and Mougiakakou, S. (2014). Classification of interstitial lung disease patterns using local DCT features and random forest. In *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6040–6043.
- Arzhaeva, Y., Tax, D., van Ginneken, B., et al. (2006). Improving computer-aided diagnosis of interstitial disease in chest radiographs by combining one-class and two-class classifiers. *Proceedings of SPIE*, 6144:614458–614458–8.
- Azevedo-Marques, P. M. d. (2001). Diagnóstico auxiliado por computador na radiologia. *Radiologia Brasileira*, 34:285–293.
- Bagci, U., Bray, M., Caban, J., Yao, J., and Mollura, D. J. (2012). Computer-assisted detection of infectious lung diseases: A review. *Computerized Medical Imaging and Graphics*, 36(1):72–84.
- Brown, M. S. et al. (2005). Computer-aided Lung Nodule Detection in CT. *Academic Radiology*, 12(6):681–686.
- Caritá, E. C., Mendes, M. A. L., and de, A.-M. P. M. (2004). Ferramentas para visualização de imagens médicas em hospital universitário. *Radiologia Brasileira*, 37:437–440.
- Chang, C.-C. and Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27. Software disponível em <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

- Chang, Y., Kim, N., Lee, Y., Lim, J., Seo, J. B., and Lee, Y. K. (2012). Fast and efficient lung disease classification using hierarchical one-against-all support vector machine and cost-sensitive feature selection. *Computers in Biology and Medicine*, 42(12):1157–1164.
- Chang, Y., Lim, J., Kim, N., Seo, J. B., and Lynch, D. A. (2013). A support vector machine classifier reduces interscanner variation in the HRCT classification of regional disease pattern in diffuse lung disease: Comparison to a Bayesian classifier. *Medical Physics*, 40(5):051912.
- Cortes, C. and Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20(3):273–297.
- Costa, Y. M. (2013). *Reconhecimento de Gêneros Musicais Utilizando Espectrogramas com Combinação de Classificadores*. PhD thesis, Universidade Federal do Paraná.
- Dash, J., Mukhopadhyay, S., Garg, M., Prabhakar, N., and Khandelwal, N. (2014). Multi-classifier framework for lung tissue classification. In *IEEE Students' Technology Symposium (TechSym)*, pages 264–269.
- Depeursinge, A. (2010). *Affine-invariant texture analysis and retrieval of 3D medical images with clinical context integration*. PhD thesis, Universidade de Genebra.
- Depeursinge, A., Iavindrasana, J., Hidki, A., Cohen, G., Geissbühler, A., Platon, A., Poletti, P.-A., Müller, H., et al. (2010). Comparative Performance Analysis of State-of-the-Art Classification Algorithms Applied to Lung Tissue Categorization. *Journal of Digital Imaging*, 23(1):18–30.
- Depeursinge, A., Van de Ville, D., Platon, A., Geissbuhler, A., Poletti, P.-A., and Müller, H. (2012). Near-Affine-Invariant Texture Learning for Lung Tissue Analysis Using Isotropic Wavelet Frames. *IEEE Transactions on Information Technology in Biomedicine*.
- Depeursinge, A., Vargas, A., Platon, A. a. A., Poletti, P.-A., and Muller, H. (2011).

- Building a Reference Multimedia Database for Interstitial Lung Diseases. *Computerized Medical Imaging and Graphics*.
- Doi, K. (1999). Computer-Aided Diagnosis and its Potential Impact on Diagnostic Radiology. *Computer-Aided Diagnosis in Medical Imaging*, page 489.
- Doi, K. (2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, 31(4-5):198–211.
- Dougherty, E. R. and Lotufo, R. A. (2003). *Hands-On Morphological Image Processing*. SPIE.
- Elicker, B., de Castro, P. C. A., Richard, W., and O., L. K. (2008). High-resolution computed tomography patterns of diffuse interstitial lung disease with clinical and pathological correlation. *Jornal Brasileiro de Pneumologia*, 34:715–744.
- Facon, J. (1996). *Morfologia Matemática: Teoria e exemplos*, volume 10. Editora Universitária Champagnat da Pontifícia Universidade Católica do Paraná, Curitiba, 1 edition.
- Flokstra, J., Drukker, C., and Kortland, K. (2014). Computed Tomography: Scientific Technological fact sheet. Disponível em: [http://www.fisme.science.uu.nl/pos/MI\\\_2\\\_3\\\_2](http://www.fisme.science.uu.nl/pos/MI\_2\_3\_2). Acesso em mai. 2014.
- Gangeh, M., Sørensen, L., Shaker, S., Kamel, M., de Bruijne, M., and Loog, M. (2010). A Texton-Based Approach for the Classification of Lung Parenchyma in CT Images. In *Medical Image Computing & Computer-Assisted Intervention*, volume 6363 of *Lecture Notes in Computer Science*, pages 595–602. Springer Berlin Heidelberg.
- Gonzalez, R. C. and Woods, R. E. (2002). *Digital Image Processing*. Pearson, Boston, MA, USA, 2nd edition.
- Guo, Z., Zhang, D., and Zhang, D. (2010). A Completed Modeling of Local Binary Pattern Operator for Texture Classification. *IEEE Transactions on Image Processing*, 19(6):1657–1663.

- Hall-Beyer, M. (2015). Properties of the GLCM - The GLCM Tutorial Home Page. University of Calgary. Disponível em: <http://www.ucalgary.ca/mhallbey/Properties>. Acesso em mai. 2015.
- Hansell, D. M., Bankier, A. A., MacMahon, H., McLoud, T. C., Müller, N. L., and Remy, J. (2008). Fleischner Society: Glossary of Terms for Thoracic Imaging. *Radiology*, 246(3):697–722.
- Haralick, R. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786–804.
- Haralick, R., Shanmugam, K., and Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6):610–621.
- Heikkilä, M., Pietikäinen, M., and Schmid, C. (2009). Description of Interest Regions with Local Binary Patterns. *Pattern Recognition*, 42(3):425–436.
- Jain, A., Duin, R., and Mao, J. (2000). Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4–37.
- Kim, N., Seo, J., Lee, Y., Lee, J., Kim, S., and 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.
- Kim, N., Seo, J. B., Sung, Y. S., Park, B.-W., Lee, Y., Park, S. H., Lee, Y. K., and Kang, S.-H. (2008). Effect of various binning methods and ROI sizes on the accuracy of the automatic classification system for differentiation between diffuse infiltrative lung diseases on the basis of texture features at HRCT. *Proceedings of SPIE*, 6914:69143N–69143N–7.
- Kittler, J., Hatef, M., Duin, R. P. W., and Matas, J. (1998). On Combining Classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239.

- Knerr, S., Personnaz, L., and Dreyfus, G. (1990). Single-layer learning revisited: a step-wise procedure for building and training a neural network. In Soulié, Françoise Fogelman and Héroult, Jeanny, editor, *Neurocomputing*, volume 68 of *NATO ASI Series*, pages 41–50. Springer Berlin Heidelberg.
- Korfiatis, P., Skiadopoulos, S., Sakellaropoulos, P., Kalogeropoulou, C., and Costaridou, L. (2007). Automated 3D Segmentation of Lung Fields in Thin Slice CT Exploiting Wavelet Preprocessing. In *Proceedings of the 12th International Conference on Computer Analysis of Images and Patterns, CAIP'07*, pages 237–244, Berlin, Heidelberg. Springer-Verlag.
- Li, Q., Cai, W., and Feng, D. D. (2013). Lung image patch classification with automatic feature learning. pages 6079–6082.
- Lim, J., Namkug, Seo, J. B., Lee, Y. K., Lee, Y., and Kang, S.-H. (2011). Regional Context-Sensitive Support Vector Machine Classifier to Improve Automated Identification of Regional Patterns of Diffuse Interstitial Lung Disease. *Journal of Digital Imaging*, 24(6):1133–1140.
- Malone, J., Rossiter, J., Prabhu, S., and Goddard, P. (2004). Identification of disease in CT of the lung using texture-based image analysis. In *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 1620–1624.
- Marques Filho, O. and Vieira Neto, H. (1999). *Processamento Digital de Imagens*. Editora Brasport, Rio de Janeiro, Brazil.
- Martinez, J. A. B. (1998). Doenças intersticiais pulmonares. *Revista MEDICINA-Ribeirão Preto*, 31(2):247–256.
- Mäenpää, T., Ojala, T., Pietikäinen, M., and Soriano, M. (2000). Robust texture classification by subsets of local binary patterns. In *Proceedings of the 15th International Conference on Pattern Recognition*, volume 3, pages 935–938 vol.3.

- Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51–59.
- Ojala, T., Pietikäinen, M., and Mäenpää, T. (2002). Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987.
- Ojansivu, V. and Heikkilä, J. (2008). Blur Insensitive Texture Classification Using Local Phase Quantization. In *Proceedings of the 3rd International Conference on Image and Signal Processing, ICISP '08*, pages 236–243, Berlin, Heidelberg. Springer-Verlag.
- Park, S. O., Seo, J. B., Kim, N., Park, S. H., Lee, Y. K., Park, B.-W., Sung, Y. S., Lee, Y., Lee, J., and Kang, S.-H. (2009). Feasibility of Automated Quantification of Regional Disease Patterns Depicted on High-Resolution Computed Tomography in Patients with Various Diffuse Lung Diseases. *Korean Journal of Radiology*, 10(5):455–463.
- Paula-Filho, P. L. (2012). *Reconhecimento de Espécies Florestais Através de Imagens Macroscópicas*. PhD thesis, Universidade Federal do Paraná.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Pereyra, L., Rangayyan, R., Ponciano-Silva, M., and Azevedo-Marques, P. (2014). Fractal analysis for computer-aided diagnosis of diffuse pulmonary diseases in HRCT images. In *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 1–6.
- Pianykh, O. S. (2012). *Digital Imaging and Communications in Medicine (DICOM) - A Practical Introduction and Survival Guide*. Springer Berlin Heidelberg.

- Platt, J. C. (1999). Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. In *Advances in Large Margin Classifiers*, pages 61–74. MIT Press.
- Ponciano-Silva, M. (2009). Processamento de consultas por similaridade em imagens médicas visando à recuperação perceptual guiada pelo usuário. Dissertação de mestrado, Instituto de Ciências Matemáticas e de Computação - USP.
- Rahtu, E., Heikkilä, J., Ojansivu, V., and Ahonen, T. (2012). Local Phase Quantization for Blur-insensitive Image Analysis. *Image and Vision Computing*, 30(8):501–512. Special Section: Opinion Papers.
- Rubin, A. S. et al. (2012). Diretrizes de Doenças Pulmonares Intersticiais da Sociedade Brasileira de Pneumologia e Tisiologia. *Jornal Brasileiro de Pneumologia*, 38(2).
- Sluimer, I., Schilham, A. M. R., Prokop, M., and van Ginneken, B. (2006). Computer analysis of computed tomography scans of the lung: a survey. *IEEE Transactions on Medical Imaging*, 25(4):385–405.
- Sluimer, I. C., Niemeijer, M., and van Ginneken, B. (2004). Lung field segmentation from thin-slice CT scans in presence of severe pathology. In *Proceedings of SPIE*, volume 5370, pages 1447–1455.
- Soille, P. (2003). *Morphological Image Analysis: Principles and Applications*. Springer, Heidelberg, 2. ed. edition.
- Song, Y., Cai, W., Zhou, Y., and Feng, D. D. (2013). Feature-Based Image Patch Approximation for Lung Tissue Classification. *IEEE Transactions on Medical Imaging*, 32(4):797–808.
- Souza Jr., A. S., de, A. N. C., Dany, J., Edson, M., Jorge, K., L., I. K., Felipe, N. L., Marcelo, F., Marcelo, C., Mário, T. F., and Pedro, D. (2002). Terminologia para a descrição de tomografia computadorizada do tórax: Sugestões iniciais para um consenso brasileiro. *Radiologia Brasileira*, 35:125–128.

- Sørensen, L., Shaker, S., and de Bruijne, M. (2010). Quantitative Analysis of Pulmonary Emphysema Using Local Binary Patterns. *IEEE Transactions on Medical Imaging*, 29(2):559–569.
- Sørensen, L., Shaker, S. B., and Bruijne, M. (2008). Texture Classification in Lung CT Using Local Binary Patterns. In *Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention - Part I, MICCAI '08*, pages 934–941, Berlin, Heidelberg. Springer-Verlag.
- Tawhai, M. H., Nash, M. P., Lin, C.-L., and Hoffman, E. A. (2009). Supine and prone differences in regional lung density and pleural pressure gradients in the human lung with constant shape. *Journal of Applied Physiology*, 107(3):912–920.
- Theodoridis, S. and Koutroumbas, K. (2008). *Pattern Recognition*. Academic Press, 4th edition.
- Vapnik, V. (1982). *Estimation of Dependences Based on Empirical Data: Springer Series in Statistics (Springer Series in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- Vapnik, V. N. (1998). *Statistical learning theory*. Wiley, 1 edition.
- Vo, K. and Sowmya, A. (2010). Multiple kernel learning for classification of diffuse lung disease using HRCT lung images. In *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3085–3088.
- Vriesmann, L. M. (2012). *Seleção Dinâmica de Subconjuntos de Classificadores: Abordagem Baseada em Acurácia Local*. PhD thesis, Pontifícia Universidade Católica do Paraná.
- Wang, L. and He, D.-C. (1990). Texture classification using texture spectrum. *Pattern Recognition*, 23(8):905–910.
- Wu, T.-F., Lin, C.-J., and Weng, R. C. (2004). Probability Estimates for Multi-class

Classification by Pairwise Coupling. *Journal of Machine Learning Research*, 5:975–1005.

Zhao, W., Xu, R., Hirano, Y., Tachibana, R., and Kido, S. (2013). Classification of diffuse lung diseases patterns by a sparse representation based method on HRCT images. In *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5457–5460.

LUIZA DRI BAGESTEIRO

**CLASSIFICAÇÃO DE PADRÕES RADIOLÓGICOS POR  
BLOCOS EM IMAGENS NÃO SEGMENTADAS DE  
TOMOGRAFIA COMPUTADORIZADA**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, Universidade Federal do Paraná.

Orientador: Daniel Weingaertner

Co-orientador: Lucas Ferrari de Oliveira

CURITIBA

2015