UNIVERSIDADE DE SÃO PAULO ESCOLA DE ENGENHARIA DE SÃO CARLOS FACULDADE DE MEDICINA DE RIBEIRÃO PRETO INSTITUTO DE QUÍMICA DE SÃO CARLOS

José Raniery Ferreira Junior

Arcabouço para Classificação, Recuperação por Conteúdo e Radiômica de Imagens Médicas: uma investigação de biomarcadores quantitativos para o câncer de pulmão

Arcabouço para Classificação, Recuperação por Conteúdo e Radiômica de Imagens Médicas: uma investigação de biomarcadores quantitativos para o câncer de pulmão

Tese apresentada ao Programa de Pós-Graduação Interunidades em Bioengenharia da Escola de Engenharia de São Carlos – Faculdade de Medicina de Ribeirão Preto e Instituto de Química de São Carlos da Universidade de São Paulo, como requisito para a obtenção do Título de Doutor em Ciências.

Área de concentração: Bioengenharia

Orientador: Prof. Dr. Paulo Mazzoncini de Azevedo Marques Coorientador: Prof. Dr. Marcel Koenigkam Santos

VERSÃO CORRIGIDA

São Carlos 2019

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da EESC/USP com os dados inseridos pelo(a) autor(a).

Ferreira-Junior, José Raniery F383a Arcabouço para Classificação, Recuperação por Conteúdo e Radiômica de Imagens Médicas: uma investigação de biomarcadores quantitativos para o câncer de pulmão / José Raniery Ferreira-Junior; orientador Paulo Mazzoncini Azevedo-Marques; coorientador Marcel Koenigkam-Santos. São Carlos, 2019.

> Tese (Doutorado) - Programa de Pós-Graduação Interunidades em Bioengenharia e Área de Concentração em Bioengenharia -- Escola de Engenharia de São Carlos; Faculdade de Medicina de Ribeirão Preto; Instituto de Química de São Carlos, da Universidade de São Paulo, 2019.

1. Câncer de pulmão. 2. Radiômica. 3. Reconhecimento de padrões. 4. Classificação de imagens médicas. 5. Recuperação de imagens baseada em conteúdo. Eduardo Graziosi Silva - CRB - 8/8907

FOLHA DE JULGAMENTO

Candidato(a): José Raniery Ferreira Junior

Título: "Arcabouço para Classificação, Recuperação por Conteúdo e Radiômica de Imagens Médicas: uma investigação de biomarcadores quantitativos para o câncer de pulmão"

Data da defesa: 19/03/2019

Comissão Julgadora

Prof(a). Dr(a). Paulo Mazzoncini de Azevedo Marques FMRP/USP

Prof(a). Dr(a). Agma Juci Machado Traina ICMC/USP

Prof(a). Dr(a). Juliana Pereira de Souza Zinader UFG

Prof(a). Dr(a). Edson Amaro Junior FM/USP

Prof(a). Dr(a). Alexandre Todorivic Fabro FMRP/USP

Prof(a). Dr(a). Letícia Rittner UNICAMP

rain

Assinatura

APROVADO

Resultado

Não

Votante

Aos meus pais, Raniery e Magal, pelo amor e pela educação que sempre me deram. À minha esposa, Priscilla, por sua "paxão" e compreensão em todos os momentos.

AGRADECIMENTOS

À toda a minha família, em especial, meus irmãos, Bruno e Lígia, meus sogros, Junior e Wirlene, e meus cunhados, Anderson e Deysi, pelo apoio dado durante todo o caminho de pós-graduação percorrido.

Aos meus orientadores, Paulo, Marcel e Yoshida, pelas oportunidades dadas, pelos ensinamentos e, principalmente, pela paciência.

Aos doutores Marcello, Tinós, Alexandre e Federico por todas as valiosas contribuições dadas durante meu período na USP.

Aos amigos do LaTIM/UFAL, Marcelo, Lucas, Ailton e Tércio, pelas sugestões e discussões sempre produtivas para a pesquisa.

Aos amigos do CCIFM, Hilton, Rafa, Lucas, Victor, Matheus, Eddy, Natália, Rodolfo, Julio, Jaque, Gui, Rodrigo, Camila, Danilo, Beth, Luciana, Ângela e D. Elza, pelos bons momentos de convivência e todo o suporte dado para o desenvolvimento do trabalho.

Aos amigos do MGH, Yasuka, Toru e Janne, por todo o auxílio prestado durante meu estágio no exterior de doutorado sanduíche.

Ao pessoal da Bioengenharia, Profa Ana Plepis, Marcia e Janete, pela ajuda nos processos burocráticos desde a seleção do doutorado até o depósito da tese.

Aos amigos de SP que fiz ou reencontrei, Diógenes, Felipe, Amanda, Samara, Karina, Juninho, Inaê, Danilo e D. Miralda, pela amizade, incentivo e bons momentos compartilhados.

Às instituições e programas CAPES (Código de Financiamento 001), USP, PAE, PDSE (processo nº 88881.134004/2016-01), AUCANI, PROAP e FAPESP (processo nº 2016/17078-0), pelas bolsas e apoios financeiros durante todo o curso e pesquisa.

E a todos aqueles que me ajudaram direta ou indiretamente, torceram por mim e acreditaram no meu trabalho.

"No meio da dificuldade encontra-se a oportunidade." Albert Einstein

RESUMO

FERREIRA JUNIOR, J.R. Arcabouço para Classificação, Recuperação por Conteúdo e Radiômica de Imagens Médicas: uma investigação de biomarcadores quantitativos para o câncer de pulmão. 2019. 183p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2019.

Sistemas de diagnóstico auxiliado por computador (do Inglês computer-aided diagnosis/ detection - CAD), recuperação de imagens baseada em conteúdo (do Inglês content-based *image retrieval* - CBIR) e radiômica (do Inglês *radiomics*) vêm sendo desenvolvidos com o objetivo de melhorar a acurácia diagnóstica, a consistência na interpretação de imagens médicas, a avaliação prognóstica do paciente e o suporte à decisão terapêutica. Estas ferramentas possuem grande potencial em reconhecer objetiva e precisamente padrões em imagens, porém ainda possuem diversas limitações. Este trabalho tentou diminuir as limitações de característica, referente à representação numérica ineficiente das imagens, e de desempenho, referente à qualidade baixa de integração sistemática, que estas ferramentas podem possuir. Alguns conjuntos de imagens de tomografia computadorizada de neoplasias pulmonares foram utilizados no desenvolvimento desta pesquisa retrospectiva. Modelos semiautomáticos de reconhecimento de padrões, utilizando segmentação por crescimento de região volumétrica, extração de atributos calibrados de imagens, classificação e recuperação por conteúdo, avaliação prognóstica e rede neural convolucional, foram desenvolvidos para compor um arcabouço de análise quantitativa de imagens médicas. Os modelos de CAD, CBIR e radiômica mostraram potencial de integração com sistemas de informação em radiologia e hospitalar para a rotina clínica. Este trabalho também desenvolveu e identificou diferentes variáveis quantitativas para a avaliação diagnóstica e prognóstica de diferentes doenças. O arcabouço pode proporcionar diversos biomarcadores computacionais de imagens e mais informações importantes para os especialistas, potencialmente auxiliando as tomadas de decisões radiológicas. Por fim, esta tese contribuiu para a identificação de uma assinatura radiômica de avaliação diagnóstica e prognóstica do câncer de pulmão para a estratificação de risco do paciente.

Palavras-chave: Câncer de pulmão. Radiômica. Reconhecimento de padrões. Classificação de imagens médicas. Recuperação de imagens baseada em conteúdo. Análise de imagens médicas.

ABSTRACT

FERREIRA JUNIOR, J.R. Framework for Classification, Content-Based Retrieval, and Radiomics of Medical Images: an investigation of quantitative biomarkers for lung cancer. 2019. 183p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2019.

Systems of computer-aided diagnosis (CAD), content-based image retrieval (CBIR), and radiomics have been developed to increase diagnostic accuracy, medical image interpretation consistency, patient prognostic assessment, and therapy decision support. Those tools have great potential to recognize objectively and precisely patterns in images, despite still having several limitations. This work aimed to decrease the feature gap, which refers to inefficient numeric representation of images, and the performance gap, which refers to low quality of systems integration, that those tools may have. Some datasets of computed tomography images of lung neoplasms were used during the development of this retrospective study. Semiautomatic models of pattern recognition, using volumetric region growing segmentation, fine-tuned image feature extraction, content-based retrieval and classification, prognosis evaluation, and convolutional neural network, were developed to build a medical image analysis framework. CAD, CBIR, and radiomics models disclosed potential to integrate with radiology and hospital information systems in clinical routine. This work also developed and identified different quantitative variables for diagnostic and prognostic assessment of different pathologies. The framework may provide several computational image biomarkers and more important information to specialists, potentially supporting imaging decision making. Finally, this thesis contributed to the identification of a radiomics signature to assess lung cancer diagnosis and prognosis for patient risk stratification.

Keywords: Lung cancer. Radiomics. Pattern recognition. Medical image classification. Content-based image retrieval. Medical image analysis.

LISTA DE FIGURAS

Figura	1 –	Detecção de tumores em estágio inicial em imagens de tomografia	
		computadorizada: (a-d) apresentam os nódulos e quantos radiologistas	
		de um total de 13 foram capazes de identificar estas lesões	22
Figura	2 -	Subtipos histopatológicos mais comuns do câncer de pulmão em imagens	
		de tomografia computadorizada: (a) ADC, (b) SCC, (c) LCC, (d) SCLC.	
		As setas indicam a localização do tumor	24
Figura	3 -	Exemplo de avaliação visual subjetiva dos contornos de um tumor: (a),	
		(b) e (c) apresentam tumores sem contornos espiculados, com contor-	
		nos finamente espiculados e com contornos grosseiramente espiculados,	
		respectivamente. As setas indicam a localização do tumor	24
Figura	4 -	Possíveis associações para a radiômica	26
Figura	5 -	Fontes de <i>Big Data</i> para geração de novas informações clínicas	27
Figura	6 –	Quantificação da heterogeneidade intratumoral. Os eixos $x \in y$ no gráfico	
		se referem a posições espaciais em coordenadas cartesianas. A escala de	
		cores se refere a um atributo extraído da imagem refletindo as subregiões	
		do tumor	28
Figura	7 –	Representação da função de uma imagem digital em escala de cinza	33
Figura	8 -	Representação de um <i>voxel.</i>	34
Figura	9 -	Aplicações de janelamento em TC de tórax: (a) janela de pulmão com	
		largura de 1.400 UH e nível de -500 UH, (b) janela de mediastino com	
		largura de 350 UH e nível de 40 UH, (c) janela de osso com largura de	
		1.000 UH e nível de 400 UH	35
Figura	10 -	Aplicação de segmentação do pulmão em imagem de TC de tórax com	
		256 níveis de cinza e limiar 115: (a) imagem de TC original, (b) resultado	
		da limiarização, (c) resultado da binarização. $\ldots \ldots \ldots \ldots \ldots \ldots$	36
Figura	11 –	Aplicação de segmentação com crescimento de região de tumor em	
		imagem de TC: (a) imagem de TC original com a semente marcada, (b)	
		estágio primário de crescimento de região, (c) estágio intermediário de	
		crescimento de região, (d) estágio final de crescimento de região. $.$	37
Figura	12 –	Processo de extração de n atributos de uma imagem de TC	38
Figura	13 –	Articulações sacroilíacas em imagens de ressonância magnética em que	
		os níveis de cinza podem ser úteis para o reconhecimento de padrões	
		inflamatórios: (a) articulações sem inflamação, (b) articulações com	
		inflamação.	39
Figura	14 -	Exemplo de histograma de uma imagem de TC com 256 níveis de cinza.	39

Figura 15 -	- Fraturas de corpos vertebrais em imagens segmentadas de ressonância	
	magnética em que a textura pode ser útil para o reconhecimento da	
	etiologia de uma fratura como sendo associada a presença de osteopo-	
	rose (benigna) ou de um câncer (maligna) na coluna: linhas (a) e (b)	
	apresentam fraturas benignas e malignas, respectivamente	41
Figura 16 –	- Aplicação do filtro Laplaciano de Gaussiano em uma imagem de TC de	
	tumor pulmonar: (a) e (b) apresentam as imagens original e filtrada,	
	respectivamente.	46
Figura 17 -	- Aplicação da transformada de Fourier em uma imagem de TC: (a) e (b)	
	apresentam as imagens original e resultante no domínio da frequência,	
	respectivamente.	47
Figura 18 –	- Exemplo de um banco de filtros de Gabor em diferentes escalas e	
	orientações.	48
Figura 19 –	- Heterogeneidade morfológica de tumores pulmonares	50
Figura 20 –	- Segmentação semi-automática de nódulos pulmonares subsólidos (mar-	
	cação em laranja) e de componentes sólidos (marcação em amarelo) em	
	imagens de TC: (a) e (b) apresentam casos em que as segmentações	
	apresentaram performance satisfatória e insatisfatória, respectivamente.	51
Figura 21 –	- Linha ortogonal traçada sobre a margem de nódulo pulmonar em imagem	
	de TC para cálculo de atributo de nitidez de borda	52
Figura 22 –	- Funcionamento de algoritmos de seleção de atributos relevantes: (a)	
	filtro, (b) wrapper, (c) embarcado	54
Figura 23 –	- Exemplo de classificação com KNN em espaço bidimensional. O caso	
	novo desconhecido será classificado como B por existirem 3 amostras	
	(dentre 5) da Classe B mais próximas. $\ldots \ldots \ldots \ldots \ldots \ldots \ldots$	56
Figura 24 –	- Mecanismo de criação de hiperplano do método SVM para classificação.	57
Figura 25 –	- Representação de hiperplano e margem obtidos pelo classificador SVM $$	
	em um problema linearmente separável	58
Figura 26 –	- Exemplo de árvore de decisão para duas classes A e B utilizando 3	
	atributos	58
Figura 27 –	- Arquitetura de uma rede neural artificial multicamadas	59
Figura 28 –	- Representação de fronteiras de decisão para redes neurais artificiais: (a)	
	MLP, (b) RBF	60
Figura 29 –	- Representações no ajuste dos dados a um modelo de aprendizado de	
	máquina: (a) subajuste, (b) apropriado, (c) sobrejuste	61
Figura 30 –	- Imagem de TC como entrada em uma rede neural convolucional	62

Figura 31 –	Exemplos de validação em aprendizagem de máquina: (a) <i>holdout</i> , (b) validação-cruzada com 5 subconjuntos, (c) aprendizagem com subcon- junto de validação, (d) aprendizagem com validação-cruzada e subcon-	
Figura 32 –	junto de validação	63
	área igual a 0,5	64
Figura 33 –	Curvas Kaplan-Meier para os grupos de pacientes da Tabela 4. Cada traco na curva do grupo Tratamento representa um paciente censurado	66
Figura 34 –	Representação dos conjuntos de casos na recuperação de imagens base-	00
Figura 35 –	ada em conteúdo	68
Figure 26	em CBIR: a <i>Função 1</i> apresenta maior eficiência do que a <i>Função 2</i> Exemple de méteo emerição (10) para explicit a performance em CBIP:	69
Figura 30 –	a <i>Função 2</i> apresenta maior eficiência do que a <i>Função 1</i>	70
Figura 37 –	Exemplos de lesões em imagens de TC de acordo com as categorias do LIDC: (a) nódulo \geq 3 mm, (b) nódulo < 3 mm, (c) não-nódulo \geq 3	
Figura 38 –	mm. As setas indicam a localização dos achados radiológicos Análise de nitidez de borda de nódulo pulmonar: (a) pontos de controle da borda, (b) segmentos de reta perpendiculares à borda, (c) máscara	72
	obtida por binarização e dilatação, (d) imagem resultante para extração de atributos.	75
Figura 39 –	Curvas ROC dos experimentos com o conjunto de amostras desbalance- ado: (a) todos os atributos, (b) atributos selecionados por estatística, (c)	
Figura 40 –	atributos selecionados por filtro, (d) atributos selecionados por <i>wrapper</i> . Curvas ROC dos experimentos com o conjunto de amostras balanceado:	79
	(a) todos os atributos, (b) atributos selecionados por estatística, (c) atributos selecionados por filtro, (d) atributos selecionados por <i>wrapper</i> .	81
Figura 41 –	Curvas ROC de cenários (número de atributos + classificador + conjunto de amostras) de classificação relevantes sem diferença estatisticamente	
	significativa na sensibilidade e especificidade	82
Figura 42 –	Eficiência na recuperação de nódulos pulmonares similares pela métrica	80
Figura 43 –	Eficiência na recuperação de nódulos pulmonares similares pela métrica	09
~	precisão vs. revocação com diferentes atributos de imagens	90

Figura 44 –	Segmentação semiautomática com o algoritmo GrowCut de tumor pul-
	monar em imagem de TC: (a) marcações internas e externas ao tumor,
	(b) resultado do crescimento de regiões, (c) remoção da região externa,
	(d) delimitação da borda da região interna
Figura 45 –	Extratores de atributos utilizados na radiômica de imagens de TC. $$. $$ 97
Figura 46 –	Conversão do sistema de coordenadas do paciente em milímetros (a)
	para o sistema de coordenadas da imagem no plano cartesiano (b) 98
Figura 47 –	Exemplos dos complexos padrões visuais tumorais em imagens de TC
	utilizados nos experimentos de radiômica do câncer de pulmão: (a)
	estadio M1, (b) estadio M0, (c) estadio N1–3, (d) estadio N0, (e) ADC,
	(f) SCC, (g) alto risco com ocorrência de óbito e sobrevida global de 55
	dias, (h) baixo risco sem ocorrência de óbito e sobrevida global de 2801
	dias
Figura 48 –	Atributos estatisticamente mais significativos associados com metástase
	distante. A significância estatística está apresentada no fim de cada barra. 109
Figura 49 –	Atributos estatisticamente mais significativos associados com metástase
	linfonodal. A significância estatística está apresentada no fim de cada
	barra
Figura 50 –	Atributos estatisticamente mais significativos associados com histopa-
C	tologia. A significância estatística está apresentada no fim de cada
	barra
Figura 51 –	Maiores performances individuais por categoria de atributo para a
0	associação com metástase distante
Figura 52 –	Maiores performances individuais por categoria de atributo para a
0	associação com metástase linfonodal
Figura 53 –	Maiores performances individuais por categoria de atributo para a
0	associação com histopatologia
Figura 54 –	Distribuição dos valores normalizados entre [0,1] de atributos significati-
0	vamente associados com metástase distante. Os atributos estão listados
	com nome, AUC e significância estatística. MD+ e MD- significam
	presenca e ausência de metástase distante, respectivamente
Figura 55 –	Distribuição dos valores normalizados entre [0,1] de atributos signifi-
0	cativamente associados com metástase linfonodal. Os atributos estão
	listados com nome. AUC e significância estatística. ML+ e ML- signifi-
	cam presenca e ausência de metástase linfonodal, respectivamente, 116
Figura 56 –	Distribuição dos valores normalizados entre [0,1] de atributos significati-
-0	vamente associados com histopatologia. Os atributos estão listados com
	nome. AUC e significância estatística

Visualização de dados por um sistema gerenciador para MongoDB. O	
Componente 1 ilustra as coleções do banco de dados desenvolvido e o	
Componente 2, um documento da coleção <i>nodules</i>	62
Arquitetura de uma CPU e uma GPU. ULA é a unidade lógica e	
aritmética	.63
Desempenho de execução da análise de similaridade de imagens em	
diferentes arquiteturas computacionais	64
Exemplos de cavidade (a) e pseudocavidade (b) de tumores pulmonares	
malignos em imagens de TC. As setas indicam a localização do achado	
radiológico	.66
Maiores performances estatísticas dos atributos para a distinção pseu-	
docavidade vs. cavidade. Os atributos de EIBSL, DPDPL, DPAL, ESIB	
e EIBSC são ênfase de intensidade baixa de sequência longa, desvio	
padrão do desvio padrão local, desvio padrão da amplitude local, ênfase	
de sequência de intensidade baixa e ênfase de intensidade baixa de	
sequência curta, respectivamente	.67
	Visualização de dados por um sistema gerenciador para MongoDB. O Componente 1 ilustra as coleções do banco de dados desenvolvido e o Componente 2, um documento da coleção <i>nodules</i>

LISTA DE TABELAS

Tabela 1 –	Variabilidade interobservador na avaliação visual de um nódulo pul-	
	monar: houve discordância em sete de oito características da lesão. Os	
	valores em negrito apresentam as discordâncias.	23
Tabela 2 –	Escala Hounsfield.	34
Tabela 3 –	Matriz de confusão.	64
Tabela 4 –	Exemplos de tempos de sobrevida para dois grupos de pacientes: um	
	grupo controle de pacientes sem terapia e um grupo que realizou algum	
	tratamento. Os pacientes censurados estão com um asterisco no tempo.	65
Tabela 5 –	Número de nódulos pulmonares utilizados no trabalho	73
Tabela 6 –	Significância estatística e ocorrência dos atributos na seleção de carac-	
	terísticas para classificação de imagens utilizando filtro e <i>wrapper</i>	77
Tabela 7 –	Performance de classificação com o conjunto de amostras desbalanceado.	78
Tabela 8 –	Performance de classificação com o conjunto de amostras balanceado	80
Tabela 9 –	Cenários onde ocorreu diferença estatisticamente significativa na sensi-	
	bilidade entre os conjuntos de amostras.	81
Tabela 10 –	Cenários de classificação sem diferença estatisticamente significativa na	
	sensibilidade e especificidade (intervalos de confiança de 95%)	82
Tabela 11 –	Significância estatística e número de ocorrências dos atributos na	
	validação-cruzada da seleção de características para recuperação de	
	imagens baseada em conteúdo.	88
Tabela 12 –	Dados clínicos do Conjunto1 de pacientes $(n = 52)$	94
Tabela 13 –	Dados clínicos do Conjunto2 de pacientes $(n = 16)$	95
Tabela 14 –	Dados clínicos do Conjunto3 de pacientes $(n = 85)$	95
Tabela 15 –	Performance da associação radiômica com metástase distante. AUCs	
	são apresentadas no formato $x \mid y$, onde x é o desempenho utilizando o	
	conjunto de validação e \boldsymbol{y} utilizando o conjunto de teste. Algumas das	
	maiores performances estão com valor em negrito	104
Tabela 16 –	Matrizes de confusão das maiores performances associativas para me-	
	tástase distante utilizando o conjunto de pacientes para somente testes.	
	Cenários são apresentados com o tipo de característica, classificador e	
	conjunto para treinamento utilizados	104
Tabela 17 –	Performance da associação radiômica com metástase linfonodal. AUCs	
	são apresentadas no formato $x \mid y,$ onde x é o desempenho utilizando o	
	conjunto de validação e \boldsymbol{y} utilizando o conjunto de teste. Algumas das	
	maiores performances estão com valor em negrito	105

Tabela 18 –	Matrizes de confusão das maiores performances associativas para metás-	
	tase linfonodal utilizando o conjunto de pacientes para somente testes.	
	Cenários são apresentados com o tipo de característica, classificador e	
	conjunto para treinamento utilizados	105
Tabela 19 –	Performance da associação radiômica com histopatologia. AUCs são	
	apresentadas no formato $x \mid y$, onde $x \notin o$ desempenho utilizando o	
	conjunto de validação e y utilizando o conjunto de teste. Algumas das	
	maiores performances estão com valor em negrito	106
Tabela 20 –	Matrizes de confusão das maiores performances associativas para his-	
	topatologia utilizando o conjunto de pacientes para somente testes.	
	Cenários são apresentados com o tipo de característica, classificador e	
	conjunto para treinamento utilizados	107
Tabela 21 –	Maiores AUCs de cada experimento da análise multivariada para a	
	associação com metástase distante	118
Tabela 22 –	Maiores AUCs de cada experimento da análise multivariada para a	
	associação com metástase linfonodal	118
Tabela 23 –	Maiores AUCs de cada experimento da análise multivariada para a	
	associação com histopatologia	119
Tabela 24 –	Características selecionadas no cenário de maior performance da análise	
	multivariada para a associação com metástase distante	119
Tabela 25 –	Características selecionadas no cenário de maior performance da análise	
	multivariada para a associação com metástas e linfonodal	120
Tabela 26 –	Características selecionadas no cenário de maior performance da análise	
	multivariada para a associação com histopatologia.	121
Tabela 27 –	Avaliação volumétrica dos tumores baseada na classificação por corte.	
	Em negrito estão os maiores valores por métrica e por tipo de imagem	
	utilizada.	128
Tabela 28 –	Comparação entre diferentes números de iterações no treinamento da	
	rede neural convolucional. Em negrito estão os maiores valores por	
	métrica e por número de iterações.	129

LISTA DE ABREVIATURAS E SIGLAS

ADC	Adenocarcinoma
ALK	do Inglês Anaplastic Lymphoma Kinase
API	do Inglês Application Programming Interface
AUC	do Inglês Area Under the (ROC) Curve
CAD	do Inglês Computer-Aided Diagnosis
CBIR	do Inglês Content-Based Image Retrieval
CNN	do Inglês Convolutional Neural Network
COM	do Inglês Cooccurrence Matrix
CPU	do Inglês Central Processing Unit
DICOM	do Inglês Digital Imaging and Communications in Medicine
DICOM-RT	do Inglês Digital Imaging and Communications in Medicine Radiation Therapy
EGFR	do Inglês Epidermal Growth Factor Receptor
GPU	do Inglês Graphics Processing Unit
HIS	do Inglês Hospital Information System
IBSI	do Inglês Image Biomarker Standardisation Initiative
IC	Intervalo de Confiança
JSON	do Inglês JavaScript Object Notation
KNN	do Inglês K-Nearest Neighbors
LCC	do Inglês Large Cell Carcinoma
LIDC	do Inglês Lung Image Database Consortium
LOG	do Inglês Laplacian of Gaussian
MDI	Momento da Diferença Inverso
MLP	do Inglês MultiLayer Perceptron

- NB do Inglês Naive Bayes
- NIDM do Inglês Neighborhood Intensity Difference Matrix
- NoSQL do Inglês Not only Structured Query Language
- NSCLC do Inglês Non-Small Cell Lung Cancer
- PACS do Inglês Picture Archiving and Communication System
- RBF do Inglês Radial Basis Function-based neural network
- ReLU do Inglês Rectified Linear Unit
- RF do Inglês Random Forest
- RGB do Inglês *Red-Green-Blue*
- RIS do Inglês Radiology Information System
- RLM do Inglês Run-Length Matrix
- RM Ressonância Magnética
- RNA Rede Neural Artificial
- ROC do Inglês Receiver Operating Characteristic
- ROI do Inglês Region of Interest
- SCC do Inglês Squamous Cell Carcinoma
- SCLC do Inglês Small Cell Lung Cancer
- SGBD Sistema Gerenciador de Banco de Dados
- SMC Segundo Momento Central
- SMOTE do Inglês Synthetic Minority Over-sampling Technique
- SVM do Inglês Support Vector Machine
- TC Tomografia Computadorizada
- TNM Tumor, (Metástase Linfo) Nodal, Metástase (Distante)
- UH Unidade Hounsfield
- VOI do Inglês Volume of Interest
- XML do Inglês eXtensible Markup Language

SUMÁRIO

1	INTRODUÇÃO	22
1.1	Contextualização	22
1.2	Motivação	24
1.3	Estado da Arte	29
1.4	Objetivos	31
1.5	Organização	31
2	ANÁLISE E RECONHECIMENTO DE PADRÕES EM IMAGENS	
	MÉDICAS	33
2.1	Considerações Iniciais	33
2.2	Imagens Digitais de Tomografia Computadorizada	33
2.3	Segmentação de Imagens Médicas	35
2.4	Extração de Atributos	37
2.4.1	Atributos de Nível de Cinza	38
2.4.2	Atributos de Textura	40
2.4.3	Atributos de Forma	49
2.5	Seleção de Atributos Relevantes	53
2.6	Classificação de Imagens Médicas e Aprendizado de Máquina	55
2.6.1	Aprendizado Profundo	60
2.6.2	Validação e Avaliação de Performance	61
2.7	Análise de Sobrevida (Tempo-para-Evento)	64
2.8	Recuperação de Imagens Médicas Baseada em Conteúdo	65
2.8.1	Avaliação de Performance	67
3	CLASSIFICAÇÃO DE NÓDULOS PULMONARES	71
3.1	Fundamentação	71
3.2	Material e Métodos	71
3.2.1	Conjunto e Segmentação de Imagens	71
3.2.2	Extração de Atributos	74
3.2.3	Seleção de Atributos	75
3.2.4	Classificação de Imagens	76
3.3	Resultados	76
3.4	Discussão e Considerações Específicas	83
4	RECUPERAÇÃO POR CONTEÚDO DE NÓDULOS PULMONARES	86
4.1	Fundamentação	86

4.2	Material e Métodos
4.2.1	Conjunto de Imagens
4.2.2	Extração de Atributos
4.2.3	Seleção de Atributos
4.2.4	Recuperação de Imagens
4.3	Resultados
4.4	Discussão e Considerações Específicas
5	RADIÔMICA DO CÂNCER DE PULMÃO
5.1	Fundamentação
5.2	Material e Métodos
5.2.1	Conjunto e Segmentação de Imagens
5.2.2	Extração de Atributos
5.2.3	Análise Associativa
5.3	Experimentos e Resultados 102
5.3.1	Primeira Etapa
5.3.2	Segunda Etapa
5.3.3	Terceira Etapa
5.3.4	Quarta Etapa
5.4	Discussão e Considerações Específicas
6	CONCLUSÃO
6.1	Considerações Finais
6.2	Direções Futuras
6.3	Contribuições Científicas
	REFERÊNCIAS
	APÊNDICES 158
	APÊNDICE A – ESQUEMA DE BANCO DE DADOS NÃO-RELACIONAL PARA O REPOSITÓRIO DE IMAGENS LIDC 159
	APÊNDICE B – OTIMIZAÇÃO DE CBIR EM UNIDADE DE PRO- CESSAMENTO GRÁFICO
	APÊNDICE C – ANÁLISE RADIÔMICA DE TEXTURA DE TU- MORES PULMONARES PARA DISTINÇÃO DE PSEUDOCAVIDADES E CAVIDADES 166

ANEXO	A – ARTIGO PUBLICADO NO PERIÓDICO <i>J DIGIT IMA-</i> <i>GING (2018) 31:451–463</i>
ANEXO	B – ARQUIVO XML DE UM EXAME DE TC DO REPO- SITÓRIO DE IMAGENS LIDC
ANEXO	C – ARTIGO PUBLICADO NO PERIÓDICO <i>INT J CARS</i> (2017) 12:509–517
ANEXO	D – ARTIGO PUBLICADO NO PERIÓDICO COMPUT METHODS PROGRAMS BIOMED (2018) 159:23–30 178
ANEXO	E – OFÍCIO COM PARECER DO COMITÊ DE ÉTICA EM PESQUISA
ANEXO	F – ARQUIVO DICOM-RT DE UM EXAME DE IMA- GENS DE TC SEGMENTADAS
ANEXO	G – ARTIGO PUBLICADO NO PERIÓDICO <i>J DIGIT IMA-GING (2016) 29:716–729</i> 182
ANEXO	H – ARTIGO PUBLICADO NO PERIÓDICO <i>J SUPER-</i> <i>COMPUT (2017) 73:3451–3467</i>

1 INTRODUÇÃO

1.1 Contextualização

O câncer é uma doença em que células anormais se dividem incontrolavelmente, podendo invadir tecidos adjacentes e se espalhar através da corrente sanguínea e do sistema linfático para outras partes do corpo (JEMAL et al., 2014). O câncer de pulmão é o que mais causa vítimas relacionadas à neoplasia em todo mundo, sendo responsável por 1 em cada 5 mortes (HOWLADER et al., 2016). O câncer de pulmão é uma doença agressiva em que geralmente a presença de sintomas é um sinal de prognóstico ruim. Apenas 15% dos pacientes conseguem sobreviver nos primeiros cinco anos pós-diagnóstico (AMERICAN CANCER SOCIETY, 2016). Na maioria das vezes, quando a doença já se encontra em estágios avançados, o diagnóstico tardio impede o tratamento curativo (HOWLADER et al., 2016).

O diagnóstico do câncer de pulmão é uma tarefa complexa e apresenta grandes desafios, como por exemplo, a detecção de nódulos pequenos em imagens radiológicas. A identificação destas lesões em estágio inicial é uma tarefa complicada porque os nódulos pulmonares podem ser pequenos e estar conectados em estruturas anatômicas complexas (Figura 1) (VASSALLO et al., 2019; FERREIRA JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2016; RUBIN et al., 2014).

Figura 1: Detecção de tumores em estágio inicial em imagens de tomografia computadorizada: (a-d) apresentam os nódulos e quantos radiologistas de um total de 13 foram capazes de identificar estas lesões.



Fonte: Rubin et al. (2014).

Outro desafio é o da interpretação das imagens médicas. Imagens radiológicas são extremamente complexas por natureza, e a interpretação delas é um desafio reconhecido mesmo para especialistas experientes (AZEVEDO-MARQUES et al., 2017; AKGÜL et al., 2011; DOI, 2007). Além disso, radiologistas são limitados pela capacidade humana de análise, sofrendo influências de fatores externos (e.g. ruído e luminosidade) e internos,

como o nível de treinamento e condições psicológicas (e.g. fadiga e pressa). Portanto, o especialista está sujeito a erros de detecção (reconhecimento de um padrão não-normal) ou de diagnóstico (identificação da natureza benigna ou maligna do achado radiológico), podendo resultar em variabilidade intra e interobservador (Tabela 1) (DEGNAN et al., 2018; FERREIRA JUNIOR; OLIVEIRA, 2017).

Tabela 1: Variabilidade interobservador na avaliação visual de um nódulo pulmonar: houve discordância em sete de oito características da lesão. Os valores em negrito apresentam as discordâncias.

Radiologista	Sutileza	Calcificação	Esfericidade	Margem	Lobulação	Espiculação	Textura	\mathbf{PM}
А	5	6	3	3	3	5	5	5
В	5	6	3	4	4	5	5	5
\mathbf{C}	5	6	5	4	1	5	4	4
D	4	3	4	4	3	5	5	5

PM é a probabilidade de malignidade.

Fonte: Kaya (2018).

O prognóstico do câncer de pulmão é outro grande desafio, pois ele varia acentuadamente de acordo com o estágio do tumor no diagnóstico (AMERICAN CANCER SOCIETY, 2016; KOENIGKAM-SANTOS et al., 2014). A escolha do tratamento frente a um tumor maligno pulmonar é tomada principalmente pelo sistema internacional de estadiamento de neoplasias, chamado de TNM. O TNM leva em consideração a extensão e tamanho do tumor primário (T), a ausência/presença e a extensão de metástase em linfonodos regionais (N) e a ausência/presença de metástase distante (M) (SOBIN; GOSPODAROWICZ; WITTEKIND, 2010).

Contudo, estudos recentes mostram que outros fatores podem influenciar o prognóstico, a tomada de decisão clínica e o impacto no tratamento do paciente, como o tipo histológico da lesão (KOENIGKAM-SANTOS et al., 2014; COOPER et al., 2011). O câncer de pulmão pode ser classificado de acordo com duas principais classificações histopatológicas na prática clínica: câncer de pulmão de células não pequenas (do Inglês *non-small cell lung cancer* - NSCLC, Figuras 2(a–c)) e câncer de pulmão de células pequenas (do Inglês *small cell lung cancer* - SCLC, Figura 2(d)). Os tumores pulmonares pertencentes à classe NSCLC compreendem cerca de 85% dos casos e ainda podem ser subclassificados em adenocarcinoma (ADC, Figura 2(a)), carcinoma de células escamosas (do Inglês *squamous cell carcinoma* - SCC, Figura 2(b)) e carcinoma de células grandes (do Inglês *large cell carcinoma* - LCC, Figura 2(c)) (BRANT; HELMS, 2012).

Alguns trabalhos mostram que outras características de tumor em imagens de tomografia computadorizada (TC) também podem influenciar no prognóstico e na tomada de decisão terapêutica (TAILOR et al., 2015; KOENIGKAM-SANTOS et al., 2014; AUSTIN et al., 2013). Tipicamente estas características descrevem a forma e a estrutura interna do tumor, porém elas são descritas de maneira subjetiva, qualitativa ou semiquantitaFigura 2: Subtipos histopatológicos mais comuns do câncer de pulmão em imagens de tomografia computadorizada: (a) ADC, (b) SCC, (c) LCC, (d) SCLC. As setas indicam a localização do tumor.



Fonte: Brant e Helms (2012).

tiva (e.g. atenuação heterogênea, contornos espiculados, diâmetro bidimensional no plano axial, etc (Figura 3) (MACMAHON et al., 2017; YIP et al., 2017; BARTHOLMAI et al., 2015; TRUONG et al., 2014)), pois são resultantes da inspeção visual das imagens pelos radiologistas. Além disso, os tumores ADC, SCC e SCLC podem possuir características visuais qualitativas próximas a lesões benignas (BARTHOLMAI et al., 2015), dificultando a tomada de decisão clínica.

Figura 3: Exemplo de avaliação visual subjetiva dos contornos de um tumor: (a), (b) e (c) apresentam tumores sem contornos espiculados, com contornos finamente espiculados e com contornos grosseiramente espiculados, respectivamente. As setas indicam a localização do tumor.



Fonte: Yip et al. (2017).

1.2 Motivação

Ferramentas de diagnóstico/detecção auxiliado por computador (do Inglês *computer-aided diagnosis/detection* - CAD) têm sido desenvolvidas para auxiliar especialistas a interpretar achados radiológicos e identificar precocemente doenças, especialmente tumores de mama e pulmão (GIGER, 2018; ERICKSON et al., 2017; GILLIES; KINAHAN;

HRICAK, 2016; AERTS, 2016). O objetivo do CAD é melhorar a acurácia e consistência da interpretação e diagnóstico por imagem, mediante o uso da sugestão de resposta fornecida por técnicas de, principalmente, processamento de imagens, visão computacional e aprendizado de máquina, entre outras (AZEVEDO-MARQUES et al., 2017; FERREIRA JUNIOR; OLIVEIRA, 2017; DOI, 2007).

Historicamente, os primeiros sistemas CAD surgiram no final da década de 1980 e eram baseados no processamento de filmes radiológicos digitalizados. Eles foram fundamentalmente concebidos para trabalhar como uma segunda opinião em programas de rastreamento populacional de câncer, buscando aumentar a acurácia diagnóstica por meio de uma segunda leitura. No modelo do CAD, a segunda leitura é feita pelo computador ao invés de um segundo radiologista. No final da década de 1990 e início dos anos 2000, ocorreu uma grande revolução tecnológica na área da radiologia e diagnóstico por imagem com a estruturação de ambientes radiológicos sem filme, totalmente digitalizados e integrados com outros sistemas de informação, dentro de um modelo conhecido como Hospital Digital (PAVLOPOULOS; DELOPOULOS, 1999). A base da radiologia sem filme é o sistema de arquivamento e comunicação de imagens (do Inglês picture archiving and communication system - PACS), responsável por receber as imagens no padrão de comunicação de imagem digital em medicina (do Inglês digital imaging and communications in medicine - DICOM) dos diversos dispositivos de aquisição, deixá-las disponíveis para que especialistas ou outros sistemas computacionais possam acessá-las e ainda armazená-las de maneira organizada em grandes bases de dados (PONCIANO-SILVA et al., 2013). A integração do PACS com os sistemas de informação clínica HIS (do Inglês hospital information system) e RIS (do Inglês radiology information system) dos hospitais possibilitou o surgimento de um outro modelo de auxílio computadorizado ao diagnóstico, baseado na recuperação de imagens visualmente similares. Com o objetivo de aumentar a acurácia diagnóstica e a aplicabilidade clínica do PACS, foram incorporadas técnicas de recuperação de imagens baseada em conteúdo (do Inglês content-based image retrieval - CBIR) nas ferramentas computacionais para a radiologia (TRAINA et al., 2017; SILVA et al., 2013).

A CBIR tem sido descrita como uma das ferramentas computacionais mais promissoras, pois ela potencialmente pode auxiliar o processo de decisão clínica, recuperando em grandes bases de dados casos similares já previamente diagnosticados (MA et al., 2017; BEDO et al., 2016; FERREIRA JUNIOR; OLIVEIRA, 2015; BUGATTI et al., 2014). Sistemas CBIR podem ser muito úteis em um ambiente clínico real, pois podem auxiliar o radiologista na tomada de decisão diagnóstica, ou aumentar sua certeza, valendo-se de um modelo de decisão baseada em exemplo (SILVA et al., 2013). Tradicionalmente, sistemas CAD executam tarefas de classificação de imagens, fornecendo uma única resposta (uma segunda opinião) ao especialista, e sistemas CBIR executam buscas de imagens médicas por similaridade, fornecendo um conjunto de casos similares a um caso desconhecido dado como entrada. Nenhum dos dois tipos de tecnologia fornece diretamente informações de auxílio ao prognóstico e à decisão terapêutica, o que limita sua aplicabilidade à rotina clínica.

Em contrapartida, a radiômica (do Inglês *radiomics*) tem sido descrita como uma extensão do CAD que associa¹ características quantitativas (atributos) de imagens médicas com desfechos clínicos do paciente (Figura 4), potencialmente aumentando o poder dos modelos de suporte à decisão clínica (GIGER, 2018; LARUE et al., 2017; GILLIES; KINAHAN; HRICAK, 2016; WU et al., 2016a). Em face dos recentes avanços em terapias alvo para a medicina de precisão (ou personalizada), a necessidade de uma abordagem simples de análise robusta de imagens médicas se tornou imperativa, e a radiômica pode prover isto por ser uma ferramenta não-invasiva, rápida e de baixo-custo para a rotina clínica (GIARDINO et al., 2017; SACCONI et al., 2017; PERMUTH et al., 2016). Além disso, a radiômica utiliza algoritmos computacionais e matemáticos avançados que descrevem a forma e a textura das lesões radiológicas em mínimos detalhes que podem passar despercebidos até por especialistas experientes na avaliação visual desarmada.



Figura 4: Possíveis associações para a radiômica.

Fonte: Autor.

¹ Neste contexto, o termo "associar" é sinônimo de "correlacionar". Estatisticamente, a correlação implica que o desfecho possui um valor contínuo (numérico), o que não necessariamente ocorre na radiômica. Logo, nesta tese será utilizado o termo "associação" para descrever o processo de reconhecimento de padrões na radiômica. Contudo, muitas vezes o termo "associação" também possui o sentido de "combinação" (na farmacologia, por exemplo), porém este não é o caso no contexto desta tese.

A análise radiômica define de forma explícita um processo de extração massiva de atributos das imagens médicas, de inserção destes atributos em bases que agreguem outros dados e informações do paciente que possam ser compartilhadas, e que permitam a geração e/ou a confirmação de hipóteses (GILLIES; KINAHAN; HRICAK, 2016). O aspecto de integração dos dados clínicos é ainda mais importante à medicina de precisão devido o potencial preditivo da mineração dos dados do paciente na atual "Era *Big Data*" (KEEK et al., 2018; LAMBIN et al., 2017). O volume de dados heterogêneos em saúde vem crescendo em um ritmo acelerado nos últimos anos, e esses dados eletrônicos estão disponíveis em volumosas quantidades nos sistemas de informação dos grandes hospitais e centros de saúde (Figura 5) (FILONENKO; SEERAM, 2018; KANSAGRA et al., 2016). Ou seja, a radiômica pode ser entendida como a translação dos conceitos de *Big Data* e medicina de precisão para o contexto do diagnóstico por imagem e está focada no desenvolvimento de ferramentas e bases de apoio à tomada de decisão clínica, que possam potencialmente melhorar o diagnóstico, o prognóstico e a precisão de uma previsão (GILLIES; KINAHAN; HRICAK, 2016).



Figura 5: Fontes de Big Data para geração de novas informações clínicas.

Fonte: Kansagra et al. (2016).

Quando o foco da radiômica é o estudo da correlação entre os padrões radiológicos e genômicos, o processo passa a receber o nome de radiogenômica, área de estudo que tem despertado grande interesse na comunidade científica (THAWANI et al., 2017; RUTMAN; KUO, 2009). Características de imagens médicas e padrões de expressão e mutações genéticas possuem associações significativas, demonstrando que a análise radiogenômica pode investigar diferentes mecanismos biológicos por meio de dispositivos matemáticos e computacionais, possibilitando a decodificação de fenótipos de tumores por meio de métodos não invasivos de obtenção de imagens (GEVAERT et al., 2017; AERTS et al., 2014).

A radiômica vem sendo ainda empregada em estudos que analisam a heterogeneidade intratumoral, que é a presença de múltiplas subregiões em um mesmo tumor, o que pode aumentar a progressão da doença e a resistência terapêutica (LEE et al., 2017a; DIVINE et al., 2016; CHEN et al., 2011). A radiômica é capaz de quantificar a complexidade espacial do tumor e identificar as subregiões heterogêneas (Figura 6) (LEE et al., 2017a).

Figura 6: Quantificação da heterogeneidade intratumoral. Os eixos x e y no gráfico se referem a posições espaciais em coordenadas cartesianas. A escala de cores se refere a um atributo extraído da imagem refletindo as subregiões do tumor.



Fonte: Lee et al. (2017a).

Neste contexto, como principal justificativa para a pesquisa desta tese, é citada a ausência de modelos integrados de classificação, recuperação por conteúdo e radiômica de imagens médicas robustos e biomarcadores quantitativos não-invasivos precisos que permitam o auxílio computadorizado ao diagnóstico, prognóstico e decisão terapêutica. O desenvolvimento de um arcabouço integrado poderia aumentar a aplicabilidade dos sistemas de informação em radiologia e contribuir para a tomada de decisão clínica, potencialmente melhorando a performance diagnóstica, a avaliação do prognóstico do paciente e acompanhamento da progressão da doença. Ferramentas computacionais de CAD, CBIR e radiômica possuem grande potencial em reconhecer objetiva e precisamente padrões em imagens médicas, porém estes sistemas ainda possuem diversas limitações. Este trabalho tentou diminuir algumas das lacunas apresentadas por esses modelos, principalmente as de característica (do Inglês *feature gap*), referente à representação numérica eficiente das imagens, e de desempenho (do Inglês *performance gap*), referente à implementação e qualidade na integração, avaliação e generalização/replicação dos modelos (AERTS, 2016; SOUZA, 2012; KUMAR et al., 2012; AKGÜL et al., 2011; DESERNO; ANTANI; LONG, 2009). O câncer de pulmão em imagens de TC foi definido como estudo de caso para o desenvolvimento da pesquisa, devido aos desafios apresentados anteriormente (Seção 1.1).

1.3 Estado da Arte

Diversos esquemas CAD para classificação de nódulos pulmonares malignos e benignos foram propostos pela literatura. Tradicionalmente, estes esquemas são compostos por métodos de segmentação de imagens de TC, extração de atributos quantitativos, seleção de características relevantes e classificadores de aprendizado de máquina (KAYA, 2018; GONG et al., 2018; OROOJI et al., 2018; CHOI et al., 2018; NISHIO; NAGASHIMA, 2017; ALILOU; OROOJI; MADABHUSHI, 2017; CARVALHO FILHO et al., 2017; FELIX et al., 2016; REEVES; XIE; JIRAPATNAKUL, 2016; DILGER et al., 2015; WU et al., 2013). Contudo, técnicas de aprendizado profundo (do Inglês *deep learning*) de máquina têm feito grandes avanços recentemente em problemas de reconhecimento de padrões (CHARTRAND et al., 2017; LITJENS et al., 2017; LECUN; BENGIO; HINTON, 2015). Estas técnicas permitem que a máquina receba como entrada dados brutos e descubra automaticamente as representações necessárias, de acordo com uma hierarquia onde características de alto nível são definidas de características de baixo nível (LECUN; BENGIO; HINTON, 2015; ERTOSUN; RUBIN, 2015). Muitas vezes, em tarefas de classificação de objetos em imagens, a extração de características profundas é realizada sem mesmo a necessidade da segmentação prévia do objeto, o que aumenta ainda mais a relevância do método (CHARTRAND et al., 2017; LITJENS et al., 2017). O Capítulo 2 apresenta todas estas terminologias com mais detalhes e as descrições de alguns desses métodos. Vários esquemas CAD basearam-se em aprendizado profundo para a classificação benigno vs. maligno de nódulos pulmonares (SHEN et al., 2019; HUSSEIN et al., 2019; NÓBREGA et al., 2018; ZHAO et al., 2018; PAUL et al., 2018; SUN; ZHENG; QIAN, 2017; TAJBAKHSH; SUZUKI, 2017; SILVA et al., 2017; NIBALI; HE; WOLLERSHEIM, 2017; SHEN et al., 2017; SONG et al., 2017a; HUSSEIN et al., 2017).

Além disso, algumas abordagens de recuperação de imagens de nódulos pulmonares malignos e benignos foram apresentadas pela literatura com o objetivo de auxiliar o diagnóstico do câncer de pulmão, principalmente utilizando a extração de atributos tradicionais das imagens de TC (WEI et al., 2018; DHARA et al., 2017a; DHARA et al., 2017b; AGGARWAL; VIG; SARDANA, 2016; LUCENA et al., 2016; FERREIRA JUNIOR; OLIVEIRA, 2015; AGGARWAL; SARDANA; VIG, 2014; LAM et al., 2007). Técnicas de aprendizado profundo não foram totalmente exploradas na recuperação de nódulos pulmonares similares (IBANEZ et al., 2017).

Por outro lado, estes estudos de análise tumoral não deveriam focar apenas na diferenciação entre maligno e benigno, mas também no desenvolvimento de modelos de suporte à tomada de decisão ao tratamento, o que potencialmente poderia melhorar a avaliação do prognóstico do paciente e progressão da doença (AZEVEDO-MARQUES et al., 2017; GILLIES; KINAHAN; HRICAK, 2016; AERTS et al., 2014; KOENIGKAM-SANTOS et al., 2014). Várias investigações radiômicas foram apresentadas pela literatura associando atributos quantitativos de imagens de TC com diversos desfechos clínicos:

- histopatologia, e.g. ADC e SCC (DIGUMARTHY et al., 2019; E et al., 2018; SAAD; CHOI, 2018; ZHU et al., 2018; WU et al., 2016b; PARMAR et al., 2015b; AERTS et al., 2014);
- grau de diferenciação histológica, e.g. ADC invasivo e minimamente invasivo (COHEN et al., 2018; CHEN et al., 2018; LE et al., 2018; SONG et al., 2017b; KO et al., 2016);
- estadiamento, e.g. metástases (LIU et al., 2018; ZHOU et al., 2018; WU et al., 2016a; HUYNH et al., 2016; COROLLER et al., 2015; PARMAR et al., 2015b; AERTS et al., 2014);
- mutações de genes, e.g. receptor do fator de crescimento epidérmico (do Inglês epidermal growth factor receptor EGFR) e quinase de linfoma anaplásico (do Inglês anaplastic lymphoma kinase ALK) (DIGUMARTHY et al., 2019; GEVAERT et al., 2017; SACCONI et al., 2017; LIU et al., 2016; RIZZO et al., 2016; AERTS et al., 2016; LIU et al., 2016);
- expressão genética (WANG et al., 2018; ZHOU et al., 2017; AERTS et al., 2014; GEVAERT et al., 2012);
- recorrência do tumor (LIM et al., 2018; HUYNH et al., 2017; CIRUJEDA et al., 2016; HUYNH et al., 2016; EMAMINEJAD et al., 2016; DEPEURSINGE et al., 2015);
- resposta ao tratamento (TANG et al., 2018; COROLLER et al., 2017; HUNTER et al., 2016; COROLLER et al., 2016; AERTS et al., 2016);
- sobrevida do paciente (SAAD; CHOI, 2018; JONG et al., 2018; LEGER et al., 2017; HUANG et al., 2016; HUYNH et al., 2016; PARMAR et al., 2015a; HAWKINS et al., 2014);
- heterogeneidade intratumoral (OROOJI et al., 2018; SON et al., 2016; GROVE et al., 2015).

Entretanto, poucos trabalhos realizaram radiômica do câncer de pulmão em TC baseada em aprendizado profundo (HOSNY et al., 2018; PAUL et al., 2016). Logo, ainda existe um vasto campo de pesquisa não explorado. Além disso, estratégias de CBIR focadas na radiômica do câncer de pulmão (e.g. baseada na presença/ausência de metástases) ainda não foram apresentadas na literatura científica até o momento da escrita deste trabalho.

1.4 Objetivos

O objetivo principal desta pesquisa foi desenvolver um arcabouço para classificação, recuperação baseada em conteúdo e radiômica de imagens médicas que possa ser aplicado no auxílio computadorizado ao diagnóstico, prognóstico e decisão terapêutica. Os objetivos específicos do trabalho foram:

- Desenvolver retrospectivamente bases de imagens de TC e de dados clínicos de pacientes com tumores pulmonares;
- Investigar métodos de classificação de imagens médicas para o desenvolvimento de um algoritmo de CAD;
- Investigar métodos de recuperação de imagens médicas baseada em conteúdo para a busca por similaridade;
- Investigar métodos de radiômica para o desenvolvimento de um algoritmo de associação de atributos de imagens médicas com desfechos clínicos;
- Avaliar a performance dos modelos de classificação, recuperação por conteúdo e radiômica no contexto do câncer de pulmão em imagens de TC.

1.5 Organização

O restante desta tese está organizado da seguinte forma:

- o Capítulo 2 aborda uma fundamentação teórica com os principais conceitos envolvidos no desenvolvimento do trabalho e estabelece definições para alguns métodos apresentados pela literatura em classificação, radiômica e recuperação por conteúdo de imagens médicas;
- o Capítulo 3 apresenta os experimentos realizados em classificação de nódulos pulmonares baseada em extração de atributos quantitativos, seleção de características relevantes e classificação de imagens;
- o Capítulo 4 apresenta os experimentos realizados em CBIR baseada em extração de atributos quantitativos, seleção de características relevantes e recuperação de nódulos pulmonares similares por conteúdo;

- o Capítulo 5 apresenta os experimentos realizados na radiômica do câncer de pulmão baseada em segmentação semi-automática de imagens de TC, extração de atributos quantitativos, seleção de características relevantes, aprendizado de máquina, aprendizado profundo e análise tempo-para-evento;
- o Capítulo 6 conclui esta tese, apresentando as considerações finais do trabalho, as principais contribuições obtidas e algumas direções para trabalhos futuros.

2 ANÁLISE E RECONHECIMENTO DE PADRÕES EM IMAGENS MÉDICAS

2.1 Considerações Iniciais

Este capítulo apresenta a fundamentação teórica necessária para o entendimento de análise computadorizada de imagens médicas e reconhecimento de padrões. Os principais conceitos envolvidos nesta tese foram divididos em sete partes: Imagens Digitais de Tomogra-fia Computadorizada (Seção 2.2), Segmentação de Imagens Médicas (Seção 2.3), Extração de Atributos (Seção 2.4), Seleção de Atributos Relevantes (Seção 2.5), Classificação de Imagens Médicas e Aprendizado de Máquina (Seção 2.6), Análise de Sobrevida (Tempopara-Evento) (Seção 2.7) e Recuperação de Imagens Médicas Baseada em Conteúdo (Seção 2.8).

2.2 Imagens Digitais de Tomografia Computadorizada

O exame de TC é a principal modalidade de imagem da rotina clínica para a identificação de nódulos pulmonares e a avaliação visual do câncer de pulmão (MACMAHON et al., 2017; BARTHOLMAI et al., 2015; TRUONG et al., 2014). Uma imagem digital de TC é uma função f(x, y) em escala de níveis de cinza discretizada em coordenadas espaciais que pode ser considerada uma matriz cujos índices de linhas e de colunas identificam um ponto de imagem (*pixel*). O valor de cada *pixel* na matriz identifica o nível de cinza no ponto (x, y) (Figura 7), em uma escala de valores inteiros que representam preto (menor valor), branco (maior valor) e tons de cinza (valores intermediários).

Figura 7: Representação da função de uma imagem digital em escala de cinza.





Imagens de TC são comumente adquiridas em forma de volume de fatias paralelas e uniformemente espaçadas. O empilhamento destas imagens, mantendo o espaço original ΔZ entre elas como espessura de corte, pode ser idealizado de forma que cada *pixel* represente digitalmente um ponto volumétrico (*voxel*) em estudo no paciente (Figura 8). Cada *voxel* tem sua intensidade definida pela unidade Hounsfield (UH), conforme a Equação 2.1:

$$UH(x, y, z) = 1000 \frac{(\mu(x, y, z) - \mu_w)}{\mu_w},$$
(2.1)

onde $\mu(x, y, z)$ é a média do coeficiente de atenuação linear de um *voxel* na posição (x, y, z), UH(x, y, z) representa as imagens de TC em escala de cinza nas mesmas coordenadas espaciais (x, y, z) e μ_w é o coeficiente de atenuação linear da água destilada (BUSHBERG et al., 2012). Quando o *voxel* na posição (x, y, z) representa a água, então este ponto apresenta UH(x, y, z) equivalente a zero (Tabela 2).

Figura 8: Representação de um voxel.



Fonte: Bushberg et al. (2012).

Ar	Gordura	Líquido	Partes Moles	Calcificações	Osso Cortical
-1000	-120 a -40	0 a 10	10 a 70	>200	>500

Fonte: Autor.

Voxels de imagens de TC normalmente são representados por 12 *bits*, o que pode gerar 4.096 possibilidades de nível de cinza. As unidades Hounsfield são utilizadas para a distribuição dos 4.096 níveis de cinza que vai gerar a imagem de TC. Contudo, a sua escala possui grande amplitude e a distribuição dos níveis de cinza entre os limites mínimo e máximo resulta em um baixa diferenciação de contraste entre os tecidos. Logo, é necessário realizar o janelamento na imagem de TC, ou seja, determinar o número UH (largura) e o seu valor central (nível) (Figura 9). Figura 9: Aplicações de janelamento em TC de tórax: (a) janela de pulmão com largura de 1.400 UH e nível de -500 UH, (b) janela de mediastino com largura de 350 UH e nível de 40 UH, (c) janela de osso com largura de 1.000 UH e nível de 400 UH.



Fonte: Autor.

2.3 Segmentação de Imagens Médicas

Segmentação é uma das principais etapas no processamento de imagens médicas e pode apresentar grandes desafios em problemas de reconhecimento de padrões. Seu principal objetivo é dividir a imagem em partes que tenham uma grande correlação com objetos contidos na imagem (SONKA; HLAVAC; BOYLE, 2008). Em outras palavras, a segmentação identifica os *pixels/voxels* da imagem que pertencem a uma região de interesse (do Inglês *region of interest* - ROI) ou, ainda, no contexto médico, separa anatomicamente as estruturas ou tecidos representados no exame de imagem. As técnicas para segmentação de imagens médicas são geralmente baseadas nas propriedades básicas de níveis de cinza de descontinuidade ou similaridade. Na primeira categoria, a abordagem é dividir a imagem baseando-se em mudanças bruscas nas intensidades de nível de cinza, como acontece nas bordas de uma imagem. Na segunda categoria, a abordagem se baseia na comparação dos níveis de cinza. Duas técnicas tradicionais para esta última estratégia são limiarização e orientada a região, que pode se estender a crescimento de região (BANKMAN, 2009; GONZALEZ; WOODS, 2007).

Limiarização é a mais simples técnica de segmentação de imagens, porém, devido o seu inexpressivo custo computacional e relativa eficiência, continua sendo utilizada em larga escala (HUANG et al., 2017; YOSHINO et al., 2017; CAO et al., 2017; FERREIRA JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2017). A limiarização pode ser definida como a transformação de uma imagem de entrada f para uma imagem segmentada de saída g, como segue a função da Equação 2.2:

$$g(i,j) = \begin{cases} f(i,j) & \text{se } f(i,j) \ge T, \\ 0 & \text{se } f(i,j) < T, \end{cases}$$
(2.2)

onde T é o valor do limiar e i, j são pontos no plano cartesiano. O limiar ou threshold entende-se como o valor de nível de cinza determinado como referência para a comparação entre os pixels. Quando a aplicação necessita de uma ROI sem níveis de cinza claros, a
função pode ser invertida de modo que os pontos da ROI sejam mantidos se $f(i, j) \leq T$ e o fundo preto se f(i, j) > T. O limiar pode ainda corresponder a um intervalo de valores $T_1 \in T_2$, sendo $T_1 < T_2$, de acordo com a Equação 2.3:

$$g(i,j) = \begin{cases} f(i,j) & \text{se } T_1 \le f(i,j) \le T_2, \\ 0 & \text{se } f(i,j) < T_1 & \text{ou } f(i,j) > T_2. \end{cases}$$
(2.3)

Muitas vezes há a necessidade da criação de máscaras binárias resultantes de uma limiarização, e quando isso acontece, a intensidade definida para a ROI é o branco, e assim, a segmentação passa a se chamar binarização (Figura 10).

Figura 10: Aplicação de segmentação do pulmão em imagem de TC de tórax com 256 níveis de cinza e limiar 115: (a) imagem de TC original, (b) resultado da limiarização, (c) resultado da binarização.



Fonte: Autor.

A limiarização pode ser a técnica apropriada para segmentação se os níveis de cinza forem claramente distintos entre ROI e fundo. Entretanto, ela é falha quando há um baixo contraste entre o objeto e o fundo, ruídos na imagem ou variação na intensidade dos níveis de cinza do fundo da imagem (BANKMAN, 2009; SONKA; HLAVAC; BOYLE, 2008). Nestes casos, é necessária a utilização de uma abordagem orientada a regiões baseada na descoberta das ROIs diretamente. A segmentação com crescimento de região considera grupos de *pixels* com intensidades de nível de cinza similares. Ela agrupa *pixels* ou sub-regiões em regiões maiores iterativamente. A técnica baseia-se na escolha de um ponto ou um conjunto de pontos na imagem/volume, chamado de semente, e, a partir dele, as regiões "crescem" anexando a cada ponto semente aqueles *pixels/voxels* que possuem propriedades similares, como nível de cinza, textura, homogeneidade, entre outros (Figura 11) (GONZALEZ; WOODS, 2007).

O crescimento de região inicia após a determinação da semente e do cálculo da propriedade feito entre ela e todo o resto da imagem. Cada *pixel* da imagem, com exceção da semente, é chamado de ponto candidato. Uma propriedade que pode ser utilizada no cálculo de similaridade, por exemplo, é a diferença absoluta entre os níveis de cinza da semente e do ponto candidato à região não exceda x% da diferença entre o maior e o

Figura 11: Aplicação de segmentação com crescimento de região de tumor em imagem de TC: (a) imagem de TC original com a semente marcada, (b) estágio primário de crescimento de região, (c) estágio intermediário de crescimento de região, (d) estágio final de crescimento de região.



Fonte: Autor.

menor nível de cinza da imagem. Se o cálculo dessa propriedade entre a semente e o ponto candidato for menor que um valor pré-definido, então esse ponto candidato pertencerá e será acrescentado à região, caso contrário, ele não pertencerá à região.

Entretanto, a segmentação com crescimento de região pode apresentar alguns desafios, como a seleção de sementes que representem adequadamente as ROIs. A seleção de um ou mais pontos iniciais deve se basear na natureza do problema. Contudo, a semente escolhida pode ser insuficiente para realizar a comparação das propriedades entre os pontos. Além disso, a definição da propriedade a ser calculada para a inclusão dos pontos candidatos à ROI também pode ser um fator desafiador para a segmentação das imagens devido a possibilidade da falta de representatividade da propriedade à natureza do problema e aos padrões presentes na imagem.

2.4 Extração de Atributos

O processo de extração de características de uma imagem consiste no cálculo de valores numéricos (atributos, descritores) que representem o seu conteúdo visual. Os atributos são obtidos por meio da execução de algoritmos, conhecidos como extratores de características, na imagem de entrada. Estes algoritmos de extração de características realizam procedimentos quantitativos nas imagens, como por exemplo, construção de histogramas, classificação de texturas, reconhecimento de formas e contornos, estimativas de área e volume, entre outros. Após a extração de características a partir destes algoritmos, os valores são armazenados em um vetor de atributos da imagem (Figura 12).

Os atributos são basicamente classificados em três grupos principais: cor, textura e forma, que juntos podem oferecer uma melhor discriminação na caracterização de imagens e posteriormente no reconhecimento de padrões. No contexto de imagens de TC, não existem componentes de cores, somente de intensidade luminosa, e os atributos, neste caso,



Figura 12: Processo de extração de n atributos de uma imagem de TC.



são extraídos dos níveis de cinza dos *pixels* das imagens. Logo, os atributos de cor serão tratados nesta tese como descritores de nível de cinza.

2.4.1 Atributos de Nível de Cinza

Descritores de nível de cinza ou intensidade são utilizados desde a criação dos primeiros sistemas de auxílio ao diagnóstico e, até hoje, são dos mais empregados devido a sua facilidade de implementação (GIGER, 2018; GILLIES; KINAHAN; HRICAK, 2016; DOI, 2007). Os atributos de níveis de cinza descrevem individualmente a distribuição dos valores dos *pixels* de uma ROI ou *voxels* de um volume de interesse (do Inglês *volume of interest* - VOI) (Figura 13). A extração de atributos de intensidade pode ser realizada diretamente dos níveis de cinza ou do histograma da imagem. O histograma de níveis de cinza é a descrição da quantidade de níveis de cinza presentes na imagem, e o seu cálculo envolve apenas a contagem de *pixels/voxels* com as intensidades na escala de cinza (Figura 14).

Atributos utilizados na extração de características de nível de cinza são de estatística de primeira ordem, como média aritmética, média geométrica, variância da população, variância da amostra, desvio padrão, energia, entropia, desvio absoluto médio, raiz do quadrado médio, assimetria (ou obliquidade) e curtose (ZHANG et al., 2015; AERTS et al., 2014). As Equações 2.4-2.13 listam esses atributos, onde x é um histograma de n níveis:

média aritmética
$$(\mu) = \frac{1}{n} \sum_{i=1}^{n} x_i,$$
 (2.4)

média geométrica =
$$\sqrt[n]{\prod_{i=1}^{n} x_i},$$
 (2.5)

variância da população =
$$\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2$$
, (2.6)

Figura 13: Articulações sacroilíacas em imagens de ressonância magnética em que os níveis de cinza podem ser úteis para o reconhecimento de padrões inflamatórios: (a) articulações sem inflamação, (b) articulações com inflamação.



Fonte: Faleiros et al. (2016).

Figura 14: Exemplo de histograma de uma imagem de TC com 256 níveis de cinza.



Fonte: Autor.

variância da amostra
$$(v) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \mu)^2,$$
 (2.7)

desvio padrão
$$(s) = \sqrt{v},$$
 (2.8)

$$energia = \sum_{i=1}^{n} x_i^2, \tag{2.9}$$

$$entropia = \sum_{i=1}^{n} x_i \log_2 x_i, \tag{2.10}$$

desvio absoluto médio =
$$\frac{1}{n} \sum_{i=1}^{n} |x_i - \mu|,$$
 (2.11)

raiz do quadrado médio =
$$\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$$
, (2.12)

assimetria =
$$\frac{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^3}{s^3}$$
, (2.13)

$$curtose = \frac{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^4}{s^4}.$$
(2.14)

Atributos de média e mediana são capazes de quantificar a densidade da imagem. Os atributos de desvio padrão, variância e desvio absoluto médio são medidas de dispersão, ou seja, quanto os níveis de cinza diferem da média. Já a entropia pode caracterizar a aleatoriedade e variações presentes na imagem (AERTS, 2016). As medidas de curtose e obliquidade são os momentos centrais mais utilizados e caracterizam os graus de achatamento e assimetria do histograma, respectivamente (AERTS et al., 2014). Contudo, utilizar somente descritores de níveis de cinza e/ou histograma não fornece informações da distribuição espacial do conteúdo de uma imagem. Além disso, estes atributos podem ser sensíveis a alterações de brilho, contraste e artefatos de compressão de dados (OLIVEIRA, 2006). Desta forma, características de textura são importantes para o complemento da representação quantitativa de ROIs e VOIs, pois levam em consideração relações de segunda ordem, ou ordens maiores da imagem (GILLIES; KINAHAN; HRICAK, 2016).

2.4.2 Atributos de Textura

Apesar de não haver um consenso sobre a definição formal de textura, a sua maior característica é a repetição de padrões sobre uma região na imagem (PARKER, 2011). Este padrão pode ser repetido exatamente igual ou como um conjunto de pequenas variações. Alguns aspectos da textura podem variar entre as regiões e não devem ser ignorados, como por exemplo, tamanho, forma, cor e orientação dos elementos presentes (GONZALEZ; WOODS, 2007). Em algumas imagens médicas, regiões que possuem intensidades similares são possíveis de se distinguir por causa das diferentes texturas (Figura 15). Atributos de textura se tornaram particularmente importantes à radiologia devido a sua capacidade de refletir os mínimos detalhes contidos dentro de uma lesão em uma imagem médica (AZEVEDO-MARQUES et al., 2017; AKGÜL et al., 2011). Figura 15: Fraturas de corpos vertebrais em imagens segmentadas de ressonância magnética em que a textura pode ser útil para o reconhecimento da etiologia de uma fratura como sendo associada a presença de osteoporose (benigna) ou de um câncer (maligna) na coluna: linhas (a) e (b) apresentam fraturas benignas e malignas, respectivamente.



Fonte: Frighetto-Pereira et al. (2016).

Existem diversas formas de caracterizar a textura de uma imagem. Tradicionalmente, as abordagens utilizadas para representar elementos texturais podem ser classificadas em três métodos principais: estatísticos, espectrais e estruturais (GONZALEZ; WOODS, 2007). Métodos estatísticos obtêm informações sobre o posicionamento dos *pixels* e distribuição espacial dos níveis de cinza por meio de algum procedimento estatístico. Métodos espectrais analisam a textura no domínio da frequência da imagem por meio de transformadas. E métodos estruturais descrevem a textura por meio da identificação de primitivas estruturais e pela forma como estão dispostas (OLIVEIRA, 2006).

Um dos métodos de análise textural estatística mais presentes na literatura foi proposto por Haralick, Shanmugam e Dinstein (1973). Estes descritores de textura são classificados como de segunda ordem por serem obtidos da matriz de coocorrência (do Inglês *cooccurrence matrix* - COM) de níveis de cinza da imagem. A COM é a representação da distribuição de probabilidade de ocorrência de pares de *pixels* com determinada intensidade e distância em uma direção. Em outras palavras, a matriz de coocorrência envolve a estimativa de uma função de probabilidade de segunda ordem discreta, que representa a probabilidade de ocorrência de pares de *pixels* com níveis de cinza i, j, dada uma distância d e uma orientação θ entre os *pixels* nas dimensões x, y (HARALICK; SHANMUGAM; DINSTEIN, 1973).

O cálculo da matriz de coocorrência também pode ser realizado em um VOI. Neste caso, a COM estende a avaliação da função de probabilidade para o eixo z, dada uma orientação ϕ , com o objetivo de examinar a probabilidade de ocorrência de pares de *voxels* entre fatias de um volume de imagens (GHONEIM et al., 2003). Funções de segunda ordem são aplicadas na COM produzindo os atributos de textura (ZHANG et al., 2015; AERTS et al., 2014; SOH; TSATSOULIS, 1999; HARALICK; SHAPIRO, 1992; HARALICK; SHANMUGAM; DINSTEIN, 1973). Alguns dos atributos de textura que podem ser calculados a partir da matriz de coocorrência estão listados nas Equações 2.15-2.29:

$$energia = \sum_{i,j} C(i,j)^2, \qquad (2.15)$$

$$entropia = -\sum_{i,j} C(i,j) \log_2 C(i,j), \qquad (2.16)$$

$$contraste = \sum_{i,j} (i-j)^2 C(i,j), \qquad (2.17)$$

$$variancia = \sum_{i,j} (i - \mu)^2 C(i, j), \qquad (2.18)$$

$$variância \ inversa = \sum_{i,j} \frac{C(i,j)}{|i-j|^2}, i \neq j$$
(2.19)

$$tend\hat{e}ncia = \sum_{i,j} (i+j-\mu_x-\mu_y)^2 C(i,j), \qquad (2.20)$$

$$matiz = \sum_{i,j} (i+j-\mu_x-\mu_y)^3 C(i,j), \qquad (2.21)$$

$$proemin\hat{e}ncia = \sum_{i,j} (i+j-\mu_x-\mu_y)^4 C(i,j),$$
(2.22)

$$correlação = \frac{\sum_{i,j} ijC(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y},$$
(2.23)

$$autocorrelação = \sum_{i,j} ijC(i,j),$$
(2.24)

$$dissimilaridade = \sum_{i,j} |i - j| C(i,j), \qquad (2.25)$$

$$homogeneidade = \sum_{i,j} \frac{C(i,j)}{1+|i-j|},$$
(2.26)

momento da diferença inverso =
$$\sum_{i,j} \frac{C(i,j)}{1+|i-j|^2}$$
, (2.27)

diferença inversa normalizada =
$$\sum_{i,j} \frac{C(i,j)}{1 + \frac{|i-j|}{n}}$$
, (2.28)

momento da diferença inversa normalizada =
$$\sum_{i,j} \frac{C(i,j)}{1 + \frac{|i-j|^2}{n^2}}$$
, (2.29)

onde C(i, j) é o elemento da matriz de coocorrência para o par i, j, n é o número de níveis de cinza, μ_x , μ_y são as médias e σ_x , σ_y são os desvios padrões.

Cada atributo de textura representa características intrínsecas à imagem. Por exemplo, a energia mede a uniformidade dos níveis de cinza em toda a região. A entropia mede o grau de aleatoriedade e dispersão de ocorrências de níveis de cinza em uma imagem. O contraste mede as variações locais de intensidade dos *pixels/voxels* e resulta em valores maiores quando existem grandes diferenças entre os níveis de cinza. Variância mede a dispersão dos níveis de cinza em relação à intensidade média da imagem. Tendência, matiz e proeminência emulam a percepção humana, medem a assimetria da região e resultam em altos valores quando a imagem não é simétrica (YIP et al., 2017). Correlação calcula a dependência linear dos níveis de cinza em uma região. A medida de homogeneidade é o inverso do contraste e resulta em grandes valores para níveis de cinza similares (PHILLIPS et al., 2017; OLIVEIRA, 2006).

Uma segunda abordagem estatística de textura é baseada na matriz de comprimento de sequência (do Inglês run-length matrix - RLM) de níveis de cinza (TANG, 1998; GALLOWAY, 1975). A RLM leva em consideração a quantidade j de vezes que uma mesma intensidade i aparece em sequências de diferentes comprimentos em uma determinada orientação θ (YANG et al., 2016). Algumas características podem ser computadas a partir do cálculo da RLM, resultando em atributos de textura extraídos da matriz de comprimento de sequência (Equações 2.30-2.40) (ZHANG et al., 2015; AERTS et al., 2014):

$$\hat{e}nfase \ de \ sequência \ curta = \frac{\sum_{i,j} \frac{p(i,j|\theta)}{j^2}}{\sum_{i,j} p(i,j|\theta)},$$
(2.30)

$$\hat{e}nfase \ de \ sequência \ longa = \frac{\sum_{i,j} j^2 p(i,j|\theta)}{\sum_{i,j} p(i,j|\theta)},$$
(2.31)

não-uniformidade de nível de cinza =
$$\frac{\sum_{i} (\sum_{j} p(i, j|\theta))^2}{\sum_{i,j} p(i, j|\theta)}$$
, (2.32)

não-uniformidade de comprimento de sequência =
$$\frac{\sum_{j} (\sum_{i} p(i, j|\theta))^2}{\sum_{i,j} p(i, j|\theta)}$$
, (2.33)

porcentagem de sequência =
$$\sum_{i,j} \frac{p(i,j|\theta)}{n}$$
, (2.34)

 $(\cdot \cdot \cdot | \mathbf{n})$

$$\hat{e}nfase \ de \ sequência \ de \ intensidade \ baixa = \frac{\sum_{i,j} \frac{p(i,j|\theta)}{i^2}}{\sum_{i,j} p(i,j|\theta)}, \tag{2.35}$$

$$\hat{e}nfase \ de \ sequência \ de \ intensidade \ alta = \frac{\sum_{i,j} i^2 p(i,j|\theta)}{\sum_{i,j} p(i,j|\theta)}, \tag{2.36}$$

$$\hat{e}nfase \ de \ intensidade \ baixa \ de \ sequência \ curta = \frac{\sum_{i,j} \frac{p(i,j|\theta)}{i^2 j^2}}{\sum_{i,j} p(i,j|\theta)},$$
(2.37)

$$\hat{e}nfase \ de \ intensidade \ alta \ de \ sequência \ curta = \frac{\sum_{i,j} \frac{p(i,j|\theta)i^2}{j^2}}{\sum_{i,j} p(i,j|\theta)},$$
(2.38)

$$\hat{e}nfase \ de \ intensidade \ baixa \ de \ sequência \ longa = \frac{\sum_{i,j} \frac{p(i,j|\theta)j^2}{i^2}}{\sum_{i,j} p(i,j|\theta)}, \tag{2.39}$$

$$\hat{e}nfase \ de \ intensidade \ alta \ de \ sequência \ longa = \frac{\sum_{i,j} p(i,j|\theta) i^2 j^2}{\sum_{i,j} p(i,j|\theta)}, \qquad (2.40)$$

onde $p(i, j|\theta)$ é o elemento da matriz de comprimento de sequência e n é o número de níveis de cinza da imagem.

De acordo com a lógica da RLM, texturas finas possuem mais sequências curtas com intensidades de nível de cinza similares, enquanto que texturas ásperas possuem mais sequências longas com diferentes níveis de cinza (DAVNALL et al., 2012). Os atributos de ênfase de sequência curta e ênfase de sequência longa medem as distribuições de sequências de uma imagem. As medidas de não-uniformidade caracterizam as similaridades entre os níveis de cinza de uma imagem e entre as sequências de comprimento de intensidades. A medida de porcentagem de sequência é a relação entre o número de sequências e o número total de sequências possíveis, logo, para imagens com estrutura de textura linear, o seu valor resultante é baixo (PHILLIPS et al., 2017; DAVNALL et al., 2012).

Outra estratégia de análise textural baseia-se na matriz de diferença de intensidade de vizinhança (do Inglês *neighborhood intensity difference matrix* - NIDM) (AMADASUN; KING, 1989). Apesar de serem uma abordagem estatística também baseada em matriz, atributos de NIDM são considerados de alta ordem pois examinam relações espaciais entre três ou mais pontos da imagem e não pares como da COM (LUBNER et al., 2017; DAVNALL et al., 2012). A NIDM é uma matriz de uma coluna que calcula a soma das diferenças entre pontos com intensidade *i* de nível de cinza e a média de intensidades da vizinhança desses pontos da imagem (YANG et al., 2016; FRIED, 2015). Similarmente à COM, a NIDM pode ser calculada com vizinhança 2D e 3D. Cinco atributos podem ser calculados a partir da construção da NIDM (Equações 2.41-2.45) (ZHANG et al., 2015; AMADASUN; KING, 1989):

$$granularidade = \left[\epsilon + \sum_{i}^{N_g} P(i)S(i)\right]^{-1},$$
(2.41)

$$contraste = \left[\frac{1}{N_a(N_a - 1)} \sum_{i,j}^{N_g} P(i)P(j)(i - j)^2\right] \left[\frac{1}{N} \sum_{i}^{N_g} S(i)\right],$$
 (2.42)

$$ocupação = \frac{\sum_{i}^{N_g} P(i)S(i)}{\sum_{i,j}^{N_g} [iP(i) - jP(j)]}, P(i) \neq 0, P(j) \neq 0$$
(2.43)

$$complexidade = \sum_{i,j}^{N_g} \frac{|i-j|[P(i)S(i)+P(j)S(j)]}{n^2[P(i)+P(j)]}, P(i) \neq 0, P(j) \neq 0$$
(2.44)

$$for_{\zeta a} = \frac{\sum_{i,j}^{N_g} [P(i) + P(j)](i-j)^2}{\epsilon + \sum_i^{N_g} S(i)}, P(i) \neq 0, P(j) \neq 0$$
(2.45)

onde P(i) é a probabilidade de ocorrência de um nível de cinza i, S(i) é um elemento na NIDM, N_g é o maior nível de cinza da imagem, N_a é o número de níveis de cinza presentes na imagem e ϵ uma constante equivalente à 1E - 6.

O atributo de granularidade é uma medida de densidade de borda da imagem. O contraste mede a variação local de níveis de cinza dentro de uma região. A característica de ocupação mede a taxa espacial de mudança de intensidades. A complexidade calcula a quantidade de informação da textura da região. E por fim, o atributo de força mede o quão distinta é a textura primitiva (PHILLIPS et al., 2017; DAVNALL et al., 2012).

Uma abordagem alternativa à estatística foi proposta por Tamura, Mori e Yamawaki (1978). As características de Tamura teoricamente correspondem a percepção visual humana, como granularidade, contraste, direcionalidade, delineamento, regularidade e rugosidade. Geralmente, os três primeiros atributos (granularidade, contraste e direcionalidade) fornecem uma melhor descrição da textura de uma região (MARVASTI et al., 2013). Contudo, a granularidade é o fator mais fundamental pois quando existem dois padrões com diferentes estruturas na imagem, a granularidade será maior para a textura de tamanho maior ou menos repetida. A direcionalidade é uma propriedade global para uma imagem que envolve a forma e localização da textura, sem levar em consideração a orientação (TAMURA; MORI; YAMAWAKI, 1978).

Uma estratégia que vem sendo recentemente utilizada para extração de atributos estatísticos de textura é baseada na aplicação do filtro Laplaciano de Gaussiano (do Inglês *Laplacian of Gaussian* - LOG), comumente usado em problemas de detecção de bordas (FAVE et al., 2015; FRIED, 2015; MARR; HILDRETH, 1980). Os atributos de LOG realçam as características texturais não visíveis na imagem original em diferentes escalas espaciais (Figura 16) (DAVNALL et al., 2012).

Um filtro LOG primeiro convolve uma determinada imagem com um kernel Gaussiano em uma determinada escala σ e depois o resultado é computado com um operador Laplaciano, de acordo com a distribuição no domínio espacial dada pela Equação 2.46

Figura 16: Aplicação do filtro Laplaciano de Gaussiano em uma imagem de TC de tumor pulmonar: (a) e (b) apresentam as imagens original e filtrada, respectivamente.



Fonte: Ferreira Junior et al. (2018).

(FRIED, 2015; DAVNALL et al., 2012):

$$\nabla^2 G(x,y) = \frac{-1}{\pi\sigma^4} \left(1 - \frac{x^2 + y^2}{2\sigma^2} \right) e^{-\frac{x^2 + y^2}{2\sigma^2}},$$
(2.46)

onde $\nabla^2 G(x, y)$ é o filtro LOG de uma imagem com coordenadas espaciais $x, y \in e^{-\frac{x^2+y^2}{2\sigma^2}}$ é a distribuição Gaussiana. Após a aplicação do filtro Laplaciano de Gaussiano, funções estatísticas de primeira ordem, como as de histograma (Seção 2.4.1), são extraídas da imagem, compondo os atributos de LOG.

Ao contrário das abordagens estatísticas, métodos espectrais de textura analisam as propriedades do domínio da frequência de uma imagem por meio da aplicação de transformadas, como de Fourier, Gabor e *wavelets*. Na transformada de Fourier, imagens são convertidas do domínio espacial para um conjunto de componentes seno e cosseno. A transformada de Fourier analisa o conteúdo da frequência sem considerar localizações espaciais e/ou temporais (DAVNALL et al., 2012). A saída da transformada representa a imagem no domínio de espectro, onde cada ponto significa cada frequência contida no domínio espacial da imagem, revelando informações sobre a periodicidade e direções da textura (Figura 17) (OLIVEIRA, 2006). Assim, ao se analisar o plano das frequências, a identificação dos picos, com sua proeminência e localização, fornece a direção principal das texturas e o período espacial fundamental dos padrões, respectivamente (CHIERICI, 2015). A transformada discreta de Fourier é dada pela Equação 2.47 (PARKER, 2011):

$$F(w) = \sum_{k=0}^{N-1} f(k) e^{\frac{2\pi j w k}{N}}.$$
(2.47)

Figura 17: Aplicação da transformada de Fourier em uma imagem de TC: (a) e (b) apresentam as imagens original e resultante no domínio da frequência, respectivamente.



Fonte: Autor.

Contudo, a computação da transformada de Fourier pode demandar muito tempo. Para simplificá-la, algumas estratégias matemáticas e computacionais foram propostas pela literatura, o que acabou gerando a transformada rápida de Fourier (PARKER, 2011; BRIGHAM, 1974; BRACEWELL, 1965). Após a aplicação da transformada (rápida) de Fourier em uma imagem, funções estatísticas de primeira ordem são extraídas de coordenadas polares dos *pixels* da imagem transformada, compondo os atributos de Fourier.

Outro método espectral de caracterização de texturas é baseado nos filtros de Gabor. Diferentes tipos de texturas podem ser descritas pela freqüência e orientação de funções senoidais resultantes da aplicação dos filtros de Gabor em uma imagem. O filtro de Gabor é baseado essencialmente em uma transformada janelada de Fourier derivada da introdução de uma função Gaussiana (DAVNALL et al., 2012). Isto permite a representação espacial, temporal e da frequência de um sinal, indicando características locais da textura, que por sua vez permitem a manipulação de parâmetros, como frequência, orientação, excentricidade e simetria (KOLOSSVÁRY et al., 2018; CHIERICI, 2015). A formulação no domínio espacial do filtro de Gabor é dada pela Equação 2.48 (BIANCONI; FERNÁNDEZ, 2007):

$$\psi(x,y) = \frac{F^2}{\pi\gamma\eta} e^{-F^2[(x'/\gamma)^2 + (y'/\eta)^2]} e^{i2\pi Fx'},$$
(2.48)

onde $x' \notin x \cos \theta + y \sin \theta$, $y' \notin -x \sin \theta + y \cos \theta$, $F \notin a$ frequência central do filtro, $\theta \notin a$ ô ângulo entre a direção da onda senoidal e o eixo x do domínio espacial, $\gamma \in \eta$ são os

desvios padrões da Gaussiana.

Após a aplicação de um banco de filtros de Gabor em uma imagem (Figura 18), funções estatísticas de primeira ordem, como média e desvio padrão, são extraídas das imagens filtradas, compondo os atributos de Gabor.

Figura 18: Exemplo de um banco de filtros de Gabor em diferentes escalas e orientações.





A transformada de Gabor é limitada pela resolução espacial da janela única, ao contrário do que acontece com a transformada *wavelet* que utiliza múltiplas funções escaláveis e transladáveis para diferentes frequências, o que pode permitir a representação da textura em uma escala mais adequada (KOLOSSVÁRY et al., 2018; CHIERICI, 2015). Existem diversas funções de *wavelet* na literatura, porém todas são derivadas de $\Psi(t)$

(Equação 2.49) (DAVNALL et al., 2012):

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), \quad a,b \in R, a \neq 0,$$
(2.49)

onde a é o parâmetro escalar que mede o grau de compressão e b é o parâmetro de translação que indica a localização temporal da *wavelet*.

A transformada de Haar é uma das funções *wavelet* mais utilizadas e, quando aplicada, pode-se obter a energia (Equação 2.9) em bandas de frequências alta e baixa em diferentes níveis do espectro (FALEIROS et al., 2017; GILLIES; KINAHAN; HRICAK, 2016; AZEVEDO-MARQUES et al., 2015; AERTS et al., 2014).

Por fim, uma abordagem alternativa a estatística e espectral de caracterização de textura é baseada em geometria estrutural fractal (MANDELBROT, 1984). Este modelo é baseado na premissa de que muitas regiões têm um padrão estatístico de rugosidade ou irregularidade em diferentes escalas (CHIERICI, 2015). As medidas fractais representam diferentes aspectos do padrão espacial da imagem, proporcionando importantes informações de heterogeneidade da região (KOLOSSVÁRY et al., 2018; DAVNALL et al., 2012).

A análise fractal pode ser realizada por meio da estimativa de sua dimensão, que é uma medida de como um objeto preenche espaços; abundância, que é uma medida de volume de espaço preenchido; e lacunaridade, que é uma medida de heterogeneidade estrutural dentro do objeto (DAVNALL et al., 2012). Existem diversas formas de calcular medidas fractais e um dos métodos mais simples é conhecido como contagem por caixas. Este método calcula quantos pontos são ocupados por um objeto em diferentes escalas. A estimativa da dimensão fractal pela contagem por caixas é o valor da inclinação do gráfico log-log do número de pontos ocupados *versus* as diferentes escalas (KOLOSSVÁRY et al., 2018).

2.4.3 Atributos de Forma

O termo forma refere-se à informação que pode ser inferida do VOI, mas que não é representada por cor e textura. Atributos de forma descrevem a borda da imagem e características geométricas extraídas do objeto segmentado, como contornos, junções, curvas e regiões polígonas (Figura 19) (ECHEGARAY et al., 2015; AKGÜL et al., 2011). Caracterizar quantitativamente formas de objetos é uma tarefa complicada porque depende da eficiência dos algoritmos de segmentação de imagens para definir o VOI (Seção 2.3). Os objetos em imagens radiológicas do pulmão, por exemplo, muitas vezes podem apresentar opacidades anexadas a eles e padrões com baixo contraste de nível de cinza, como o vidro fosco (Figura 20), o que pode acarretar em baixa performance de segmentação e, consequentemente, de caracterização por atributos de forma.

Tradicionalmente, atributos de forma podem ser divididos em dois tipos principais: contorno ou região. Nos descritores baseados em contorno, as características são obtidas



Figura 19: Heterogeneidade morfológica de tumores pulmonares.

Fonte: Aerts et al. (2014).

analisando as coordenadas da borda do objeto. Enquanto que nos descritores baseados em região, as características são extraídas da região interior do contorno fechado (OLIVEIRA, 2006). Alguns atributos baseados em forma foram utilizados na caracterização de lesões em imagens radiológicas, como máximo diâmetro 3D (maior distância Euclidiana entre *voxels* de um VOI), volume (número de *voxels* de uma região multiplicado pelo tamanho do *voxel*), área da superfície, compacidade, disproporção esférica, esfericidade e densidade (Equações 2.50-2.55) (ZHANG et al., 2015; AERTS et al., 2014):

área da superfície
$$(A) = \sum_{i=1}^{N} \frac{1}{2} |a_i b_i \times a_i c_i|,$$
 (2.50)

$$compacidade1 = \frac{V}{\sqrt{\pi}A^{\frac{2}{3}}},\tag{2.51}$$

$$compacidade2 = 36\pi \frac{V^2}{A^3},\tag{2.52}$$

disproporção esférica =
$$\frac{A}{4\pi R^2}$$
, (2.53)

$$esfericidade = \frac{\pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}}{A},$$
 (2.54)

$$densidade = \frac{A}{V},\tag{2.55}$$

Figura 20: Segmentação semi-automática de nódulos pulmonares subsólidos (marcação em laranja) e de componentes sólidos (marcação em amarelo) em imagens de TC:
(a) e (b) apresentam casos em que as segmentações apresentaram performance satisfatória e insatisfatória, respectivamente.



Fonte: Cohen et al. (2016).

onde V é o volume, R é o raio esférico, N é o número total de triângulos conectados que cobrem a superfície do objeto, a, b e c são os vértices destes triângulos. Os atributos de máximo diâmetro 3D, volume e área da superfície fornecem informações do tamanho do objeto. Medidas de compacidade, disproporção esférica, esfericidade e densidade descrevem o quanto o VOI é esférico, arredondado ou alongado (ZHANG et al., 2015; AERTS et al., 2014).

Descritores que medem a nitidez da borda de uma região são uma subcategoria especial pertencente aos atributos de forma. Um atributo de nitidez de borda é importante para diferenciar lesões em termos de malignidade pois neoplasias malignas crescem e invadem os tecidos vizinhos (LEVMAN; MARTEL, 2011). Os primeiros atributos criados para quantificar margens foram propostos por Gilhuijs, Giger e Bick (1998), em imagens de ressonância magnética (RM) de mama. Estes atributos são baseados na média e variância do gradiente espacial da borda da região (Equações 2.56 e 2.57) (XU et al., 2012):

média do gradiente de borda =
$$\max_{i=0, \dots, M-1} \left\{ \frac{\operatorname{mean}_r ||\nabla I_m(r)||}{\operatorname{mean}_r I_m(r)} \right\},$$
(2.56)

variância do gradiente de borda =
$$\frac{\operatorname{var}_r ||\nabla I_m(r)||}{[\operatorname{mean}_r I_m(r)]^2},$$
 (2.57)

onde $I_m(r)$ é a intensidade de nível de cinza de um *voxel*, e a amplitude do vetor r em $I_m(\cdot)$ está limitada a superfície da região.

Levman e Martel (2011) também desenvolveram um método para quantificar a nitidez da borda de lesões da mama em imagens de RM. Contudo, os *voxels* utilizados na

abordagem dos autores pertencem tanto ao interior quanto ao extremos vizinhos exteriores à região, ao contrário do método de Gilhuijs, Giger e Bick (1998). O atributo de nitidez de borda de Levman e Martel (2011) é definido pela Equação 2.58 (XU et al., 2012):

$$nitidez \ de \ borda = \frac{\overline{I(r_i) - I(N(r_i))}}{d}, \tag{2.58}$$

onde $I(r_i)$ é a intensidade de cinza de *voxels* do interior da região, $N(r_i)$ é o operador de vizinhança tridimensional que fornece as posições dos *voxels* vizinhos a r_i que são externos à região e d é o termo de normalização.

Xu et al. (2012) definiram a nitidez de borda de nódulos pulmonares em imagens de TC em duas características. A primeira mede a diferença de nível de cinza entre os *pixels* dos tecidos do pulmão e da lesão. A segunda mede a transição de intensidade dos *pixels* sobre a borda do nódulo. Os dois atributos são extraídos de uma função sigmóide aplicada em um vetor de intensidades formado pelos *pixels* presentes em linhas ortogonais traçadas sobre a borda do nódulo pulmonar (Figura 21). Para cada linha L, o problema pode ser formulado de acordo com a Equação 2.59 (XU et al., 2012):

$$\arg\min_{S,W,x_o,L_o} \sum_{x} \left[L(x) - L_o - \frac{S}{1 + e^{-\frac{x - x_o}{W}}} \right],$$
(2.59)

onde x é a distância ao longo da normal, x_o é a intersecção entre o ponto na borda e a normal, L(x) é a intensidade ao longo da normal em x, e L_o é a intensidade do offset.

Figura 21: Linha ortogonal traçada sobre a margem de nódulo pulmonar em imagem de TC para cálculo de atributo de nitidez de borda.



Fonte: Xu et al. (2012).

2.5 Seleção de Atributos Relevantes

Dada esta grande gama de atributos que podem ser extraídos, alguns problemas podem ocorrer na consequente análise quantitativa das imagens. Muitas vezes as características podem ser irrelevantes para uma determinada tarefa, e.g. atributos de forma para o reconhecimento de padrões benignos e malignos de fraturas de vértebras em imagens de RM (Figura 15) (FRIGHETTO-PEREIRA et al., 2016). Em outras, os dados podem apresentar redundância, inserindo ruído ou inconsistências ao vetor de características da imagem (Figura 12) (CATALDO et al., 2014). Outro problema bem comum é conhecido por maldição da dimensionalidade, onde a quantidade de atributos é desproporcionalmente maior que o número de amostras, o que pode dificultar o processo de análise estatística e mineração dos dados devido à alta complexidade do espaço *n*-dimensional (LARUE et al., 2017; AERTS, 2016).

Desta forma, é necessário realizar uma seleção das características mais relevantes do vetor de atributos da imagem. Diversos algoritmos foram criados com o objetivo de reduzir a dimensionalidade espacial do vetor. Basicamente, estes métodos podem ser classificados em três tipos principais: filtro, *wrapper* e embarcado (CONSTANZO et al., 2017; WITTEN; FRANK; HALL, 2011). Os filtros avaliam a relevância dos atributos por meio de heurísticas baseadas em características dos dados, como correlação e informação mútua (PENG; LONG; DING, 2005; HALL, 1999; KONONENKO, 1994). Os algoritmos *wrapper* avaliam o potencial dos atributos em conjunto com um método de aprendizado de máquina pré-definido (Seção 2.6) (KOHAVI; JOHN, 1997; JOHN; KOHAVI; PFLEGER, 1994). Os métodos embarcados avaliam os atributos durante o processo de treinamento de classificadores de aprendizado de máquina e levam em consideração a performance preditiva do classificador na seleção das características mais relevantes (Figura 22). Alguns destes algoritmos que realizam a seleção interna dos atributos são redes neurais artificiais, árvores de decisão e florestas aleatórias, explorados em seções seguintes desta tese.

O filtro baseado em correlação busca por subconjuntos de atributos que são altamente correlacionados com uma classe (desfecho), sem haver intercorrelação entre eles (HALL, 1999). Para isso, ele avalia o potencial de subconjuntos de características considerando a habilidade individual preditiva de cada atributo junto com o grau de redundância entre eles (FRANK; HALL; WITTEN, 2016). Após a execução de um filtro baseado em correlação em um conjunto de vetores de atributos de imagens, são criados novos vetores para cada imagem com um subconjunto de atributos selecionados.

Contudo, existem filtros que não realizam a eliminação de características e consequente seleção do melhor subconjunto, mas sim o ranqueamento de atributos de acordo com a relevância de cada um atribuída por uma heurística, como os algoritmos Relief e ReliefF (KONONENKO, 1994; KIRA; RENDELL, 1992). Basicamente estes métodos estimam os atributos de acordo com o potencial dos seus valores em distinguir amostras

Figura 22: Funcionamento de algoritmos de seleção de atributos relevantes: (a) filtro, (b) *wrapper*, (c) embarcado.



Fonte: Autor.

mais próximas, utilizando uma métrica de distância, e.g. Euclidiana. Uma das diferenças principais entre esses algoritmos é que, enquanto o Relief lida apenas com problemas binários, o ReliefF permite lidar com problemas multiclasse. Para atribuir uma relevância (peso) a um atributo, o filtro ReliefF escolhe aleatoriamente amostras e atualiza seu peso baseado nas amostras mais similares de classes iguais e diferentes. Para reduzir a dimensionalidade espacial do vetor inicial, após a atribuição da relevância aos atributos basta selecionar uma quantidade de características ranqueadas pelo peso para a criação dos novos vetores de atributos de imagens.

O algoritmo *wrapper* avalia subconjuntos de atributos utilizando um esquema de aprendizado de máquina e validação-cruzada (Seção 2.6.2) para estimar a precisão deste esquema para um conjunto de características selecionadas (KOHAVI; JOHN, 1997; JOHN; KOHAVI; PFLEGER, 1994). Como cada classificador apresenta um mecanismo diferente de treinamento, o método *wrapper* utiliza diferentes algoritmos de indução para estimar o mérito dos subconjuntos de atributos (HALL, 1999). Uma das desvantagens apresentadas pelos métodos *wrapper* é que, por utilizar um processo iterativo de aprendizado de máquina e validação-cruzada, eles podem possuir um grande custo computacional, principalmente para bases de *Big Data*, como as existentes nos dias atuais (MORRIS et al., 2018). Além disso, eles possuem menor capacidade de generalização do que os filtros por estarem acomplados a um classificador específico (HALL, 1999).

2.6 Classificação de Imagens Médicas e Aprendizado de Máquina

Uma das áreas mais estudadas em inteligência artificial e classificação de imagens médicas que continua atraindo muitos pesquisadores é a de aprendizado de máquina. Esta técnica permite a identificação de padrões a partir de casos anteriores, assim como ocorre com a inteligência humana (ERICKSON et al., 2017). No contexto radiológico, a junção de atributos extraídos das imagens e algoritmos de aprendizado de máquina pode permitir o reconhecimento robusto de padrões não existentes visualmente nas imagens ou que podem passar despercebidos mesmo por especialistas experientes (CHOY et al., 2018; AERTS, 2016).

Métodos de aprendizado de máquina têm sido aplicados para classificação de diferentes modalidades de imagens médicas, utilizando uma variedade de atributos, para diversas doenças (GIGER, 2018; CHOY et al., 2018; ERICKSON et al., 2017). Para a avaliação do câncer de pulmão em imagens de TC, alguns trabalhos mostraram a eficácia destes métodos em conjunto com características quantitativas no suporte a decisões clínicas, seja para CAD ou para radiômica (Seção 1.3).

A modelagem de um método de aprendizado de máquina envolve a criação de uma função para um conjunto de dados (os vetores de atributos, no caso de classificação de imagens médicas) de treinamento, fazendo-se uso de um mecanismo de inferência lógica. Cada algoritmo utiliza um mecanismo diferente de inferência indutiva para realizar o treinamento e construção do modelo preditivo. Quando os grupos, conhecidos como classes, para a decisão final do modelo estão estabelecidos, o processo de treinamento é supervisionado. Quando não existe classe definida para cada amostra de treinamento, o processo é não-supervisionado. Neste último caso, o algoritmo visa a formação de agrupamentos (do Inglês *clustering*) de amostras para o reconhecimento de padrões. Para o treinamento supervisionado, existem dois tipos de tarefas que o algoritmo pode realizar: classificação, quando a classe possui um valor discreto (e.g. benigno *vs.* maligno), e regressão, quando a classe possui um valor contínuo¹.

¹ Computacionalmente, a radiômica é uma tarefa supervisionada, ou seja, de classificação ou regressão. Contudo, como ela é um processo de associação entre atributo e padrão (muitas vezes não presente visualmente na imagem, e.g. resposta do tratamento ou recorrência da lesão) visando uma informação prognóstica, nesta tese ela não será denominada como uma tarefa de classificação ou regressão. Tradicionalmente, o CAD que pode ser denominado como uma tarefa de classificação de imagens porque visa uma resposta diagnóstica baseada em um padrão presente visualmente na imagem (e.g. patologia da lesão).

Diversos classificadores foram apresentados pela literatura de aprendizagem de máquina. Alguns dos mais utilizados são *k-nearest neighbors* (KNN), *naive Bayes* (NB), *support vector machine* (SVM), redes neurais artificiais, árvores de decisão e florestas aleatórias (LEGER et al., 2017; PARMAR et al., 2015a; WITTEN; FRANK; HALL, 2011).

O KNN é um dos métodos mais simples da literatura. Ele é um classificador nãoparamétrico que escolhe a classe majoritária entre k amostras mais próximas para classificar um caso novo com classe desconhecida (AHA; KIBLER; ALBERT, 1991). O cálculo de proximidade entre amostras é realizado por alguma métrica de distância utilizando os vetores de atributos das imagens. A Seção 2.8 traz alguns exemplos de funções de distância mais detalhadamente. A quantidade de amostras mais próximas é primordial para a eficiência de classificação do KNN (Figura 23), logo, deve-se fazer o ajuste fino deste classificador no valor definido para k.

Figura 23: Exemplo de classificação com KNN em espaço bidimensional. O caso novo desconhecido será classificado como B por existirem 3 amostras (dentre 5) da Classe B mais próximas.



Fonte: Kuhn e Johnson (2013).

O NB é um classificador probabilístico baseado no teorema de Bayes e supõe a independência dos atributos (JOHN; LANGLEY, 1995). Apesar desta suposição, ele tem se mostrado eficiente quando os atributos não apresentam independência, o que acontece na maioria dos casos, inclusive em estudos de classificação de imagens médicas e radiômica (EMAMINEJAD et al., 2016; WU et al., 2016b; FRIGHETTO-PEREIRA et al., 2016). Além disso, o NB possui a vantagem de não requerer grandes quantidades de amostras de treinamento para estimar as probabilidades de cada atributo pertencer a uma classe (WU et al., 2016b). Para realizar esta estimativa, o NB supõe que o atributo possui uma distribuição normal (ou Gaussiana), que obedece à Equação 2.60 (WITTEN; FRANK; HALL, 2011):

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$
(2.60)

onde x é o valor do atributo de um caso novo com classe desconhecida, $\mu \in \sigma$ são a média e o desvio padrão, respectivamente, dos valores do atributo das amostras de treinamento para uma determinada classe. A probabilidade final de um caso novo pertencer a esta classe é dada pelo produto das estimativas de todos os atributos do vetor da imagem.

O SVM é um dos classificadores mais tradicionais da literatura de aprendizagem de máquina. Ele tenta encontrar um hiperplano para separar o conjunto de dados em classes discretas (Figura 24) (VAPNIK, 2013). Para isso, ele tenta maximizar a distância dos vetores de atributos das amostras para o hiperplano, ou seja, aumentando a margem de decisão entre as classes (Figura 25) (KAVZOGLU; COLKESEN, 2009).

Figura 24: Mecanismo de criação de hiperplano do método SVM para classificação.



Fonte: Constanzo et al. (2017).

Modelos de árvore de decisão para aprendizado de máquina são compostos por estruturas de dados chamadas de folhas (que indicam uma classe) e nós de decisão (que especificam alguma condição para um atributo) (QUINLAN, 1993). Em uma árvore de decisão, a classificação de um caso desconhecido inicia-se pela raiz da árvore que também é um nó de decisão, atravessando-a pelos nós intermediários até uma folha ser alcançada para a atribuição de uma classe para este caso novo (Figura 26) (HAWKINS et al., 2014).

Figura 25: Representação de hiperplano e margem obtidos pelo classificador SVM em um problema linearmente separável.



Fonte: Kavzoglu e Colkesen (2009).





Fonte: Autor.

O método da floresta aleatória (do Inglês *random forest* - RF) é um dos métodos mais eficientes já criados para tarefas de classificação e regressão. Ele é um classificador que combina várias árvores de decisão para aumentar a performance preditiva (BREIMAN, 2001). Cada árvore depende dos valores de vetores de atributos aleatórios que foram gerados independentemente. Após a criação das árvores, um caso novo desconhecido é classificado em cada árvore de forma independente, e ele tem sua classificação final determinada pela classe majoritária escolhida dentre todas as árvores da floresta.

Redes neurais artificiais (RNAs) são um dos mais tradicionais métodos de aprendizado de máquina e são bastante utilizados em tarefas de classificação de imagens médicas (GIGER, 2018; CHOY et al., 2018; ERICKSON et al., 2017). As RNAs foram propostas tendo como referência estruturas do sistema nervoso humano (HAYKIN, 1994). Estas redes são modelos matemáticos descritos em forma de camadas, em que camada é composta por neurônios. O tipo de RNA mais conhecido é o *perceptron* de multicamadas (do Inglês *multilayer perceptron* - MLP). Tradicionalmente, o MLP possui uma camada de entrada (cujos neurônios correspondem aos atributos do vetor da imagem), uma camada de saída (cujos neurônios correspondem às classes) e um conjunto de camadas ocultas intermediárias (cujos neurônios correspondem a funções de ativação) (Figura 27). Cada neurônio de uma camada oculta recebe como entrada uma soma ponderada das saídas dos neurônios da camada anterior. A diferença entre os valores de saída produzidos e os desejados para cada neurônio da camada de saída indica o erro apresentado pela rede. O valor deste erro é então utilizado no ajuste dos pesos das camadas anteriores da RNA. Este processo iterativo de ajuste de pesos com fase para frente e para trás da rede neural é conhecido por backpropagation.





Fonte: Gerônimo et al. (2013).

Um subtipo de RNA que historicamente possui boa performance preditiva é conhecido como RBF (do Inglês *radial basis function*), devido ao fato dos neurônios da camada oculta possuírem funções de base radial como função de ativação (MOODY; DARKEN, 1989). Tradicionalmente, as redes RBF possuem três camadas: as de entrada e saída iguais do MLP, e apenas uma intermediária oculta com funções radiais. A diferença principal entre redes neurais do tipo MLP e RBF é que a primeira utiliza hiperplanos para gerar as fronteiras de decisão, e a segunda utiliza hiperelipsóides (Figura 28) (SCHWENKER; KESTLER; PALM, 2001). Diversos métodos foram propostos para gerar as fronteiras de decisão hiperelipsoidais, sendo os mais utilizados: *k-means* de agrupamento, mapa auto-organizável de Kohonen e algoritmos genéticos (NERUDA; KUDOVÁ, 2005; SCHWENKER; KESTLER; PALM, 2001).

Figura 28: Representação de fronteiras de decisão para redes neurais artificiais: (a) MLP, (b) RBF.





Um dos grandes desafios em aprendizagem de máquina é quando ocorre o sobreajuste do modelo ao conjunto de dados utilizado no treinamento (Figura 29). Caso o modelo treinado esteja sobreajustado, pode ocorrer dele não conseguir generalizar a solução de um problema, e assim, ele não consegue classificar corretamente um caso novo desconhecido. Uma das soluções possíveis para reduzir este problema é realizar a redução da dimensionalidade do vetor de atributos da imagem por meio da seleção das características mais relevantes (Seção 2.5), de modo que o número de atributos seja menor que o número de amostras utilizadas no treinamento (ERICKSON et al., 2017). Outra estratégia é a de empregar um conjunto separado de amostras para utilizar somente para testes e avaliar a performance final do modelo neste conjunto, abordagem esta explicada posteriomente na Seção 2.6.2.

2.6.1 Aprendizado Profundo

Uma das áreas de inteligência artificial mais recentes que vem ganhando atenção da comunidade científica é a de aprendizado profundo. Esta técnica pertence a uma classe de algoritmos do tipo de aprendizagem de representação em que eles aprendem uma composição de características, das mais simples às mais complexas, que refletem uma hierarquia de estruturas nos dados (CHARTRAND et al., 2017; LITJENS et al., 2017). Historicamente, métodos tradicionais de aprendizado de máquina possuem limitações no processamento de dados em sua forma mais bruta, como os *pixels* das imagens. E devido a isto, por décadas existiu a necessidade de se desenvolver extratores de atributos para





Fonte: Erickson et al. (2017).

representar as imagens e servir de entrada para os classificadores (LECUN; BENGIO; HINTON, 2015). Logo, pesquisadores viram essa necessidade como uma oportunidade de desenvolver algoritmos que integrassem os processos de extração de atributos e classificação de imagens. Além disso, estes métodos podem fornecer um novo nível de representação às imagens não oferecido pelas técnicas tradicionais de visão computacional (Seção 2.4).

O método de aprendizado profundo que mais ganhou notoriedade no processamento e análise de imagens é a rede neural convolucional (do Inglês *convolutional neural network* -CNN). Uma CNN é composta basicamente por três tipos de camadas: convolucional, *pooling* e totalmente conectada (Figura 30). A camada convolucional tem como objetivo detectar elementos locais distintos, como linhas, bordas, texturas, etc. Para isto, ela aplica diversos filtros, também chamados de *kernels*, sobre toda a imagem, gerando múltiplos mapas de características. A camada *pooling* reduz a quantidade destes mapas de características, servindo como método de redução de dimensionalidade espacial da imagem, sem perder robustez de representação. As operações mais comuns neste tipo de camada são o valor máximo (*pooling* máximo) e a média de ativação (*pooling* médio). A camada totalmente conectada serve para integrar todas as características extraídas pelas camadas anteriores convolucionais e *pooling* e representadas pelos mapas (LEE et al., 2017b). Normalmente, esta última camada da CNN é composta por uma rede neural do tipo MLP para realizar a classificação final da imagem, dada pela predição da classe com maior probabilidade.

2.6.2 Validação e Avaliação de Performance

Dado um conjunto de amostras de imagens, um classificador de aprendizado de máquina deve utilizar no mínimo dois subconjuntos diferentes dele para realizar o treinamento do modelo e a predição para fins de validação deste modelo preditivo. A validação *holdout* é a mais simples estratégia para aprendizagem de máquina e é baseada na separação destes subconjuntos em treino e teste (Figura 31(a)). Contudo, ela não é



Figura 30: Imagem de TC como entrada em uma rede neural convolucional.

Fonte: Chartrand et al. (2017).

muito utilizada na radiologia pois nem todas as amostras são testadas, e consequentemente, nem todos os pacientes são avaliados (KEEK et al., 2018; LARUE et al., 2017). Uma estratégia muito utilizada para validar algoritmos preditivos é a validação-cruzada. Nela, as amostras são separadas em n subconjuntos, sendo n-1 para treino e 1 para teste (Figura 31(b)). A validação-cruzada é realizada até que todas as amostras do conjunto inicial sejam testadas. Quando n é igual ao número total do conjunto de amostras, a validação-cruzada é chamada de *leave-one-out*. Entretanto, um problema das estratégias de validação-cruzada é que elas tendem a produzir resultados superestimados na avaliação de performance do processo de reconhecimento de padrões (LAMBIN et al., 2017; AERTS, 2016). As abordagens de validação mais apropriadas, visando aumentar as chances de generalização da solução de um problema (e diminuição do risco de sobreajuste do modelo), são baseadas em três subconjuntos de amostras: um para treino, um para validação e um independente somente para teste² (Figura 31(c)) (ERICKSON et al., 2017; CHARTRAND et al., 2017). Enquanto o subconjunto de validação é utilizado na avaliação preliminar e ajuste fino do modelo treinado com o primeiro subconjunto de amostras, um terceiro subconjunto independente é utilizado somente para testes, para obter a performance final do modelo. Uma outra estratégia que pode ser empregada é separar o conjunto de amostras em dois subconjuntos: no primeiro realiza-se uma validação-cruzada para o ajuste fino do classificador e no segundo subconjunto realiza-se somente os testes de avaliação de performance do modelo preditivo (Figura 31(d)).

Algumas das métricas que podem ser utilizadas na avaliação de performance do modelo preditivo são acurácia, sensibilidade, especificidade e área sob a curva ROC

² Em alguns casos, os termos "validação" e "teste" podem ser invertidos, dependendo da área de atuação em que o pesquisador se encontra. Na estatística/saúde, é comum encontrar o termo "conjunto de validação" como sendo o subconjunto independente. Nas engenharias/exatas, é comum encontrar o termo "conjunto de teste" como sendo o subconjunto independente. Nesta tese, será utilizada esta última nomenclatura.

Figura 31: Exemplos de validação em aprendizagem de máquina: (a) holdout, (b) validaçãocruzada com 5 subconjuntos, (c) aprendizagem com subconjunto de validação, (d) aprendizagem com validação-cruzada e subconjunto de validação.



Fonte: Autor.

(do Inglês receiver operating characteristic). Acurácia refere-se à proporção de amostras classificadas corretamente (Equação 2.61). Sensibilidade refere-se à proporção de amostras positivas que foram classificadas corretamente (Equação 2.62). Especificidade refere-se à proporção de amostras negativas que foram classificadas corretamente (Equação 2.63). A curva ROC é definida como um gráfico de um teste diagnóstico onde no eixo y é apresentada a sensibilidade ou taxa de verdadeiro positivo e no eixo x é apresentada a taxa de falso positivo representada por 1 - especificidade (Figura 32). O valor da AUC (do Inglês area under the ROC curve) corresponde à área da região sob a curva ROC.

$$acurácia = \frac{VP + VN}{VP + VN + FP + FN},$$
(2.61)

$$sensibilidade = \frac{VP}{VP + FN},\tag{2.62}$$

$$especificidade = \frac{VN}{VN + FP},\tag{2.63}$$

onde VP, VN, FP e FN são o número de verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos, respectivamente, de acordo com a matriz de confusão apresentada na Tabela 3.

Figura 32: Exemplo de curvas ROC: a curva 1 representa um teste com performance perfeita de área igual a 1; a curva 2 representa um teste com performance intermediária de área igual a 0,75; e a curva 3 é um teste aleatório com área igual a 0,5.



Fonte: Autor.

Tabela 3: Matriz de confusão.

		Predição	
		Positivo	Negativo
al	Positivo	VP	FN
Re	Negativo	FP	VN

VP, VN, FP e FN são verdadeiros positivos e negativos, falsos positivos e negativos, respectivamente. Fonte: Autor.

2.7 Análise de Sobrevida (Tempo-para-Evento)

A análise de sobrevida é derivada de um procedimento analítico específico que constitui uma variável que relaciona o tempo entre o início da observação em um estudo até a ocorrência de um determinado evento (FERREIRA; PATINO, 2016). Apesar do nome "sobrevida", este evento não necessariamente está vinculado somente a óbito, mas a qualquer desfecho clínico, como por exemplo, tempo para recidiva, tempo para hospitalização, entre

outros (GEORGE; SEALS; ABAN, 2014). Historicamente, a sobrevida global é considerada o desfecho temporal de maior importância para a rotina clínica, pois é objetiva, inequívoca e facilmente calculável (CHEEMA; BURKES, 2013; MACHADO et al., 2010). Neste caso, o evento de interesse é o óbito do paciente por qualquer natureza. Na análise de sobrevida, os pacientes que não desenvolveram o evento ou tiveram perda de acompanhamento são considerados *censurados*, pois o tempo exato da sobrevida é desconhecido (FERREIRA; PATINO, 2016).

A estimativa de Kaplan-Meier é o método estatístico mais utilizado para análises do tipo tempo-para-evento (KAPLAN; MEIER, 1958). Uma vez definido o evento de interesse e as razões para censura dos pacientes, é possível construir curvas probabilísticas que descrevem a taxa de ocorrência do desfecho ao longo do tempo, podendo servir para comparação de diferentes grupos de pacientes (MACHADO et al., 2010). Para calcular as probabilidades de ocorrência de um evento, considere, por exemplo, dois grupos de 21 pacientes cada, sendo um grupo controle de pacientes que não recebeu tratamento e um grupo de pacientes que realizou terapia com uma droga x, e os tempos da Tabela 4 para a realização do evento. A probabilidade de ocorrência do evento, ou probabilidade de sobrevida, S(t) em um instante t computacionalmente pode ser dada pela Equação 2.64:

$$S(t) = S(t-1) \times \frac{N_s}{N_r},$$
(2.64)

onde S(t-1) é a probabilidade de sobrevida no instante anterior à t, N_r é o número de pacientes ("em risco") que restam no estudo eliminando os casos que foram censurados no instante anterior à $t \in N_s$ é o número de pacientes ("sobreviventes") que não sofreram o evento até o instante t. As curvas Kaplan-Meier para os grupos Controle e Tratamento do exemplo anterior estão na Figura 33.

Tabela 4: Exemplos de tempos de sobrevida para dois grupos de pacientes: um grupo controle de pacientes sem terapia e um grupo que realizou algum tratamento. Os pacientes censurados estão com um asterisco no tempo.

Grupo	Tempos (semanas)
Controle	1, 2, 3, 3, 3, 4, 4, 5, 6, 8, 8, 9, 10, 11, 11, 12, 13, 15, 17, 20, 25
Tratamento	$3^*, 5, 6, 7, 7, 9^*, 10^*, 11, 12^*, 13, 15, 17^*, 18^*, 20^*, 21, 24, 28^*, 30^*, 32^*, 33^*, 34^*$

Fonte: Autor.

2.8 Recuperação de Imagens Médicas Baseada em Conteúdo

Métodos de CBIR surgiram na medicina devido à necessidade de buscar informações contidas na imagem e não apenas as armazenadas nos sistemas de informação em saúde e prontuário eletrônico do paciente (AZEVEDO-MARQUES; RANGAYYAN, 2013; OLIVEIRA; CIRNE; AZEVEDO-MARQUES, 2007). O conceito de CBIR refere-se à busca

Figura 33: Curvas Kaplan-Meier para os grupos de pacientes da Tabela 4. Cada traço na curva do grupo Tratamento representa um paciente censurado.



Fonte: Autor.

por imagens cujos conteúdos são similares a um caso de referência utilizando informações derivadas das próprias imagens, representadas por seu conteúdo intrínseco (vetor de atributos), ao invés de índices de textos associados ou anotações externas (SOUZA, 2012).

A CBIR despertou rápido interesse na comunidade médica em virtude da sua capacidade de recuperar imagens já diagnosticadas, permitindo ao especialista confirmar uma hipótese diagnóstica de uma imagem em estudo ou buscar informações prognósticas de casos com padrão visual semelhante a esta imagem de referência. Ainda no contexto clínico, a CBIR também pode servir para o monitoramento da progressão de uma determinada doença (OLIVEIRA, 2006). Além de suporte à tomada de decisão clínica, sistemas de recuperação por conteúdo podem auxiliar no ensino (MA et al., 2017; DHARA et al., 2017a; AZEVEDO-MARQUES; RANGAYYAN, 2013), podendo resultar no aprimoramento da qualidade educacional em radiologia por meio da utilização de repositórios de imagens para servir de base para medicina baseada em evidências.

A similaridade entre imagens em sistemas de CBIR pode ser medida por uma função de distância computada entre os vetores de atributos da imagem de referência e das imagens armazenadas nos sistemas PACS ou de informação em radiologia (LUCENA et al., 2016; KUMAR et al., 2013). A escolha da métrica de distância depende do tipo de imagem médica utilizada, dos descritores extraídos e de sua representação (AKGÜL et al., 2011). As principais medidas de similaridade utilizadas em sistemas de recuperação por conteúdo pertencem a um conjunto de funções de distância conhecido como família de distâncias Minkowski (PEDROSA, 2015). Esta família de medidas é definida por uma função L_p entre dois vetores x e y de tamanho n, de acordo com a Equação 2.65:

$$L_p(x,y) = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p}.$$
(2.65)

As distâncias pertencentes à família Minkowski são variantes da função L_p , de acordo com o valor atribuído a p. A função de distância L_1 , conhecida como Manhattan, utiliza a soma das diferenças absolutas entre os valores dos atributos dos vetores das imagens. A distância Manhattan d_M entre dois vetores $x \in y$ de tamanho n está definida na Equação 2.66:

$$d_M(x,y) = \sum_{i=1}^n |x_i - y_i|.$$
(2.66)

A função de distância L_2 , conhecida como Euclidiana, é a medida de similaridade mais tradicional empregada em sistemas de CBIR. A distância Euclidiana d_E entre os dois vetores $x \in y$ de tamanho n está definida na Equação 2.67. As funções $L_1 \in L_2$ são medidas inversamente proporcionais à similaridade de imagens. Ou seja, quanto menor a distância entre os vetores de atributos de duas imagens, maior a similaridade entre elas e vice-versa.

$$d_E(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$
(2.67)

2.8.1 Avaliação de Performance

Existem diversas formas de avaliar a performance de uma ferramenta de recuperação por conteúdo. Os métodos mais tradicionais e comuns utilizados na avaliação de eficiência em CBIR são as métricas de precisão e revocação (AZEVEDO-MARQUES; RAN-GAYYAN, 2013; KUMAR et al., 2013). Precisão é a proporção de imagens recuperadas que são relevantes para uma consulta por similaridade (Equação 2.68). Revocação indica a proporção de imagens relevantes nos sistemas de armazenamento que foram recuperadas ao responder uma consulta, de acordo com a Equação 2.69:

$$precisão = \frac{I_R}{I_O},\tag{2.68}$$

$$revocação = \frac{I_R}{I_T},\tag{2.69}$$

onde I_R é o número de imagens relevantes recuperadas, I_O é o número de imagens recuperadas e I_T é o número de imagens relevantes no total presentes nos sistemas de armazenamento de imagens médicas (Figura 34). A relevância de uma imagem pode ser determinada pela classe a que ela pertence (e.g. maligno ou benigno) ou por um especialista. Contudo, para esta tese, a relevância será determinada apenas pela classe a que pertence a imagem recuperada.



Figura 34: Representação dos conjuntos de casos na recuperação de imagens baseada em conteúdo.



Normalmente, as métricas de precisão e revocação são apresentadas como um gráfico em que a primeira é calculada tendo como base o cálculo da segunda até 100% do seu valor (Figura 35). Quanto mais próxima do topo do gráfico *precisão vs. revocação* a função se encontra, maior performance ela terá.

A eficiência de um sistema de CBIR também pode ser avaliada para os primeiros casos recuperados. Esta métrica é conhecida como precisão(N), onde N é o número de imagens retornadas (Figura 36). A razão para avaliar a performance apenas nos primeiros casos recuperados é que são eles os mais similares a um caso desconhecido em estudo e que podem ajudar o especialista na decisão clínica (AZEVEDO-MARQUES; RANGAYYAN, 2013; OLIVEIRA; FERREIRA, 2013). O cálculo da precisão para os N primeiros casos recuperados é realizado da seguinte maneira: dadas duas classes A e B, considere um caso de referência da classe A. Se o primeiro caso retornado pertencer também à classe A, logo a precisão para a primeira imagem é 100% (de acordo com a Equação 2.68, 1 imagem relevante recuperada dividido por 1 imagem recuperada). Se o segundo caso retornado pertencer à classe B, logo a precisão para a segunda imagem é 50% (1 imagem relevante recuperada dividido por 2 imagens recuperadas). Se o terceiro caso retornado pertencer à classe recuperadas, e assim por diante, até que os N casos mais similares sejam recuperados.

Figura 35: Exemplo de gráfico *precisão vs. revocação* para avaliar a performance em CBIR: a *Função 1* apresenta maior eficiência do que a *Função 2*.



Fonte: Autor.

Figura 36: Exemplo de gráfico precisão(10) para avaliar a performance em CBIR: a *Função* 2 apresenta maior eficiência do que a *Função* 1.



Fonte: Autor.

3 CLASSIFICAÇÃO DE NÓDULOS PULMONARES

3.1 Fundamentação

O objetivo deste capítulo é apresentar técnicas implementadas para a extração de características quantitativas de imagens médicas, seleção de atributos relevantes e classificação de imagens para o consequente desenvolvimento de um método de reconhecimento de padrões que possa ser aplicado no auxílio computadorizado ao diagnóstico. Para isso, busca-se também desenvolver retrospectivamente uma base de imagens médicas com o intuito da realização de experimentos para validar os métodos propostos.

Neste capítulo são apresentados os resultados obtidos de experimentos para classificação de imagens de TC de tórax com nódulos pulmonares (Seção 3.2.1), utilizando atributos de textura e nitidez de borda (Seção 3.2.2), selecionados por diferentes algoritmos (Seção 3.2.3). A classificação foi realizada em relação à probabilidade de malignidade do nódulo por métodos de aprendizado de máquina (Seção 3.2.4). Os resultados de classificação de nódulos pulmonares foram publicados no periódico *Journal of Digital Imaging* (Fator de Impacto 2017: 1,536, Qualis CAPES 2013–2016 na área Engenharias IV: A2) em forma de artigo científico (Anexo A) (FERREIRA JR; OLIVEIRA; AZEVEDO-MARQUES, 2018).

3.2 Material e Métodos

3.2.1 Conjunto e Segmentação de Imagens

As imagens utilizadas nesta etapa do trabalho foram de TC de tórax no padrão DICOM provenientes do repositório público LIDC (do Inglês *Lung Image Database Consortium, Image Database Resource Initiative*), pertencente ao *The Cancer Imaging Archive* do *National Cancer Institute, National Institutes of Health* dos Estados Unidos (ARMATO III et al., 2011). O LIDC é uma base de dados referência para a comunidade de pesquisa em engenharia e informática de imagens médicas, tendo sido utilizada em diversos trabalhos (KAYA, 2018; DHARA et al., 2017a; CAO et al., 2017; SILVA et al., 2017; NIBALI; HE; WOLLERSHEIM, 2017; LUCENA et al., 2016). Atualmente, ela é composta por 1.018 exames de TC com 244.527 imagens de 1.010 pacientes (ARMATO III et al., 2018).

As lesões em cada exame foram identificadas por 4 radiologistas experientes do LIDC em um processo de interpretação de imagens de TC que requisitou dos especialistas a leitura das imagens e a marcação manual de nódulos. Inicialmente, as lesões identificadas pelos radiologistas foram categorizadas em três tipos:

- nódulos ≥ 3 milímetros: lesões que possuem tamanho entre 3 e 30 mm, independente de histologia (Figura 37(a));
- nódulos < 3 milímetros: lesões que possuem tamanho menor que 3 mm, sem evidências radiológicas da natureza do achado (Figura 37(b));
- não-nódulos ≥ 3 milímetros: lesões que possuem tamanho maior que 3 mm e não possuem características consistentes de um nódulo pulmonar (Figura 37(c)). Não-nódulos < 3 mm foram desconsiderados.
- Figura 37: Exemplos de lesões em imagens de TC de acordo com as categorias do LIDC: (a) nódulo \geq 3 mm, (b) nódulo < 3 mm, (c) não-nódulo \geq 3 mm. As setas indicam a localização dos achados radiológicos.



Fonte: Armato III et al. (2011).

Posteriormente, os nódulos ≥ 3 mm da primeira categoria foram avaliados pelos radiologistas em características qualitativas de calcificação, estrutura interna, lobulação, margem, esfericidade, espiculação, sutileza, textura e probabilidade de malignidade. Para esta última característica, os especialistas atribuíram um valor inteiro entre 1 e 5 para cada nódulo, de acordo com as seguintes condições:

- 1 para probabilidade alta em ser benigno;
- 2 para probabilidade moderada em ser benigno;
- 3 para probabilidade indeterminada;
- 4 para probabilidade moderada em ser maligno;
- 5 para probabilidade alta em ser maligno.

Além disso, os nódulos ≥ 3 mm foram marcados manualmente pelos radiologistas em cada corte do volume. Esta marcação corresponde ao primeiro *pixel* externo ao nódulo, não sobrepondo a borda da lesão. Nódulos < 3 mm e não-nódulos ≥ 3 mm tiveram apenas
o centro tridimensional de massa marcado. Todas as informações de cada exame do LIDC estão armazenadas em um arquivo XML (do Inglês *eXtensible Markup Language*) que descreve o VOI dos nódulos identificados com as posições cartesianas (x, y, z) traçadas manualmente pelos radiologistas, as classificações das lesões de acordo com as três categorias do LIDC, as nove características qualitativas dos nódulos ≥ 3 mm e as informações do exame de TC (Anexo B).

Nesta etapa do projeto, todos os nódulos foram segmentados utilizando as marcações presentes no arquivo XML do exame de TC realizadas pelos radiologistas do LIDC. Apenas a leitura do radiologista que identificou mais nódulos ≥ 3 mm foi considerada. As leituras de 3 radiologistas foram descartadas de cada exame com o objetivo de evitar redundância de amostras nos experimentos. A segmentação manual foi realizada da seguinte forma: cada corte do volume da lesão possui um número n de coordenadas cartesianas (x, y)correspondendo à marcação do nódulo na imagem. Em cada fatia, um polígono de nlados foi criado, sendo cada lado a união de duas coordenadas cartesianas. Para *pixels* não inseridos no polígono foi atribuído o tom preto na imagem segmentada, e para *pixels* inseridos no polígono foi atribuído o tom de cinza da coordenada cartesiana na imagem de TC original. Os nódulos foram segmentados para imagens de 8 *bits* e dimensões 256 × 256.

O conjunto total de nódulos pulmonares segmentados possui 1.944 casos nas 5 classificações de probabilidade de malignidade. Contudo, nódulos com probabilidade no valor 3 foram descartados por possuírem malignidade indeterminada, resultando em 1.171 nódulos pulmonares (Tabela 5). Para este trabalho, casos com probabilidade de malignidade 1 e 2 foram considerados benignos e casos com probabilidade de malignidade 4 e 5, malignos. Para prevenir viés para a classe majoritária (nódulos com maior probabilidade de benignidade) e potencialmente aumentar a sensibilidade dos métodos, os experimentos foram realizados com um conjunto de imagens balanceado e com o original desbalanceado. O conjunto de nódulos inicial foi randomicamente balanceado, resultando em 600 amostras, sendo 300 para cada classe.

Tabela 5: Número de nódulos pulmonares utilizados no trabalho.

Probabilidade de Malignidade	1	2	4	5
Número de Nódulos	273	472	266	160
Total por Classe	745		42	26

Fonte: Autor.

Como o repositório de imagens do LIDC não está organizado em um banco de dados, este trabalho propôs um esquema estruturado de armazenamento para ele. O Apêndice A apresenta o desenvolvimento de uma base de imagens para o LIDC baseado neste esquema de banco de dados.

3.2.2 Extração de Atributos

Após a segmentação manual das imagens de TC, atributos de textura baseados em matriz de coocorrência foram extraídos dos VOIs dos nódulos (Seção 2.4.2). O cálculo da COM utilizou distância 1 entre os *voxels* da imagem e entre as fatias do volume e orientações θ iguais a 0°, 45°, 90° e 135°. Os atributos extraídos da COM foram energia, entropia, contraste, variância, matiz, proeminência, correlação, homogeneidade e momento da diferença inverso (MDI) (HARALICK; SHANMUGAM; DINSTEIN, 1973). A utilização de nove atributos e quatro orientações angulares permitiu a criação de 36 atributos de textura de COM para cada volume de imagens.

Além da extração de características texturais, foi realizada uma análise da nitidez de borda do nódulo nas imagens originais de TC, inicialmente proposta em (FERREIRA JUNIOR; OLIVEIRA, 2015). Foi adotada a abordagem de criação de linhas ortogonais traçadas sobre a borda do nódulo pulmonar para esta análise (Seção 2.4.3) (XU et al., 2012). Em seguida, atributos estatísticos foram calculados das intensidades dos *pixels* das linhas ortogonais em todas as imagens do volume. Os seguintes passos foram realizados para implementar a análise da nitidez de borda: vinte pontos de controle foram selecionados automaticamente a partir do primeiro *pixel* marcado pelo radiologista do LIDC na imagem de TC original (Figura 38(a)). Em seguida, segmentos de reta foram desenhados em cada um dos 20 pontos de controle perpendicularmente à borda da lesão (Figura 38(b)). Uma máscara foi criada para eliminar os segmentos de reta que cruzavam a parede do pulmão para assim incluir apenas *pixels* que pertenciam ao tecido do nódulo e do pulmão. A máscara foi gerada por um algoritmo de binarização (Seção 2.3) com aplicação da operação morfológica matemática de dilatação¹ (Figura 38(c)). Após a exclusão dos segmentos de reta não inseridos no pulmão (Figura 38(d)), todos os *pixels* contidos nos segmentos de reta restantes do volume do nódulo tiveram seus níveis de cinza armazenados em um array ordenado em ordem crescente. Em seguida, foram calculados neste array de intensidades atributos estatísticos de amplitude (diferença entre os extremos do array), soma dos valores, soma dos quadrados, soma dos logs, médias aritmética e geométrica, variâncias da população e da amostra, desvio padrão, curtose, assimetria e segundo momento central (SMC). Por fim, a análise da nitidez de borda resultou na criação de mais 12 características quantitativas que, adicionadas às características de textura, permitiram a criação de um vetor de atributos para cada nódulo de 48 dimensões.

A extração de atributos foi realizada com a linguagem de programação Java v1.8 e a biblioteca gráfica OpenCV v2.4 com *binding* JavaCV v0.1, em um ambiente de desenvolvimento Linux com sistema operacional Ubuntu v13.04.

¹ Morfologia matemática consiste em extrair a geometria de uma imagem utilizando a transformação de um elemento estruturante. Uma das operações básicas de morfologia matemática é a dilatação, que é uma função não linear que adiciona vetorialmente objetos do elemento estruturante em uma imagem (GONZALEZ; WOODS, 2007).

Figura 38: Análise de nitidez de borda de nódulo pulmonar: (a) pontos de controle da borda, (b) segmentos de reta perpendiculares à borda, (c) máscara obtida por binarização e dilatação, (d) imagem resultante para extração de atributos.



Fonte: Autor.

3.2.3 Seleção de Atributos

Com o objetivo de potencialmente aumentar a performance de classificação das imagens eliminando atributos irrelevantes e/ou redundantes e diminuir as chances de sobreajuste do modelo, foi realizada uma seleção das características mais relevantes do vetor de atributos das imagens. Três métodos foram empregados nesta etapa do projeto: análise de significância estatística, filtro baseado em correlação e *wrapper* (Seção 2.5) (ALMEIDA; RANGAYYAN; AZEVEDO-MARQUES, 2015; HALL, 1999; KOHAVI; JOHN, 1997).

A análise estatística foi realizada com valores-p para testar a significância dos atributos estatisticamente entre as classes. Valor- $p \ge 0.05$ implica que as diferenças nos valores do atributo não são significativas para distinguir as classes. $0.01 \le$ valor-p < 0.05implicam na existência de significância estatística nas diferenças nos valores do atributo. E valor-p < 0.01 implica em alta significância do atributo para distinguir as classes. A análise estatística foi feita pelo teste paramétrico t-Student não pareado com a ferramenta GraphPad Prism v6.

Os métodos filtro e *wrapper* foram executados com uma validação-cruzada de 10 subconjuntos utilizando a ferramenta Weka v3.6.12 (FRANK; HALL; WITTEN, 2016). Desta forma, após a execução dos 10 subconjuntos, uma lista de atributos e ocorrências na seleção é gerado. A relevância do atributo pode ser medida pelo número de ocorrências em que é selecionado na validação-cruzada. Por exemplo, se um atributo é selecionado em 10 subconjuntos, então ele deve possuir alta relevância. E se um atributo é selecionado em nenhum subconjunto da validação-cruzada, então ele deve possuir baixa (ou nenhuma) relevância na caracterização das imagens. O limiar de ocorrências escolhido foi 4; logo, todos os atributos que foram selecionados em pelo menos 5 subconjuntos da validação-cruzada foram utilizados. Os classificadores utilizados internamente pelo método *wrapper* de seleção de atributos foram os mesmos do processo de classificação das imagens, explicado em seguida.

3.2.4 Classificação de Imagens

Seis algoritmos de aprendizado de máquina foram utilizados na classificação das imagens: KNN, SVM, NB, rede neural RBF, árvore de decisão J48 e RF (Seção 2.6) (AHA; KIBLER; ALBERT, 1991; VAPNIK, 2013; JOHN; LANGLEY, 1995; MOODY; DARKEN, 1989; QUINLAN, 1993; BREIMAN, 2001). O KNN utilizou a distância Euclidiana como função de distância entre amostras, e o valor de k foi definido como 9. O SVM utilizou um vetor do tipo *nu-support* com função de base radial como *kernel*. A RBF utilizou agrupamento com o algoritmo *k-means* no treinamento da rede. O RF utilizou 100 árvores de decisão para realizar a classificação. A ferramenta Weka v3.6.12 foi utilizada na tarefa de classificação dos nódulos pulmonares (FRANK; HALL; WITTEN, 2016).

Os classificadores utilizaram diferentes vetores de atributos das imagens para realizar a classificação binária. O primeiro vetor foi composto por todos os 48 atributos de textura e nitidez de borda. O segundo vetor foi composto por atributos selecionados pela análise de significância estatística e que possuem valor-p < 0,05. O terceiro e quarto vetores foram compostos por atributos selecionados pelo filtro baseado em correlação e pelo *wrapper*, respectivamente. Antes da classificação das imagens, todos os atributos foram normalizados entre 0 e 1 devido os diferentes intervalos em seus valores, o que poderia influenciar na classificação de métodos que utilizam alguma métrica de distância, como o KNN, por exemplo.

Uma validação-cruzada com 10 subconjuntos foi empregada nos experimentos. Para avaliar a eficiência dos classificadores, foram utilizadas as medidas de acurácia, sensibilidade, especificidade e AUC (Seção 2.6.2).

3.3 Resultados

A Tabela 6 apresenta os resultados de seleção de atributos dos três métodos para redução de dimensionalidade vetorial (Seção 3.2.3). A análise de significância estatística foi realizada individualmente para cada atributo extraído (Seção 3.2.2). O filtro baseado em correlação foi executado em cada conjunto de amostras (desbalanceado e balanceado) (Seção 3.2.1). E o *wrapper* foi executado em cada conjunto de amostras para cada um dos seis classificadores (Seção 3.2.4).

A Tabela 7 apresenta os resultados de classificação utilizando o conjunto de amostras desbalanceado, e as Figuras 39(a–d) ilustram a performance diagnóstica usando curvas ROC e diferentes vetores de atributos das imagens.

Categoria	Atributo	valor- p	Filtro	KNN	SV	M N	В	RE	F	\mathbf{J}_{4}	18	R	F
	Energia em 0°	0,1215	0 0	0 0	1	0 3	5	1	1	1	1	1	1
	Entropia em 0°	< 0,0001	$8 \mid 3$	$2 \mid 0$	3	2 0	0	0	0	1	1	0	7
	Contraste em 0°	< 0,0001	10 10	0 9	5	5 8	0	8	9	1	0	2	3
	Homogeneidade em 0°	< 0,0001	$0 \mid 0$	2 1	4	1 0	0	0	0	2	1	0	4
	Correlação em 0°	< 0,0001	$0 \mid 0$	$1 \mid 0$	$3 \mid$	3 3	5	0	0	1	0	0	2
	Matiz em 0°	< 0,0001	$0 \mid 0$	$3 \mid 0$	0	1 0	0	0	0	2	2	3	0
	Proeminência em 0°	< 0,0001	10 4	$4 \mid 0$	0	2 0	0	0	0	0	0	1	0
	Variância em 0°	< 0,0001	$0 \mid 0$	$2 \mid 0$	$1 \mid$	0 0	0	$2 \mid$	3	1	0	0	0
	$MDI em 0^{\circ}$	< 0,0001	9 3	5 5	$3 \mid$	4 0	0	6	0	3	3	$5 \mid$	3
	Energia em 45°	0,0017	$6 \mid 1$	$0 \mid 0$	$2 \mid$	0 4	$\mid 0$	$2 \mid$	0	1	0	0	1
	Entropia em 45°	< 0,0001	$0 \mid 0$	$0 \mid 0$	$1 \mid$	2 0	$\mid 0$	0	0	0	2	$2 \mid$	5
	Contraste em 45°	< 0,0001	$6 \mid 1$	0 5	$3 \mid$	1 0	1	7	5	3	0	5	4
	Homogeneidade em 45°	< 0,0001	$0 \mid 0$	$1 \mid 0$	$3 \mid$	1 0	0	0	0	0	2	1	1
	Correlação em 45°	< 0,0001	$0 \mid 0$	$0 \mid 0$	$3 \mid$	5 2	0	0	0	0	0	0	3
	Matiz em 45°	< 0,0001	$0 \mid 0$	$2 \mid 0$	1	2 0	0	0	0	1	0	$2 \mid$	0
	Proeminência em 45°	< 0,0001	$4 \mid 0$	0 0	1	1 0	0	0	0	0	0	0	1
	Variância em 45º	< 0,0001	0 0	0 0	0	1 0	0	0	0	0	1	2	2
Textura	$MDI em 45^{\circ}$	< 0,0001	10 5	$3 \mid 0$	1	3 0	0	3	0	2	0	3	1
	Energia em 90°	0,0424	0 2	0 0	0	6 2	5	4	5	1	3	3	0
	Entropia em 90°	< 0,0001	$2 \mid 0$	$1 \mid 0$	2	1 0	0	0	0	1	0	1	2
	Contraste em 90°	< 0,0001	1 1	0 2	1	5 1	6	7	4	0	3	0	0
	Homogeneidade em 90°	< 0,0001	0 0	1 0	1	3 0	0	0	0	1	0	1	0
	Correlação em 90 ⁹	< 0,0001	0 0	5 0	1	4 7	5	5	1	0	1	2	4
	Matiz em 90 ^o	< 0,0001	0 0	$0 \mid 0$	0	$\frac{3}{0}$		0	0	0	2	1	0
	Proeminencia em 90 ^e	< 0,0001	10 10	$0 \mid 0$	1	0 0		0	0	1		1	0
	Variancia em 90 [™]	< 0,0001		0 0 1 2	1	$\frac{3}{4}$ 0		1	1	5		1	3 1
	$MDI \text{ em } 90^{=}$	< 0,0001	9 8	1 3		4 0 0 C		0 10	0	1			1
	Energia em 135 ^e	0,0030	10 3		0	0 6	8	10	3	1	2	0	2
	Entropia em 135 ⁼	< 0,0001	0 0		1	3 0		0	4	1	1 1		1
	Unitraste em 155-	< 0,0001	0 0	$0 \mid i$ $0 \mid 1$	4	$\frac{3}{4}$ 0		0	4	ം റ	1 0	0	ა ე
	Compleção em 1259	< 0,0001	$0 \mid 2$		4	4 0			0	2		2	3
	Correlação em 155 ⁻	< 0,0001		$0 \mid 0$ $2 \mid 0$	2	$\begin{array}{ccc} 0 & 2 \\ 2 & 0 \end{array}$			0	1		0	0
	Droominâncie om 1259	< 0,0001		2 0 1 0	1	2 0			0	1		1	2
	Variância om 135 ⁹	< 0,0001			2	$\frac{0}{2}$ 0		0	0	1		1	1
	MDL em 135º	< 0,0001	8 5	6 2	2	2 0 5 0			0	3		5	1
	Amplitude	0.3367	$\frac{0}{10}$	0 2 0 10	10	10 8	$\frac{10}{10}$	10	10	8	10	10	4
	Soma dos valores	< 0.0001			7	5 2	10	9	0	2	10	3	1
	Soma dos quadrados	< 0.0001	$0 \mid 1$	1 1	6	7 0	12	0	1	2		6	1
	Soma dos <i>logs</i>	< 0,0001		0 2	3	5 7	0	0	0	2		4	3
	Média aritmética	0.2876	10 ± 1	0 0	10	6 6	6	8	9	1		4	3
	Média geométrica	0.0073			6	3 8	5	0	3	2		2	2
Nitidez de Borda	Variância da população	0.6129	2 1	0 2	5	2 7	5	1	0	1	1	1	-
	Variância da amostra	0.6491	4 3	0 1	3	3 8	4	3	0	1	0	4	3
	Desvio padrão	0,6491	0 0	0 0	4	4 5	5	2	1	0	0	0	4
	Curtose	< 0,0001	10 10	0 3	1	5 1	8	4	6	1	0	8	3
	Assimetria	< 0,0001	6 3	0 0	3	5 3	8	10	2	2	1	3	2
	SMC	< 0,0001	$0 \mid 1$	0 3	$2 \mid$	1 0	0	0	0	0	0	1	0

Tabela 6: Significância estatística e ocorrência dos atributos na seleção de características para classificação de imagens utilizando filtro e *wrapper*.

Ocorrências são apresentadas em pares $x \mid y$, onde $x \in y$ correspondem ao número de ocorrências do atributo na validação-cruzada com o conjunto de amostras desbalanceado e balanceado, respectivamente.

Fonte: Autor.

		KNN	SVM	NB	RBF	J48	RF
	AUC	0,827	0,713	0,769	$0,\!695$	0,735	0,858
Todog og stributog	Sensibilidade	$0,\!596$	$0,\!493$	$0,\!244$	0,209	$0,\!617$	0,702
Todos os atributos	Especificidade	0,859	0,933	$0,\!965$	0,966	$0,\!831$	$0,\!856$
	Acurácia	0,763	0,773	0,703	$0,\!691$	0,753	0,800
Atributor	AUC	0,794	0,713	0,749	0,702	0,747	$0,\!847$
Attributos	Sensibilidade	0,563	$0,\!491$	$0,\!242$	$0,\!350$	$0,\!636$	$0,\!685$
selecionados	Especificidade	0,866	0,936	$0,\!965$	$0,\!874$	$0,\!834$	$0,\!843$
por estatística	Acurácia	0,756	0,774	0,702	$0,\!683$	0,762	$0,\!786$
Atributor	AUC	0,811	0,707	0,778	0,709	0,771	$0,\!846$
Attibutos	Sensibilidade	$0,\!559$	$0,\!495$	$0,\!223$	$0,\!392$	$0,\!650$	$0,\!655$
selectollados	Especificidade	0,885	0,919	$0,\!977$	$0,\!940$	$0,\!850$	$0,\!854$
por mitro	Acurácia	0,766	0,765	0,703	0,740	0,777	$0,\!781$
Atributor	AUC	0,816	0,730	0,792	0,752	0,710	$0,\!843$
selecionados	Sensibilidade	$0,\!664$	0,531	0,502	0,505	$0,\!575$	$0,\!662$
	Especificidade	0,866	0,929	0,921	0,907	0,906	$0,\!860$
por wrapper	Acurácia	0,792	0,784	0,769	0,761	0,786	0,788

Tabela 7: Performance de classificação com o conjunto de amostras desbalanceado.

Valores em negrito são os maiores por linha.

Fonte: Autor.

Todos os classificadores obtiveram maior performance que um teste diagnóstico aleatório, independentemente do vetor de atributos empregado, utilizando o conjunto de amostras desbalanceado. A menor performance foi obtida pela rede RBF, mas ainda assim maior que o *baseline* de 0,50 de AUC. O classificador RF obteve a maior performance diagnóstica com o vetor completo de atributos (Figura 39(a)).

A performance classificatória não foi melhorada com a seleção de atributos estatisticamente significativos em comparação com o vetor completo de características. O RF obteve a maior performance de classificação neste cenário (Figure 39(b)), porém nenhuma diferença estatisticamente significativa foi encontrada em comparação com todos os atributos combinados (intervalo de confiança de 95% - considere este nível de confiança para o restante da análise estatística).

O filtro baseado em correlação também não foi capaz de melhorar a eficiência da classificação para a maioria dos métodos. Entretanto, a acurácia da rede RBF com os 17 atributos selecionados pelo filtro foi aumentada em 5 pontos percentuais em comparação com todos os atributos extraídos, com diferença estatisticamente significativa na sensibilidade. A AUC da árvore de decisão J48 obteve aumento de performance de aproximadamente 0,3, porém nenhuma diferença estatisticamente significativa foi encontrada. A Figura 39(c) ilustra a performance dos classificadores utilizando o vetor de atributos selecionados pelo filtro. O *wrapper* conseguiu aumentar a eficiência diagnóstica para a maioria dos classificadores, seja em AUC ou acurácia. As maiores diferenças (estatisticamente significativas) em relação ao vetor completo de atributos foram encontradas pela rede RBF (usando 12 atributos) e NB (usando 10 atributos). A Figura 39(d) ilustra a performance diagnóstica ao utilizar atributos selecionados pelo *wrapper*.

Figura 39: Curvas ROC dos experimentos com o conjunto de amostras desbalanceado: (a) todos os atributos, (b) atributos selecionados por estatística, (c) atributos selecionados por filtro, (d) atributos selecionados por *wrapper*.



Fonte: Autor.

A Tabela 8 apresenta os resultados de classificação utilizando o conjunto de amostras balanceado, e as Figuras 40(a–d) ilustram a performance diagnóstica usando curvas ROC e diferentes vetores de atributos das imagens.

Todos os classificadores também obtiveram maior performance que um teste diagnóstico aleatório, independentemente do vetor de atributos empregado, utilizando o conjunto de amostras balanceado. A menor performance foi obtida pelo SVM com os atributos selecionados pelo filtro, mas ainda assim maior que o *baseline* de 0,50 de AUC. O classificador RF obteve a maior performance diagnóstica com o vetor completo de atributos (Figura 40(a)).

		KNN	SVM	NB	RBF	J48	RF
	AUC	0,806	0,777	0,734	$0,\!685$	0,720	0,856
Todog og stributog	Sensibilidade	$0,\!663$	0,700	$0,\!247$	0,643	$0,\!680$	0,750
10008 08 attributos	Especificidade	0,813	0,853	$0,\!970$	$0,\!627$	$0,\!697$	0,770
	Acurácia	0,738	0,777	$0,\!608$	$0,\!635$	$0,\!688$	0,760
Atributor	AUC	0,791	0,747	0,711	$0,\!667$	0,747	0,856
Attibutos	Sensibilidade	$0,\!683$	$0,\!677$	0,240	0,593	0,707	0,780
selecionados	Especificidade	0,770	0,817	$0,\!970$	$0,\!623$	0,733	0,740
por estatística	Acurácia	0,727	0,747	$0,\!605$	$0,\!608$	0,720	0,760
Atributor	AUC	$0,\!820$	$0,\!627$	0,765	0,736	0,746	$0,\!819$
Attibutos	Sensibilidade	0,703	$0,\!643$	0,247	$0,\!497$	$0,\!660$	0,733
selectofiados	Especificidade	0,793	$0,\!610$	$0,\!973$	$0,\!900$	0,773	0,723
por intro	Acurácia	0,748	$0,\!627$	$0,\!610$	$0,\!698$	0,717	0,728
Atributor	AUC	0,796	0,773	0,752	0,784	0,751	0,809
Atributos	Sensibilidade	$0,\!687$	$0,\!697$	$0,\!607$	$0,\!600$	$0,\!693$	0,753
selectoffactos	Especificidade	0,790	$0,\!850$	0,777	0,893	$0,\!823$	0,740
por wrapper	Acurácia	0,738	0,773	$0,\!692$	0,747	0,758	0,747

Tabela 8: Performance de classificação com o conjunto de amostras balanceado.

Valores em negrito são os maiores por linha.

Fonte: Autor.

A eficiência classificatória também não foi melhorada com a seleção dos atributos estatisticamente significativos em comparação com todos os atributos combinados. A J48 ainda foi capaz de aumentar a AUC, porém nenhuma diferença estatisticamente significativa foi encontrada. A Figura 40(b) ilustra as curvas ROC dos classificadores usando atributos selecionados pela análise de significância estatística.

O filtro baseado em correlação foi capaz de aumentar a performance dos classificadores na maioria dos cenários utilizando o conjunto de amostras balanceado. A rede RBF conseguiu aumentar a performance com os 8 atributos selecionados, com diferença estatisticamente significativa na sensibilidade e especificidade. A Figura 40(c) ilustra a performance dos classificadores com os atributos selecionados pelo filtro utilizando curvas ROC.

O método *wrapper* conseguiu aumentar a eficiência diagnóstica para a maioria dos classificadores em relação ao vetor completo de atributos. O classificador NB com 13 atributos selecionados aumentou a performance com diferenças estatisticamente significativas na sensibilidade e especificidade. A J48 com 2 atributos aumentou a eficiência classificatória com diferença estatisticamente significativa na especificidade. A Figura 40(d) ilustra a performance com curvas ROC destes métodos com atributos selecionados pelo *wrapper*.

Figura 40: Curvas ROC dos experimentos com o conjunto de amostras balanceado: (a) todos os atributos, (b) atributos selecionados por estatística, (c) atributos selecionados por filtro, (d) atributos selecionados por *wrapper*.



Fonte: Autor.

Aumentos estatisticamente significativos na sensibilidade (importante em problemas de classificação maligno vs. benigno) foram identificados em alguns cenários quando foi realizado o balanceamento no conjunto de amostras, utilizando diferentes vetores de atributos (Tabela 9). A rede RBF com todos os atributos obteve o maior aumento de sensibilidade, com média de mais 43 pontos percentuais, em comparação ao conjunto de amostras desbalanceado.

Tabela 9: Cenários onde ocorreu diferença estatisticamente significativa na sensibilidade entre os conjuntos de amostras.

	KNN	SVM	NB	RBF	J48	RF
Todos os atributos		Х		Х		
Atributos selecionados por estatística	Х	Х		Х		
Atributos selecionados por filtro	Х	Х				
Atributos selecionados por <i>wrapper</i>		Х			Х	

Fonte: Autor.

O RF com todos os atributos combinados e o conjunto desbalanceado obteve a maior performance geral de classificação com AUC de 0,858. Contudo, não foi identificada diferença estatisticamente significativa na sensibilidade e especificidade em comparação com alguns cenários de classificadores, atributos selecionados pelo *wrapper* e conjunto de amostras empregados. A Tabela 10 e a Figura 41 apresentam a sensibilidade, especificidade e curvas ROC para alguns destes cenários de classificação (os mais relevantes de acordo com alta performance e baixo número de atributos selecionados pelo *wrapper*).

Número de atributos	Classificador	Conjunto	Sensibilidade	Especificidade
48	RF	Desbalanceado	$0,\!656\!-\!0,\!744$	0,829–0,880
6	RF	Desbalanceado	$0,\!615\!-\!0,\!706$	$0,\!833 – 0,\!884$
5	KNN	Balanceado	$0,\!630\!-\!0,\!738$	0,739-0,834
3	KNN	Desbalanceado	$0,\!617\!-\!0,\!709$	$0,\!839\!-\!0,\!889$
2	J48	Balanceado	$0,\!637\!-\!0,\!744$	0,774-0,864

Tabela 10: Cenários de classificação sem diferença estatisticamente significativa na sensibilidade e especificidade (intervalos de confiança de 95%).

Fonte: Autor.

Figura 41: Curvas ROC de cenários (número de atributos + classificador + conjunto de amostras) de classificação relevantes sem diferença estatisticamente significativa na sensibilidade e especificidade.



Fonte: Autor.

O reconhecimento computadorizado de padrões em imagens médicas é clinicamente importante pois pode auxiliar o processo de decisão de distinguir um nódulo pulmonar de acordo com a sua malignância (GONG et al., 2018; OROOJI et al., 2018; NISHIO; NAGASHIMA, 2017). Entretanto, ele depende da extração de atributos representativos das imagens para caracterizar os nódulos pulmonares, selecionar os atributos mais relevantes para melhor discriminar as lesões e classificar as imagens com algoritmos de aprendizado de máquina eficientes utilizando os atributos extraídos. Neste capítulo, foi proposto o reconhecimento de padrões em imagens de TC com atributos de textura de segunda-ordem, que são tradicionais e ainda relevantes para a literatura científica (KAYA, 2018; CHOI et al., 2018; DILGER et al., 2015), e nitidez de borda, que são importantes para o diagnóstico do câncer de pulmão e possuem menos dependência da marcação ground truth do nódulo pelo radiologista (FERREIRA JUNIOR; OLIVEIRA, 2015; XU et al., 2012; LEVMAN; MARTEL, 2011).

Selecionar os atributos estatisticamente significativos não apresentou relevantes melhoras de performance de classificação e redução de dimensionalidade do espaço vetorial. Apenas 6 atributos foram excluídos do vetor de características das imagens (amplitude, média aritmética, variância da população, variância da amostra, desvio padrão e energia em 0°), o que representou uma redução de apenas 13% do espaço multidimensional. Desta forma, a seleção de atributos por análise estatística pode não apresentar relevância alguma, seja computacional ou diagnóstica, para classificação por aprendizado de máquina.

Contudo, os métodos de seleção de atributos *wrapper* e filtro baseado em correlação foram capazes de melhorar a performance em diversos cenários de classificação, ou pelo menos, reduzir a complexidade e dimensionalidade do espaço vetorial. Por exemplo, o filtro selecionou 17 atributos no conjunto desbalanceado de amostras (amplitude, média aritmética, curtose, assimetria, entropia em 0°, contraste em 0°, proeminência em 0°, MDI em 0°, energia em 45°, contraste em 45°, MDI em 45°, proeminência em 90°, MDI em 90°, energia em 135°, contraste em 135°, proeminência em 135° e MDI em 135°), o que correspondeu uma redução de 65% do espaço multidimensional, e 8 atributos no conjunto balanceado (amplitude, curtose, contraste em 0°, MDI em 45°, proeminência em 90°, MDI em 90°, contraste em 135° e MDI em 135°), o que correspondeu uma redução de 83% do espaço multidimensional.

É importante notar que como o *wrapper* utiliza um classificador internamente no mecanismo de seleção de atributos, cada método de aprendizado de máquina utilizou um vetor de atributos diferente para realizar a classificação das imagens. Por exemplo, com o NB o *wrapper* selecionou 13 atributos no conjunto de amostras balanceado (amplitude, média aritmética, média geométrica, variância da população, desvio padrão, curtose, assimetria, energia em 0° , correlação em 0° , energia em 90° , contraste em 90° , correlação

em 90° e energia em 135°) e 10 atributos no conjunto desbalanceado (amplitude, soma dos logs, média aritmética, média geométrica, variância da população, variância da amostra, desvio padrão, contraste em 0° , correlação em 90° e energia em 135°). E com eles, o NB obteve aumento de performance em relação ao vetor completo de atributos, com diferenças estatisticamente significativas na sensibilidade e especificidade em ambos os conjuntos de amostras. O menor número de atributos selecionados pelo wrapper foi obtido pela árvore de decisão J48 com ambos os conjuntos de amostras (2 atributos - amplitude e MDI em 135° para o conjunto balanceado e amplitude e variância em 90° para o desbalanceado), o que correspondeu uma redução de 96% do espaço multidimensional.

Apesar do *wrapper* conseguir reduzir mais efetivamente a dimensionalidade do vetor de atributos das imagens em comparação com o filtro, o primeiro leva muito mais tempo que o segundo para selecionar os atributos. Isto acontece porque para cada subconjunto de atributos examinado pelo *wrapper* um modelo é treinado para cada um dos 10 subconjuntos de amostras da validação-cruzada. Desta forma, alguns *wrappers* levaram horas para selecionar os atributos do vetor completo de características, enquanto que o filtro escolhe os atributos mais relevantes em apenas alguns segundos.

O balanceamento do número de amostras por classe resolveu parcialmente o problema de baixa sensibilidade de alguns classificadores na avaliação com o conjunto desbalanceado. Em alguns casos, recomenda-se balancear as amostras para o treinamento e classificação de imagens dos nódulos pulmonares, como por exemplo, para o SVM com qualquer conjunto de atributos de textura e nitidez de borda (Tabela 9).

Como abordado anteriormente, o RF com todos os 48 atributos combinados e o conjunto de amostras desbalanceado obteve a maior eficiência entre todos os cenários de classificação. Isto aconteceu pelo fato do RF possuir o mecanismo de seleção de atributos embarcado no treinamento (Seção 2.5). Logo, a redução da dimensionalidade do vetor de atributos não interferiu na performance do método da floresta aleatória. Ademais, a J48 com apenas 2 atributos e o conjunto balanceado obteve performance de classificação equivalente ao obtido pelo RF (Tabela 10) e ainda conseguiu a redução de 96% do espaço vetorial. Além destes 2 atributos (amplitude e MDI em 135°), é importante ressaltar o atributo MDI em 0° para distinguir os nódulos pulmonares devido a sua seleção nos outros três cenários de classificação que não apresentaram diferenças estatisticamente significativas em comparação com o cenário de maior eficiência diagnóstica.

Uma limitação ocorrida nesta etapa da pesquisa foi a falta de diagnóstico clínico dos nódulos pulmonares. O LIDC possui um número pequeno de lesões com diagnóstico final por biópsia, ressecção cirúrgica ou acompanhamento radiológico por dois anos (ARMATO III et al., 2018). Nesta etapa, foi priorizado o número de amostras para a análise, logo, apenas a avaliação radiológica foi utilizada como padrão ouro (probabilidade de malignidade) para a definição das classes (DHARA et al., 2017a; SILVA et al., 2017; NIBALI; HE;

WOLLERSHEIM, 2017; LUCENA et al., 2016). Além disso, a segmentação das imagens foi realizada a partir da marcação manual feita pelos radiologistas do LIDC. Caso uma ferramenta seja desenvolvida para a classificação das imagens, a segmentação manual pode demandar muito tempo do especialista. Logo, é necessária a investigação de um método de segmentação (semi-)automática das imagens, como o proposto em (EGGER et al., 2013). Outro ponto a ser citado é a baixa quantidade de atributos extraídos em comparação com os estudos de radiômica. As imagens poderiam ser quantificadas por outras características de forma e textura (Seções 2.4.3 e 2.4.2), além de atributos de primeira ordem (Seção 2.4.1). Por último, existem classificadores mais recentes e potencialmente mais robustos que poderiam melhorar a performance diagnóstica e ainda realizar a extração automática de atributos (Seção 2.6.1).

Este capítulo apresentou a classificação de nódulos pulmonares em termos de potencial malignidade baseada em atributos de matriz de coocorrência e nitidez de borda de imagens de TC. Todos os atributos combinados com um classificador de floresta aleatória apresentaram a maior performance para reconhecimento de padrões e resultados promissores para o diagnóstico auxiliado por computador. Entretanto, uma árvore de decisão com apenas dois atributos (a diferença entre o tom de cinza mais claro e o mais escuro de linhas perpendiculares ao nódulo e o momento de diferença inverso calculado da matriz de coocorrência de níveis de cinza à 135°) foi uma solução de baixo custo computacional para uma ferramenta CAD, com performance estatisticamente equivalente à primeira para a predição do potencial de malignidade de nódulos pulmonares e classificação de imagens de TC.

4 RECUPERAÇÃO POR CONTEÚDO DE NÓDULOS PULMONARES

4.1 Fundamentação

O objetivo deste capítulo é apresentar um método implementado de recuperação de imagens médicas baseada em conteúdo que possa ser aplicado na busca por similaridade de imagens de TC e potencialmente servir como ferramenta computadorizada para auxiliar a rotina radiológica.

Neste capítulo são apresentados os resultados obtidos de experimentos para recuperação de imagens de TC de tórax de nódulos pulmonares (Seção 4.2.1) baseada em atributos de textura e nitidez de borda (Seção 4.2.2), selecionados após análises estatísticas e computacionais (Seção 4.2.3). A recuperação das imagens foi realizada em relação à probabilidade de malignidade do nódulo (Seção 4.2.4). Os resultados de CBIR de nódulos pulmonares foram publicados no periódico *International Journal of Computer Assisted Radiology and Surgery* (Fator de Impacto 2017: 1,961, Qualis CAPES 2010–2012 na área Engenharias IV: B1) em forma de artigo científico (Anexo C) (FERREIRA JR; AZEVEDO-MARQUES; OLIVEIRA, 2017).

4.2 Material e Métodos

4.2.1 Conjunto de Imagens

As imagens de TC de tórax utilizadas nesta etapa da pesquisa foram as mesmas do Capítulo 3 de Classificação de Nódulos Pulmonares, provenientes do repositório LIDC (Seção 3.2.1) (ARMATO III et al., 2018; ARMATO III et al., 2011). O número total de nódulos segmentados e utilizados nos experimentos de CBIR foi 618, igualmente distribuídos em 309 benignos (probabilidades moderada ou alta para benignidade) e 309 malignos (probabilidades moderada ou alta para malignidade).

4.2.2 Extração de Atributos

As características quantitativas extraídas das imagens de TC dos nódulos pulmonares foram as mesmas do Capítulo 3 (Seção 3.2.2). A extração de atributos de textura de matriz de coocorrência e de nitidez de borda permitiu a criação de um vetor de atributos para cada volume de imagens de 48 dimensões, sendo 36 da primeira categoria e 12 da segunda. O cálculo dos atributos foi realizado com a linguagem de programação Java v1.8 e a biblioteca gráfica OpenCV v2.4 com *binding* JavaCV v0.1, em um ambiente de desenvolvimento Linux com sistema operacional Ubuntu v13.04.

4.2.3 Seleção de Atributos

Dois métodos foram empregados na redução da dimensionalidade do vetor de atributos das imagens e eliminação de características irrelevantes que poderiam atribuir algum ruído à recuperação das imagens: análise de significância estatística e filtro baseado em correlação, ambos também utilizados no Capítulo 3 (Seção 3.2.3). A análise estatística foi realizada pelo teste paramétrico t-Student não pareado com a ferramenta GraphPad Prism v6. O filtro foi executado com uma validação-cruzada de 10 subconjuntos utilizando a ferramenta Weka v3.6.12 (FRANK; HALL; WITTEN, 2016). O limiar de ocorrências escolhido foi 2; logo, todos os atributos que foram selecionados em pelo menos 3 subconjuntos da validação-cruzada foram utilizados.

4.2.4 Recuperação de Imagens

A distância Euclidiana foi empregada como métrica de similaridade entre os vetores de atributos para a recuperação das imagens. Logo, quanto menor a distância Euclidiana entre os vetores, maior a similaridade entre os nódulos pulmonares (Seção 2.8).

Para avaliar o efeito da seleção de atributos relevantes em CBIR, cada nódulo foi caracterizado por três vetores de características das imagens. O primeiro vetor foi composto por todos os 48 atributos de textura e nitidez de borda combinados. O segundo vetor foi composto apenas por atributos estatisticamente significativos. E o terceiro vetor foi composto por atributos selecionados pelo filtro baseado em correlação.

Antes da recuperação dos nódulos pulmonares mais similares, todos os atributos foram normalizados entre 0 e 1 devido os diferentes intervalos nos valores deles, o que poderia influenciar no cálculo da distância Euclidiana. A performance da CBIR foi avaliada pelas medidas precisão(10), precisão e revocação (Seção 2.8.1) com o método *leave-one-out*. A relevância da imagem na recuperação foi determinada pela classe a que pertence o nódulo.

Como a análise de similaridade entre as imagens pode demandar muito tempo de processamento, principalmente com a utilização de bases *Big Data*, este trabalho propôs a otimização do processo de análise de similaridade entre imagens baseada em unidade de processamento gráfico (do Inglês *graphics processing unit* - GPU). O Apêndice B apresenta uma avaliação de otimização de CBIR em GPU.

4.3 Resultados

A Tabela 11 apresenta os resultados da análise de significância estatística e seleção de atributos com o filtro baseado em correlação.

Categoria	Atributo	Ocorrências	valor- p
	Energia em 0°	0	0,1215
	Entropia em 0°	2	< 0,0001
	Contraste em 0°	10	< 0,0001
	Homogeneidade em 0°	0	< 0,0001
	Correlação em 0°	0	< 0,0001
	Matiz em 0°	0	< 0,0001
	Proeminência em 0°	2	< 0,0001
	Variância em 0°	0	< 0,0001
	$MDI em 0^{\circ}$	1	< 0,0001
	Energia em 45°	0	0,0017
	Entropia em 45°	0	< 0,0001
	Contraste em 45°	0	< 0,0001
	Homogeneidade em 45°	0	< 0,0001
	Correlação em 45°	0	< 0,0001
	Matiz em 45°	0	< 0,0001
	Proeminência em 45°	0	< 0,0001
	Variância em 45°	1	< 0,0001
	MDI em 45°	4	< 0,0001
Textura	Energia em 90°	1	0,0424
	Entropia em 90°	0	< 0,0001
	Contraste em 90°	1	< 0,0001
	Homogeneidade em 90°	0	< 0,0001
	Correlação em 90°	0	< 0,0001
	Matiz em 90°	0	< 0,0001
	Proeminência em 90°	9	< 0,0001
	Variância em 90°	3	< 0,0001
	$MDI em 90^{\circ}$	5	< 0,0001
	Energia em 135°	1	0,0030
	Entropia em 135°	0	< 0,0001
	Contraste em 135°	1	< 0,0001
	Homogeneidade em 135°	1	< 0,0001
	Correlação em 135°	0	< 0,0001
	Matiz em 135°	0	< 0,0001
	Proeminência em 135°	0	< 0,0001
	Variância em 135°	0	< 0,0001
	$MDI em 135^{\circ}$	7	< 0,0001
	Amplitude	10	0,3367
	Soma dos valores	0	< 0,0001
	Soma dos quadrados	1	< 0,0001
	Soma dos logs	0	< 0,0001
	Média aritmética	2	0,2876
Nitidan de Devil	Média geométrica	0	0,0073
Nitidez de Borda	Variância da população	1	0,6129
	Variância da amostra	0	0,6491
	Desvio padrão	0	0,6491
	Curtose	10	< 0,0001
	Assimetria	0	< 0,0001
	SMC	0	< 0,0001

Tabela 11: Significância estatística e número de ocorrências dos atributos na validação-cruzada da seleção de características para recuperação de imagens baseada em conteúdo.

Fonte: Autor.

Sete atributos de nitidez de borda apresentaram significância estatística para distinguir as duas classes de nódulos pulmonares: soma de valores, soma dos quadrados, soma dos *logs*, média geométrica, curtose, assimetria e segundo momento central. A textura foi mais eficiente para a caracterização das imagens, apresentando 35 atributos estatisticamente significativos.

O filtro baseado em correlação selecionou apenas 2 atributos de nitidez de borda para a recuperação das imagens: amplitude e curtose. Apesar da amplitude não apresentar significância estatística, este atributo apareceu em 10 subconjuntos da validação-cruzada da seleção de atributos. Seis atributos de textura apresentaram relevância para distinguir os nódulos pulmonares de acordo com o filtro baseado em correlação: contraste em 0° , proeminência em 90° , variância em 90° e MDI em 45° , 90° e 135° . As Figuras 42 e 43 apresentam os gráficos de *precisão(10)* e *precisão vs. revocação*, respectivamente, para a recuperação de imagens utilizando os vetores de atributos combinados, estatisticamente significativos e selecionados pelo filtro baseado em correlação.

Figura 42: Eficiência na recuperação de nódulos pulmonares similares pela métrica *preci-*são(10) com diferentes atributos de imagens.



Fonte: Autor.

Os atributos selecionados apresentaram precisão média maior para os 10 primeiros casos recuperados (0,780 \pm 0,089, média \pm desvio padrão) em comparação com todas as características combinadas (0,773 \pm 0,089) e os atributos estatisticamente significativos (0,761 \pm 0,094). A maior diferença na precisão foi registrada nos 3 primeiros casos

Figura 43: Eficiência na recuperação de nódulos pulmonares similares pela métrica *precisão* vs. revocação com diferentes atributos de imagens.



Fonte: Autor.

recuperados quando o vetor de atributos selecionados obteve uma média de 2,5 pontos percentuais a mais que os outros dois vetores de características das imagens.

Os atributos selecionados também apresentaram maior precisão média $(0,600 \pm 0,140)$ em comparação com todas as características combinadas $(0,587 \pm 0,141)$ e os atributos estatisticamente significativos $(0,586 \pm 0,140)$ usando o parâmetro de avaliação precisão pela revocação. A maior diferença de precisão foi registrada em 30% da revocação quando o vetor de atributos selecionados obteve uma média de 3,5 pontos percentuais a mais que os outros dois vetores das imagens. Entretanto, de 70% a 100% da revocação, os três vetores de atributos obtiveram performance equivalente na recuperação das imagens por conteúdo.

4.4 Discussão e Considerações Específicas

A importância da CBIR como ferramenta de auxílio computadorizado ao processo clínico de interpretação de imagens e diagnóstico de doenças vem crescendo com o passar dos anos (TRAINA et al., 2017; FERREIRA JUNIOR; OLIVEIRA, 2017; BEDO et al., 2016; SILVA et al., 2013; KUMAR et al., 2013). Neste capítulo, foi apresentada a avaliação de um algoritmo de recuperação de imagens baseada em atributos de textura e nitidez de borda de nódulos pulmonares presentes em imagens de TC. O propósito do

desenvolvimento deste algoritmo de CBIR foi complementar as tarefas de classificação e reconhecimento de padrões das imagens (apresentadas no capítulo anterior desta tese) baseado na busca por casos similares de uma imagem de referência que o especialista pode ter dificuldades de interpretar ou que esteja buscando por mais informações clínicas.

Os 8 atributos selecionados (amplitude, curtose, contraste em 0° , MDI em 45° , proeminência, variância e MDI em 90° e MDI em 135°) foram capazes de aumentar a performance da recuperação das imagens por conteúdo. O filtro baseado em correlação, além de ter aumentado a eficiência da busca por similaridade de imagens médicas, conseguiu reduzir a complexidade e dimensionalidade do espaço vetorial em 83%. Os aumentos de precisão principalmente na recuperação dos 3 primeiros casos podem indicar que estes atributos de nitidez de borda e textura podem auxiliar o diagnóstico do câncer de pulmão baseado na similaridade de imagens de casos passados.

É importante notar também a orientação 90º para a matriz de coocorrência como sendo relevante na recuperação de imagens baseada em textura de segunda-ordem. Primeiro, pelo fato de possuir o maior número de ocorrências na seleção de atributos (19 ocorrências, Tabela 11). E, segundo, porque apresentou 3 atributos (proeminência, variância e MDI) dos 6 de textura do vetor final de atributos das imagens para recuperação de nódulos pulmonares baseada em conteúdo.

Algumas limitações desta etapa da pesquisa já foram discutidas anteriormente no Capítulo 3 de classificação de nódulos pulmonares, como a falta de diagnóstico clínico final dos nódulos pulmonares e a segmentação manual das imagens de TC (Seção 3.4). Outra limitação é que a recuperação dos casos foi baseada apenas nos atributos extraídos das imagens. Contudo, estas características isoladamente podem não ser suficientes para prover a similaridade das lesões para o ponto de vista dos especialistas. O julgamento subjetivo seria necessário para a recuperação das imagens que fossem realmente relevantes para auxiliar a tomada de decisão diagnóstica dos especialistas. Uma possibilidade de melhorar isto é considerar uma medida de similaridade psicofísica que inclua os atributos das imagens e o julgamento subjetivo de um grupo de radiologistas (NISHIMURA et al., 2013; MURAMATSU et al., 2008). Outra possibilidade é a inclusão de um módulo de realimentação de relevância para reduzir a lacuna semântica, por meio de interações do usuário com o sistema para ponderar os casos recuperados de acordo com a relevância avaliada pelo próprio especialista (BEDO et al., 2016; BUGATTI et al., 2014; AZEVEDO-MARQUES et al., 2008).

Este capítulo apresentou a recuperação de imagens de nódulos pulmonares em termos de potencial malignidade baseada em atributos de matriz de coocorrência e nitidez de borda. Dois atributos estatísticos de nitidez de borda (amplitude e curtose) e seis de textura de segunda-ordem (contraste em 0° , proeminência em 90° , variância em 90° e MDI em 45° , 90° e 135°) apresentaram a maior eficiência na busca por similaridade de imagens

de TC e resultados promissores para o diagnóstico diferencial de nódulos pulmonares. A recuperação de imagens médicas baseada em conteúdo pôde representar uma solução de baixo custo computacional e boa eficiência para uma ferramenta de auxílio ao diagnóstico do câncer de pulmão.

5 RADIÔMICA DO CÂNCER DE PULMÃO

5.1 Fundamentação

O objetivo deste capítulo é apresentar técnicas implementadas para a segmentação de imagens médicas, extração de características quantitativas, seleção de atributos relevantes e análise radiômica associativa para o consequente desenvolvimento de um método de reconhecimento de padrões que possa ser aplicado no auxílio computadorizado à decisão terapêutica e ao prognóstico do paciente. Para isto, busca-se também desenvolver retrospectivamente uma base de imagens médicas e dados clínicos com o intuito da realização de experimentos para validar os métodos propostos.

Neste capítulo são apresentados os resultados obtidos de experimentos para radiômica de imagens de TC de tórax segmentadas semiautomaticamente com tumores pulmonares malignos (Seção 5.2.1), utilizando atributos de nível de cinza, textura e forma (Seção 5.2.2). A associação foi realizada em relação à presença de metástase distante, metástase linfonodal, ao tipo histopatológico do tumor e ao tempo de sobrevida do paciente utilizando métodos estatísticos, de aprendizado de máquina e profundo (Seção 5.2.3). Parte dos resultados de radiômica do câncer de pulmão foi publicada no periódico *Computer Methods and Programs in Biomedicine* (Fator de Impacto 2017: 2,674, Qualis CAPES 2013– 2016 na área Engenharias IV: A1) em forma de artigo científico (Anexo D) (FERREIRA JUNIOR et al., 2018).

5.2 Material e Métodos

5.2.1 Conjunto e Segmentação de Imagens

Esta etapa da pesquisa foi aprovada pelo Comitê de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, bem como a solicitação de dispensa de aplicação do Termo de Consentimento Livre e Esclarecido (parecer nº 1.996.131, Anexo E).

O conjunto de imagens e dados clínicos foi desenvolvido retrospectivamente de modo progressivo durante o projeto. Logo, diferentes versões do conjunto de pacientes foram obtidas para a realização dos experimentos de radiômica. O conjunto total possui 126 casos de câncer de pulmão confirmados histologicamente ou cirurgicamente. Os pacientes foram diagnosticados e tratados no Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto ou no Hospital Estadual de Américo Brasiliense durante os anos de 2010 a 2017. Destes 126 casos, 19 não tiveram as imagens obtidas com o protocolo padrão de injeção de meio de contraste iodado endovenoso para TC (o que influenciaria no processo de caracterização das imagens), 4 apresentaram outras opacidades anexas ao tumor (o que influenciaria no processo de segmentação das imagens) e 2 não apresentaram todos os dados clínicos disponíveis. Logo, 25 casos foram excluídos da análise e 101 identificados e incluídos, todos apresentando imagens de TC com contraste e dados clínicos disponíveis. As informações de idade, gênero, história de tabagismo e histórico de outro tumor primário foram obtidas dos arquivos do prontuário eletrônico do paciente. Três versões do conjunto de pacientes foram desenvolvidas durante esta etapa da pesquisa: a primeira (denominada de Conjunto1) apresenta 52 pacientes (Tabela 12), a segunda (denominada de Conjunto2) apresenta 16 pacientes (Tabela 13), e a terceira (denominada de Conjunto3) apresenta os 52 casos do Conjunto1 acrescidos de 33 pacientes, totalizando 85 casos (Tabela 14).

Característica	Valor
Idade (anos)	$66,6 \pm 9,1 \ (41-85)^a$
$G\hat{e}nero (F M)$	22 30
Tabagista $(S N)$	43 9
Histórico de outro câncer primário (S N)	21 31
Histopatologia $(ADC SCC outrob)$	30 15 7
Estágio T $(1 2 3 4)$	21 25 3 3
Estágio N $(0 1 2 3)$	29 12 9 2
Estágio M $(0 1)$	44 8
Localização $(SD MD ID SE IE^{c})$	11 2 14 14 11
Posição $(C P^d)$	2 50
Diâmetro (milímetros)	$30.9 \pm 13.1 \ (10-74)^a$

Tabela 12: Dados clínicos do Conjunto1 de pacientes (n = 52).

^{*a*}Média \pm Desvio Padrão (Mínimo-Máximo). ^{*b*}1 NSCLC apresentando ambos os subtipos ADC e SCC, 1 NSCLC sem especificação, 1 LCC, 1 SCLC e 3 carcinoides. ^{*c*}Lobos superior direito, médio direito, inferior direito, superior esquerdo ou inferior esquerdo. ^{*d*}Central ou periférico.

Fonte: Autor.

As imagens de TC dos três conjuntos de pacientes foram obtidas nos aparelhos multidetectores Brilliance Big Bore, Philips e Aquilion Prime, Toshiba antes de qualquer tratamento do paciente, após a administração intravenosa de meio de contraste iodado (taxa de 3,0 ml/s), volume de 80–100 ml de acordo com o peso do paciente, utilizando injeção única em *bolus* rápido seguida de solução salina (30 ml, mesma taxa). As imagens foram reconstruídas utilizando aquisições volumétricas com matriz de 512 \times 512 *pixels*, espessura de corte de 0,5–1,25 mm e espaçamento de *pixel* de 0,58–1,07 mm. Outros parâmetros típicos foram kVp de 120 e mAs de 39–464.

Como a segmentação manual de imagens médicas é uma tarefa intensa que demanda muito tempo e trabalho do especialista (ECHEGARAY et al., 2015), todos os casos utilizados nesta etapa da pesquisa foram semiautomaticamente segmentados pelo algoritmo de crescimento de região volumétrica GrowCut da plataforma de análise e visualização de imagens 3D-Slicer v4.3.1 (FEDOROV et al., 2012). O método de segmentação GrowCut foi

Característica	Valor
Idade (anos)	$64.8 \pm 9.1 \ (41-79)^a$
$G\hat{e}nero (F M)$	6 10
Tabagista $(S N)$	15 1
Histórico de outro câncer primário (S N)	3 13
Histopatologia $(ADC SCC outrob)$	7 6 3
Estágio T $(1 2 3 4)$	6 6 1 3
Estágio N $(0 1 2 3)$	7 3 3 3
Estágio M $(0 1)$	11 5
Localização $(SD MD ID SE IE^c)$	5 1 5 3 2
Posição $(C P^d)$	2 14
Diâmetro (milímetros)	$42,6 \pm 23,0 \ (11-98)^a$

Tabela 13: Dados clínicos do Conjunto2 de pacientes (n = 16).

^{*a*}Média \pm Desvio Padrão (Mínimo-Máximo). ^{*b*}3 SCLC. ^{*c*}Lobos superior direito, médio direito, inferior direito, superior esquerdo ou inferior esquerdo. ^{*d*}Central ou periférico.

Fonte: Autor.

Tabela 14: Dados clínicos do Conjunto3 de pacientes (n = 85).

Característica	Valor
Idade (anos)	$67.0 \pm 9.2 \ (41-85)^a$
$G\hat{e}nero (F M)$	39 46
Tabagista $(S N)$	69 16
Histórico de outro câncer primário $(S N)$	29 56
Histopatologia $(ADC SCC outrob)$	51 21 13
Estágio T $(1 2 3 4)$	33 41 5 6
Estágio N $(0 1 2 3)$	46 17 17 5
Estágio M $(0 1)$	68 17
Localização $(SD MD ID SE IE^c)$	26 2 22 17 18
Posição $(C P^d)$	3 82
Diâmetro (milímetros)	$30,4 \pm 11,8 \ (10-74)^a$

^{*a*}Média \pm Desvio Padrão (Mínimo-Máximo). ^{*b*}1 NSCLC apresentando ambos os subtipos ADC e SCC, 1 NSCLC sem especificação, 1 LCC, 1 NSCLC neuroendócrino, 4 SCLC e 5 carcinoides. ^{*c*}Lobos superior direito, médio direito, inferior direito, superior esquerdo ou inferior esquerdo. ^{*d*}Central ou periférico.

Fonte: Autor.

primeiro validado com tumores de glioblastoma multiforme em imagens de RM (EGGER et al., 2013) e depois com tumores de pulmão em imagens de TC (VELAZQUEZ et al., 2013; PARMAR et al., 2014). O resultado da segmentação semiautomática com o GrowCut apresenta forte correlação com o diâmetro do tumor pulmonar na patologia (r = 0,89, intervalo de confiança de 95%, 0,81–0,94) e alta concordância com delineamentos manuais (frações de sobreposição > 0,90) (VELAZQUEZ et al., 2013). Além disso, atributos extraídos de imagens segmentadas com o GrowCut são mais robustos (intervalos significativamente menores entre observadores, valor-p < 0,001) e possuem significativamente maior reprodutibilidade em relação a atributos extraídos de imagens segmentadas manualmente (valor-p < 0,001) (PARMAR et al., 2014).

Para realizar a segmentação semiautomática das imagens no 3D-Slicer, duas regiões foram marcadas, uma interna e uma externa ao tumor (Figura 44(a)), em três cortes do exame de TC (um corte para cada plano anatômico – axial, sagital e coronal) utilizando uma janela de pulmão com nível de -500 UH e largura de 1400 UH. Em seguida, o algoritmo GrowCut detectou tridimensionalmente o tecido tumoral (Figura 44(b)). Após isto, a região externa ao tumor foi removida (Figura 44(c)), e a sua borda delimitada (Figura 44(d)). Finalmente, a marcação do tumor foi exportada como um arquivo estruturado do tipo DICOM-RT (do Inglês *digital imaging and communications in medicine radiation therapy*, Anexo F) (PINTER et al., 2012), a ser utilizado no processo de extração de atributos, explicado na sequência.

Figura 44: Segmentação semiautomática com o algoritmo GrowCut de tumor pulmonar em imagem de TC: (a) marcações internas e externas ao tumor, (b) resultado do crescimento de regiões, (c) remoção da região externa, (d) delimitação da borda da região interna.





5.2.2 Extração de Atributos

As imagens segmentadas de TC foram caracterizadas por 2.465 atributos calculados por três extratores diferentes (Figura 45). Características de intensidade de nível de cinza (número de atributos: 53), histograma (51), matriz de coocorrência (550 em 2D e 1.540 em 3D), matriz de comprimento de sequência (33), matriz de diferença de intensidade de vizinhança (10 em 2D e 10 em 3D), filtros de Laplaciano de Gaussiano (12) e forma (18) foram extraídas utilizando as imagens no formato DICOM junto do arquivo DICOM-RT com a ferramenta Ibex v1.0 (ZHANG et al., 2015). Atributos de textura de Tamura (36) e filtros de Gabor (120) foram calculados após a criação de imagens segmentadas a partir das marcações presentes no arquivo DICOM-RT com um algoritmo que utiliza a biblioteca de programação Lire-JFeatureLib v1.6.3 (LUX; MARQUES, 2013). Este algoritmo primeiro converte as marcações originais do DICOM-RT no sistema de coordenadas do paciente (Anexo F, *tag* (3006,0050)) para o sistema de coordenadas da imagem (Figura 46), de acordo com a Equação 5.1, para depois segmentar a imagem de TC e, por fim, realizar a extração dos atributos. Características quantitativas de transformada de Fourier (14), wavelet de Haar (16) e estimativa da dimensão fractal (2) foram computadas em imagens segmentadas e redimensionadas para valores potência de 2 para aplicar a transformada rápida de Fourier e a transformada de Haar e obter o espectro das imagens por meio da API (do Inglês application programming interface) da ferramenta ImageJ v1.50i (FALEIROS et al., 2017; SCHNEIDER; RASBAND; ELICEIRI, 2012).



Figura 45: Extratores de atributos utilizados na radiômica de imagens de TC.

Fonte: Autor.

$$i = \frac{x - x_0}{S}$$
 e $j = \frac{y - y_0}{S}$, (5.1)

onde i, j são as posições do ponto no sistema de coordenadas da imagem no plano cartesiano, x, y são as posições do ponto no sistema de coordenadas do paciente em milímetros, x_0, y_0 são as posições do primeiro *pixel* da imagem no sistema de coordenadas do paciente (dadas pela *tag* (0020,0032)) e S é o espaçamento do *pixel* dado pela *tag* (0028,0030).

Os atributos de histograma, COM, RLM, NIDM, Tamura, Gabor, Fourier, *wavelet* e fractal foram extraídos de imagens convertidas para 8 *bits* (256 níveis de cinza) para prevenir que as matrizes e o histograma fossem esparsamente produzidos e reduzir potencial ruído na textura das imagens (YANG et al., 2016; FAVE et al., 2015; FRIED, 2015). Os atributos de Tamura, Gabor, Fourier, *wavelet* e fractal foram calculados em 2D em cada imagem segmentada, e os volumes foram caracterizados pela média dos valores dos atributos extraídos em todos os cortes e também pelo valor do atributo no corte central.





Fonte: Autor.

Antes do cálculo dos atributos, foi feito um ajuste fino nos extratores com o objetivo de melhorar a caracterização das imagens por meio do aumento de configurações diferentes nos parâmetros dos atributos.

Os 53 atributos de nível de cinza extraídos foram energia, máximo global, média, mediana, mínimo global, variância, desvio padrão, curtose, assimetria, amplitude, raiz do quadrado médio, desvio absoluto médio, desvio absoluto mediano, amplitude interquartil, percentil, quantil, máxima entropia local, média da entropia local, mediana da entropia local, mínima entropia local, desvio padrão da entropia local, máxima amplitude local, média da amplitude local, mediana da amplitude local, mínima amplitude local, desvio padrão da amplitude local, máximo desvio padrão local, média do desvio padrão local, mediana do desvio padrão local, mínimo desvio padrão local e desvio padrão do desvio padrão local (ZHANG et al., 2015; AERTS et al., 2014). O ajuste do atributo de percentil variou de acordo com o intervalo 5–95 com incremento de 5, e o de quantil variou em 0,025, 0,25, 0,5, 0,75 e 0,975. Os atributos de entropia local utilizaram vizinhança de 9 *pixels*, e os de amplitude local e desvio padrão local utilizaram vizinhança de 5 *pixels* (YANG et al., 2016).

Os 51 atributos de histograma de níveis de cinza de 256 *bins* foram entropia, uniformidade, amplitude interquartil, curtose, assimetria, desvio absoluto médio, desvio absoluto mediano, percentil, área de percentil, quantil e amplitude. Os atributos de percentil e área de percentil variaram de acordo com o intervalo 5–95 com incremento de 5, e o de quantil variou em 0,025, 0,25, 0,5, 0,75 e 0,975 (ZHANG et al., 2015).

Os atributos de matriz de coocorrência foram calculados tanto em 2D quanto em 3D. A COM em 2D computou a ocorrência de pares de intensidade corte-a-corte, e a COM do volume foi a soma das ocorrências em todos os cortes. As distâncias variaram entre 1 e 5 pixels, e todas 4 orientações θ (0°, 45°, 90°, 135°) foram empregadas, resultando em 20 COMs em 2D. Vinte e dois atributos foram calculados das 20 matrizes, resultando em 440 características. Também foi calculada a média do valor dos 22 atributos em todas as orientações θ para cada uma das 5 distâncias, produzindo mais 110 características e resultando nos 550 atributos de COM em 2D. A COM em 3D computou diretamente a ocorrência dos pares de intensidade na orientação θ para as dimensões x, y e na orientação ϕ para a dimensão z. As distâncias variaram entre 1 e 5 voxels, e todas as 13 orientações $\theta \in \phi \ (\theta = 0^{\circ} \in \phi = 0^{\circ}, 45^{\circ}, 90^{\circ} \in 135^{\circ}; \theta = 45^{\circ} \in \phi = 45^{\circ}, 90^{\circ} \in 135^{\circ}; \theta = 90^{\circ} \in \phi = 45^{\circ}, \theta = 135^{\circ}; \theta = 135$ 45° , 90° e 135° ; $\theta = 135^{\circ}$ e $\phi = 45^{\circ}$, 90° e 135°) for am empregadas, resultando em 65 COMs em 3D. Os 22 atributos foram calculados das 65 matrizes, resultando em 1.430 características. Também foi calculada a média do valor dos 22 atributos em todas as orientações $\theta \in \phi$ para cada uma das 5 distâncias, produzindo mais 110 características e, finalmente, resultando nos 1.540 atributos de COM em 3D. Os 22 atributos calculados das matrizes foram autocorrelação, contraste, correlação, diferença inversa normalizada, dissimilaridade, energia, entropia, entropia da diferença, entropia da soma, homogeneidade, matiz, média da soma, duas medidas de informação de correlação, momento da diferença inversa normalizada, momento da diferença inverso, probabilidade máxima, proeminência, tendência, variância, variância inversa e variância da soma (ZHANG et al., 2015; AERTS et al., 2014; SOH; TSATSOULIS, 1999; HARALICK; SHAPIRO, 1992; HARALICK; SHANMUGAM; DINSTEIN, 1973).

Para a matriz de comprimento de sequência, onze atributos foram computados de 3 matrizes calculadas em 0° e 90° , e depois somadas, resultando nos 33 atributos de RLM. Os atributos extraídos foram ênfase de sequência curta, ênfase de sequência longa, não-uniformidade de nível de cinza, não-uniformidade de comprimento de sequência, porcentagem de sequência, ênfase de sequência de intensidade baixa, ênfase de sequência de intensidade alta, ênfase de intensidade baixa de sequência curta, ênfase de intensidade alta de sequência curta, ênfase de intensidade baixa de sequência longa e ênfase de intensidade alta de sequência longa (ZHANG et al., 2015; TANG, 1998; GALLOWAY, 1975).

A matriz de diferença de intensidade de vizinhança foi calculada tanto em 2D quanto em 3D, similarmente à COM. Os atributos de NIDM foram calculados com vizinhanças de tamanho 3 e 5 *pixels* e *voxels*, resultando em 4 matrizes. Cinco atributos foram computados das 4 matrizes, resultando nos 20 atributos de NIDM, 10 em 2D e 10 em 3D. Os 5 atributos foram granularidade, contraste, ocupação, complexidade e força (ZHANG et al., 2015; AMADASUN; KING, 1989). As características de textura de Tamura extraídas foram granularidade, contraste e direcionalidade, sendo esta última subdividida em 16 características (LUX; MARQUES, 2013; TAMURA; MORI; YAMAWAKI, 1978). Os volumes de imagens foram caracterizados pelo valor da média dos cortes e pelo valor do corte central, resultando nos 36 atributos de Tamura.

Os atributos de LOG foram calculados com dois filtros diferentes: um fino com tamanho 5 e escala 1 e um médio com tamanho 7 e escala 1,5 (ZHANG et al., 2015; FAVE et al., 2015). Seis medidas estatísticas foram extraídas das imagens filtradas (entropia, média, desvio padrão, uniformidade, curtose e assimetria), resultando nos 12 atributos de LOG.

A transformada rápida de Fourier foi aplicada para obter atributos do domínio da frequência das imagens (SCHNEIDER; RASBAND; ELICEIRI, 2012). Sete medidas foram calculadas do espectro da imagem: média, variância, desvio padrão, curtose, assimetria, valor máximo e coeficiente de variação (FALEIROS et al., 2017). A média dos cortes e o corte central caracterizaram os volumes das imagens, resultando nos 14 atributos de Fourier.

Filtros de Gabor foram aplicados nas imagens para obter as energias de cada banda de frequência no domínio do espectro, de acordo com 5 escalas e 6 orientações (LUX; MARQUES, 2013). Para cada imagem filtrada, a média e o desvio padrão foram calculados e, desta forma, 60 características foram extraídas. Como os volumes foram caracterizados pela média dos cortes e pelo corte central, isto resultou nos 120 atributos de Gabor.

A transformada *wavelet* de Haar foi aplicada para obter a energia em 3 bandas de alta e 1 de baixa frequência nos níveis 2 e 3 do espectro das imagens, o que resultou em 8 características (FALEIROS et al., 2017; SCHNEIDER; RASBAND; ELICEIRI, 2012). A média dos cortes e o corte central caracterizaram os volumes das imagens, resultando nos 16 atributos de *wavelet*.

A estimativa da dimensão fractal foi calculada pelo método de contagem por caixas, utilizando tamanhos máximo e mínimo de caixa de 24 e 6 *pixels*, respectivamente, e limiares de 50, 100, 150 e 200 para otimizar a busca por padrões em diferentes escalas (FALEIROS et al., 2017; SCHNEIDER; RASBAND; ELICEIRI, 2012). Como este método resulta em apenas uma característica e os volumes foram caracterizados pela média dos cortes e pelo corte central, isto resultou nos 2 atributos de fractal.

Os 18 atributos de forma extraídos a partir da marcação da borda das imagens foram área da superfície, duas medidas de compacidade, convexidade, disproporção esférica, esfericidade, extensão média, volume, densidade, número de objetos, número e tamanho de *voxels*, orientação, massa, máximo diâmetro 3D, redondeza, volumes de *convex hull* 2D e 3D (ZHANG et al., 2015; AERTS et al., 2014; LEGLAND; KIÊU; DEVAUX, 2007).

5.2.3 Análise Associativa

A análise radiômica do câncer de pulmão foi realizada em quatro etapas. Nas duas primeiras, a associação foi em relação à presença ou ausência de metástase distante (estadio M1 vs. estadio M0), presença ou ausência de metástase linfonodal (estadio N1-3 vs. estadio N0) e subtipo histopatológico de NSCLC (ADC vs. SCC, por possuírem os maiores números de amostras nos conjuntos de pacientes e por serem os tipos mais frequentes de câncer de pulmão). A terceira etapa analisou apenas os padrões de metástase distante. Na quarta etapa foi realizada uma avaliação prognóstica por meio da correlação entre os atributos e a sobrevida global dos pacientes (Figura 47). Uma análise preliminar de textura das imagens de TC também foi realizada neste trabalho com o objetivo de distinguir padrões de cavidade e pseudocavidade dos tumores pulmonares (Apêndice C). As metástases foram avaliadas de acordo com o estadiamento clínico do TNM por TC, cintilografia, tomografia por emissão de pósitrons ou cirurgia. A histopatologia foi confirmada por biópsia ou ressecção cirúrgica. As seções seguintes descrevem os experimentos realizados e os resultados obtidos em cada etapa da análise deste capítulo. É importante ressaltar que a radiômica não pretende substituir os especialistas, mas sim auxiliá-los em tarefas específicas fornecendo novas ferramentas e informações adicionais e, portanto, esta tese não focou em comparações radiômica vs. radiologista.

Figura 47: Exemplos dos complexos padrões visuais tumorais em imagens de TC utilizados nos experimentos de radiômica do câncer de pulmão: (a) estadio M1, (b) estadio M0, (c) estadio N1–3, (d) estadio N0, (e) ADC, (f) SCC, (g) alto risco com ocorrência de óbito e sobrevida global de 55 dias, (h) baixo risco sem ocorrência de óbito e sobrevida global de 2801 dias.



Fonte: Autor.

5.3 Experimentos e Resultados

5.3.1 Primeira Etapa

Nesta etapa da pesquisa, foram utilizados o Conjunto1 de pacientes para validação (Tabela 12), o Conjunto2 de pacientes para somente teste (Tabela 13) e o extrator de atributos Ibex v1.0 (Figura 45) (ZHANG et al., 2015). Três abordagens de caracterização foram empregadas nos experimentos: uma utilizando os atributos das imagens de TC, uma utilizando dados clínicos e uma utilizando atributos e dados clínicos combinados. As informações clínicas obtidas do prontuário eletrônico do paciente foram idade (anos), sexo (masculino ou feminino), tabagista (sim ou não) e histórico de outro câncer primário (sim ou não). Estes dados foram incorporados às informações radiológicas do tumor de localização (lobo direito superior, direito médio, direito inferior, esquerdo superior ou esquerdo inferior), posição (central ou periférico) e diâmetro no plano axial (milímetros), compondo os dados clínicos do paciente.

O algoritmo ReliefF selecionou os dados mais relevantes dos vetores de características das imagens (Seção 2.5) (KONONENKO, 1994; KIRA; RENDELL, 1992). Após a execução do método, as características foram listadas em ordem decrescente (da mais relevante à menos), de acordo com a relevância atribuída pelo ReliefF. O número de características selecionadas variou de 1 até as 100 melhores ranqueadas. O método ReliefF e sua função de ranqueamento estão implementados na API da ferramenta Weka v3.6.12 (FRANK; HALL; WITTEN, 2016).

Três algoritmos tradicionais de aprendizado de máquina associaram os atributos e dados clínicos selecionados pelo ReliefF com os desfechos: NB, KNN e RBF (Seção 2.6) (JOHN; LANGLEY, 1995; AHA; KIBLER; ALBERT, 1991; MOODY; DARKEN, 1989). Os valores de k para o KNN e o de neurônios da camada oculta da rede RBF variaram de 1 até o número de amostras disponíveis da classe majoritária. O KNN utilizou a distância Euclidiana como medida de cálculo de vizinhança entre amostras, e a RBF utilizou agrupamento com o algoritmo k-means no treinamento da RNA. Devido os atributos terem valores com intervalos diferentes e alguns classificadores utilizarem uma medida de distância no mecanismo de treinamento, os dados quantitativos foram normalizados entre 0 e 1, e os dados qualitativos foram binarizados antes da análise. Todos os classificadores também estão implementados na API da ferramenta Weka v3.6.12 (FRANK; HALL; WITTEN, 2016).

Como o número de amostras por classe do conjunto de pacientes para validação estava muito desbalanceado para a análise por aprendizado de máquina (por exemplo, 44 pacientes sem metástase distante vs. 8 pacientes com metástase distante), foram realizados dois diferentes testes de balanceamento. O primeiro consistiu em excluir amostras da classe majoritária, utilizando o filtro *SpreadSubsample* da ferramenta Weka v3.6.12 (FRANK; HALL; WITTEN, 2016). Desta forma, o conjunto de pacientes para validação apresentou 8 amostras por padrão de metástase distante, 23 amostras por padrão de metástase linfonodal e 15 amostras por padrão de histopatologia. O segundo teste de balanceamento consistiu em criar amostras sintéticas para a classe minoritária, utilizando a técnica SMOTE (do Inglês *synthetic minority over-sampling technique*), também implementada na ferramenta Weka v3.6.12 (FRANK; HALL; WITTEN, 2016; EMAMINEJAD et al., 2016; CHAWLA et al., 2002). Desta forma, o conjunto de pacientes para validação apresentou 44 amostras por padrão de metástase distante, 29 amostras por padrão de metástase linfonodal e 30 amostras por padrão de histopatologia.

Os classificadores de aprendizado de máquina foram avaliados pelas métricas AUC, sensibilidade e especificidade (Seção 2.6.2). O Conjunto1 de pacientes foi primeiro avaliado com uma validação cruzada *leave-one-out*. Depois de obtido o modelo de maior performance com o conjunto de pacientes para validação, este modelo foi avaliado com o conjunto independente de pacientes (Conjunto2) para somente teste.

As Tabelas 15, 17 e 19 apresentam as maiores performances associativas para os padrões de metástase distante, metástate linfonodal e histopatologia, respectivamente, utilizando diferentes tipos de características (atributos de imagens, dados clínicos e estes dois combinados), classificadores (NB, KNN e RBF) e conjuntos de amostras (nãobalanceado e balanceado com exclusão e criação de amostras) para validação e teste.

A rede RBF com todas as características combinadas obteve a maior performance associativa para metástase distante, em nível de validação com AUC de 0,97, utilizando amostras sintetizadas pelo método SMOTE. Contudo, vários cenários obtiveram desempenho estatisticamente equivalentes, como por exemplo, a própria RBF também com todas as características combinadas utilizando, porém, o conjunto balanceado de pacientes com exclusão de amostras. Por outro lado, a maior performance associativa utilizando o conjunto para somente testes foi obtida pelo NB com apenas atributos de imagens e as amostras não-balanceadas como treinamento e pelo KNN com dados clínicos e as amostras balanceadas com o SMOTE (AUC de 0,71). Entretanto, o primeiro cenário obteve sensibilidade de 0,80 (intervalo de confiança de 95% (IC): 0,30-0,99) e especificidade de 0,46 (IC: 0,18-0,75), enquanto que o segundo obteve sensibilidade de 0,40 (IC: 0,07-0,83) e especificidade de 0,91 (IC: 0,57-0,99). As matrizes de confusão destes dois cenários estão na Tabela 16.

O classificador NB com todas as características combinadas obteve a maior performance associativa para metástase linfonodal, em nível de validação com AUC de 0,89, utilizando amostras sintetizadas pelo método SMOTE. Contudo, vários cenários obtiveram desempenho estatisticamente equivalente, como por exemplo, o KNN e a rede RBF utilizando o mesmo conjunto balanceado de amostras. O KNN utilizando dados clínicos obteve a maior performance associativa para o conjunto de somente testes com AUC

	periori	mances estad com valo	i em negrito.	
		Não-balanceado	Balanceado-	Balanceado+
		(8 MD+ vs. 44 MD-)	(8 MD+ vs. 8 MD-)	(44 MD+ vs. 44 MD-)
tos ens	NB	$0,69 \mid 0,71$	$0,78 \mid 0,46$	$0,88 \mid 0,58$
ribut imag	KNN	$0,75 \mid 0,36$	$0,88 \mid 0,47$	0,96 0,55
At de j	RBF	$0,84 \mid 0,39$	$0,\!95 \mid 0,\!57$	0,95 0,38
s s	NB	$0,\!43 \mid 0,\!43$	$0,31 \mid 0,60$	$0,90 \mid 0,55$
)ado ínice	KNN	$0,77 \mid 0,63$	$0,54 \mid 0,56$	$0,85 \mid 0,71$
cl T	RBF	$0,78 \mid 0,62$	$0,58 \mid 0,53$	$0,84 \mid 0,65$
sop	NB	$0,69 \mid 0,67$	$0,78 \mid 0,46$	0,95 0,60
odo. bina	KNN	$0,75 \mid 0,36$	$0,88 \mid 0,47$	$0,94 \mid 0,64$
L	RBF	$0,84 \mid 0,39$	$0,95 \mid 0,57$	0,97 0,36

Tabela 15: Performance da associação radiômica com metástase distante. AUCs são apresentadas no formato $x \mid y$, onde x é o desempenho utilizando o conjunto de validação e y utilizando o conjunto de teste. Algumas das maiores performances estão com valor em negrito.

Balanceado+ e Balanceado- representam os conjuntos de pacientes após balanceamento por criação e exclusão de amostras, respectivamente. MD+ e MD- significam presença e ausência de metástase distante, respectivamente.

Fonte: Autor.

Tabela 16: Matrizes de confusão das maiores performances associativas para metástase distante utilizando o conjunto de pacientes para somente testes. Cenários são apresentados com o tipo de característica, classificador e conjunto para treinamento utilizados.

Atributos, NB, Não-balanceado			Dad	Dados clínicos, KNN, Balanceado+				
	Prediç	ão			Prediçã	ão		
	MD+	MD-			MD+	MD-		
HD+ ھ	4	1	۲ <u>ج</u>	MD+	2	3		
ස් MD-	6	5	$^{\rm R}$	MD-	1	10		

Balanceado+ representa o conjunto de pacientes após balanceamento por criação de amostras sintéticas. MD+ e MD- significam presença e ausência de metástase distante, respectivamente.

Fonte: Autor.

de 0,76, sensibilidade de 0,67 (IC: 0,31–0,91) e especificidade de 0,86 (IC: 0,42–0,99). Entretanto, outros cenários também obtiveram performance equivalente, como a RBF utilizando os atributos das imagens com AUC de 0,75, sensibilidade de 0,67 (IC: 0,31–0,91) e especificidade de 0,71 (IC: 0,30–0,95). As matrizes de confusão destes dois cenários estão na Tabela 18.

		com valor em negrito.		
		Não-balanceado	Balanceado-	Balanceado+
		(23 ML+ vs. 29 ML-)	(23 ML+ vs. 23 ML-)	(29 ML+ vs. 29 ML-)
Atributos de imagens	NB	$0,84 \mid 0,73$	$0,86 \mid 0,73$	$0,86 \mid 0,73$
	KNN	$0,82 \mid 0,71$	$0,85 \mid 0,72$	$0,85 \mid 0,71$
	RBF	$0,82 \mid 0,56$	$0,82 \mid 0,68$	$0,85 \mid 0,75$
Dados clínicos	NB	$0,84 \mid 0,70$	$0,82 \mid 0,68$	$0,86 \mid 0,71$
	KNN	$0,84 \mid 0,68$	$0,83 \mid 0,71$	0,85 0,76
	RBF	$0,84 \mid 0,70$	$0,81 \mid 0,67$	$0,86 \mid 0,73$
Todos combinados	NB	$0,85 \mid 0,73$	$0,85 \mid 0,73$	0,89 0,71
	KNN	0,87 0,68	$0,85 \mid 0,72$	0,88 0,73
	RBF	$0,83 \mid 0,70$	$0,87 \mid 0,62$	$0,87 \mid 0,73$

Tabela 17: Performance da associação radiômica com metástase linfonodal. AUCs são apresentadas no formato $x \mid y$, onde x é o desempenho utilizando o conjunto de validação e y utilizando o conjunto de teste. Algumas das maiores performances estão com valor em negrito.

Balanceado+ e Balanceado- representam os conjuntos de pacientes após balanceamento por criação e exclusão de amostras, respectivamente. ML+ e ML- significam presença e ausência de metástase linfonodal, respectivamente.

Fonte: Autor.

Tabela 18: Matrizes de confusão das maiores performances associativas para metástase linfonodal utilizando o conjunto de pacientes para somente testes. Cenários são apresentados com o tipo de característica, classificador e conjunto para treinamento utilizados.

Dados clí	nicos, K	NN, Balanceado+	Atributos	, RBF, I	Balanceado+
	Prediç	ão		Prediç	ão
	ML+	ML-		ML+	ML-
ה ML+	6	3	ਯੂ ML+	6	3
ਔ ML-	1	6	ਔ ML-	2	5

Balanceado+ representa o conjunto de pacientes após balanceamento por criação de amostras sintéticas. ML+ e ML- significam presença e ausência de metástase linfonodal, respectivamente.

Fonte: Autor.

A maior performance associativa para histopatologia apresentou AUC de 0,92 em nível de validação obtida por todas as características combinadas, a rede RBF e o conjunto balanceado com exclusão de amostras. Contudo, vários cenários obtiveram desempenho estatisticamente equivalente, e.g. NB com apenas dados clínicos e o conjunto balanceado com amostras sintetizadas. As maiores performances em nível de teste foram obtidas pelo NB com os atributos das imagens e os conjuntos não-balanceado e balanceado

		0		
		Não-balanceado	Balanceado-	Balanceado+
		(30 ADC <i>vs.</i> 15 SCC)	(15 ADC vs. 15 SCC)	(30 ADC vs. 30 SCC)
Atributos de imagens	NB	$0,60 \mid 0,81$	$0,71 \mid 0,55$	$0,62 \mid 0,81$
	KNN	$0,73 \mid 0,54$	$0,71 \mid 0,74$	$0,76 \mid 0,64$
	RBF	$0,83 \mid 0,55$	$0,75 \mid 0,41$	$0,80 \mid 0,45$
Dados clínicos	NB	$0,80 \mid 0,60$	$0,80 \mid 0,66$	0,90 0,67
	KNN	$0,80 \mid 0,60$	$0,78 \mid 0,69$	$0,89 \mid 0,66$
	RBF	$0,82 \mid 0,52$	$0,80 \mid 0,68$	$0,91 \mid 0,57$
Todos combinados	NB	$0,68 \mid 0,81$	$0,85 \mid 0,52$	$0,85 \mid 0,74$
	KNN	$0,73 \mid 0,61$	$0,84 \mid 0,67$	0,90 0,74
	RBF	$0,86 \mid 0,71$	$0,92 \mid 0,59$	$0,87 \mid 0,67$

Tabela 19: Performance da associação radiômica com histopatologia. AUCs são apresentadas no formato $x \mid y$, onde x é o desempenho utilizando o conjunto de validação e y utilizando o conjunto de teste. Algumas das maiores performances estão com valor em negrito.

Balanceado+ e Balanceado- representam os conjuntos de pacientes após balanceamento por criação e exclusão de amostras, respectivamente.

Fonte: Autor.

com amostras sintéticas, e também com todas as características combinadas utilizando o conjunto não-balanceado de pacientes. Entretanto, este último cenário obteve taxa de verdadeiro positivo para ADC de 1,00, enquanto que os atributos das imagens obtiveram 0,86 (conjunto não-balanceado, IC: 0,42–0,99) e 0,71 (balanceado com amostras sintéticas, IC: 0,30–0,95). As matrizes de confusão destes três cenários estão na Tabela 20.

Tabela 20: Matrizes de confusão das maiores performances associativas para histopatologia utilizando o conjunto de pacientes para somente testes. Cenários são apresentados com o tipo de característica, classificador e conjunto para treinamento utilizados.

Todos combinados, RBF, Balanceado-		Dados clínicos, RBF, Balanceado+				
Predição		Predição				
ADC	SCC		ADC	SCC		
ਕ ADC 7	0	ਰ ADC	6	1		
$\stackrel{\circ}{ m H}$ SCC 2	4	ar SCC	2	4		
Todos combinados, KNN, Balanceado+						
Predic	ção					
ADC	\mathbf{SCC}					
Fa ADC 5	2					
x SCC 4	2					

Balanceado+ e Balanceado- representam os conjuntos de pacientes após balanceamento por criação e exclusão de amostras, respectivamente.

Fonte: Autor.

5.3.2 Segunda Etapa

Nesta fase do trabalho, foram utilizados o Conjunto3 de pacientes (Tabela 14) e os três extratores de atributos investigados na radiômica do câncer de pulmão (Figura 45) (ZHANG et al., 2015; LUX; MARQUES, 2013; SCHNEIDER; RASBAND; ELICEIRI, 2012). Esta segunda etapa da pesquisa visou principalmente aumentar o número de amostras para a análise, verificar a viabilidade dos outros dois extratores de atributos das imagens de TC não utilizados nos experimentos da primeira etapa e avaliar o potencial individual de todos os atributos extraídos.

A associação radiômica foi realizada por análises univariada e multivariada e avaliada pela AUC. O Teste U de Mann-Whitney realizou a análise univariada avaliando a diferença estatística nos atributos entre as classes. A AUC e o valor-p foram calculados para cada atributo individualmente nos três desfechos clínicos estudados na primeira etapa, utilizando a API Apache Commons Mathematics v3.6.1. A análise multivariada foi realizada com a rede neural RBF, o mesmo algoritmo de avaliação da etapa anterior com o método ReliefF para seleção de características relevantes e os dados clínicos do paciente combinados com os atributos para a composição do vetor de características das imagens.

Novamente, devido o número de amostras por classe ser bastante desbalanceado para a análise por aprendizado de máquina (por exemplo, 68 pacientes sem metástase distante vs. 17 pacientes com metástase distante), diferentes experimentos foram realizados para diminuir esta limitação. As mesmas estratégias da primeira etapa compostas pela exclusão e criação de amostras para as classes majoritária e minoritária, respectivamente, foram replicadas nesta fase. Estes experimentos foram avaliados após validações cruzadas *leave-one-out* e com 10 subconjuntos. Além disso, outra estratégia foi criar amostras sintéticas com o SMOTE e utilizar somente elas para o treinamento e as amostras reais somente para a validação. E um último experimento consistiu em dividir aleatoriamente o conjunto de amostras em dois subconjuntos para treinamento e validação, e antes de treinar a RBF, balancear as classes e aumentar a amostragem, utilizando o SMOTE, para potencialmente melhorar a performance do classificador.

A análise univariada identificou 29, 1.994 e 7 atributos com significância estatística para associação com metástase distante, metástase linfonodal e histopatologia, respectivamente. As Figuras 48–50 apresentam os 20 atributos com maior performance individual de associação radiômica. A seguinte notação foi empregada para a significância estatística: *** para valor-p < 0,001, ** para $0,001 \le \text{valor-} p < 0,01$ e * para $0,01 \le \text{valor-} p < 0,05$ (FRIGHETTO-PEREIRA et al., 2016).

Os atributos de *wavelet* e forma obtiveram as associações estatisticamente mais significativas para metástase distante, com 8 e 7 atributos dentre os 20 de maior performance, respectivamente. Os atributos de *wavelet*, forma e COM obtiveram as associações




Fonte: Autor.

estatisticamente mais significativas para metástase linfonodal, com 8, 6 e 5 atributos dentre os 20 de maior performance, respectivamente. Variações do atributo probabilidade máxima da COM em 2D e 3D obtiveram as associações estatisticamente mais significativas para histopatologia. Os maiores desempenhos para cada categoria de atributo estão nas Figuras 51–53. As Figuras 54–56 apresentam *boxplots* dos 7 atributos com maior performance de associação radiômica. Como é observado principalmente pelos atributos de *wavelet* para ambas as metástases, as diferenças nos grupos são maiores quando a AUC e/ou a significância estatística aumenta.

A análise multivariada conseguiu aumentar a performance da univariada, exceto na associação com metástase linfonodal (valor-p < 0.05). O algoritmo de aprendizado de máquina obteve as maiores AUCs de 0.92 (com 13 características selecionadas e 15 neurônios na camada oculta da RBF), 0.84 (com 11 características selecionadas e 16 neurônios na camada oculta da RBF) e 0.88 (com 45 características e 20 neurônios na camada oculta da RBF) e 0.88 (com 45 características e 20 neurônios na camada oculta da RBF) para a associação com metástase distante, metástase linfonodal e histopatologia, respectivamente, excluindo desempenhos onde pode ter havido viés devido à utilização de amostras sintéticas no conjunto de validação. As Tabelas 21–23 apresentam





Fonte: Autor.

o desempenho completo da análise multivariada. As Figuras 57–59 apresentam mapas de calor retratando a avaliação pelo algoritmo de aprendizado de máquina dos cenários de maiores performances.

As características selecionadas pelo método ReliefF foram marcadamente determinadas por diversas categorias. Características de intensidade (quantidade: 3), clínicas (2), RLM (2), NIDM (2), histograma (1), COM (1), forma (1) e fractal (1) obtiveram a maior performance para a associação com metástase distante. Características de COM (5), clínicas (2), forma (2), Gabor (1) e fractal (1) obtiveram o maior desempenho para metástase linfonodal. Características de COM (19), clínicas (7), forma (6), Tamura (5), intensidade (3), Fourier (3), Gabor (1) e fractal (1) obtiveram a maior performance para histopatologia. As listas completas das características relevantes selecionadas para a análise multivariada estão nas Tabelas 24–26.



COM3D_MaxProbability_d1_0135_0135 COM2D_MaxProbability_d1_045 COM3D_MaxProbability_d1_ θ 0_ Φ 0 $COM3D_MaxProbability_d1_\theta135_\Phi90$ COM3D_MaxProbability_d1_090_0135 COM3D_MaxProbability_d1_ θ 0_ Φ 135 COM3D_MaxProbability_d1_00_045 COM3D_MaxProbability_d2_00_Ф135 COM3D_MaxProbability_d4_090_0135 COM3D_MaxProbability_d1_ 045_0135 COM3D_MaxProbability_d2_ 045_0135 COM2D_MaxProbability_d5_00 COM3D_MaxProbability_d5_ θ 90_ Φ 90 $COM3D_MaxProbability_d3_\theta90_\Phi135$ COM3D_MaxProbability_d2_ $\theta 0_{\Phi} 0$ COM3D_MaxProbability_d1_0135_045 COM3D_MaxProbability_average_d1 $COM3D_MaxProbability_d2_\theta0_\Phi45$ COM3D_MaxProbability_d2_0135_0135 COM2D_MaxProbability_d1_090



Fonte: Autor.

Figura 51: Maiores performances individuais por categoria de atributo para a associação com metástase distante.

feature_id AUC

0.656

0.654

0.654

0.651

0.651

0.650

0.644

0.644

0.644 0.643

0.596

0.587

0.562

0.557

0.553

0.551

0.549

1450

340

1490

315

1420

1520

934

1440

1519

1460

2270

2271

2266

2276

2267

2274

2269

feature_id AUC

significância

*

-

_

-

-

-

significância

-

-

-

Intensidade						
feature_id	feature_id AUC					
2175	0.663	*				
2124	0.647	-				
2135	0.619	-				
2142	0.615	-				
2136	0.610	-				
2137	0.606	-				
2126	0.606	-				
2162	0.587	-				
2128	0.587	-				
2160	0.587	-				

NIDM						
feature_id	AUC	significância				
2246	0.663	*				
2228	0.646	-				
2233	0.646	-				
2239	0.644	-				
2230	0.644	-				
2244	0.640	-				
2238	0.621	-				
2236	0.615	-				
2241	0.609	-				
2231	0.602	-				

Fourier							
feature_id	AUC	significância					
2450	0.663	*					
2440	0.645	-					
2435	0.638	-					
2438	0.636	-					
2452	0.634	-					
2434	0.622	-					
2454	0.606	-					
2455	0.606	-					
2436	0.600	-					
2451	0.592	-					

Histograma						
feature_id	AUC	significância				
2226	0.663	*				
2194	0.587	-				
2212 0.587		-				
2192	0.587	-				
2193	0.587	-				
2223 0.587		-				
2208	0.584	-				
2210	0.584	-				
2211	0.584	-				
2213	0.584	-				

Tamura							
feature_id	feature_id AUC signific						
2425	0.669	*					
2345	*						
2418	0.663	*					
2346	0.661	*					
2424	0.661	*					
2347	0.659	*					
2348	0.652	-					
2426	0.650	-					
2427	0.646	-					
2423	0.643	-					

Wavelet							
feature_id	AUC	significância					
2443	0.704	**					
2446	0.704	* *					
2442	0.700	*					
2444	0.700	*					
2447	0.691	*					
2448	0.671	*					
2441	0.670	*					
2445	0.669	*					
2462	0.659	*					
2464	0.656	*					

2277	0.538	-
2268	0.528	-
2275	0.523	-
	Form	а
feature_id	AUC	significância
2250	0.684	*
2255	0.676	*
2262	0.673	*
2264	0.671	*
2251	0.670	*
2253	0.666	*
2252	0.664	*
2248	0.654	-
2254	0.646	-
2257	0.642	-

	2109	0.645	-
	2110	0.642	-
	2093	0.640	-
	2091	0.638	-
	2092	0.638	-
I	2101	0.620	-
	2098	0.606	-
	2113	0.601	-
	2116	0.598	-
		Gabo	r
	feature_id	AUC	significância

RLM

feature_id AUC significância

0.646

2111

Gabor						
feature_id	AUC	significância				
2357	0.624	-				
2359	0.624	-				
2361	0.624	-				
2363	0.624	-				
2365	0.624	-				
2367	0.624	-				
2369	0.624	-				
2371	0.624	-				
2373	0.624	-				
2375	0.624	-				
2373	0.021					

Fractal							
feature_id	AUC	significância					
2449	0.663	*					
2465	0.629	-					

Figura	52:	Maiores	performances	individuais	por	categoria	de	atributo	para	a	associa	ıção
		com met	ástase linfono	odal.								

2270 2271

2266

2267

2269

2268

2276

2274

0.728

0.725

0.716

0.712

0.682

0.681

0.652

0.640

Intensidade						
feature_id	AUC	significância				
2124	0.734	***				
2177	0.713	***				
2133	0.700	**				
2135	0.698	**				
2154	0.693	**				
2131	0.691	**				
2136	0.690	**				
2155	0.690	**				
2171	0.690	**				
2156	0.689	**				

NIDM			
feature_id	AUC	significância	
2239	0.725	***	
2228	0.711	***	
2238	0.706	**	
2244	0.702	**	
2233	0.696	**	
2231	0.693	**	
2229	0.679	**	
2243	0.666	**	
2234	0.656	*	
2236	0.649	*	

Fourier			
feature_id	AUC	significância	
2450	0.723	***	
2452	0.718	***	
2434	0.713	***	
2454	0.672	**	
2435	0.671	**	
2438	0.671	**	
2436	0.664	**	
2451	0.659	*	
2453	0.630	*	
2455	0.581	-	

Histograma			
feature_id	AUC	significância	
2227	0.713	***	
2180	0.700	**	
2186	0.693	**	
2205	0.692	**	
2187	0.690	**	
2206	0.690	**	
2222	0.690	**	
2188	0.689	**	
2207	0.689	**	
2185	0.686	**	

Tamura			
feature_id AUC		significância	
2340	0.745	***	
2418	0.734	***	
2345	0.729	***	
2343	0.726	***	
2425	0.721	***	
2423	0.718	***	
2347	0.717	***	
2426	0.717	***	
2351	0.716	***	
2350	0.714	* * *	

Wavelet			
feature_id	AUC	significância	
2446	0.768	***	
2447	0.756	***	
2448	0.756	***	
2442	0.744	***	
2443	0.742	***	
2463	0.739	***	
2464	0.737	***	
2462	0.737	***	
2444	0.736	***	
2459	0.736	***	

2277	0.621	-			
2273	0.601	-			
	Form	а			
feature_id	AUC	significância			
2251	0.755	***			
2262	0.751	***			
2253	0.750	***			
2264	0.746	***			
2252	0.744	***			
2254	0.741	***			
2257	0.724	***			
2248	0.714	***			
2249	0.714	***			
2260	0.714	***			

COM

2291	0.695	**
2293	0.695	**
2295	0.695	**
2297	0.695	**
	Fracta	al
feature_id	Fracta AUC	al significância
feature_id 2465	Fracta AUC 0.586	al significância -

2279

2281

2283

2285

2287

2289

feature_id	AUC	significância		feature_id	AUC	significância
1433	0.743	***		2093	0.729	***
1504	0.742	***		2091	0.729	***
1503	0.740	***		2092	0.728	***
1434	0.739	***		2111	0.712	***
1432	0.739	***		2109	0.711	***
345	0.735	***		2110	0.708	**
1470	0.735	***		2101	0.695	**
1527	0.735	***		2100	0.692	**
1502	0.734	***		2116	0.684	**
344	0.734	***		2115	0.683	**
	LOG				Gabo	r
feature_id	AUC	significância		feature_id	AUC	significância
2270	0.728	***		2279	0.695	**

**

**

*

*

	IVE IVI	
feature_id	AUC	significância
2093	0.729	***
2091	0.729	***
2092	0.728	***
2111	0.712	***
2109	0.711	***
2110	0.708	**
2101	0.695	**
2100	0.692	**
2116	0.684	**
2115	0.683	**

0.695

0.695

0.695

0.695

0.695

0.695

**

**

**

**

**

Figura 53: Maiores performances individuais por categoria de atributo para a associação com histopatologia.

Intensidade			
feature_id	feature_id AUC		
2131	0.639	-	
2177	0.623	-	
2150	0.622	-	
2154	0.620	-	
2155	0.620	-	
2171	0.620	-	
2125	0.619	-	
2158	0.618	-	
2153	0.616	-	
2176	0.615	-	

NIDM			
feature_id	AUC	significância	
2247	0.631	-	
2243	0.615	-	
2237	0.614	-	
2229	0.602	-	
2239	0.591	-	
2238	0.590	-	
2242	0.589	-	
2231	0.585	-	
2236	0.583	-	
2241	0.579	-	

Fourier			
feature_id AUC significância			
2453	0.631	-	
2451	0.629	-	
2454	0.615	-	
2455	0.598	-	
2456	0.573	-	
2437	0.567	-	
2436	0.556	-	
2438	0.554	-	
2435	0.552	-	
2434	0.537	-	

Histograma			
feature_id	AUC	significância	
2227	0.623	-	
2182	0.622	-	
2205	0.620	-	
2206	0.620	-	
2186	0.620	-	
2187	0.620	-	
2222	0.620	-	
2190	0.618	-	
2209	0.618	-	
2185	0.616	-	

Tamura				
feature_id	AUC	significância		
2354	0.587	-		
2338	0.579	-		
2427	0.577	-		
2351	0.576	-		
2349	0.576	-		
2353	0.576	-		
2416	0.575	-		
2428	0.574	-		
2422	0.573	-		
2341	0.570	-		

Wavelet				
feature_id	AUC	significância		
2445	0.601	-		
2441	0.596	-		
2457	0.580	-		
2461	0.569	-		
2447	0.564	-		
2448	0.558	-		
2443	0.556	-		
2444	0.554	-		
2442	0.547	-		
2460	0.537	-		

2271	0.528	-
2275	0.523	-
	Form	а
feature_id	AUC	significância
2250	0.599	-
2248	0.572	-
2263	0.567	-
2257	0.559	-
2253	0.556	-
2249	0.550	-
2260	0.550	-
2261	0.550	-
2256	0.549	-
2264	0.546	-

COM

0.660

0.652

0.652

0.652

0.651

0.651

0.649

0.647

0.646

0.646 LOG

feature_id AUC significância

0.595

0.580

0.572

0.570

0.563

0.547

0.541

0.538

-

-

-

-

feature_id AUC

1806

436

1756

1766

1776

1786

1781

1787

1779

1801

2268

2267

2266

2273

2274

2272

2270

2269

2202	0.011				
2284	0.611	-			
2286	0.611	-			
2288	0.611	-			
2290	0.611	-			
2292	0.611	-			
2294	0.611	-			
2296	0.611	-			
· · ·					
Fractal					
feature_id	AUC	significância			

Fractal			
feature_id	AUC	significância	
2465	0.628	-	
2449	0.592	-	

			RLM	
significância		feature_id	AUC	
*		2094	0.616	
*		2095	0.615	
*		2096	0.614	
*	1	2100	0.612	
*		2098	0.611	
*		2102	0.609	
*		2112	0.609	
-		2115	0.609	
-		2097	0.608	

feature_id

2278

2280

2282

RLM				
feature_id	AUC	significância		
2094	0.616	-		
2095	0.615	-		
2096	0.614	-		
2100	0.612	-		
2098	0.611	-		
2102	0.609	-		
2112	0.609	-		
2115	0.609	-		
2097	0.608	-		
2113	0.607	-		

Gabo

0.611

0.611

0.611

AUC significância

_

Figura 54: Distribuição dos valores normalizados entre [0,1] de atributos significativamente associados com metástase distante. Os atributos estão listados com nome, AUC e significância estatística. MD+ e MD- significam presença e ausência de metástase distante, respectivamente.



Fonte: Autor.

Figura 55: Distribuição dos valores normalizados entre [0,1] de atributos significativamente associados com metástase linfonodal. Os atributos estão listados com nome, AUC e significância estatística. ML+ e ML- significam presença e ausência de metástase linfonodal, respectivamente.





Figura 56: Distribuição dos valores normalizados entre [0,1] de atributos significativamente associados com histopatologia. Os atributos estão listados com nome, AUC e significância estatística.



Fonte: Autor.

Tabela 21: Maiores AUCs de cada experimento da análise multivariada para a associação com metástase distante.

Validação emuzada	Não-balanceado	Balanceado-	Balanceado+	
10 sub serientes	(68 MD- vs. 17 MD+)	(17 MD- vs. 17 MD+)	(68 MD- vs. 68 MD+)	
10 subconjunios	0,86	$0,\!86$	0,96	
Validação emuzada	Não-balanceado	Balanceado-	Balanceado+	
	(68 MD- vs. 17 MD+)	(17 MD- vs. 17 MD+)	(68 MD- vs. 68 MD+)	
leave-one-out	0,86	0,92	0,95	
Amostras sintéticas para treinamento,		(68 MD- $vs.$ 68 MD+ para treinamento e		
		68 MD- vs. 17 MD+ para validação)		
Amostras reais para vano	laçao	0,94		
Validação haldout	A (34 MD- vs. 34 MD+)	A (68 MD- vs. 68 MD+)	A (102 MD- vs. 102 MD+)	
(subconjuntos A e B)	0,76	0,76	0,79	
	B (34 MD- vs. 34 MD+)	B (68 MD- vs. 68 MD+)	B (102 MD- vs. 102 MD+)	
com amostras sinteticas	0,74	0,83	0,78	

Balanceado+ e Balanceado- representam os conjuntos de pacientes após balanceamento por criação e exclusão de amostras, respectivamente. MD+ e MD- significam presença e ausência de metástase distante, respectivamente.

Fonte: Autor.

Tabela 22: Maiores AUCs de cada experimento da análise multivariada para a associação com metástase linfonodal.

V-1: 1	Não-balanceado	Balanceado-	Balanceado+	
validação cruzada	(46 ML- vs. 39 ML+)	(39 ML- vs. 39 ML+)	(46 ML- vs. 46 ML+)	
10 subconjuntos	0,76	0,78	0,76	
Validação aruzada	Não-balanceado	Balanceado-	Balanceado+	
Vanuação cruzada	(46 ML- vs. 39 ML+)	(39 ML- vs. 39 ML+)	(46 ML- vs. 46 ML+)	
leave-one-oui	0,78	0,78	0,76	
Amostras sintéticas para treinamento,		(46 ML- $vs.$ 46 ML+ para treinamento e		
		46 ML- $vs.$ 39 ML+ para	validação)	
Amostras reais para vanc	laçao	0,93		
Validação holdout	A (23 ML- $vs.$ 23 ML+)	A (46 ML- vs. 46 ML+)	A (69 ML- vs. 69 ML+)	
(subconjuntos A e B)	0,84	0,80	0,81	
	B (23 ML- vs. 23 ML+)	B (46 ML- vs. 46 ML+)	B (69 ML- vs. 69 ML+)	
com amostras sintericas	0,75	0,78	0,76	

Balanceado+ e Balanceado- representam os conjuntos de pacientes após balanceamento por criação e exclusão de amostras, respectivamente. ML+ e ML- significam presença e ausência de metástase linfonodal, respectivamente.

Tabela 23:	Maiores	AUCs d	e cada	experimento	da	análise	multivariada	para	a a	issocia	ção
	com hist	topatolog	gia.								

Validação cruzada	Não-balanceado	Balanceado-	Balanceado+					
10 subconjuntos	(51 ADC vs. 21 SCC)	(21 ADC vs. 21 SCC)	(51 ADC vs. 51 SCC)					
10 subconjuntos	0,78	$0,\!83$	$0,\!93$					
Validação cruzada	Não-balanceado	Balanceado-	Balanceado+					
logue one out	(51 ADC vs. 21 SCC)	(21 ADC vs. 21 SCC)	(51 ADC vs. 51 SCC)					
ieuve-one-oui	0,77	$0,\!87$	$0,\!92$					
Amostras sintáticas para	trainamento	(51 ADC vs. 51 SCC para treinamento e						
Amostras sinteticas para	tremamento,	51 ADC vs. 21 SCC para validação)						
Amostras reais para vanc	laçao	0,95						
Validação haldout	A (25 ADC vs. 25 SCC)	A (50 ADC vs. 50 SCC)	A (75 ADC vs. 75 SCC)					
(anhaqiao nonooni	$0,\!66$	0,73	0,79					
(subconjuntos A e B)	B (26 ADC vs. 26 SCC)	B (52 ADC vs. 52 SCC)	B (78 ADC vs. 78 SCC)					
com amostras sinteticas	$0,\!74$	0,84	0,88					

Balanceado+ e Balanceado- representam os conjuntos de pacientes após balanceamento por criação e exclusão de amostras, respectivamente.

Fonte: Autor.

Tabela 24: Características selecionadas no cenário de maior performance da análise multivariada para a associação com metástase distante.

Relevância no intervalo [-1,1]	Característica
0,03824	Clinical_Gender
0,03243	Intensity_Global_Minimum
0,02941	Clinical_Smoker_status
0,02803	Shape_Orientation
0,02639	NIDM2D_Complexity_N3
0,02390	Intensity_Range
0,02390	Histogram_Range
0,01887	$COM3D_Correlation_d2_\theta 90_\phi 45$
0,01845	Intensity_Global_Maximum
0,01651	RLM_Long_Run_Low_Gray-level_Emphasis_ $\theta 90$
0,01466	NIDM3D_Busyness_N5
0,01381	Fractal_M
0,01374	RLM_Long_Run_Low_Gray-level_Emphasis_average

Tabela 25: Características selecionadas no cenário de maior performance da análise multivariada para a associação com metástase linfonodal.

Relevância no intervalo [-1,1]	Característica
0,07278	Clinical_Diameter
0,06441	Fractal_M
0,05952	Clinical_Age
0,05724	$COM3D_IMC2_d5_\theta135_\phi45$
0,05600	$COM3D_IMC2_d1_\theta 45_\phi 45$
0,05386	Shape_Convex
0,05180	$COM3D_IMC2_d5_\theta135_\phi135$
0,05042	Shape_Convex_Hull_Volume3D
0,04975	$COM2D_IMC2_d2_\theta90$
0,04967	$COM3D_IMC2_d1_\theta 45_\phi 90$
0,04917	Gabor_Mean6_M

Relevância no intervalo [-1,1]	Característica
0,46667	Clinical_Upper_right_lobe
0,31090	Clinical_Gender
$0,\!29551$	Clinical_Lower_right_lobe
0,28269	Clinical_History_of_other_primary_tumor
0,20577	Clinical_Upper_left_lobe
$0,\!19876$	Shape_Orientation
$0,\!12980$	COM3D_Cluster_Shade_d5_ θ 135_ ϕ 135
$0,\!10487$	Fractal_C
0,09766	Intensity_Local_Range_Maximum
0,09655	Fourier_Mean_M
0,09402	Clinical_Age
0,09285	$COM3D_Correlation_d2_\theta90_\phi45$
0,08994	Shape_Voxel_Size
0,08960	Intensity_Local_Std_Std
0,08932	Intensity_Local_Std_Maximum
0,08799	$COM3D_IMC2_d3_\theta 45_\phi 90$
0,08660	$COM3D_IMC2_d3_\theta 90_\phi 135$
0,08659	Tamura_D14_C
$0,\!08524$	Tamura_D14_M
$0,\!08496$	COM3D_Correlation_d3_ θ 90_ ϕ 45
0,08078	Tamura_Contrast_M
0,08077	$COM3D_IMC2_d5_\theta 90_\phi 135$
0,08008	COM3D_Cluster_Shade_d4_ θ 135_ ϕ 135
0,07999	$COM3D_IMC2_d5_\theta 135_\phi 90$
0,07986	$COM3D_IMC2_d4_\theta135_\phi45$
0,07905	Fourier_Max_M
0,07814	$COM3D_Correlation_d4_\theta90_\phi45$
0,07641	Fourier_Kurtosis_C
0,07625	$COM3D_Cluster_Shade_d5_\theta90_\phi135$
0,07594	Shape_Convex
0,07584	Shape_Max_3D_Diameter
0,07548	Shape_Roundness
0,07510	COM3D_Correlation_d1_ θ 0_ ϕ 135
0,07501	$Gabor_Std3_M$
0,07455	COM3D_Correlation_d5_ θ 45_ ϕ 45
0,07356	$COM3D_Correlation_d2_\theta0_\phi90$
0,07283	Tamura_D17_C
0,07276	COM3D_Correlation_d5_ θ 90_ ϕ 90
0,07242	$COM3D_Correlation_d4_\theta90_\phi90$
0,07210	Shape_Surface_Area_Density
0,07184	$COM3D_IMC2_d5_\theta 135_\phi 45$
0,07177	$COM2D_Correlation_d3_\theta 45$
0,07173	Tamura_D13_C
0,07168	$COM2D_IMC2_d2_\theta 135$
0,07138	Clinical_Diameter

Tabela 26: Características selecionadas no cenário de maior performance da análise multivariada para a associação com histopatologia.

Figura 57: Mapa de calor retratanto a performance pela AUC da análise multivariada na associação com metástase distante. As colunas representam o número de características utilizadas. As linhas representam o número de neurônios da camada oculta da RBF utilizados. O mapa é apresentado até o algoritmo obter a maior performance com 13 características e 15 neurônios.

0.46	0.48	0.61	0.60	0.61	0.62	0.60	0.61	0.57	0.60	0.61	0.56
0.58	0.57	0.64	0.76	0.64	0.55	0.64	0.62	0.53	0.53	0.63	0.58
0.71	0.62	0.64	0.64	0.53	0.57	0.53	0.52	0.45	0.62	0.69	0.69
0.53	0.61	0.78	0.58	0.57	0.62		0.54	0.61		0.64	0.68
0.72	0.58	0.66	0.59	0.52	0.51	0.55	0.62	0.68	0.39	0.64	0.60
0.64	0.57	0.62	0.45	0.64	0.55	0.52	0.66	0.62	0.56	0.64	0.59
0.67	0.63	0.53	0.52	0.62	0.54	0.61	0.64	0.62	0.64	0.58	0.62
0.79	0.75	0.63	0.56	0.58	0.57	0.64	0.51	0.46	0.60	0.62	0.59
0.71	0.67	0.59	0.59	0.54	0.57	0.65	0.54	0.56	0.70	0.63	0.65
0.62	0.58	0.62	0.75	0.68	0.69	0.72	0.53	0.61	0.73	0.80	0.78
0.66	0.59	0.60	0.74	0.63	0.71	0.62	0.49	0.49	0.63	0.71	0.75
0.71	0.72	0.65	0.80	0.70	0.58	0.67	0.63	0.57	0.69	0.70	0.71
0.70	0.71	0.61	0.76	0.65	0.58	0.70	0.67	0.60	0.71	0.84	0.85
0.74	0.74	0.69	0.82	0.68	0.69	0.75	0.73	0.69	0.64	0.84	0.91
0.67	0.71	0.68	0.75	0.77	0.77	0.80	0.80	0.74	0.74	0.85	0.92

Figura 58: Mapa de calor retratanto a performance pela AUC da análise multivariada na associação com metástase linfonodal. As colunas representam o número de características utilizadas. As linhas representam o número de neurônios da camada oculta da RBF utilizados. O mapa é apresentado até o algoritmo obter a maior performance com 11 características e 16 neurônios.

0.78	0.75	0.68	0.72	0.73	0.74	0.75	0.75	0.75	0.75	0.76
0.78	0.73	0.64	0.63	0.70	0.74	0.75	0.73	0.54	0.75	0.62
0.73	0.70	0.70	0.60		0.60	0.57	0.73	0.65	0.70	0.66
0.61	0.70	0.63	0.56	0.60	0.70	0.65	0.60	0.63	0.57	0.56
0.75	0.63	0.64		0.61	0.72	0.67	0.76	0.68	0.78	0.65
0.60	0.72	0.59	0.60	0.54		0.71	0.68	0.58	0.72	0.68
0.62	0.68	0.66	0.60	0.56	0.65	0.60	0.70		0.66	0.66
0.66	0.72		0.62	0.62	0.67	0.66	0.39		0.66	0.67
0.66	0.74	0.68		0.56	0.56	0.55	0.64		0.59	0.69
0.77	0.68		0.64	0.58		0.56	0.58	0.66	0.65	0.60
0.76	0.70	0.62	0.67	0.56	0.57	0.58	0.60	0.60	0.67	0.63
0.74	0.65	0.68		0.57			0.62			0.68
0.77	0.62	0.66	0.56		0.55	0.27	0.61	0.60	0.55	0.72
0.74	0.73		0.64	0.62	0.58		0.63	0.68	0.69	0.76
0.73	0.68	0.65	0.64	0.63	0.58	0.37	0.63	0.72	0.69	0.77
0.72	0.72	0.70	0.59	0.70	0.59	0.69	0.79	0.78	0.78	0.84

Figura 59: Mapa de calor retratanto a performance pela AUC da análise multivariada na associação com histopatologia. As colunas representam o número de características utilizadas. As linhas representam o número de neurônios da camada oculta da RBF utilizados. O mapa é apresentado de 26 características e 1 neurônio até o algoritmo obter a maior performance com 45 características e 20 neurônios.

0.63	0.60	0.62	0.71	0.67	0.69	0.79	0.32		0.66	0.65	0.71	0.70	0.69	0.70	0.67	0.71	0.78	0.72	0.80
0.62	0.61	0.22	0.20	0.21	0.19	0.20	0.35		0.84	0.37	0.82	0.34	0.82	0.81	0.79	0.80	0.80	0.80	0.81
0.23	0.20	0.24						0.63	0.30	0.64	0.36	0.70	0.32	0.23	0.69	0.29	0.32	0.74	0.74
0.36		0.56	0.34	0.79	0.69	0.38	0.38	0.37	0.37	0.35	0.35	0.34	0.45	0.47	0.45	0.48		0.61	0.62
0.44		0.34	0.35	0.40	0.37	0.36	0.40										0.47	0.49	0.41
0.43		0.45		0.44	0.42	0.50	0.54	0.41	0.46							0.39	0.45		0.37
0.41	0.49	0.49				0.44		0.30		0.46		0.51	0.52	0.51	0.49	0.62			0.34
0.33					0.39					0.48	0.39	0.49			0.51	0.61			0.37
				0.45	0.42					0.44	0.40	0.52	0.61	0.57	0.57	0.59	0.47	0.47	0.58
				0.45	0.50					0.57	0.59	0.46	0.59	0.58	0.57	0.58	0.47	0.47	0.58
		0.36	0.42							0.49		0.41	0.41	0.43					0.33
		0.42	0.46																0.33
0.40	0.38	0.39	0.44	0.45	0.44	0.45	0.45	0.49	0.33	0.42	0.43	0.42	0.42	0.40				0.40	0.37
0.39		0.39	0.45	0.44				0.42	0.39	0.43	0.42	0.44	0.45	0.44	0.44	0.45	0.43	0.43	0.43
0.42		0.39	0.45	0.44	0.43	0.43	0.50	0.50	0.47	0.42		0.42	0.42	0.42	0.42	0.74	0.76	0.82	0.82
0.36			0.45	0.44	0.43		0.49	0.48	0.49	0.42		0.42	0.42	0.42	0.42	0.74	0.76	0.82	0.82
0.36	0.38	0.36	0.45	0.49	0.67	0.45	0.48	0.48	0.49		0.42	0.42	0.71	0.71	0.74	0.73	0.74	0.82	0.82
0.40		0.40	0.47	0.45	0.47	0.45	0.48	0.47	0.49		0.45	0.46	0.74	0.73	0.76	0.76	0.73	0.78	0.79
0.44	0.45	0.43	0.48	0.48	0.45	0.43	0.47	0.47	0.50	0.43	0.48	0.48	0.75	0.75	0.80	0.81	0.82	0.87	0.88

123

5.3.3 Terceira Etapa

Nesta terceira etapa de experimentos do capítulo, assim como na segunda, também foi utilizado o Conjunto3 de pacientes para os experimentos. Contudo, ao contrário do que foi realizado nas duas primeiras fases, foi aplicada uma abordagem baseada em aprendizado profundo para a associação radiômica, o que excluiu a necessidade de utilização de extratores de atributos (Seção 2.6.1).

Janelas dos tumores tendo como base as marcações de borda contidas no arquivo DICOM-RT foram recortadas para servir como entrada para uma rede neural convolucional (Figuras 60(a–b)). Para aumentar a robustez dos experimentos, foram extraídas imagens com 1 canal em escala de níveis de cinza e com 3 canais em RGB (do Inglês *red-green-blue*) (Figuras 60(c–d)).

Figura 60: Recorte de janela do tumor usada como entrada para a rede neural convolucional:
(a) tumor na imagem de TC com a marcação de borda, (b) região a ser extraída pela janela baseada na marcação de borda, (c) e (d) resultados do recorte de janela em escala de níveis de cinza e em RGB, respectivamente.



Fonte: Autor.

Uma CNN de quatro camadas foi utilizada para realizar o aprendizado profundo nas imagens de TC das janelas dos tumores. Destas quatro camadas, três são convolucionais com função de unidade de retificação linear (do Inglês *rectified linear unit* - ReLU), seguidas de um *pooling* máximo (para a primeira camada convolucional) ou *pooling* médio (para as duas últimas), e uma camada final totalmente conectada de classificação binária (Seção 2.6.1) (STELMASZCZYK et al., 2017). As dimensões de entrada da rede convolucional foram $1 \times 32 \times 32$ e $3 \times 32 \times 32$ quando imagens em escala de níveis de cinza e em RGB foram utilizadas, respectivamente.

Uma validação cruzada com 2 subconjuntos foi empregada nestes experimentos. O subconjunto de treinamento foi submetido a uma etapa de aumento de amostras, explicada a seguir, e ainda subdividido em treinamento (90% das imagens) e validação (10% das imagens). O subconjunto de testes da validação cruzada utilizou as janelas originais dos tumores. Devido a grande variedade no tamanho dos tumores, a separação dos subconjuntos de treinamento e testes levou em consideração o diâmetro das lesões para não ocorrer o

problema de um subconjunto possuir apenas tumores grandes e o outro possuir apenas tumores pequenos, o que poderia diminuir o desempenho da classificação. O *framework* Caffe v1.0 foi utilizado no treinamento, validação e teste da CNN (JIA et al., 2014).

Devido à necessidade de grande amostragem para o treinamento da rede convolucional, as imagens das janelas dos tumores foram submetidas a uma etapa de préprocessamento com algoritmos de recortes de *patches* e de rotação em 90°, 180° e 270°. Duas estratégias foram utilizadas para a extração de *patches* de tamanho 32×32 : uma com aplicação de deslocamento de 1 *pixel* nos eixos x e y, e uma com maximização de não-sobreposição de *patches* (Figura 61).

Figura 61: Estratégias de extração de *patches* das imagens: (a) deslocamento de 1 *pixel* nos eixos $x \in y$, (b) maximização de não-sobreposição de *patches* e (c) caso perfeito para a estratégia descrita em (b).





Como o treinamento da CNN foi realizado em duas dimensões, o resultado da classificação foi dado por imagem. Logo, diferentes abordagens foram empregadas para avaliar volumetricamente os tumores tendo como referência o resultado da classificação por corte. As abordagens de avaliação volumétrica baseadas na classe ou na probabilidade por classe obtidas da camada totalmente conectada da rede convolucional foram:

- a) classe obtida pelo corte central;
- b) classe majoritária entre todos os cortes (caso ocorra empate, considerar positivo);
- c) classe majoritária entre todos os cortes (caso ocorra empate, considerar negativo);
- d) classe majoritária entre os 3 cortes centrais;
- e) classe majoritária entre os 5 cortes centrais;
- f) classe majoritária entre os 7 cortes centrais;

g) probabilidade máxima por classe entre todos os cortes;

- h) probabilidade máxima por classe entre os 3 cortes centrais;
- i) probabilidade máxima por classe entre os 5 cortes centrais;
- j) probabilidade máxima por classe entre os 7 cortes centrais;
- k) média da probabilidade por classe entre todos os cortes;
- 1) média da probabilidade por classe entre os 3 cortes centrais;
- m) média da probabilidade por classe entre os 5 cortes centrais;
- n) média da probabilidade por classe entre os 7 cortes centrais;
- o) produto da probabilidade por classe entre todos os cortes;
- p) produto da probabilidade por classe entre os 3 cortes centrais;
- q) produto da probabilidade por classe entre os 5 cortes centrais;
- r) produto da probabilidade por classe entre os 7 cortes centrais.

Os algoritmos de recorte de janela, recorte de *patches* e rotação de imagens foram desenvolvidos na linguagem de programação Java v1.8 com a API da ferramenta ImageJ v1.50i. O algoritmo de aprendizado profundo foi avaliado pelas métricas de acurácia, sensibilidade e especificidade.

A acurácia média para o subconjunto de validação de imagens com 1 canal, utilizando a extração de *patches* com deslocamento de *pixel*, foi de 0,989 \pm 0,002. Utilizando imagens com 3 canais e o mesmo método de extração de *patches*, a performance obtida foi de 0,975 \pm 0,016. A acurária média para o subconjunto de validação de imagens com 1 canal, utilizando a extração de *patches* com maximização de não-sobreposição, foi de 0,936 \pm 0,010. Utilizando imagens com 3 canais e o mesmo método de extração de *patches*, a performance obtida foi de 0,938 \pm 0,009.

Em contrapartida, o desempenho obtido para o subconjunto de testes foi exatamente o mesmo para os métodos de extração de *patches*. A maior performance obtida com as imagens em escala de níveis de cinza foi 0,624 de acurácia, 0,471 de sensibilidade e 0,662 de especificidade. A maior performance obtida com as imagens em RGB foi 0,753 de acurácia, 0,647 de sensibilidade e 0,779 de especificidade. A Tabela 27 apresenta todos os resultados das abordagens aplicadas para a avaliação volumétrica.

Estes resultados foram obtidos com treinamento após 100 mil iterações da rede convolucional. Com o objetivo de aumentar a robustez dos experimentos, foi realizado o aumento no número de iterações da rede para 200 mil, e com isso, foi obtido um aumento

Abordagem de	Image	ens em escala de	e cinza	Imagens em BGB						
avaliação										
volumétrica	Acurácia	Sensibilidade	Especific.	Acurácia	Sensibilidade	Especific.				
a)	$0,\!576$	0,294	$0,\!647$	$0,\!682$	$0,\!471$	0,735				
b)	0,518	$0,\!176$	$0,\!603$	0,753	$0,\!647$	0,779				
c)	0,529	$0,\!176$	$0,\!618$	0,729	0,529	0,779				
d)	$0,\!600$	$0,\!471$	$0,\!632$	$0,\!694$	$0,\!471$	0,750				
e)	0,588	$0,\!412$	$0,\!632$	0,706	0,529	0,750				
f)	0,588	0,412	$0,\!632$	0,718	0,588	0,750				
g)	0,529	0,235	$0,\!603$	$0,\!682$	0,412	0,750				
h)	$0,\!612$	$0,\!471$	$0,\!647$	$0,\!694$	0,529	0,735				
i)	$0,\!612$	$0,\!471$	$0,\!647$	0,706	0,529	0,750				
j)	$0,\!600$	$0,\!471$	$0,\!632$	0,706	$0,\!471$	0,765				
k)	$0,\!553$	$0,\!176$	$0,\!647$	0,718	$0,\!471$	0,779				
l)	$0,\!624$	$0,\!471$	$0,\!662$	$0,\!682$	$0,\!471$	0,735				
m)	$0,\!576$	0,412	$0,\!618$	$0,\!694$	0,529	0,735				
n)	$0,\!600$	0,412	$0,\!647$	0,706	$0,\!588$	0,735				
o)	0,565	$0,\!176$	$0,\!662$	0,718	0,412	0,794				
p)	$0,\!612$	0,412	$0,\!662$	0,706	0,529	0,750				
$\mathbf{q})$	0,576	0,412	$0,\!618$	0,694	0,529	0,735				
r)	$0,\!588$	0,412	$0,\!632$	0,706	$0,\!588$	0,735				

Tabela 27: Avaliação volumétrica dos tumores baseada na classificação por corte. Em negrito estão os maiores valores por métrica e por tipo de imagem utilizada.

Fonte: Autor.

na sensibilidade da CNN para 0,882. Contudo, a especificidade caiu para 0,397 e, como existem mais casos negativos para metástase distante, a acurácia foi afetada com um decréscimo para 0,494. A Tabela 28 apresenta os resultados na comparação entre 100 e 200 mil iterações no treinamento da rede convolucional. Ainda foi realizado um experimento com o balanceamento de amostras, porém a performance permaneceu inalterada.

Tabela 28: Comparação entre diferentes números de iterações no treinamento da rede neural convolucional. Em negrito estão os maiores valores por métrica e por número de iterações.

Abordagem de avaliação		100 mil iteraçõe	es	200 mil iterações					
volumétrica	Acurácia	Sensibilidade	Especific.	Acurácia	Sensibilidade	Especific.			
a)	$0,\!682$	0,471	0,735	0,424	0,765	0,338			
b)	0,753	$0,\!647$	0,779	$0,\!447$	0,824	0,353			
c)	0,729	0,529	0,779	$0,\!459$	0,824	0,368			
d)	$0,\!694$	$0,\!471$	0,750	$0,\!435$	0,824	0,338			
e)	0,706	0,529	0,750	$0,\!447$	0,882	0,338			
f)	0,718	$0,\!588$	0,750	$0,\!447$	0,882	0,338			
g)	$0,\!682$	0,412	0,750	$0,\!494$	0,882	0,397			
h)	$0,\!694$	0,529	0,735	$0,\!435$	0,824	0,338			
i)	0,706	0,529	0,750	0,412	0,882	$0,\!294$			
j)	0,706	$0,\!471$	0,765	$0,\!388$	0,882	0,265			
k)	0,718	$0,\!471$	0,779	$0,\!482$	0,882	0,382			
1)	$0,\!682$	$0,\!471$	0,735	$0,\!435$	0,824	0,338			
m)	$0,\!694$	0,529	0,735	$0,\!459$	0,882	0,353			
n)	0,706	$0,\!588$	0,735	$0,\!424$	0,882	0,309			
o)	0,718	0,412	$0,\!794$	$0,\!482$	0,882	0,382			
p)	0,706	0,529	0,750	$0,\!424$	0,824	0,324			
q)	$0,\!694$	0,529	0,735	$0,\!435$	0,882	0,324			
r)	0,706	0,588	0,735	0,388	0,882	0,265			

5.3.4 Quarta Etapa

Na quarta fase experimental do capítulo, todos os 101 pacientes identificados para a análise radiômica, descritos nas Tabelas 13 e 14, foram utilizados. Além disso, os três extratores de atributos investigados anteriormente na segunda etapa foram utilizados para a caracterização quantitativa das imagens de TC (Figura 45) (ZHANG et al., 2015; LUX; MARQUES, 2013; SCHNEIDER; RASBAND; ELICEIRI, 2012). Esta última etapa da pesquisa visou avaliar o prognóstico dos pacientes com câncer de pulmão por meio de uma análise de sobrevida global de todos os 2465 atributos extraídos.

A avaliação prognóstica foi baseada na análise de Kaplan-Meier com o cálculo das probabilidades de tempo para o evento de óbito por qualquer natureza (Seção 2.7). Pacientes vivos ou com perda de acompanhamento foram censurados para os cálculos das probabilidades de sobrevida global. Os grupos de risco foram divididos de acordo com a mediana dos valores dos atributos. O Teste *log-rank* avaliou a diferença estatística entre as curvas Kaplan-Meier para a identificação dos atributos com maior potencial prognóstico. Foi utilizado o pacote *survival* da linguagem de programação científica R v3.3.3 para a realização da análise de tempo-para-evento.

Vinte e oito pacientes foram à óbito (dezesseis por câncer de pulmão) durante o estudo, e setenta e três foram censurados. O tempo médio de sobrevida global foi de 1819,4 dias (IC: 1481,2–2157,5). Apenas um atributo apresentou diferença estatisticamente significativa nas curvas Kaplan-Meier (valor-p < 0,05): média de Fourier. O grupo de pacientes de alto risco foi identificado pelos valores altos da média de Fourier (maior que a mediana 109,0613458807), apresentando 18 eventos, 33 censuras, tempo médio de sobrevida de 1465,4 dias (IC: 985,2–1945,6). O grupo de pacientes de baixo risco foi identificado pelos valores baixos do atributo (menor ou igual a mediana 109,0613458807), apresentando 10 eventos, 40 censuras, tempo médio de sobrevida de 2164,8 dias (IC: 1745,4–2584,1) e razão de risco de 0,47 (IC: 0,22–0,99). A Figura 62 apresenta as curvas Kaplan-Meier do atributo média de Fourier.

A Figura 63 apresenta exemplos de casos dos grupos de risco estratificados pela média de Fourier e a quantificação da heterogeneidade intratumoral pela distribuição tridimensional dos níveis de cinza e pelo cálculo da energia local da imagem com aplicação de janela de 5×5 *pixels*. É possível notar nas figuras uma maior heterogeneidade intratumoral no caso de alto risco pela maior quantidade de picos no gráfico tridimensional da distribuição dos níveis de cinza e pela presença de regiões menos homogêneas no mapa da energia local em comparação com o caso de baixo risco.

Figura 62: Curvas Kaplan-Meier do atributo média de Fourier associado com sobrevida global. Cada traço nas curvas representa um paciente censurado.



Fonte: Autor.

5.4 Discussão e Considerações Específicas

A radiômica tem crescido promissoramente como uma abordagem de desenvolvimento de biomarcadores de imagens e modelos preditivos, aumentando a precisão diagnóstica, avaliação prognóstica e auxiliando a decisão terapêutica para a medicina personalizada (GIARDINO et al., 2017; GILLIES; KINAHAN; HRICAK, 2016; WU et al., 2016a). É importante ressaltar a inexistência de relação causa-efeito na radiômica. O foco da área é descobrir variáveis quantitativas que sejam capazes de reconhecer padrões nas imagens. Neste capítulo, foram investigados atributos de imagens de TC na associação radiômica com padrões de metástase distante, metástase linfonodal, histopatologia e sobrevida global de pacientes com câncer de pulmão. A identificação precoce do desenvolvimento de metástases, a subclassificação patológica do tumor e a estratificação de risco do paciente poderiam permitir aos especialistas a adoção rápida do tratamento apropriado (e.g. quimioterapia para casos com metástase distante) dirigido a um tipo histológico específico (e.g. inibidores de tirosina quinase do EGFR para ADC avançado) para o paciente com um determinado prognóstico (e.g. caso de alto risco e baixa probabilidade de sobrevida) (SHROFF et al., 2017; COROLLER et al., 2015; MOK et al., 2009).

Figura 63: Quantificação da heterogeneidade intratumoral em casos estratificados pelo atributo média de Fourier: (a) imagem original segmentada, (b) distribuição tridimensional dos níveis de cinza, (c) mapa refletindo a energia local do tumor.



Fonte: Autor.

As análises radiômicas identificaram diversos atributos com potencial diagnóstico e prognóstico. Reconhece-se que um radiologista poderia obter uma performance maior na detecção de metástase linfonodal do que a obtida neste trabalho com a leitura de um exame de TC de tórax. Porém a performance do especialista poderia potencialmente ser melhorada com a inclusão de uma ferramenta de radiômica na rotina clínica. Além disso, a avaliação de metástase distante é tradicionalmente realizada em tumores secundários espalhados em locais distantes que precisariam de outros exames de imagem, como TC ou RM de cabeça, cintilografia de osso, ou ainda ultrassonografia ou TC abdominal (TAMURA et al., 2015). Existem poucas evidências da presenca de metástases baseada apenas no tumor primário (ZHOU et al., 2018; COROLLER et al., 2015). Logo, uma abordagem que proporcione mais informações sobre a existência de metástase distante baseada na avaliação do tumor primário em um exame de TC de tórax realizado no primeiro momento poderia reduzir consideravelmente o tempo para a adoção de um exame de imagem complementar, podendo resultar em uma decisão mais rápida para a terapia adequada ao paciente. Ademais, a radiômica da histopatologia poderia fornecer mais informações clínicas importantes que potencialmente aumentaria a confiança do radiologista na predição do resultado de uma biópsia ou ressecção cirúrgica. Por fim, a avaliação prognóstica do paciente pela radiômica poderia proporcionar um avanço aos modelos clínicos atuais baseados em, por exemplo, informações radiológicas subjetivas e

exames genômicos invasivos de alto custo (AERTS, 2016; VELAZQUEZ, 2014), por ser objetiva, analítica não-interpretativa lógica, não-invasiva e de baixo custo.

Este trabalho realizou novas descobertas para a radiômica do câncer de pulmão em comparação com o que foi encontrado na literatura. Por exemplo, os atributos de energia de *wavelet*, medidas de informação de correlação da COM e ocupação da NIDM nunca foram associados com metástases até o momento da escrita desta tese. A direcionalidade de Tamura e a estimativa da dimensão fractal estão associadas com a histopatologia e ambas as metástases, e isto também nunca foi relatado na literatura. O atributo de média de Fourier, associado com a sobrevida global dos pacientes, nunca foi investigado antes. Além disso, nenhum trabalho foi encontrado analisando padrões de metástase distante no tumor primário com uma abordagem de aprendizado profundo. Por fim, esta tese também confirmou alguns achados da comunidade científica, como o atributo de volume de forma (COROLLER et al., 2015) e o gênero do paciente (ZHOU et al., 2018; COROLLER et al., 2015) associados com metástase distante e o atributo de probabilidade máxima da COM (ZHU et al., 2018), o diâmetro clínico e o gênero do paciente (DIGUMARTHY et al., 2019) associados com histopatologia.

Embora o aprendizado de máquina tenha aumentado a performance dos atributos na maioria dos casos, estas características quantitativas poderiam ser facilmente incluídas em uma ferramenta clínica com o seu cálculo automático após a segmentação semiautomática das imagens. Por exemplo, o atributo de amplitude de primeira ordem, que está associado com metástase distante, poderia ser implementado simplesmente com a subtração da maior pela menor intensidade de nível de cinza ou frequência no histograma da imagem. Atributos são mais interpretáveis do que modelos de aprendizado de máquina por serem projetados para propósitos específicos, e.g. atributos de *wavelet* analisam a textura no domínio da frequência da imagem, e assim, podem descrever informações não visíveis primariamente, como os padrões metastáticos (quanto maior o valor do atributo de *wavelet*, maior a chance de presença de metástases); atributos de forma, como área da superfície, massa e extensão média, também estão associados com metástases e descrevem o tamanho das regiões/volumes de interesse (quanto maior o valor do atributo, maior é o tumor e a chance de possuir estágio avançado apresentando metástases); e o atributo de probabilidade máxima da COM está relacionado com o par de níveis de cinza mais frequente da imagem e é capaz de distinguir os subtipos de NSCLC (quanto maior o valor do atributo, maior a chance do tumor ser SCC), indicando padrões mais homogêneos na imagem de TC para SCC e mais heterogêneos para ADC. Além disso, técnicas de aprendizado de máquina podem ser difíceis de interpretar por especialistas clínicos, sendo muitas vezes conhecidas como "caixas pretas", especialmente as RNAs, o que faz a aceitação delas na comunidade clínica desafiadora (LITJENS et al., 2017; AERTS, 2016). Desta forma, recomenda-se o uso dos atributos tanto isolados quanto combinados com algoritmos de aprendizado de máquina como biomarcadores de imagens.

Os atributos foram extraídos de todos os tradicionais níveis conhecidos da imagem: intensidade (primeira ordem), textura (segunda e mais alta ordens) e forma, descrevendo as distribuições individuais dos *pixels*, os relacionamentos espaciais dos *voxels* e a composição geométrica das regiões. Entretanto, atributos de textura local e rugosidade da forma, por exemplo, também poderiam ser calculados para investigar outras propriedades texturais e marginais da imagem, como heterogeneidade local, irregularidade de forma e transição de tecidos (GUO et al., 2016; DHARA et al., 2016). Uma vez realizada esta investigação ainda mais abrangente de caracterização quantitativa, será importante uma avaliação específica do ajuste fino feito nos atributos para melhorar a reprodutibilidade de pesquisa em radiômica. Com esta avaliação, será possível mapear os relacionamentos entre categorias e parâmetros dos atributos com os padrões das imagens, potencialmente melhorando a associação radiômica e o entendimento da natureza dos desfechos clínicos.

Vários atributos selecionados para as análises multivariadas não apresentaram previamente significância estatística, e diversos atributos estatisticamente significativos não foram selecionados para o aprendizado de máquina, o que era esperado pois significância estatística não implica em relevância para aprendizagem de máquina (Capítulo 3). Logo, foi necessário um método de seleção de características robusto o suficiente para diminuir a dimensionalidade vetorial e ainda deixá-lo capaz de caracterizar as imagens representativamente para a associação radiômica. Diferentemente dos capítulos anteriores, neste foi utilizado, quando necessário, o algoritmo ReliefF (WU et al., 2016b; PARMAR et al., 2015a; HAWKINS et al., 2014), mas outros métodos poderiam selecionar mais efetivamente as características, e.g. filtros baseados em informação mútua (PENG; LONG; DING, 2005). Uma outra possibilidade é a conjugação de diferentes métodos para a diminuição, primeiro, da irrelevância e, depois, da redundância das características utilizando, por exemplo, um filtro para reduzir a dimensionalidade do vetor e posteriormente um *wrapper* para identificar as informações mais relevantes do subconjunto filtrado de características para um determinado classificador (KOHAVI; JOHN, 1997).

A principal vantagem do algoritmo de aprendizado profundo utilizado neste trabalho foi a não dependência direta da segmentação dos VOIs e de extratores de atributos como pré-requisitos. Além disso, diversas estratégias de avaliação volumétrica baseada em classificação por corte foram desenvolvidas, aumentando a relevância do método proposto na pesquisa. A etapa de associação radiômica com CNN apresentou inicialmente algumas limitações, como o pequeno número de pacientes. Contudo, algoritmos de aumento de amostragem fizeram com que esta limitação fosse diminuída. A abordagem de extração de *patches* por maximização de não-sobreposição foi suficiente para realizar a análise com aprendizado profundo. Por outro lado, a extração de *patches* com aplicação de deslocamento de *pixel* aumentou ainda mais o número de casos para os experimentos, porém foi obtida performance equivalente à primeira abordagem, devido ao fato da CNN utilizada possuir poucas camadas e, portanto, os padrões não eram totalmente reconhecidos mesmo com mais imagens para o treinamento da rede. É importante salientar que a CNN pode ter sofrido queda de desempenho devido à utilização de imagens pré-processadas para o conjunto de validação. Logo, outras investigações são necessárias para aumentar a performance de associação utilizando aprendizado profundo. Melhorias no projeto de pré-processamento de imagens, modelagem da arquitetura da rede convolucional e extensão para três dimensões podem fazer com que a performance seja maior, aumentando a associação radiômica com os padrões clínicos.

A análise de sobrevida conseguiu identificar apenas um atributo associado com o prognóstico dos pacientes com câncer de pulmão. Isto pode ter ocorrido pelo fato de ter havido poucos eventos para a análise de sobrevida global, visto o acompanhamento curto dos pacientes. Além disso, existe uma grande heterogeneidade nos tumores da amostragem, cada um implicando em um prognóstico diferente para o paciente. Logo, a variabilidade existente nos estadios clínicos, tipos histológicos e tratamentos empregados, por exemplo, pode ter influenciado na identificação de outros atributos prognósticos.

É importante ressaltar que muitas das ferramentas computacionais empregadas nesta análise radiômica possuem código-fonte aberto para desenvolvimento, como o 3D-Slicer para segmentação de imagens (FEDOROV et al., 2012), Ibex para extração de atributos (ZHANG et al., 2015) e Weka para mineração de dados e aprendizado de máquina (FRANK; HALL; WITTEN, 2016). Isto é importante pela necessidade de criar *benchmarks* para os estudos em radiômica (LARUE et al., 2017; GILLIES; KINAHAN; HRICAK, 2016). Além disso, estas ferramentas citadas podem ser utilizadas por pesquisadores sem habilidades e experiência com programação de sistemas, pois possuem interface gráfica de usuário, aumentando ainda mais a relevância dos métodos para a comunidade científica.

Estas fases do trabalho apresentaram algumas limitações. Como já foi relatado, a pesquisa apresentou um conjunto de amostras relativamente pequeno. Como as análises foram realizadas apenas com imagens de TC com contraste para melhorar a segmentação e caracterização quantitativa, isto limitou a casuística. Contudo, a escolha por estas imagens poderia aumentar a performance associativa porque os tecidos estariam mais visualmente distinguíveis e os atributos mais representativos. Além disso, TC com contraste é o exame de referência da rotina clínica para a avaliação do câncer de pulmão e deveria ser utilizado como padrão de aquisição de imagens por outros grupos de pesquisa para aumentar a reprodutibilidade, já que existe uma falta de padronização metodológica para conduzir estudos em radiômica (LARUE et al., 2017). Existem bases públicas disponíveis na literatura, porém elas não foram utilizadas porque não possuem os mesmos dados clínicos que os deste trabalho, em alguns casos os dados são inconsistentes e/ou as imagens não foram adquiridas com o mesmo protocolo de TC. Devido ao aspecto retrospectivo, outros desfechos clínicos não puderam ser investigados, como a presença de mutações genômicas. O grau de diferenciação tumoral estava disponível para apenas um pequeno número de pacientes e, portanto, também não foi investigado.

Outras limitações foram que os atributos poderiam ter sido implementados de acordo com as diretrizes da IBSI (do Inglês *Image Biomarker Standardisation Initiative*), em que cerca de 20 diferentes grupos internacionais de pesquisa contribuíram para chegar em um consenso para implementação de extratores de atributos de imagens (VALLIÈRES et al., 2018; ZWANENBURG et al., 2018). As versões das ferramentas utilizadas no cálculo das características quantitativas ainda não tinham sido validadas pela IBSI durante a investigação e desenvolvimento dos extratores de atributos. Além disso, algumas variáveis clínicas eram subjetivas, o que necessitaria leitura interobservador e testes de confiabilidade intra e interobservador. Por fim, análises de estabilidade e reprodutibilidade dos atributos não foram realizadas, o que proporcionaria mais evidências de robustez para as variáveis quantitativas e modelos computacionais.

Este capítulo apresentou a radiômica de tumores pulmonares malignos em imagens de TC. Diversos atributos e modelos de inteligência artificial demonstraram potencial diagnóstico e prognóstico para serem utilizados como biomarcadores quantitativos de imagens para o câncer de pulmão. A abordagem radiômica descrita neste trabalho pode oferecer grandes avanços na predição de metástases, histopatologia em forma de biópsia virtual e sobrevida de pacientes, em nível macroscópico não-invasivamente, proporcionando informações adicionais para a rotina clínica e potencialmente melhorando o diagnóstico, prognóstico e decisão terapêutica.

6 CONCLUSÃO

6.1 Considerações Finais

A radiologia vem passando por importantes atualizações com a revolução tecnológica que acontece no mundo: primeiro com a digitalização dos ambientes radiológicos; depois com a evolução das técnicas de visão computacional e inteligência artificial no desenvolvimento de sistemas CAD e CBIR; e mais recentemente com o amadurecimento dos modelos computacionais de suporte a decisões clínicas por meio da radiômica. Este trabalho investigou diferentes métodos de CAD, CBIR e radiômica para servir como arcabouço para a análise de imagens médicas, auxiliar as tomadas de decisão clínica e contribuir para o avanço na área radiológica. O principal estudo de caso do arcabouço foi o câncer de pulmão em imagens de TC.

Um modelo semiautomático de reconhecimento de padrões de imagens médicas foi desenvolvido neste trabalho baseado em segmentação por crescimento de região 3D, extração de atributos calibrados, classificação e recuperação por conteúdo de imagens e avaliação prognóstica do paciente. Um segundo modelo semiautomático baseado em aprendizado profundo e CNN também foi proposto com o objetivo de diminuir os processos do arcabouço para a análise computadorizada das imagens. Os modelos possuem potencial para compor uma ferramenta que poderia ser integrada aos sistemas RIS na rotina clínica. Uma vez que a ferramenta fornecesse o resultado das avaliações diagnóstica e prognóstica de uma lesão, o paciente poderia ser colocado em uma lista de prioridades. Em situações que as saídas fossem, por exemplo, de um tumor maligno de pulmão com subtipo SCC, baixa probabilidade de sobrevida e presenças de metástases linfonodal e distante, a prioridade seria máxima devido os fatores de risco apresentados pelo caso. A integração com sistemas HIS poderia ainda permitir o encaminhamento mais rápido deste paciente com prioridade máxima para a realização dos exames complementares de rotina e tratamento.

A relevância desta tese pode ser estendida para um aspecto multidisciplinar. O trabalho foi capaz de construir modelos computacionais que poderiam trazer avanços aos sistemas RIS e HIS, mostrou o potencial de diferentes métodos tecnológicos na avaliação quantitativa de imagens médicas e ainda a possibilidade de exclusão de técnicas tradicionais por métodos modernos de inteligência artificial. Além disso, é enfatizado o potencial da CBIR paralela ao CAD e à radiômica para o suporte clínico com a busca por informações de casos similares aos especialistas. Este trabalho também mostrou a possibilidade e importância do ajuste fino nos diferentes processos de análise quantitativa de imagens, como de extração de atributos e aprendizado de máquina. Quando não há possibilidade de ajuste nos extratores de atributos na etapa pré-classificação, o ajuste fino no algoritmo de aprendizado de máquina pré-treinamento torna-se mandatório para a otimização de

performance do classificador.

Do ponto de vista radiológico, o arcabouço pode proporcionar diversos biomarcadores quantitativos de imagens para diferentes doenças, mais informações importantes para os especialistas, potencialmente auxiliando as tomadas de decisões da rotina clínicoradiológica. Por fim, esta tese ainda descobriu uma assinatura radiômica de avaliação diagnóstica e prognóstica do câncer de pulmão para a estratificação de risco dos pacientes.

6.2 Direções Futuras

Para a evolução do arcabouço e dos biomarcadores desenvolvidos, alguns trabalhos futuros são sugeridos abaixo:

- detecção automática de anormalidades nos exames de imagens médicas utilizando aprendizado profundo;
- segmentação de imagens com crescimento de região utilizando função de *convex hull* integrado com o algoritmo GrowCut;
- extração de atributos de textura local e irregularidade de forma;
- recuperação de imagens por conteúdo utilizando outros desfechos clínicos, como o tipo histopatológico por exemplo, e avaliação de relevância por especialista;
- recuperação de imagens por conteúdo utilizando aprendizado profundo e CNN como extratores de atributos;
- associação radiômica com outros tipos histopatológicos, além de ADC e SCC, e desfechos clínicos, como recorrência do tumor e genômica;
- associação radiômica utilizando aprendizado profundo com outros desfechos clínicos, além de metástase distante;
- avaliação da heterogeneidade intratumoral;
- síntese de imagens de TC a partir de exames de RM utilizando redes neurais adversárias generativas e, assim, reduzir a exposição à radiação para o paciente;
- desenvolvimento e validação clínica de uma ferramenta final para o especialista.

6.3 Contribuições Científicas

As principais contribuições diretas desta tese em forma de publicações científicas estão listadas abaixo, em ordem cronológica decrescente:

- Ferreira Junior JR, Koenigkam-Santos M, Cipriano FEG, Fabro AT, Azevedo-Marques PM. Radiomics-based features for pattern recognition of lung cancer histopathology and metastases. *Computer Methods and Programs in Biomedicine*, v. 159, p. 23–30, 2018;
- Ferreira Jr JR, Oliveira MC, Azevedo-Marques PM. Characterization of pulmonary nodules based on features of margin sharpness and texture. *Journal of Digital Imaging*, v. 31, n. 4, p. 451–463, 2018;
- Ferreira Junior JR, Koenigkam-Santos M, Tenorio APM, Faleiros MC, Cipriano FEG, Fabro AT, Nappi J, Yoshida H, Azevedo-Marques PM. Radiomics association of quantitative CT features with lung cancer patterns. *International Congress and Exhibition on Computer Assisted Radiology and Surgery*, p. S107–S109, 2018;
- Ferreira Jr JR, Azevedo-Marques PM, Oliveira MC. Selecting relevant 3D image features of margin sharpness and texture for lung nodule retrieval. *International Journal of Computer Assisted Radiology and Surgery*, v. 12, n. 3, p. 509–517, 2017;
- Ferreira Junior JR, Cipriano FEG, Fabro AT, Koenigkam-Santos M, Azevedo-Marques PM. Radiomics-based recognition of metastatic and histopathological patterns of lung cancer. *Thematic Conference on Computational Vision and Medical Image Processing*, p. 613–623, 2017;
- Koenigkam-Santos M, Ferreira-Junior JR, Cipriano FEG, Faccio A, Fabro AT, Azevedo-Marques PM. Radiômica do câncer de pulmão. *Jornada Paulista de Radiologia*, 2017;
- Ferreira Junior JR, Oliveira MC, Azevedo-Marques PM. Cloud-based NoSQL open database of pulmonary nodules for computer-aided lung cancer diagnosis and reproducible research. *Journal of Digital Imaging*, v. 29, n. 6, p. 716–729, 2016;
- Ferreira Junior JR, Oliveira MC, Azevedo-Marques PM. Integrating 3D image descriptors of margin sharpness and texture on a GPU-optimized similar pulmonary nodule retrieval engine. *The Journal of Supercomputing*, v. 73, n. 8, p. 3451–3467, 2016;
- Ferreira Jr J, Oliveira MC, Azevedo-Marques PM. Pulmonary nodule classification with 3D features of texture and margin sharpness. *International Congress and Exhibition on Computer-Assisted Radiology and Surgery*, p. S272, 2016;
- Ferreira Junior JR, Tinós R, Koenigkam-Santos M, Azevedo-Marques PM. Classificação histopatológica de tumores malignos pulmonares utilizando características radiômicas. *Congresso Brasileiro de Informática em Saúde*, p. 199–200, 2016.

Algumas contribuições parciais foram obtidas a partir da utilização de módulos do arcabouço, principalmente na associação radiômica de atributos de RM de articulações sacroilíacas com espondiloartrite e padrões inflamatórios de sacroiliíte, aumentando a relevância e generalização da proposta desta tese. Estas publicações científicas estão listadas abaixo em ordem cronológica decrescente:

- Matias Júnior I, Medeiros P, Freitas RL, Vicente-César H, Ferreira Junior JR, Machado HR, Menezes-Reis R. Effective parameters for gait analysis in experimental models for evaluating peripheral nerve injuries in rats. *Neurospine*, 2019, DOI:10.14245/ns.1836080.040;
- Tenorio APM, Faleiros MC, Raniery Ferreira Junior J, Dalto VF, Assad RL, Yoshida H, Nogueira-Barbosa MH, Azevedo-Marques PM. Radiomics association of MRI texture features with spondyloarthritis and sacroiliitis. *International Congress and Exhibition on Computer-Assisted Radiology and Surgery*, p. S258–S259, 2018;
- Faleiros MC, Raniery Ferreira Junior J, Tenorio APM, Dalto VF, Assad RL, Nogueira-Barbosa MH, Azevedo-Marques PM. Features selection analysis to quantify sacroiliitis in magnetic resonance imaging. *International Congress and Exhibition on Computer* Assisted Radiology and Surgery, p. S106–S107, 2018;
- Francisco V, Koenigkam-Santos M, Wada DT, Ferreira Junior JR, Fabro AT, Cipriano FEG, Quatrina SG, Azevedo-Marques PM. Computer-aided diagnosis of lung cancer in magnetic resonance imaging exams. *Congresso Brasileiro de Engenharia Biomédica*, 2018;
- Faleiros MC, Ferreira Junior JR, Jens EZ, Dalto VF, Nogueira-Barbosa MH, Azevedo-Marques PM. Pattern recognition of inflammatory sacroiliitis in magnetic resonance imaging. *Thematic Conference on Computational Vision and Medical Image Proces*sing, v. 27, p. 639–644, 2017;
- Faleiros MC, Ferreira Junior JR, Zavala EJR, Dalto VF, Nogueira-Barbosa MH, Azevedo-Marques PM. Reconhecimento computadorizado de padrões inflamatórios de sacroiliíte em imagens de ressonância magnética. Workshop de Informática Médica, 2017;
- Faleiros MC, Zavala EJR, Raniery Ferreira Junior J, Dalto VF, Assad RL, Louzada Junior P, Nogueira-Barbosa MH, Azevedo-Marques PM. Computer-aided classification of inflammatory sacroiliitis in magnetic resonance imaging. *International Congress and Exhibition on Computer-Assisted Radiology and Surgery*, p. S154–S155, 2017;

- Lucena DJF, Ferreira Junior JR, Machado AP, Oliveira MC. Automatic weighing attribute to retrieve similar lung cancer nodules. *BMC Medical Informatics and Decision Making*, v. 16, n. 79, p. 135–149, 2016;
- Felix A, Oliveira M, Machado A, Raniery J. Using 3D texture and margin sharpness features on classification of small pulmonary nodules. *Conference on Graphics*, *Patterns and Images*, p. 394–400, 2016;
- Faleiros MC, Ferreira Junior JR, Dalto VF, Nogueira-Barbosa MH, Azevedo-Marques PM. Avaliação computadorizada de sacroiliíte em imagens de ressonância magnética. Congresso Brasileiro de Informática em Saúde, p. 85–94, 2016.

REFERÊNCIAS

AERTS, H. The potential of radiomic-based phenotyping in precision medicine: a review. **JAMA Oncology**, v. 2, n. 12, p. 1636–1642, 2016.

AERTS, H. et al. Defining a radiomic response phenotype: a pilot study using targeted therapy in NSCLC. Scientific Reports, v. 6, n. 33860, 2016.

AERTS, H. et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. **Nature Communications**, v. 5, n. 4006, 2014.

AGGARWAL, P.; SARDANA, H.; VIG, R. Content based image retrieval approach in creating an effective feature index for lung nodule detection with the inclusion of expert knowledge and proven pathology. **Current Medical Imaging Reviews**, v. 10, n. 3, p. 178–204, 2014.

AGGARWAL, P.; VIG, R.; SARDANA, H. Lung cancer detection using fusion of medical knowledge and content based image retrieval for LIDC dataset. Journal of Medical Imaging and Health Informatics, v. 6, n. 2, p. 297–311, 2016.

AHA, D.; KIBLER, D.; ALBERT, M. Instance-based learning algorithms. Machine Learning, v. 6, n. 1, p. 37–66, 1991.

AKGÜL, C. B. et al. Content-based image retrieval in radiology: current status and future directions. Journal of Digital Imaging, v. 24, n. 2, p. 208–222, 2011.

ALILOU, M.; OROOJI, M.; MADABHUSHI, A. Intra-perinodular textural transition (Ipris): A 3D descriptor for nodule diagnosis on lung CT. In: **Proceedings of International Conference on Medical Image Computing and Computer-**Assisted Intervention. Quebec City, Canada: MICCAI, 2017. p. 647–655.

ALMEIDA, E.; RANGAYYAN, R. M.; AZEVEDO-MARQUES, P. M. Gaussian mixture modeling for statistical analysis of features of high-resolution CT images of diffuse pulmonary diseases. In: **Proceedings of the 2015 IEEE International Symposium on Medical Measurements and Applications**. Torino, Italy: IEEE, 2015. p. 1–5.

AMADASUN, M.; KING, R. Textural features corresponding to textural properties. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 19, n. 5, p. 1264–1274, 1989.

AMERICAN CANCER SOCIETY. **Cancer Facts & Figures 2016**. Atlanta, USA: ACS, 2016. Disponível em <www.bit.ly/2fk6DvY>. Acesso em: 23 Jan. 2019.

ARMATO III, S. et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. Medical Physics, v. 38, p. 915–931, 2011.

ARMATO III, S. et al. **Data From LIDC-IDRI. The Cancer Imaging Archive**. Rockville, USA: TCIA, 2018. Disponível em <doi.org/10.7937/K9/TCIA.2015. LO9QL9SX>. Acesso em: 23 Jan. 2019. AUSTIN, J. et al. Radiologic implications of the 2011 classification of adenocarcinoma of the lung. **Radiology**, v. 266, n. 1, p. 62–71, 2013.

AZEVEDO-MARQUES, P. M. et al. Medical Image Analysis and Informatics: Computer-Aided Diagnosis and Therapy. Boca Raton, USA: CRC Press, 2017.

AZEVEDO-MARQUES, P. M.; RANGAYYAN, R. M. Content-based Retrieval of Medical Images: Landmarking, Indexing, and Relevance Feedback. Williston, USA: Morgan & Claypool Publishers, 2013.

AZEVEDO-MARQUES, P. M. et al. Reducing the semantic gap in content-based image retrieval in mammography with relevance feedback and inclusion of expert knowledge. **International Journal of Computer Assisted Radiology and Surgery**, v. 3, n. 1-2, p. 123–130, 2008.

AZEVEDO-MARQUES, P. M. et al. Classification of vertebral compression fractures in magnetic resonance images using spectral and fractal analysis. In: **Proceedings of 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society**. Milan, Italy: IEEE, 2015. p. 723–726.

BANKER, K. MongoDB in Action. Shelter Island, USA: Manning Publications Co., 2011.

BANKMAN, I. H. Handbook of Medical Image Processing and Analysis. Burlington, USA: Academic Press, 2009.

BARTHOLMAI, B. et al. Pulmonary nodule characterization, including computer analysis and quantitative features. Journal of Thoracic Imaging, v. 30, n. 2, p. 139–156, 2015.

BEDO, M. V. N. et al. Endowing a content-based medical image retrieval system with perceptual similarity using ensemble strategy. **Journal of Digital Imaging**, v. 29, n. 1, p. 22–37, 2016.

BIANCONI, F.; FERNÁNDEZ, A. Evaluation of the effects of Gabor filter parameters on texture classification. **Pattern Recognition**, v. 40, n. 12, p. 3325–3335, 2007.

BRACEWELL, R. N. The Fourier Transform and Its Applications. New York, USA: McGraw-Hill, 1965.

BRANT, W.; HELMS, C. **Fundamentals of Diagnostic Radiology**. Philadelphia, USA: Lippincott Williams & Wilkins, 2012.

BREIMAN, L. Random forests. Machine Learning, v. 45, n. 1, p. 5–32, 2001.

BRIGHAM, E. O. **The Fast Fourier Transform**. Englewood Cliffs, USA: Prentice Hall, 1974.

BUGATTI, P. H. et al. PRoSPer: Perceptual similarity queries in medical CBIR systems through user profiles. Computers in Biology and Medicine, v. 45, p. 8–19, 2014.

BUSHBERG, J. et al. The Essential Physics of Medical Imaging. Philadelphia, USA: Lippincott Williams & Wilkins, 2012.

CAO, P. et al. A $\ell_{2,1}$ norm regularized multi-kernel learning for false positive reduction in lung nodule CAD. Computer Methods and Programs in Biomedicine, v. 140, p. 211–231, 2017.

CARVALHO FILHO, A. O. et al. Computer-aided diagnosis system for lung nodules based on computed tomography using shape analysis, a genetic algorithm, and SVM. Medical & Biological Engineering & Computing, v. 55, n. 8, p. 1129–1146, 2017.

CATALDO, S. et al. Subclass discriminant analysis of morphological and textural features for HEp-2 staining pattern classification. **Pattern Recognition**, v. 47, n. 7, p. 2389–2399, 2014.

CHARTRAND, G. et al. Deep Learning: A primer for radiologists. **Radiographics**, v. 37, n. 7, p. 2113–2131, 2017.

CHAWLA, N. et al. SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, v. 16, p. 321–357, 2002.

CHEEMA, P.; BURKES, R. Overall survival should be the primary endpoint in clinical trials for advanced non-small-cell lung cancer. **Current Oncology**, v. 20, n. 2, p. e150, 2013.

CHEN, L. et al. Tissue-specific compartmental analysis for dynamic contrast-enhanced MR imaging of complex tumors. **IEEE Transactions on Medical Imaging**, v. 30, n. 12, p. 2044–2058, 2011.

CHEN, X. et al. A radiomics signature in preoperative predicting degree of tumor differentiation in patients with non-small cell lung cancer. Academic Radiology, v. 25, n. 12, p. 1548–1555, 2018.

CHIERICI, C. Classificação de texturas com diferentes orientações baseada em descritores locais. Escola de Engenharia de São Carlos, Universidade de São Paulo, 2015.

CHOI, W. et al. Radiomics analysis of pulmonary nodules in low-dose CT for early detection of lung cancer. **Medical Physics**, v. 45, n. 4, p. 1537–1549, 2018.

CHOY, G. et al. Current applications and future impact of machine learning in radiology. Radiology, v. 288, n. 2, p. 318–328, 2018.

CIRUJEDA, P. et al. A 3-D Riesz-covariance texture model for prediction of nodule recurrence in lung CT. **IEEE Transactions on Medical Imaging**, v. 35, n. 12, p. 2620–2630, 2016.

COHEN, J. et al. CT-texture analysis of subsolid nodules for differentiating invasive from in-situ and minimally invasive lung adenocarcinoma subtypes. **Diagnostic and Interventional Imaging**, v. 99, n. 5, p. 291–299, 2018.

COHEN, J. G. et al. Software performance in segmenting ground-glass and solid components of subsolid nodules in pulmonary adenocarcinomas. **European Radiology**, v. 26, n. 12, p. 4465–4474, 2016.

CONSTANZO, J. et al. Radiomics in precision medicine for lung cancer. **Translational** Lung Cancer Research, v. 6, n. 6, p. 635, 2017.
COOPER, W. et al. What's new in non-small cell lung cancer for pathologists: the importance of accurate subtyping, EGFR mutations and ALK rearrangements. **Pathology**, v. 43, n. 2, p. 103–115, 2011.

COROLLER, T. et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. **Radiotherapy and Oncology**, v. 114, n. 3, p. 345–350, 2015.

COROLLER, T. P. et al. Radiomic-based pathological response prediction from primary tumors and lymph nodes in NSCLC. Journal of Thoracic Oncology, v. 12, n. 3, p. 467–476, 2017.

COROLLER, T. P. et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. **Radiotherapy and Oncology**, v. 119, n. 3, p. 480–486, 2016.

DAVNALL, F. et al. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? **Insights Into Imaging**, v. 3, n. 6, p. 573–589, 2012.

DEGNAN, A. J. et al. Perceptual and interpretive error in diagnostic radiology – causes and potential solutions. Academic Radiology, 2018. DOI:10.1016/j.acra.2018.11.006.

DEPEURSINGE, A. et al. Predicting adenocarcinoma recurrence using computational texture models of nodule components in lung CT. Medical Physics, v. 42, n. 4, p. 2054–2063, 2015.

DESERNO, T. M.; ANTANI, S.; LONG, R. Ontology of gaps in content-based image retrieval. Journal of Digital Imaging, v. 22, n. 2, p. 202–215, 2009.

DHARA, A. K. et al. Content-based image retrieval system for pulmonary nodules: Assisting radiologists in self-learning and diagnosis of lung cancer. Journal of Digital Imaging, v. 30, n. 1, p. 63–77, 2017.

DHARA, A. K. et al. A study of retrieval accuracy of pulmonary nodules based on external attachment. In: **Proceedings of Medical Imaging**. Orlando, USA: SPIE, 2017. v. 10134, p. 101343T.

DHARA, A. K. et al. Differential geometry-based techniques for characterization of boundary roughness of pulmonary nodules in CT images. International Journal of Computer Assisted Radiology and Surgery, v. 11, n. 3, p. 337–349, 2016.

DIGUMARTHY, S. et al. Can CT radiomic analysis in NSCLC predict histology and EGFR mutation status? **Medicine**, v. 98, n. 1, p. e13963, 2019.

DILGER, S. et al. Improved pulmonary nodule classification utilizing lung parenchyma texture features. In: **Proceedings of Medical Imaging**. Orlando, USA: SPIE, 2015. v. 9414, p. 94142T.

DIVINE, M. et al. A population-based Gaussian mixture model incorporating 18F-FDG PET and diffusion-weighted MRI quantifies tumor tissue classes. Journal of Nuclear Medicine, v. 57, n. 3, p. 473–479, 2016.

DOI, K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. **Computerized Medical Imaging and Graphics**, v. 31, n. 4-5, p. 198–211, 2007.

E, L. et al. Radiomics for classification of lung cancer histological subtypes based on nonenhanced computed tomography. Academic Radiology, 2018. DOI:10.1016/j.acra.2018.10.013.

ECHEGARAY, S. et al. Core samples for radiomics features that are insensitive to tumor segmentation: method and pilot study using CT images of hepatocellular carcinoma. **Journal of Medical Imaging**, v. 2, n. 4, p. 041011, 2015.

EGGER, J. et al. GBM volumetry using the 3D Slicer medical image computing platform. Scientific Reports, v. 3, p. 1364, 2013.

EMAMINEJAD, N. et al. Fusion of quantitative image and genomic biomarkers to improve prognosis assessment of early stage lung cancer patients. **IEEE Transactions on Biomedical Engineering**, v. 63, n. 5, p. 1034–1043, 2016.

ERICKSON, B. et al. Machine learning for medical imaging. **Radiographics**, v. 37, n. 2, p. 505–515, 2017.

ERTOSUN, M. G.; RUBIN, D. L. Automated grading of gliomas using deep learning in digital pathology images: A modular approach with ensemble of convolutional neural networks. In: **Proceedings of Annual Symposium of the American Medical Informatics Association**. San Francisco, USA: AMIA, 2015. p. 1899–1908.

FALEIROS, M. C. et al. Avaliação computadorizada de sacroiliíte em imagens de ressonância magnética. In: Anais do XV Congresso Brasileiro de Informática em Saúde. Goiânia, GO: SBIS, 2016. p. 85–94.

FALEIROS, M. C. et al. Pattern recognition of inflammatory sacroiliitis in magnetic resonance imaging. In: Proceedings of VI ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing. Porto, Portugal: Lecture Notes in Computational Vision and Biomechanics, 2017. v. 27, p. 639–644.

FAVE, X. et al. Can radiomics features be reproducibly measured from CBCT images for patients with non-small cell lung cancer? **Medical Physics**, v. 42, n. 12, p. 6784–6797, 2015.

FEDOROV, A. et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magnetic Resonance Imaging, v. 30, n. 9, p. 1323–1341, 2012.

FELIX, A. et al. Using 3D texture and margin sharpness features on classification of small pulmonary nodules. In: **Proceedings of 29th Conference on Graphics, Patterns and Images**. São José dos Campos, SP: IEEE, 2016. p. 394–400.

FERREIRA, J. C.; PATINO, C. M. O que é análise de sobrevida e quando devo utilizá-la? **Jornal Brasileiro de Pneumologia**, v. 42, n. 1, p. 77–77, 2016.

FERREIRA JR, J. R.; AZEVEDO-MARQUES, P. M.; OLIVEIRA, M. C. Selecting relevant 3D image features of margin sharpness and texture for lung nodule retrieval. **International Journal of Computer Assisted Radiology and Surgery**, v. 12, n. 3, p. 509–517, 2017.

FERREIRA JR, J. R.; OLIVEIRA, M. C.; AZEVEDO-MARQUES, P. M. Characterization of pulmonary nodules based on features of margin sharpness and texture. **Journal of Digital Imaging**, v. 31, n. 4, p. 451–463, 2018.

FERREIRA JUNIOR, J. R. et al. Radiomics-based features for pattern recognition of lung cancer histopathology and metastases. Computer Methods and Programs in Biomedicine, v. 159, p. 23–30, 2018.

FERREIRA JUNIOR, J. R.; OLIVEIRA, M. C. Evaluating margin sharpness analysis on similar pulmonary nodule retrieval. In: **Proceedings of 28th IEEE International Symposium on Computer-Based Medical Systems**. São Carlos, SP: IEEE, 2015. p. 60–65.

FERREIRA JUNIOR, J. R.; OLIVEIRA, M. C. Unidade de processamento gráfico no diagnóstico auxiliado por computador. **Journal of Health Informatics**, v. 9, n. 3, p. 103–108, 2017.

FERREIRA JUNIOR, J. R.; OLIVEIRA, M. C.; AZEVEDO-MARQUES, P. M. Cloud-based NoSQL open database of pulmonary nodules for computer-aided lung cancer diagnosis and reproducible research. **Journal of Digital Imaging**, v. 29, n. 6, p. 716–729, 2016.

FERREIRA JUNIOR, J. R.; OLIVEIRA, M. C.; AZEVEDO-MARQUES, P. M. Integrating 3D image descriptors of margin sharpness and texture on a GPU-optimized similar pulmonary nodule retrieval engine. **The Journal of Supercomputing**, v. 73, n. 8, p. 3451–3467, 2017.

FILONENKO, E.; SEERAM, E. Big data: The next era of informatics and data science in medical imaging: A literature review. Journal of Clinical and Experimental Radiology, v. 1, n. 1, p. 1000102, 2018.

FRANK, E.; HALL, M.; WITTEN, I. **The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques**. Burlington, USA: Morgan Kaufmann, 2016.

FRIED, D. Investigation of quantitative image features from pretreatment CT and FDG-PET scans in stage III NSCLC patients undergoing definitive radiation therapy. Tese (Doutorado) — Faculty of The University of Texas Health Science Center, 2015.

FRIGHETTO-PEREIRA, L. et al. Shape, texture and statistical features for classification of benign and malignant vertebral compression fractures in magnetic resonance images. Computers in Biology and Medicine, v. 73, p. 147–156, 2016.

GALLOWAY, M. Texture analysis using gray level run lengths. Computer Graphics and Image Processing, v. 4, n. 2, p. 172–179, 1975.

GASTER, B. et al. Heterogeneous Computing with OpenCL. Waltham, USA: Morgan Kaufmann, 2012.

GEORGE, B.; SEALS, S.; ABAN, I. Survival analysis and regression models. Journal of Nuclear Cardiology, v. 21, n. 4, p. 686–694, 2014.

GERÔNIMO, T. M. et al. MLP and ANFIS applied to the prediction of hole diameters in the drilling process. In: Artificial Neural Networks - Architectures and Applications. London, UK: InTech, 2013. GEVAERT, O. et al. Predictive radiogenomics modeling of EGFR mutation status in lung cancer. Scientific Reports, v. 7, n. 41674, 2017.

GEVAERT, O. et al. Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data—methods and preliminary results. **Radiology**, v. 264, n. 2, p. 387–396, 2012.

GHONEIM, D. M. et al. Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas. Magnetic Resonance Imaging, v. 21, n. 9, p. 983–987, 2003.

GIARDINO, A. et al. Role of imaging in the era of precision medicine. Academic Radiology, v. 24, n. 5, p. 639–649, 2017.

GIGER, M. Machine learning in medical imaging. Journal of the American College of Radiology, v. 15, n. 3, p. 512–520, 2018.

GILHUIJS, K. G.; GIGER, M. L.; BICK, U. Computerized analysis of breast lesions in three dimensions using dynamic magnetic-resonance imaging. Medical Physics, v. 25, n. 9, p. 1647–1654, 1998.

GILLIES, R.; KINAHAN, P.; HRICAK, H. Radiomics: Images are more than pictures, they are data. **Radiology**, v. 278, n. 2, p. 563–77, 2016.

GONG, J. et al. Computer-aided diagnosis of lung cancer: the effect of training data sets on classification accuracy of lung nodules. **Physics in Medicine & Biology**, v. 63, n. 3, p. 035036, 2018.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. New Jersey, USA: Prentice Hall, 2007.

GROVE, O. et al. Quantitative computed tomographic descriptors associate tumor shape complexity and intratumor heterogeneity with prognosis in lung adenocarcinoma. **Plos One**, v. 10, n. 3, p. e0118261, 2015.

GUO, Y. et al. A comprehensive performance evaluation of 3D local feature descriptors. International Journal of Computer Vision, v. 116, n. 1, p. 66–89, 2016.

HALL, M. A. Correlation-based feature selection for machine learning. Tese (Doutorado) — Department of Computer Science, The University of Waikato, 1999.

HARALICK, R.; SHANMUGAM, K.; DINSTEIN, I. Textural features for image classification. **IEEE Transactions on Systems, Man and Cybernetics**, n. 6, p. 610–621, 1973.

HARALICK, R.; SHAPIRO, L. Computer and Robot Vision. Boston, USA: Addison-Wesley, 1992.

HAWKINS, S. et al. Predicting outcomes of nonsmall cell lung cancer using CT image features. **IEEE Access**, v. 2, p. 1418–1426, 2014.

HAYKIN, S. Neural networks: a comprehensive foundation. Upper Saddle River, USA: Prentice Hall PTR, 1994.

HOSNY, A. et al. Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study. **Plos Medicine**, v. 15, n. 11, p. e1002711, 2018.

HOWLADER, N. et al. **SEER Cancer Statistics Review**, **1975–2013**. Bethesda, USA: National Cancer Institute, 2016. Disponível em <www.seer.cancer.gov/csr/1975_2013/>. Acesso em: 23 Jan. 2019.

HUANG, P. et al. Added value of computer-aided CT image features for early lung cancer diagnosis with small pulmonary nodules: a matched case-control study. **Radiology**, v. 286, n. 1, p. 286–295, 2017.

HUANG, Y. et al. Radiomics signature: A potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. **Radiology**, v. 281, n. 3, p. 947–957, 2016.

HUNTER, L. A. et al. NSCLC tumor shrinkage prediction using quantitative image features. Computerized Medical Imaging and Graphics, v. 49, p. 29–36, 2016.

HUSSEIN, S. et al. TumorNet: Lung nodule characterization using multi-view convolutional neural network with Gaussian process. In: **Proceedings of 14th IEEE International Symposium on Biomedical Imaging**. Melbourne, Australia: IEEE, 2017. p. 1007–1010.

HUSSEIN, S. et al. Lung and pancreatic tumor characterization in the deep learning era: Novel supervised and unsupervised learning approaches. **IEEE Transactions on Medical Imaging**, 2019. DOI:10.1109/TMI.2019.2894349.

HUYNH, E. et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. **Radiotherapy and Oncology**, v. 120, n. 2, p. 258–266, 2016.

HUYNH, E. et al. Associations of radiomic data extracted from static and respiratory-gated CT scans with disease recurrence in lung cancer patients treated with SBRT. **Plos One**, v. 12, n. 1, p. e0169172, 2017.

IBANEZ, D. P. et al. Deep learning for pulmonary nodule CT image retrieval—an online assistance system for novice radiologists. In: **Proceedings of 2017 IEEE International Conference on Data Mining Workshops**. New Orleans, USA: IEEE, 2017. p. 1112–1121.

JEMAL, A. et al. **The Cancer Atlas**. Atlanta, USA: American Cancer Society, 2014. Disponível em <www.cancer.org/canceratlas>. Acesso em: 23 Jan. 2019.

JIA, Y. et al. Caffe: convolutional architecture for fast feature embedding. In: **Proceedings of the 22nd International Conference on Multimedia**. Orlando, USA: ACM, 2014. p. 675–678.

JOHN, G.; LANGLEY, P. Estimating continuous distributions in Bayesian classifiers. In: **Proceedings of Eleventh Conference on Uncertainty in Artificial Intelligence**. Montreal, Canada: ACM, 1995. p. 338–345.

JOHN, G. H.; KOHAVI, R.; PFLEGER, K. Irrelevant features and the subset selection problem. In: **Proceedings of the Eleventh International Conference, Rutgers University**. New Brunswick, USA: Elsevier, 1994. p. 121–129.

JONG, E. E. et al. Applicability of a prognostic CT-based radiomic signature model trained on stage I-III non-small cell lung cancer in stage IV non-small cell lung cancer. Lung Cancer, v. 124, p. 6–11, 2018.

KANSAGRA, A. et al. Big data and the future of radiology informatics. Academic Radiology, v. 23, n. 1, p. 30–42, 2016.

KAPLAN, E.; MEIER, P. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association, v. 53, n. 282, p. 457–481, 1958.

KAVZOGLU, T.; COLKESEN, I. A kernel functions analysis for support vector machines for land cover classification. International Journal of Applied Earth Observation and Geoinformation, v. 11, n. 5, p. 352–359, 2009.

KAYA, A. Cascaded classifiers and stacking methods for classification of pulmonary nodule characteristics. **Computer Methods and Programs in Biomedicine**, v. 166, p. 77–89, 2018.

KEEK, S. A. et al. A review on radiomics and the future of the ragnostics for patient selection in precision medicine. **The British Journal of Radiology**, v. 91, n. 1091, p. 20170926, 2018.

KIRA, K.; RENDELL, L. A practical approach to feature selection. In: **Proceedings** of Ninth International Workshop on Machine Learning. Aberdeen, Scotland: Elsevier, 1992. p. 249–256.

KIRK, D. B.; WEN-MEI, W. H. **Programming massively parallel processors: a hands-on approach**. Cambridge, USA: Morgan Kaufmann, 2016.

KO, J. et al. Lung adenocarcinoma: correlation of quantitative CT findings with pathologic findings. **Radiology**, v. 280, n. 3, p. 931–939, 2016.

KOENIGKAM-SANTOS, M. et al. Morphological computed tomography features of surgically resectable pulmonary squamous cell carcinomas: Impact on prognosis and comparison with adenocarcinomas. **European Journal of Radiology**, v. 83, n. 7, p. 1275–1281, 2014.

KOHAVI, R.; JOHN, G. Wrappers for feature subset selection. Artificial Intelligence, v. 97, n. 1-2, p. 273–324, 1997.

KOLOSSVÁRY, M. et al. Cardiac computed tomography radiomics: A comprehensive review on radiomic techniques. Journal of Thoracic Imaging, v. 33, n. 1, p. 26–34, 2018.

KONONENKO, I. Estimating attributes: analysis and extensions of RELIEF. In: **Proceedings of European Conference on Machine Learning**. Catania, Italy: Springer, 1994. p. 171–182.

KUHN, M.; JOHNSON, K. Applied predictive modeling. New York, USA: Springer, 2013.

KUMAR, A. et al. Content-based medical image retrieval: A survey of applications to multidimensional and multimodality data. **Journal of Digital Imaging**, v. 26, n. 6, p. 1025–1039, 2013.

KUMAR, V. et al. Radiomics: the process and the challenges. Magnetic Resonance Imaging, v. 30, n. 9, p. 1234–1248, 2012.

LAM, M. O. et al. BRISC—an open source pulmonary nodule image retrieval framework. Journal of Digital Imaging, v. 20, n. 1, p. 63–71, 2007.

LAMBIN, P. et al. Radiomics: the bridge between medical imaging and personalized medicine. **Nature Reviews Clinical Oncology**, v. 14, n. 12, p. 749, 2017.

LARUE, R. T. et al. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. **The British Journal of Radiology**, v. 90, n. 1070, p. 20160665, 2017.

LE, V. et al. Quantitative CT analysis of pulmonary nodules for lung adenocarcinoma risk classification based on an exponential weighted grey scale angular density distribution feature. **Computer Methods and Programs in Biomedicine**, v. 160, p. 141–151, 2018.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015.

LEE, G. et al. Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art. **European Journal of Radiology**, v. 86, p. 297–307, 2017.

LEE, J. et al. Deep learning in medical imaging: general overview. Korean Journal of Radiology, v. 18, n. 4, p. 570–584, 2017.

LEGER, S. et al. A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. Scientific Reports, v. 7, n. 1, p. 13206, 2017.

LEGLAND, D.; KIÊU, K.; DEVAUX, M. Computation of Minkowski measures on 2D and 3D binary images. **Image Analysis & Stereology**, v. 26, n. 2, p. 83–92, 2007.

LEVMAN, J. E.; MARTEL, A. L. A margin sharpness measurement for the diagnosis of breast cancer from magnetic resonance imaging examinations. Academic Radiology, v. 18, n. 12, p. 1577–1581, 2011.

LIM, J. et al. Can quantitative volumetric analysis predict tumor recurrence in the patients with mucinous adenocarcinoma of the lung after surgical resection? Academic Radiology, 2018. DOI:10.1016/j.acra.2018.06.010.

LITJENS, G. et al. A survey on deep learning in medical image analysis. Medical Image Analysis, v. 42, p. 60–88, 2017.

LIU, L. Computing infrastructure for big data processing. Frontiers of Computer Science, v. 7, n. 2, p. 165–170, 2013.

LIU, Y. et al. Radiomic features are associated with EGFR mutation status in lung adenocarcinomas. Clinical Lung Cancer, v. 17, n. 5, p. 441–448, 2016.

LIU, Y. et al. Prediction of pathological nodal involvement by CT-based Radiomic features of the primary tumor in patients with clinically node-negative peripheral lung adenocarcinomas. **Medical Physics**, v. 45, n. 6, p. 2518–2526, 2018.

LIU, Y. et al. CT features associated with epidermal growth factor receptor mutation status in patients with lung adenocarcinoma. **Radiology**, v. 280, n. 1, p. 271–280, 2016.

LUBNER, M. G. et al. CT texture analysis: definitions, applications, biologic correlates, and challenges. **Radiographics**, v. 37, n. 5, p. 1483–1503, 2017.

LUCENA, D. et al. Automatic weighing attribute to retrieve similar lung cancer nodules. **BMC Medical Informatics and Decision Making**, v. 16, n. 79, p. 135–149, 2016.

LUX, M.; MARQUES, O. Visual information retrieval using Java and LIRE. Williston, USA: Morgan & Claypool Publishers, 2013.

MA, L. et al. A new method of content based medical image retrieval and its applications to CT imaging sign retrieval. **Journal of Biomedical Informatics**, v. 66, p. 148–158, 2017.

MACHADO, K. K. et al. Sobrevida global e outros desfechos clínicos em câncer de mama: situação atual e controvérsias. **Revista da Associação Médica Brasileira**, v. 56, n. 5, p. 514–516, 2010.

MACMAHON, H. et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. **Radiology**, v. 284, n. 1, p. 228–243, 2017.

MANDELBROT, B. The Fractal Geometry of Nature. San Francisco, USA: Times Books, 1984.

MARR, D.; HILDRETH, E. Theory of edge detection. **Proceedings of the Royal** Society of London, Series B, Biological Sciences, v. 207, n. 1167, p. 187–217, 1980.

MARVASTI, N. B. et al. Clinical experience sharing by similar case retrieval. In: **Proceedings of the 1st International Workshop on Multimedia Indexing and Information Retrieval for Healthcare**. Barcelona, Spain: ACM, 2013. p. 67–74.

MOK, T. S. et al. Gefitinib or carboplatin–paclitaxel in pulmonary adenocarcinoma. New England Journal of Medicine, v. 361, n. 10, p. 947–957, 2009.

MOODY, J.; DARKEN, C. Fast learning in networks of locally-tuned processing units. **Neural Computation**, v. 1, n. 2, p. 281–294, 1989.

MORRIS, M. A. et al. Reinventing radiology: big data and the future of medical imaging. Journal of Thoracic Imaging, v. 33, n. 1, p. 4–16, 2018.

MURAMATSU, C. et al. Investigation of psychophysical similarity measures for selection of similar images in the diagnosis of clustered microcalcifications on mammograms. **Medical Physics**, v. 35, n. 12, p. 5695–5702, 2008.

NERUDA, R.; KUDOVÁ, P. Learning methods for radial basis function networks. Future Generation Computer Systems, v. 21, n. 7, p. 1131–1142, 2005.

NIBALI, A.; HE, Z.; WOLLERSHEIM, D. Pulmonary nodule classification with deep residual networks. International Journal of Computer Assisted Radiology and Surgery, v. 12, n. 10, p. 1799–1808, 2017.

NISHIMURA, K. et al. Psychophysical similarity measure based on multi-dimensional scaling for retrieval of similar images of breast masses on mammograms. In: **Proceedings** of Medical Imaging. Lake Buena Vista, USA: SPIE, 2013. v. 8670, p. 86701R.

NISHIO, M.; NAGASHIMA, C. Computer-aided diagnosis for lung cancer: usefulness of nodule heterogeneity. Academic Radiology, v. 24, n. 3, p. 328–336, 2017.

NÓBREGA, R. V. M. et al. Lung nodule malignancy classification in chest computed tomography images using transfer learning and convolutional neural networks. **Neural Computing and Applications**, p. 1–18, 2018. DOI:10.1007/s00521-018-3895-1.

OLIVEIRA, M. C. Grids Computacionais para recuperação de imagens médicas a partir de conteúdo: um estudo de viabilidade. Tese (Doutorado) — Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 2006.

OLIVEIRA, M. C.; CIRNE, W.; AZEVEDO-MARQUES, P. M. Towards applying content-based image retrieval in the clinical routine. Future Generation Computer Systems, v. 23, n. 3, p. 466–474, 2007.

OLIVEIRA, M. C.; FERREIRA, J. R. A Bag-of-Tasks approach to speed up the lung nodules retrieval in the BigData age. In: **Proceedings of 15th IEEE International Conference on E-Health Networking, Application & Services**. Lisbon, Portugal: IEEE, 2013. p. 632–636.

OROOJI, M. et al. Combination of computer extracted shape and texture features enables discrimination of granulomas from adenocarcinoma on chest computed tomography. **Journal of Medical Imaging**, v. 5, n. 2, p. 024501, 2018.

PACHECO, P. An introduction to parallel programming. Burlington, USA: Morgan Kaufmann, 2011.

PARKER, J. R. Algorithms for Image Processing and Computer Vision. Indianapolis, USA: Wiley Publishing, 2011.

PARMAR, C. et al. Machine learning methods for quantitative radiomic biomarkers. Scientific Reports, v. 5, n. 13087, 2015.

PARMAR, C. et al. Radiomic feature clusters and prognostic signatures specific for lung and head & neck cancer. **Scientific Reports**, v. 5, n. 11044, 2015.

PARMAR, C. et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. **Plos One**, v. 9, n. 7, p. e102107, 2014.

PAUL, R. et al. Deep feature transfer learning in combination with traditional features predicts survival among patients with lung adenocarcinoma. Tomography: A Journal for Imaging Research, v. 2, n. 4, p. 388–395, 2016.

PAUL, R. et al. Predicting malignant nodules by fusing deep features with classical radiomics features. **Journal of Medical Imaging**, v. 5, n. 1, p. 011021, 2018.

PAVLOPOULOS, S. A.; DELOPOULOS, A. N. Designing and implementing the transition to a fully digital hospital. **IEEE Transactions on Information Technology** in Biomedicine, v. 3, n. 1, p. 6–19, 1999.

PEDROSA, G. V. Caracterização e recuperação de imagens usando dicionários visuais semanticamente enriquecidos. Tese (Doutorado) — Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2015.

PENG, H.; LONG, F.; DING, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 27, n. 8, p. 1226–1238, 2005.

PERMUTH, J. B. et al. Combining radiomic features with a miRNA classifier may improve prediction of malignant pathology for pancreatic intraductal papillary mucinous neoplasms. **Oncotarget**, v. 7, n. 52, p. 85785, 2016.

PHILLIPS, I. et al. Clinical applications of textural analysis in non-small cell lung cancer. **The British Journal of Radiology**, v. 91, n. 1081, p. 20170267, 2017.

PINTER, C. et al. SlicerRT: Radiation therapy research toolkit for 3D Slicer. Medical Physics, v. 39, n. 10, p. 6332–6338, 2012.

PONCIANO-SILVA, M. et al. A differential method for representing spinal MRI for perceptual-CBIR. In: **Proceedings of Iberoamerican Congress on Pattern Recognition**. Berlin, Germany: Lecture Notes in Computer Science, 2013. v. 8258, p. 464–471.

QUINLAN, J. R. C4.5: programs for machine learning. San Francisco, USA: Morgan Kaufmann Publishers, 1993.

REEVES, A. P.; XIE, Y.; JIRAPATNAKUL, A. Automated pulmonary nodule CT image characterization in lung cancer screening. International Journal of Computer Assisted Radiology and Surgery, v. 11, n. 1, p. 73–88, 2016.

RIZZO, S. et al. CT radiogenomic characterization of EGFR, K-RAS, and ALK mutations in non-small cell lung cancer. **European Radiology**, v. 26, n. 1, p. 32–42, 2016.

RUBIN, G. et al. Characterizing search, recognition, and decision in the detection of lung nodules on CT scans: elucidation with eye tracking. **Radiology**, v. 274, n. 1, p. 276–286, 2014.

RUTMAN, A. M.; KUO, M. D. Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. **European Journal of Radiology**, v. 70, n. 2, p. 232–241, 2009.

SAAD, M.; CHOI, T. Computer-assisted subtyping and prognosis for non-small cell lung cancer patients with unresectable tumor. Computerized Medical Imaging and Graphics, v. 67, p. 1–8, 2018.

SACCONI, B. et al. Analysis of CT features and quantitative texture analysis in patients with lung adenocarcinoma: a correlation with EGFR mutations and survival rates. **Clinical Radiology**, v. 72, n. 6, p. 443–450, 2017.

SCHNEIDER, C.; RASBAND, W.; ELICEIRI, K. NIH Image to ImageJ: 25 years of image analysis. **Nature Methods**, v. 9, n. 7, p. 671–675, 2012.

SCHWENKER, F.; KESTLER, H. A.; PALM, G. Three learning phases for radial-basis-function networks. **Neural Networks**, v. 14, n. 4-5, p. 439–458, 2001.

SHEN, S. et al. An interpretable deep hierarchical semantic convolutional neural network for lung nodule malignancy classification. **Expert Systems with Applications**, v. 128, p. 84–95, 2019.

SHEN, W. et al. Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. **Pattern Recognition**, v. 61, p. 663–673, 2017.

SHROFF, G. S. et al. Targeted therapy and imaging findings. Journal of Thoracic Imaging, v. 32, n. 5, p. 313–322, 2017.

SILVA, G. L. et al. Lung nodules diagnosis based on evolutionary convolutional neural network. Multimedia Tools and Applications, v. 76, n. 18, p. 19039–19055, 2017.

SILVA, M. P. et al. Does a CBIR system really impact decisions of physicians in a clinical environment? In: Proceedings of 26th IEEE International Symposium on Computer-Based Medical Systems. Porto, Portugal: IEEE, 2013. p. 41–46.

SOBIN, L.; GOSPODAROWICZ, M.; WITTEKIND, C. **TNM Classification of Malignant Tumours**. Hoboken, USA: Wiley-Blackwell, 2010.

SOH, L.; TSATSOULIS, C. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. **IEEE Transactions on Geoscience and Remote Sensing**, v. 37, n. 2, p. 780–795, 1999.

SON, J. et al. Quantitative CT analysis of pulmonary ground-glass opacity nodules for distinguishing invasive adenocarcinoma from non-invasive or minimally invasive adenocarcinoma: the added value of using iodine mapping. **European Radiology**, v. 26, n. 1, p. 43–54, 2016.

SONG, Q. et al. Using deep learning for classification of lung nodules on computed tomography images. Journal of Healthcare Engineering, v. 2017, n. 8314740, 2017.

SONG, S. H. et al. Imaging phenotyping using radiomics to predict micropapillary pattern within lung adenocarcinoma. **Journal of Thoracic Oncology**, v. 12, n. 4, p. 624–632, 2017.

SONKA, M.; HLAVAC, V.; BOYLE, R. Image Processing, Analysis, and Machine Vision. Ludhiana, India: CL-Engineering, 2008.

SOUZA, J. P. Modelo de Qualidade para Desenvolvimento e Avaliação da Viabilidade Clínica de Sistemas de Recuperação de Imagens Médicas Baseada em Conteúdo. Tese (Doutorado) — Escola de Engenharia de São Carlos, Universidade de São Paulo, 2012.

STELMASZCZYK, A. et al. **CIFAR-10 Model for Caffe**. Berkeley, USA: BVLC, 2017. Disponível em <www.github.com/BVLC/caffe/tree/master/examples/cifar10>. Acesso em: 23 Jan. 2019.

SUN, W.; ZHENG, B.; QIAN, W. Automatic feature learning using multichannel roi based on deep structured algorithms for computerized lung cancer diagnosis. **Computers in Biology and Medicine**, v. 89, p. 530–539, 2017.

TAILOR, T. et al. The pseudocavitation sign of lung adenocarcinoma: a distinguishing feature and imaging biomarker of lepidic growth. **Journal of Thoracic Imaging**, v. 30, n. 5, p. 308–313, 2015.

TAJBAKHSH, N.; SUZUKI, K. Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs. **Pattern Recognition**, v. 63, p. 476–486, 2017.

TAMURA, H.; MORI, S.; YAMAWAKI, T. Textural features corresponding to visual perception. **IEEE Transactions on Systems, Man and Cybernetics**, v. 8, n. 6, p. 460–473, 1978.

TAMURA, T. et al. Specific organ metastases and survival in metastatic non-small-cell lung cancer. Molecular and Clinical Oncology, v. 3, n. 1, p. 217–221, 2015.

TANG, C. et al. Development of an immune-pathology informed radiomics model for non-small cell lung cancer. **Scientific Reports**, v. 8, n. 1, p. 1922, 2018.

TANG, X. Texture information in run-length matrices. **IEEE Transactions on Image Processing**, v. 7, n. 11, p. 1602–1609, 1998.

THAWANI, R. et al. Radiomics and radiogenomics in lung cancer: A review for the clinician. Lung Cancer, v. 115, p. 34–41, 2017.

TIWARI, S. Professional NoSQL. Hoboken, USA: John Wiley and Sons, Inc., 2011.

TRAINA, A. et al. Making content-based medical image retrieval systems for computer-aided diagnosis: From theory to application. In: AZEVEDO-MARQUES, P. M. et al. (Ed.). Medical Image Analysis and Informatics: Computer-Aided Diagnosis and Therapy. Boca Raton, USA: CRC Press, 2017. cap. 20, p. 467–490.

TRUONG, M. et al. Update in the evaluation of the solitary pulmonary nodule. **Radiographics**, v. 34, n. 6, p. 1658–1679, 2014.

VALLIÈRES, M. et al. Responsible radiomics research for faster clinical translation. Journal of Nuclear Medicine, v. 59, n. 2, p. 189, 2018.

VAPNIK, V. The nature of statistical learning theory. Red Bank, USA: Springer, 2013.

VASSALLO, L. et al. A cloud-based computer-aided detection system improves identification of lung nodules on computed tomography scans of patients with extra-thoracic malignancies. **European Radiology**, v. 29, n. 1, p. 144–152, 2019.

VELAZQUEZ, E. et al. Volumetric CT-based segmentation of NSCLC using 3D-Slicer. Scientific Reports, v. 3, p. 3529, 2013.

VELAZQUEZ, E. R. Multifactorial decision support systems in radiation oncology: clinical predictors and radiomics. Tese (Doutorado) — Maastricht University, 2014.

WANG, T. et al. Association of CT-based imaging features and genomic data in non-small cell lung cancer. In: **Proceedings of 2018 International Conference on Image and Video Processing, and Artificial Intelligence**. Shanghai, China: SPIE, 2018. v. 10836, p. 108360P.

WEI, G. et al. Content-based image retrieval for lung nodule classification using texture features and learned distance metric. **Journal of Medical Systems**, v. 42, n. 1, p. 13, 2018.

WITTEN, I.; FRANK, E.; HALL, M. Data Mining: Practical Machine Learning Tools and Techniques. Burlington, USA: Morgan Kaufmann Publishers, 2011.

WU, H. et al. Combination of radiological and gray level co-occurrence matrix textural features used to distinguish solitary pulmonary nodules by computed tomography. **Journal of Digital Imaging**, v. 26, n. 4, p. 797–802, 2013.

WU, J. et al. Early-stage non–small cell lung cancer: Quantitative imaging characteristics of 18F fluorodeoxyglucose PET/CT allow prediction of distant metastasis. **Radiology**, v. 281, n. 1, p. 270–278, 2016.

WU, W. et al. Exploratory study to identify radiomics classifiers for lung cancer histology. **Frontiers in Oncology**, v. 6, n. 71, 2016.

XU, J. et al. Quantifying the margin sharpness of lesions on radiological images for content-based image retrieval. **Medical Physics**, v. 39, p. 5405–5418, 2012.

YANG, J. et al. Uncertainty analysis of quantitative imaging features extracted from contrast-enhanced CT in lung tumors. **Computerized Medical Imaging and Graphics**, v. 48, p. 1–8, 2016.

YIP, S. S. et al. Associations between radiologist-defined semantic and automatically computed radiomic features in non-small cell lung cancer. **Scientific Reports**, v. 7, n. 1, p. 3519, 2017.

YOSHINO, Y. et al. Automatic classification of lung nodules on MDCT images with the temporal subtraction technique. **International Journal of Computer Assisted Radiology and Surgery**, v. 12, n. 10, p. 1789–1798, 2017.

ZHANG, L. et al. IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics. **Medical Physics**, v. 42, n. 3, p. 1341–1353, 2015.

ZHAO, X. et al. Agile convolutional neural network for pulmonary nodule classification using CT images. International Journal of Computer Assisted Radiology and Surgery, v. 13, p. 585–595, 2018.

ZHOU, H. et al. Diagnosis of distant metastasis of lung cancer: Based on clinical and radiomic features. **Translational Oncology**, v. 11, n. 1, p. 31–36, 2018.

ZHOU, M. et al. Non-small cell lung cancer radiogenomics map identifies relationships between molecular and imaging phenotypes with prognostic implications. **Radiology**, v. 286, n. 1, p. 307–315, 2017.

ZHU, X. et al. Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. **European Radiology**, v. 28, n. 7, p. 2772–2778, 2018.

ZWANENBURG, A. et al. Image biomarker standardisation initiative. **arXiv preprint arXiv:1612.07003**, 2018.

Apêndices

APÊNDICE A – ESQUEMA DE BANCO DE DADOS NÃO-RELACIONAL PARA O REPOSITÓRIO DE IMAGENS LIDC

O LIDC é um repositório de imagens médicas referência para a comunidade científica em engenharia e informática biomédica (KAYA, 2018; DHARA et al., 2017a; CAO et al., 2017; SILVA et al., 2017; NIBALI; HE; WOLLERSHEIM, 2017; LUCENA et al., 2016). Contudo, ele não está organizado em um banco de dados estruturado (Seção 3.2.1) (ARMATO III et al., 2018; ARMATO III et al., 2011). Neste trabalho, foi realizada a transferência dos dados clínicos contidos no arquivo XML de cada exame de TC (Anexo B) para uma base estruturada, junto das imagens dos nódulos para facilitar o acesso às informações. Além disso, no LIDC não existe nenhum tipo de atributo computadorizado que caracterize as imagens. Logo, atributos de textura baseados em matriz de coocorrência foram extraídos das imagens segmentadas (Seção 3.2.2) e também adicionados a este banco de imagens. Após a implementação do banco de dados proposto para o LIDC, as informações clínicas do nódulo e do exame, as marcações e classificações feitas pelo radiologista, e as imagens de TC originais e segmentadas foram armazenadas em um único banco de nódulos pulmonares. O esquema de banco de dados desenvolvido para o LIDC foi publicado no periódico Journal of Digital Imaging (Fator de Impacto 2017: 1,536, Qualis CAPES 2013–2016 na área Engenharias IV: A2) em forma de artigo científico (Anexo G) (FERREIRA JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2016).

Uma abordagem não-relacional ou NoSQL (do Inglês Not only Structured Query Language) foi utilizada para o esquema de banco de dados para facilitar a manutenção e evolução da base no futuro. NoSQL é um termo utilizado, principalmente em *Big Data*, para bancos de dados que não seguem os princípios estabelecidos pelos populares sistemas gerenciadores de bancos de dados (SGBDs) relacionais, representando uma coleção de diversos conceitos sobre armazenamento e manipulação de dados (BANKER, 2011). A abordagem NoSQL utilizada foi orientada a documentos, devido à sua consolidação como uma das estruturas de dados não-relacionais mais robustas e por possuir um esquema mais adequado ao armazenamento das informações do XML. Um banco de dados orientado a documentos é composto por um conjunto de coleções, e cada coleção possui um conjunto de documentos, que por sua vez são compostos por pares de chave-valor, listas e/ou outros documentos embarcados (TIWARI, 2011). Uma das implementações mais utilizadas para documentos em bases NoSQL é o JSON (do Inglês *JavaScript Object Notation*) (Figura 64).

O SGBD NoSQL utilizado neste trabalho foi o MongoDB v2.4.6 devido ao potencial de processamento paralelo, alta performance na recuperação de informações do banco, grande escalabilidade em instâncias de bancos de dados e por possuir a tecnologia GridFS (BANKER, 2011), importante para armazenar as imagens no banco desenvolvido. A Figura 64: Exemplo de documento JSON possuindo uma chave (*title*), uma lista (*topics*), um documento embarcado (*publisher*) e seus respectivos valores.

```
{
  title : "Journal of Digital Imaging",
  topics : ["CAD","DICOM","CBIR"],
  publisher : {
    name : "Springer",
    country : "US"
  }
}
```

Fonte: Autor.

especificação GridFS permite o MongoDB armazenar qualquer arquivo em formato binário utilizando duas coleções: *files*, que armazena os metadados do arquivo (tamanho em *bytes*, por exemplo), e *chunks*, que armazena o arquivo em formato de documento binário (Figura 65).





Fonte: Ferreira Junior, Oliveira e Azevedo-Marques (2016).

O esquema não-relacional de banco de dados proposto possui quatro coleções principais de documentos (Figura 66): exams, nodules, images.files e images.chunks. A coleção exams armazena os dados do exame de TC e da leitura das imagens. A coleção nodules armazena as informações do nódulo, incluindo as classificações do radiologista, os atributos de textura e os dados das ROIs (Figura 67). As coleções images.files e images.chunks pertencem ao GridFS e armazenam as imagens e seus metadados. O banco de dados foi implantado em uma infraestrutura de computação em nuvem para garantir a disponibilidade dos dados. A plataforma utilizada para o armazenamento foi a Morpheus, um serviço em nuvem que permite aos usuários o desenvolvimento, armazenamento e compartilhamento de instâncias de banco de dados em MongoDB.

Figura 66: Modelagem de dados do banco não-relacional orientado a documentos de nódulos pulmonares.



Fonte: Autor.

Figura 67: Visualização de dados por um sistema gerenciador para MongoDB. O Componente 1 ilustra as coleções do banco de dados desenvolvido e o Componente 2, um documento da coleção *nodules*.

E 🛛 Morpheus (1)	• db.nodules.find() ×		
DelicDB			
- Collections (5)	E 162.252.108.127:12279 E PUBLICDB		
± System	db.nodules.find()		
exams images chunks	nodules 🚯 1.544 sec.		
	Key	Value	
+ nodules	(1) Object/d("5408adbee4b0a74b0a1219c5")	(14 fields)	
 General Functions General Users 	id	ObjectId("5408adbee4b0a74b0a1219c5") 2	
	🗆 examID	Objectld("5408a10de4b034c7f369b6e1")	
	III noduleName	Nodule 001	
	"" subtlety	3	
	internal Structure	1	
	" calcification	6	
	"" sphericity	4	
	margin	5	
	E lobulation	1	
	in spiculation	1	
	iiii texture	5	
	malignancy	3	
	🕀 🖾 textureAttributes	{ 36 fields }	
	🖃 💷 rois	Array [2]	
	- \cdots O	{ 6 fields }	
	imageZposition	-147.544998	
	imageSOP_UID	1.3.6.1.4.1.14519.5.2.1.6279.6001.397214514135704746947495121412	
	inclusion	TRUE	
	🖂 originalImage	Objectld("540751dee4b06232da76e9be")	
	🔲 nodulelmage	Objectld("540751dfe4b06232da76e9c2")	
	🗄 💷 edgeMaps	Array [24]	
	÷ 😐 1	{ 6 fields }	

Fonte: Autor.

APÊNDICE B – OTIMIZAÇÃO DE CBIR EM UNIDADE DE PROCESSAMENTO GRÁFICO

A análise de similaridade de imagens é uma tarefa paralelizável, pois pode ocorrer a execução simultânea de uma mesma função (medida de distância) sobre elementos de um conjunto de dados (vetores de atributos das imagens). Como o volume de dados em saúde vem crescendo em um ritmo acelerado nos últimos anos, especialmente na radiologia (FILONENKO; SEERAM, 2018; KANSAGRA et al., 2016), neste trabalho foi realizada a otimização da análise de similaridade de imagens em arquiteturas paralelas com o objetivo de criar um modelo CBIR de alto desempenho computacional. Os resultados da otimização do tempo de execução na comparação de imagens em um algoritmo CBIR foram publicados no periódico *The Journal of Supercomputing* (Fator de Impacto 2017: 1,532, Qualis CAPES 2013–2016 na área Engenharias IV: B2) em forma de artigo científico (Anexo H) (FERREIRA JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2017).

Uma aplicação é dita paralela quando pode ser dividida em uma ou várias tarefas de execução, conhecidas como threads, que são executadas simultaneamente em núcleos de processamento distintos. Estes núcleos podem estar presentes principalmente em unidades de processamento central (do Inglês central processing unit - CPU) e, mais recentemente, nas GPUs (GASTER et al., 2012). Os dispositivos de um só núcleo de processamento (single-core) executam sequencialmente múltiplas threads. Dispositivos multi-core fazem uso de paralelismo devido a pluralidade no número de seus processadores. Arquiteturas many-core diferem de arquiteturas multi-core pelo número massivo de núcleos de processamento que as primeiras possuem sobre as segundas (KIRK; WEN-MEI, 2016). Enquanto a arquitetura multi-core é conhecida pelo reduzido número de processadores, e.g. CPUs dual-core, quad-core, octa-core, entre outras, a arquitetura many-core possui centenas ou milhares de núcleos de processamento, como nas GPUs (Figura 68) (LIU, 2013).

Figura 68: Arquitetura de uma CPU e uma GPU. ULA é a unidade lógica e aritmética.



Fonte: Kirk e Wen-Mei (2016).

O ganho de desempenho de uma aplicação paralela em relação à sua execução sequencial pode ser medido pela métrica *speedup*, dada pela divisão entre o tempo da execução sequencial pelo tempo da execução paralela (PACHECO, 2011).

Neste trabalho, três arquiteturas foram utilizadas para executar a distância Euclidiana entre os vetores de atributos das imagens: uma *single-core* (representada por um núcleo de processamento da CPU), uma *multi-core* (representada por todos os processadores da CPU) e uma *many-core* (representada por uma GPU). A avaliação de paralelismo considerou todos os 1.171 casos do conjunto de imagens desenvolvido na primeira etapa da pesquisa (Tabela 5). Os resultados de tempo de execução foram obtidos com o aumento no número de imagens que foram comparadas com uma mesma imagem de referência. Este número de imagens variou da seguinte forma: 1, 2, 4, 8, 16, 32, 64, 128, 256, 512 e 1.171.

Os dispositivos utilizados nos experimentos foram uma GPU Nvidia GeForce GT 640 954 Mhz com 1 Gb de memória e 384 núcleos de processamento e uma CPU Intel Core i5-3350P 3.10 GHz com 8 Gb de memória RAM e 4 processadores. O ambiente de desenvolvimento para ambos os dispositivos foi Linux com sistema operacional Ubuntu v13.04 e linguagem de programação Java v1.8. O *framework* de computação paralela foi OpenCL v1.2, para garantir a portabilidade do código-fonte nas diferentes arquiteturas, com o *binding* jOCL v0.1.9 para o Java. Os resultados obtidos para uma média de 10 execuções estão na Figura 69.





Fonte: Autor.

164

O tempo de execução da solução sequencial na arquitetura *single-core* aumentou consideravelmente após a comparação de 32 imagens utilizando o vetor completo de atributos. Já o desempenho do dispositivo *multi-core* começou a cair após a comparação de 128 casos. Depois da comparação de todas as imagens com o caso de referência, a solução paralela da arquitetura *multi-core* obteve *speedup* de 3,8x. A solução paralela na GPU se mostrou a mais estável, como esperado, e ao final da execução da análise de similaridade de todas as imagens obteve *speedup* de 23,7x.

APÊNDICE C – ANÁLISE RADIÔMICA DE TEXTURA DE TUMORES PULMONARES PARA DISTINÇÃO DE PSEUDOCAVIDADES E CAVIDADES

Características qualitativas de imagens de TC têm sido utilizadas como fatores preditivos do câncer de pulmão, influenciando o diagnóstico e o prognóstico do paciente (TAILOR et al., 2015; KOENIGKAM-SANTOS et al., 2014; AUSTIN et al., 2013). Duas destas características são cavidades e pseudocavidades presentes na lesão primária do tumor maligno (Figura 70). Devido à subjetividade da avaliação visual qualitativa destas características, neste trabalho foi realizada uma análise quantitativa de imagens de TC para distinguir tumores pulmonares com cavidades e pseudocavidades.

Figura 70: Exemplos de cavidade (a) e pseudocavidade (b) de tumores pulmonares malignos em imagens de TC. As setas indicam a localização do achado radiológico.



Fonte: Autor.

Dois radiologistas realizaram independentemente a avaliação visual das imagens de TC do Conjunto3 de pacientes (Tabela 14) para a identificação dos achados radiológicos. Quando houve discordância nas avaliações, os especialistas entraram em acordo. Os tumores foram caracterizados por 2.447 atributos de intensidade e textura extraídos das imagens de TC segmentadas semiautomaticamente (Seção 5.2.1) (FEDOROV et al., 2012; EGGER et al., 2013; VELAZQUEZ et al., 2013). As características quantitativas calculadas foram de nível de cinza, histograma, COM, RLM, NIDM, LOG, Tamura, Fourier, Gabor, wavelet de Haar e dimensão fractal (Seção 5.2.2) (ZHANG et al., 2015; LUX; MARQUES, 2013; SCHNEIDER; RASBAND; ELICEIRI, 2012). O Teste U de Mann-Whitney foi utilizado na avaliação estatística da associação dos atributos com os padrões radiológicos de cavidade vs. pseudocavidade.

Trinta e um pacientes foram incluídos na análise (idade média de 67,8 \pm 7,9 anos, 17 homens, 26 fumantes ativos ou passivos), sendo 25 apresentando tumor com pseudocavidade e 6 apresentando tumor com cavidade, de acordo com a classificação dos radiologistas. A análise identificou 227 atributos estatisticamente relevantes para distinguir as imagens. O atributo de Matiz (Equação 2.21) da matriz de coocorrência de níveis de cinza obteve a maior AUC com valor de 0,973 e valor-p < 0,001 (Figura 71). Atributos de nível de cinza (número de atributos / número total de atributos deste tipo: porcentagem, 6/53: 11%), histograma (3/51: 5%), COM (205/2090: 9%), RLM (9/33: 27%) e NIDM (4/20: 20%) obtiveram associações estatisticamente significativas com os padrões radiológicos.

Figura 71: Maiores performances estatísticas dos atributos para a distinção pseudocavidade vs. cavidade. Os atributos de EIBSL, DPDPL, DPAL, ESIB e EIBSC são ênfase de intensidade baixa de sequência longa, desvio padrão do desvio padrão local, desvio padrão da amplitude local, ênfase de sequência de intensidade baixa e ênfase de intensidade baixa de sequência curta, respectivamente.



Fonte: Autor.

Anexos

ANEXO A – ARTIGO PUBLICADO NO PERIÓDICO J DIGIT IMAGING (2018) 31:451–463

J Digit Imaging (2018) 31:451–463 https://doi.org/10.1007/s10278-017-0029-8

CrossMark

Characterization of Pulmonary Nodules Based on Features of Margin Sharpness and Texture

José Raniery Ferreira Jr¹ ^(D) · Marcelo Costa Oliveira² · Paulo Mazzoncini de Azevedo-Marques¹

Published online: 18 October 2017 © Society for Imaging Informatics in Medicine 2017

Abstract Lung cancer is the leading cause of cancer-related deaths in the world, and one of its manifestations occurs with the appearance of pulmonary nodules. The classification of pulmonary nodules may be a complex task to specialists due to temporal, subjective, and qualitative aspects. Therefore, it is important to integrate computational tools to the early pulmonary nodule classification process, since they have the potential to characterize objectively and quantitatively the lesions. In this context, the goal of this work is to perform the classification of pulmonary nodules based on image features of texture and margin sharpness. Computed tomography scans were obtained from a publicly available image database. Texture attributes were extracted from a co-occurrence matrix obtained from the nodule volume. Margin sharpness attributes were extracted from perpendicular lines drawn over the borders on all nodule slices. Feature selection was performed by different algorithms. Classification was performed by several machine learning classifiers and assessed by the area under the receiver operating characteristic curve, sensitivity, specificity, and accuracy. Highest classification performance was obtained by a random forest algorithm with all 48 extracted features. However, a decision tree using only two selected features

José Raniery Ferreira Jr jose.raniery@usp.br obtained statistically equivalent performance on sensitivity and specificity.

Keywords Lung cancer · Pulmonary nodule · Image classification · Pattern recognition

Introduction

Lung cancer is the most common cause of cancer-related deaths, with a 5-year overall survival rate of only 15% [1]. The evaluation of pulmonary nodules is clinically important because they may be an early manifestation of lung cancer [2]. The diagnosis of lung cancer may be a complex task to radiologists and it presents some challenges. One of them is to classify pulmonary nodules in diagnostic imaging. Nodule classification in malignant or benign depends on several aspects [3]: for instance, its growth rate and change in size from two time-separated computed tomography (CT) scans; and subjective, qualitative aspects of the lesion, e.g., "moderate heterogeneity," "highly spiculated," "large necrotic core" [4].

To aid radiologists in the diagnosis of lung cancer, it is important to integrate the computer-based assistance into the processes of imaging pattern recognition and pulmonary nodule classification [5, 6]. The purpose of computer-aided diagnosis (CAD) is to improve the accuracy and consistency of medical image diagnosis through computational support used as reference [7]. In particular, the automation of pattern classification process may considerably reduce the time and effort required by the analysis and, at the same time, improve its repeatability and reliability [8].

Some works have performed computer-aided classification of pulmonary nodules, using different classifiers and CT image features to improve the diagnosis of lung cancer

¹ Center of Imaging Sciences and Medical Physics, Ribeirão Preto Medical School, University of São Paulo, Av. dos Bandeirantes, 3900, Monte Alegre, Ribeirão Preto, São Paulo 14049-900, Brazil

² Institute of Computing, Federal University of Alagoas, Av. Lourival Melo Mota, Cidade Universitária, Maceió, Alagoas 57072-900, Brazil

ANEXO B – ARQUIVO XML DE UM EXAME DE TC DO REPOSITÓRIO DE IMAGENS LIDC

lidc xml			
<pre><?xmlversion="1.0"encoding="UTF-8"?></pre>			
<lidcreadmessageuid="1.3.6.1.4.1.1307390687803.0"< td=""></lidcreadmessageuid="1.3.6.1.4.1.1307390687803.0"<>			
xmlns="http://www.nih.gov"			
<pre>xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"</pre>			
<pre>xsi:schemaLocation="http://www.nih.gov"></pre>			
<responseheader></responseheader>			
<version>1.8.1</version>			
<messageid>-421198203</messageid>			
<daterequest>2007-11-01</daterequest>			
<timerequest>12:30:44</timerequest>			
<requestingsite>removed</requestingsite>			
<servicingsite>removed</servicingsite>			
<taskdescription>Secondunblindedread</taskdescription>			
<ctimagefile>removed</ctimagefile>			
<seriesinstanceuid>1.3.6.1.4.1.14519.03192</seriesinstanceuid>			
<dateservice>2008-08-18</dateservice>			
<timeservice>02:05:51</timeservice>			
<responsedescription>1-Readingcomplete</responsedescription>			
<studyinstanceuid>1.3.6.1.4.1.14519.30178</studyinstanceuid>			
<readingsession></readingsession>			
<annotationversion>3.12</annotationversion>			
<servicingradiologistid>540461523</servicingradiologistid>			
<unblindedreadnodule></unblindedreadnodule>			
<noduleid>Nodule001</noduleid>			
<characteristics></characteristics>			
<subtlety>5</subtlety>			
<internalstructure>1</internalstructure>			
<calcification>6</calcification>			
<sphericity>3</sphericity>			
<margin>3</margin>			
<lobulation>3</lobulation>			
<spiculation>4</spiculation>			
<texture>5</texture>			
<malignancy>5</malignancy>			
<roi></roi>			
<imagezposition>-125.000000</imagezposition>			
<imagesop_uid>1.3.6.1.4.1.14519.6001.08719</imagesop_uid>			
<inclusion>TRUE</inclusion>			
<edgemap></edgemap>			
<xcoord>312</xcoord>			
<ycoord>355</ycoord>			
<edgemap></edgemap>			
<xcoord>312</xcoord>			
<ycoord>355</ycoord>			
<roi></roi>			

```
<imageZposition>-122.500000</imageZposition>
<imageSOP_UID>1.3.6.1.4.1.14519.6001.53475</imageSOP_UID>
<inclusion>TRUE</inclusion>
<edgeMap>
<xCoord>314</xCoord>
<yCoord>346</yCoord>
</edgeMap>
. . .
<edgeMap>
<xCoord>314</xCoord>
<yCoord>346</yCoord>
</edgeMap>
</roi>
<roi>
<imageZposition>-120.000000</imageZposition>
<imageSOP_UID>1.3.6.1.4.1.14519.5.2.1.34793</imageSOP_UID>
<inclusion>TRUE</inclusion>
<edgeMap>
<xCoord>312</xCoord>
<yCoord>346</yCoord>
</edgeMap>
• • •
<edgeMap>
<xCoord>312</xCoord>
<yCoord>346</yCoord>
</edgeMap>
</roi>
<roi>
<imageZposition>-117.500000</imageZposition>
<imageSOP_UID>1.3.6.1.4.1.14519.5.2.1.88091</imageSOP_UID>
<inclusion>TRUE</inclusion>
<edgeMap>
<xCoord>308</xCoord>
<yCoord>340</yCoord>
</edgeMap>
. . .
<edgeMap>
<xCoord>308</xCoord>
<yCoord>340</yCoord>
</edgeMap>
</roi>
</unblindedReadNodule>
<unblindedReadNodule>
<noduleID>Nodule002</noduleID>
<roi>
<imageZposition>-150.000000</imageZposition>
<imageSOP_UID>1.3.6.1.4.1.98497029348077255</imageSOP_UID>
<inclusion>TRUE</inclusion>
<edgeMap>
<xCoord>187</xCoord>
<yCoord>166</yCoord>
</edgeMap>
</roi>
</unblindedReadNodule>
<nonNodule>
<nonNoduleID>_Non-nodule001</nonNoduleID>
<imageZposition>-185.000000</imageZposition>
<imageSOP_UID>1.3.6.1.4.1.2657524067855634242</imageSOP_UID>
<100118>
<xCoord>396</xCoord>
```

<yCoord>216</yCoord> </locus> </nonNodule> </readingSession> <readingSession> <annotationVersion>3.12</annotationVersion> <servicingRadiologistID>1428348137</servicingRadiologistID> <unblindedReadNodule> <noduleID>IL057_127364</noduleID> <characteristics> <subtlety>5</subtlety> <internalStructure>1</internalStructure> <calcification>6</calcification> <sphericity>4</sphericity> <margin>4</margin> <lobulation>5</lobulation> <spiculation>5</spiculation> <texture>5</texture> <malignancy>5</malignancy> </characteristics> <roi> <imageZposition>-107.5</imageZposition> <imageSOP_UID>1.3.6.1.4.1.28141667906503</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>310</xCoord> <yCoord>378</yCoord> </edgeMap> . . . <edgeMap> <xCoord>310</xCoord> <yCoord>378</yCoord> </edgeMap> </roi> <roi> <imageZposition>-110.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.2027094237779</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>329</xCoord> <yCoord>379</yCoord> </edgeMap> . . . <edgeMap> <xCoord>329</xCoord> <yCoord>379</yCoord> </edgeMap> </roi> <roi> <imageZposition>-112.5</imageZposition> <imageSOP_UID>1.3.6.1.4.1.2611512339602</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>315</xCoord> <yCoord>384</yCoord> </edgeMap> . . . <edgeMap> <xCoord>315</xCoord> <yCoord>384</yCoord>

</edgeMap> </roi> <roi> <imageZposition>-115.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.297813206491522</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>311</xCoord> <yCoord>385</yCoord> </edgeMap> . . . <edgeMap> <xCoord>311</xCoord> <yCoord>385</yCoord> </edgeMap> </roi> </unblindedReadNodule> <nonNodule> <nonNoduleID>IL057_127367</nonNoduleID> <imageZposition>-195.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.19545811240897787</imageSOP_UID> <locus> <xCoord>408</xCoord> <yCoord>209</yCoord> </locus> </nonNodule> </readingSession> <readingSession> <annotationVersion>3.12</annotationVersion> <servicingRadiologistID>-750896469</servicingRadiologistID> <unblindedReadNodule> <noduleID>MI014_12118</noduleID> <roi> <imageZposition>-175.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.315606855383999143703</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>392</xCoord> <yCoord>317</yCoord> </edgeMap> </roi> </unblindedReadNodule> <unblindedReadNodule> <noduleID>MI014_12119</noduleID> <roi> <imageZposition>-175.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.3156068553839991437</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>404</xCoord> <yCoord>273</yCoord> </edgeMap> </roi> </unblindedReadNodule> <unblindedReadNodule> <noduleID>MI014_12120</noduleID> <roi> <imageZposition>-150.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.313544823773855095</imageSOP_UID> <inclusion>TRUE</inclusion>

<edgeMap> <xCoord>189</xCoord> <yCoord>167</yCoord> </edgeMap> </roi> </unblindedReadNodule> <unblindedReadNodule> <noduleID>MI014_12127</noduleID> <characteristics> <subtlety>5</subtlety> <internalStructure>1</internalStructure> <calcification>6</calcification> <sphericity>3</sphericity> <margin>2</margin> <lobulation>3</lobulation> <spiculation>3</spiculation> <texture>5</texture> <malignancy>5</malignancy> </characteristics> <roi> <imageZposition>-107.5</imageZposition> <imageSOP_UID>1.3.6.1.4.1.28141667906503662</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>306</xCoord> <yCoord>381</yCoord> </edgeMap> . . . <edgeMap> <xCoord>306</xCoord> <yCoord>381</yCoord> </edgeMap> </roi> <roi> <imageZposition>-110.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.202709423777326874</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>321</xCoord> <yCoord>381</yCoord> </edgeMap> . . . <edgeMap> <xCoord>321</xCoord> <yCoord>381</yCoord> </edgeMap> </roi> <roi> <imageZposition>-112.5</imageZposition> <imageSOP_UID>1.3.6.1.4.1.2611512339602690987</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>311</xCoord> <yCoord>383</yCoord> </edgeMap> . . . <edgeMap> <xCoord>311</xCoord> <yCoord>383</yCoord> </edgeMap>

</roi> <roi> <imageZposition>-115.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.2978132064915229131</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>309</xCoord> <yCoord>388</yCoord> </edgeMap> . . . <edgeMap> <xCoord>309</xCoord> <yCoord>388</yCoord> </edgeMap> </roi> </unblindedReadNodule> <nonNodule> <nonNoduleID>MI014_12121</nonNoduleID> <imageZposition>-165.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.4835143194725792701</imageSOP_UID> <locus> <xCoord>213</xCoord> <yCoord>157</yCoord> </locus> </nonNodule> <nonNodule> </readingSession> <readingSession> <annotationVersion>3.12</annotationVersion> <servicingRadiologistID>2103845659</servicingRadiologistID> <unblindedReadNodule> <noduleID>0</noduleID> <characteristics> <subtlety>5</subtlety> <internalStructure>1</internalStructure> <calcification>6</calcification> <sphericity>5</sphericity> <margin>4</margin> <lobulation>1</lobulation> <spiculation>5</spiculation> <texture>4</texture> <malignancy>4</malignancy> </characteristics> <roi> <imageZposition>-105.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.257699431143</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>311</xCoord> <yCoord>361</yCoord> </edgeMap> . . . <edgeMap> <xCoord>311</xCoord> <yCoord>361</yCoord> </edgeMap> </roi> <roi> <imageZposition>-107.5</imageZposition> <imageSOP_UID>1.3.6.1.4.1.2814166799872</imageSOP_UID>

<inclusion>TRUE</inclusion> <edgeMap> <xCoord>313</xCoord> <yCoord>355</yCoord> </edgeMap> . . . <edgeMap> <xCoord>313</xCoord> <yCoord>355</yCoord> </edgeMap> </roi> <roi> <imageZposition>-110.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.2027094237774</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>316</xCoord> <yCoord>356</yCoord> </edgeMap> . . . <edgeMap> <xCoord>316</xCoord> <yCoord>356</yCoord> </edgeMap> </roi> <roi> <imageZposition>-112.5</imageZposition> <imageSOP_UID>1.3.6.1.4.1.6115123396023</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>322</xCoord> <yCoord>355</yCoord> </edgeMap> . . . <edgeMap> <xCoord>322</xCoord> <yCoord>355</yCoord> </edgeMap> </roi> </unblindedReadNodule> <unblindedReadNodule> <noduleID>1</noduleID> <roi> <imageZposition>-175.0</imageZposition> <imageSOP_UID>1.3.6.1.4.1.31560685538392</imageSOP_UID> <inclusion>TRUE</inclusion> <edgeMap> <xCoord>392</xCoord> <yCoord>316</yCoord> </edgeMap> </roi> </unblindedReadNodule> </readingSession> </LidcReadMessage>

ANEXO C – ARTIGO PUBLICADO NO PERIÓDICO INT J CARS (2017) 12:509–517

Int J CARS (2017) 12:509–517 DOI 10.1007/s11548-016-1471-7

ORIGINAL ARTICLE

Selecting relevant 3D image features of margin sharpness and texture for lung nodule retrieval

José Raniery Ferreira Jr. $^1 \odot$ · Paulo Mazzoncini de Azevedo-Marques 1 · Marcelo Costa Oliveira^2

Received: 29 April 2016 / Accepted: 11 August 2016 / Published online: 23 August 2016 © CARS 2016

Abstract

Purpose Lung cancer is the leading cause of cancer-related deaths in the world. Its diagnosis is a challenge task to specialists due to several aspects on the classification of lung nodules. Therefore, it is important to integrate content-based image retrieval methods on the lung nodule classification process, since they are capable of retrieving similar cases from databases that were previously diagnosed. However, this mechanism depends on extracting relevant image features in order to obtain high efficiency. The goal of this paper is to perform the selection of 3D image features of margin sharpness and texture that can be relevant on the retrieval of similar cancerous and benign lung nodules.

Methods A total of 48 3D image attributes were extracted from the nodule volume. Border sharpness features were extracted from perpendicular lines drawn over the lesion boundary. Second-order texture features were extracted from a cooccurrence matrix. Relevant features were selected by a correlation-based method and a statistical significance analysis. Retrieval performance was assessed according to the nodule's potential malignancy on the 10 most similar cases and by the parameters of precision and recall.

Results Statistical significant features reduced retrieval

performance. Correlation-based method selected 2 margin sharpness attributes and 6 texture attributes and obtained higher precision compared to all 48 extracted features on similar nodule retrieval.

Conclusion Feature space dimensionality reduction of 83 % obtained higher retrieval performance and presented to be a computationaly low cost method of retrieving similar nodules for the diagnosis of lung cancer.

Keywords Lung cancer · Pulmonary nodule · Contentbased image retrieval · Image feature extraction · Image feature selection

Introduction

Cancer is a pathology characterized as an uncontrolled growth of abnormal cells that invades neighboring tissues. Lung cancer accounts for one-third of all cancer-related deaths in the United States, the highest mortality of all cancers [1]. The evaluation of lung nodules is clinically important because they may be an early manifestation of lung cancer. A lung nodule is defined as a round opacity that is at least moderately well marginated and no larger than 3 cm in its maximum diameter [2].

Imaging diagnosis of lung cancer may be a complex task to radiologists due to the lung nodule classification process. The evaluation of a lung nodule depends on a variety of factors including the growth, size, location, attenuation and morphology of the nodule [3]. However, there is considerable overlap in the imaging characteristics of benign and cancerous lung nodules [2].

In order to aid specialists in the lung cancer imaging diagnosis, it is important to integrate computer-based assistance into the image interpretation and lung nodule classifica-

Springer

CrossMark

[⊠] José Raniery Ferreira Jr. jose.raniery@usp.br

¹ Center of Imaging Sciences and Medical Physics, Internal Medicine Department, Ribeirao Preto Medical School, University of Sao Paulo (USP), Av. dos Bandeirantes, 3900, Campus USP, Monte Alegre, Ribeirão Preto, São Paulo 14049-900, Brazil

² Lab of Telemedicine and Medical Informatics, University Hospital Prof. Alberto Antunes, Institute of Computing, Federal University of Alagoas (UFAL), Av. Lourival Melo Mota, s/n, Campus A. C. Simões, Cidade Universitária, Maceió, Alagoas 57072-900, Brazil

ANEXO D – ARTIGO PUBLICADO NO PERIÓDICO COMPUT METHODS PROGRAMS BIOMED (2018) 159:23–30

Computer Methods and Programs in Biomedicine 159 (2018) 23-30



Radiomics-based features for pattern recognition of lung cancer histopathology and metastases



José Raniery Ferreira Junior^{a,}*, Marcel Koenigkam-Santos^b, Federico Enrique Garcia Cipriano^b, Alexandre Todorovic Fabro^b, Paulo Mazzoncini de Azevedo-Marques^b

^a São Carlos School of Engineering, University of São Paulo, São Carlos, SP 13566-590, Brazil ^b Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP 14049-900, Brazil

ARTICLE INFO

Article history: Received 26 July 2017 Revised 16 February 2018 Accepted 22 February 2018

Keywords: Lung cancer Metastasis prediction Pattern recognition Quantitative image analysis Radiomics

ABSTRACT

Background and Objectives: lung cancer is the leading cause of cancer-related deaths in the world, and its poor prognosis varies markedly according to tumor staging. Computed tomography (CT) is the imaging modality of choice for lung cancer evaluation, being used for diagnosis and clinical staging. Besides tumor stage, other features, like histopathological subtype, can also add prognostic information. In this work, radiomics-based CT features were used to predict lung cancer histopathology and metastases using machine learning models. Methods: local image datasets of confirmed primary malignant pulmonary tumors were retrospectively evaluated for testing and validation. CT images acquired with same protocol were semiautomatically segmented. Tumors were characterized by clinical features and computer attributes of intensity, histogram, texture, shape, and volume. Three machine learning classifiers used up to 100 selected features to perform the analysis. Results: radiomics-based features yielded areas under the receiver operating characteristic curve of 0.89, 0.97, and 0.92 at testing and 0.75, 0.71, and 0.81 at validation for lymph nodal metastasis, distant metastasis, and histopathology pattern recognition, respectively. Conclusions: the radiomics characterizion approach presented great potential to be used in a computational model to aid lung cancer histopathological subtype diagnosis as a "virtual biopsy" and metastatic prediction for therapy decision support without the necessity of a whole-body imaging scanning.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Lung cancer accounts for one-third of all cancer-related deaths in the United States, with the highest mortality of all cancers [1]. The prognosis of lung cancer is still poor and varies markedly according to tumor staging at diagnosis. Tumor stage at presentation, as designated by the tumor-node-metastasis system (TNM describes the anatomical extent of disease based on assessment of three components: extent of the primary tumor (T), absence or presence and extent of regional lymph node metastasis (N), and absence or presence of distant metastasis (M)), is the most important prognostic factor and may determine therapy [2].

However, studies showed that clinical decision making may be influenced by other tumor aspects, such as, the histopathological subtype of the lesion [3]. In clinical practice, lung cancer can be classified in two main categories: non-small cell lung cancer

https://doi.org/10.1016/j.cmpb.2018.02.015 0169-2607/© 2018 Elsevier B.V. All rights reserved. (NSCLC) and small cell lung cancer (SCLC). NSCLC comprehends 85% of the cases and is mainly subclassified in adenocarcinoma (ADC), squamous cell carcinoma (SCC), and large cell carcinoma (LCC).

Besides tumor histology and staging, it has been described that other computed tomography (CT) features may also influence prognosis and response to therapy [2,4,5]. However, those features are typically described subjectively, qualitatively or semiquantitatively, e.g. non-solid, semi-solid, or solid nodules; and single 2D measure of greatest diameter on axial plane. Furthermore, subtypes of ADC, SCC, and SCLC may present visual CT features that are similar to benign lesions [6].

Computer-aided detection/diagnosis (CAD) tools have been developed to aid specialists interpret medical imaging findings and identify early diseases, especially breast and lung tumors [7–9]. The purpose of CAD is to improve the accuracy and consistency of medical image diagnosis through computational support used as reference [10]. Traditionally, CAD systems provide a single answer (presence of a cancer, for instance) as a second opinion to special-

^{*} Corresponding author.

E-mail address: jose.raniery@usp.br (J.R. Ferreira Junior).

ANEXO E – OFÍCIO COM PARECER DO COMITÊ DE ÉTICA EM PESQUISA



HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO



Ribeirão Preto, 04 de abril de 2017.

Oficio nº 996/2017 CEP/MGV

PROCESSO HCRP nº 3733/2017

Prezado Pesquisador,

O trabalho intitulado **"MODELO RADIÔMICO DE PREDIÇÃO DIAGNÓSTICA E PROGNÓSTICA DO CÂNCER DE PULMÃO",** foi analisado pelo Comitê de Ética em Pesquisa, em sua 447ª Reunião Ordinária realizada em 03/04/2017, e enquadrado na categoria: **APROVADO**, bem como a solicitação de dispensa de aplicação do Termo de Consentimento Livre e Esclarecido.

De acordo com Carta Circular nº 003/2011/CONEP/CNS, datada de 21/03/2011, o sujeito de pesquisa ou seu representante, quando for o caso, deverá rubricar todas as folhas do Termo de Consentimento Livre e Esclarecido – TCLE – apondo sua assinatura na última do referido Termo; o pesquisador responsável deverá da mesma forma, rubricar todas as folhas do Termo de Consentimento Livre e Esclarecido – TCLE – apondo sua assinatura na última do referido Termo.

Este Comitê segue integralmente a Conferência Internacional de Harmonização de Boas Práticas Clínicas (IGH-GCP), bem como a Resolução nº 466/2012 CNS/MS.

Lembramos que devem ser apresentados a este CEP, o Relatório Parcial e o Relatório Final da pesquisa.

Atenciosamente.

DR^a MARCIA GUIMARÃES VILLANOVA Coordenadora do Comitê de Ética em Pesquisa do HCRP e da FMRP-USP

Ilustríssimo Senhor **MARCEL KOENIGKAM SANTOS** Depto. de Clínica Médica – Centro de Ciências das Imagens e Física Médica

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO Campus Universitário – Monte Alegre 14048-900 Ribeirão Preto SP FVA-0002733; IRB-0002786 Registro Plataforma Brasil /CONEP nº 5440 (016) 3602-2228 cep@hcrp.usp.br

www.hcrp.usp.br

ANEXO F – ARQUIVO DICOM-RT DE UM EXAME DE IMAGENS DE TC SEGMENTADAS

	rtss.dcm
0002,0002	Media Storage SOP Class UID: 1.2.840.10008.5.1.4.1.1.481.3
0002,0003	Media Storage SOP Inst UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.117796
0002,0010	Transfer Syntax UID: 1.2.840.10008.1.2.1
0002,0012	Implementation Class UID: 1.2.276.0.7230010.3.0.3.6.1
0002,0013	Implementation Version Name: OFFIS_DCMTK_361
0008,0012	Instance Creation Date: 20110101
0008,0013	Instance Creation Time: 120000
0008,0014	Instance Creator UID: 1.2.826.0.1.3680043.8.274.1.1
0008,0016	SOP Class UID: 1.2.840.10008.5.1.4.1.1.481.3
0008,0018	SOP Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.117796
0008,0020	Study Date: 20110101
0008,0030	Study Time: 120000
0008,0050	Accession Number:
0008,0060	Modality: RTSTRUCT
0008,0070	Manufacturer: Plastimatch
0008,0080	Institution Name:
0008,0090	Referring Physician's Name:
0008,1010	Station Name:
0008,103E	Series Description:
0008,1090	- Manufacturer's Model Name: Plastimatch
0010,0010	Patient's Name:
0010,0020	Patient ID:
0010,0030	Patient's Birth Date:
0010,0040	Patient's Sex:
0018,1020	Software Versionss: 1.5.16-beta 4567
0020,000D	Study Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.117792
0020,000E	Series Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.117797
0020,0010	Study ID: 10001
0020,0011	Series Number: 103
0020,0013	Image Number: 1
3006,0002	: AutoSS
3006,0004	: AutoSS
3006,0008	: 20110101
3006,0009	: 120000
0020,0052	>Frame of Reference UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.117793
0008,1150	>Referenced SOP Class UID: 1.2.840.10008.3.1.2.3.2
0008,1155	>Referenced SOP Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.117792
0020,000E	>Series Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.117794
0008,1150	>Referenced SOP Class UID: 1.2.840.10008.5.1.4.1.1.2
0008,1155	>Referenced SOP Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.117800
0008,1150	>Referenced SOP Class UID: 1.2.840.10008.5.1.4.1.1.2
0008,1155	>Referenced SOP Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.118238
3006,0022	>: 1
3006,0024	>: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.117793
3006,0026	>: Model1_9_9_Contour_IndexedLabelmap
3006,002A	>: 0\0\0
0008,1150	>Referenced SOP Class UID: 1.2.840.10008.5.1.4.1.1.2
0008,1155	>Referenced SOP Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.118009
3006,0042	: CLOSED_PLANAR
3006,0046	: 5
3006,0050 ---: 108.371\170.552\194\107.923\170.104\194\108.371\169.655\194\108.819\170.104\194\108.371 \170.552\194 0008,1150 >Referenced SOP Class UID: 1.2.840.10008.5.1.4.1.1.2 0008,1155 >Referenced SOP Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.118010 3006,0042 ---: CLOSED_PLANAR 3006,0046 ---: 5 3006,0050 ---: 104.785\166.966\195\104.337\166.518\195\104.785\166.069\195\105.233\166.518\195\104.785 \166.966\195 0008,1150 >Referenced SOP Class UID: 1.2.840.10008.5.1.4.1.1.2 0008,1155 >Referenced SOP Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.118010 3006,0042 ---: CLOSED_PLANAR 3006,0046 ---: 21 3006,0050 ---: 111.061\170.552\195\110.164\170.552\195\109.268\170.552\195\108.371\170.552\195\107.475 $\label{eq:170.552} 195 \\ 106.578 \\ 170.552 \\ 195 \\ 105.682 \\ 170.552 \\ 195 \\ 105.233 \\ 170.104 \\ 195 \\ 105.233 \\ 169.207 \\ 195 \\ 105.682 \\ 105.6$ \168.759\195\106.13\168.311\195\106.578\167.862\195\107.475\167.862\195\108.371\167.862\195\108.819 $\label{eq:168.311} 195 \\ 109.268 \\ 168.759 \\ 195 \\ 109.716 \\ 169.207 \\ 195 \\ 110.164 \\ 169.655 \\ 195 \\ 111.061 \\ 169.655 \\ 195 \\ 111.509 \\ 111.500 \\ 111.5$ \170.104\195\111.061\170.552\195 0008,1150 >Referenced SOP Class UID: 1.2.840.10008.5.1.4.1.1.2 0008,1155 >Referenced SOP Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.118010 3006,0042 ---: CLOSED_PLANAR 3006.0046 ---: 13 $3006,0050 \quad ---: \quad 117.336 \\ 172.345 \\ 195 \\ 116.439 \\ 172.345 \\ 195 \\ 115.543 \\ 172.345 \\ 195 \\ 114.646 \\ 172.345 \\ 195 \\ 114.198 \\ 1$ $\label{eq:111.896} 114.198 \\ 171 \\ 195 \\ 114.646 \\ 170.552 \\ 195 \\ 115.543 \\ 170.552 \\ 195 \\ 116.439 \\ 170.552 \\ 195 \\ 116.888 \\ 171 \\ 171 \\ 195 \\ 116.888 \\ 195 \\ 115 \\ 195 \\ 105 \\ 1$ $\label{eq:195} \\ 117.336 \\ 171.448 \\ 195 \\ 117.784 \\ 171.896 \\ 195 \\ 117.336 \\ 172.345 \\ 195 \\ 195 \\ 117.336 \\ 172.345 \\ 195 \\ 105$ 0008,1150 >Referenced SOP Class UID: 1.2.840.10008.5.1.4.1.1.2 0008,1155 >Referenced SOP Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.118011 3006,0042 ---: CLOSED_PLANAR 3006,0046 ---: 79 3006,0050 ---: 120.025\174.138\196\119.129\174.138\196\118.232\174.138\196\117.784\173.689\196\117.336 \172.345\196\112.854\172.345\196\111.957\172.345\196\111.509\171.896\196\111.061\171.448\196\110.164 \171.448\196\109.268\171.448\196\108.371\171.448\196\107.923\171\196\107.475\170.552\196\106.578\170.552 \196\106.13\171\196\105.682\171.448\196\104.785\171.448\196\104.337\171\196\103.889\170.552\196\103.44 \170.104\196\102.992\169.655\196\102.544\169.207\196\102.096\168.759\196\101.647\168.311\196\101.647 \165.173\196\104.337\164.725\196\104.785\164.276\196\105.233\163.828\196\105.682\163.38\196\106.13 $\label{eq:162.932} $$ 162.483 196 107.026 162.035 196 107.475 161.587 196 108.371 161.587 196 109.268 162.483 196 107.026 162.035 196 107.475 161.587 196 108.371 161.587 196 109.268 162.483 196 199.268 162.483 196 199.268 162.483 196 199.268 162.483 196 199.268 162.483 196 199.268 162.483 196 199.268 162.483 196 199.268 162.483 196 199.268 162.483 196 199.268 162.483 196 199.268 19$ $\label{eq:164.276} \end{tabular} 1164.276 \end{tabular} 111.509 \end{tabular} 165.621 \end{tabular} 111.509 \end{tabular} 111.509$ \167.862\196\110.612\168.311\196\110.612\169.207\196\111.061\169.655\196\111.957\169.655\196\112.854 $\label{eq:169.655} $$ 196 113.302 169.207 196 113.75 168.759 196 114.198 168.311 196 114.646 167.862 196 115.095 $$ 168.55 196 115.005 $$ 168.55 196 115.005 $$ 168.55 196 115.005 $$ 168.55 115.005 $$ 168.55 115.005 $$ 168.55 196 115.005 $$ 168.55 115.005 $$ 168.55 115.005 $$ 168.55 115.005 $$ 168.55 115.005 $$ 168.55 115.005 $$ 168.55 115.005 $$ 168.55 115.005 $$ 168.55 115.005 $$ 168.55 115.005 $$ 168.55 115.005 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$ 168.55 150 $$$ \168.311\196\115.543\168.759\196\115.991\169.207\196\116.439\169.655\196\117.336\169.655\196\117.784 \170.104\196\118.232\170.552\196\118.681\171\196\119.129\171.448\196\119.577\171.896\196\120.025 \172.345\196\120.474\172.793\196\120.474\173.689\196\120.025\174.138\196 . . . 0008,1150 >Referenced SOP Class UID: 1.2.840.10008.5.1.4.1.1.2 0008,1155 >Referenced SOP Instance UID: 1.2.826.0.1.3680043.8.274.1.1.8323329.18437.1486514204.118028 3006,0042 ---: CLOSED_PLANAR 3006,0046 ---: 9 3006,0050 ---: 110.164\175.034\213\109.268\175.034\213\108.819\174.586\213\108.819\173.689\213\109.268 $\label{eq:173.241} $$ 110.164 173.241 213 110.612 173.689 213 110.612 174.586 213 110.164 175.034 213 110.612 174.586 110.164 175.034 110.612 110.61$ 3006,0084 ---: 1 3006,0082 >---: 1

3006,0084 >---: 1 3006,0085 >---: Model1_9_

ANEXO G – ARTIGO PUBLICADO NO PERIÓDICO J DIGIT IMAGING (2016) 29:716–729

J Digit Imaging (2016) 29:716–729 DOI 10.1007/s10278-016-9894-9

CrossMark

Cloud-Based NoSQL Open Database of Pulmonary Nodules for Computer-Aided Lung Cancer Diagnosis and Reproducible Research

José Raniery Ferreira Junior¹ • Marcelo Costa Oliveira¹ • Paulo Mazzoncini de Azevedo-Marques²

Published online: 20 July 2016 © Society for Imaging Informatics in Medicine 2016

Abstract Lung cancer is the leading cause of cancer-related deaths in the world, and its main manifestation is pulmonary nodules. Detection and classification of pulmonary nodules are challenging tasks that must be done by qualified specialists, but image interpretation errors make those tasks difficult. In order to aid radiologists on those hard tasks, it is important to integrate the computer-based tools with the lesion detection, pathology diagnosis, and image interpretation processes. However, computer-aided diagnosis research faces the problem of not having enough shared medical reference data for the development, testing, and evaluation of computational methods for diagnosis. In order to minimize this problem, this paper presents a public nonrelational document-oriented cloud-based database of pulmonary nodules characterized by 3D texture attributes, identified by experienced radiologists and classified in nine different subjective characteristics by the same specialists. Our goal with the development of this database is to improve computer-aided lung cancer diagnosis and pulmonary nodule detection and classification research through the deployment of this database in a cloud Database as a Service framework. Pulmonary nodule data was provided by the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), image descriptors were

José Raniery Ferreira Junior jrfj@ic.ufal.br

¹ Lab of Telemedicine and Medical Informatics, University Hospital Prof. Alberto Antunes, Institute of Computing, Federal University of Alagoas, Av. Lourival Melo Mota, Cidade Universitária, 57072-900 Maceió, Alagoas, Brazil

² Center of Imaging Sciences and Medical Physics, Internal Medicine Department, Ribeirão Preto Medical School, University of São Paulo, Av. dos Bandeirantes, Monte Alegre, Ribeirão Preto, São Paulo, Brazil acquired by a volumetric texture analysis, and database schema was developed using a document-oriented Not only Structured Query Language (NoSQL) approach. The proposed database is now with 379 exams, 838 nodules, and 8237 images, 4029 of them are CT scans and 4208 manually segmented nodules, and it is allocated in a MongoDB instance on a cloud infrastructure.

Keywords Lung cancer · Pulmonary nodule · Lung Image Database Consortium · Image Database Resource Initiative · Computer-aided diagnosis · Computer-aided detection · 3D texture analysis · NoSQL · Document-oriented nonrelational database · MongoDB · Cloud computing · Database as a Service · Reproducible research

Abbreviations

3DTA	3D texture attributes
API	Application programming interface
BSON	Binary JavaScript Object Notation
CAD	Computer-aided diagnosis
CBIR	Content-based image retrieval
CR	Computed radiography
CT	Computed tomography
DBaaS	Database as a Service
DBMS	Database management system
DICOM	Digital Imaging and Communications in Medicine
DX	Digital radiography
FDA	Food and Drug Administration
IDRI	Image Database Resource Initiative
GLCM	Gray-level co-occurrence matrix
GUI	Graphical user interface
HIPAA	Health Insurance Portability and Accountability Act
IP	Internet protocol
JSON	JavaScript Object Notation

181

🖄 Springer

ANEXO H – ARTIGO PUBLICADO NO PERIÓDICO J SUPERCOMPUT (2017) 73:3451–3467

J Supercomput (2017) 73:3451–3467 DOI 10.1007/s11227-016-1818-4



Integrating 3D image descriptors of margin sharpness and texture on a GPU-optimized similar pulmonary nodule retrieval engine

José Raniery Ferreira Junior¹ · Marcelo Costa Oliveira¹ · Paulo Mazzoncini de Azevedo-Marques²

Published online: 20 July 2016 © Springer Science+Business Media New York 2016

Abstract Due to the difficulty to diagnose lung cancer, it is important to integrate content-based image retrieval methods with the pulmonary nodule classification process, since they are capable of retrieving similar cases from large image databases that were previously diagnosed. The goal of this paper is to evaluate an integrated image feature vector, composed of 3D attributes of margin sharpness and texture, on similar pulmonary nodule retrieval, and to optimize the runtime of nodule comparison process with a graphics processing unit (GPU). Retrieval efficiency was evaluated on the ten most similar cases, on different multiprocessor architectures. Results showed that integrated attributes obtained higher efficiency on similar nodule retrieval, with an increase of up to 2.6 percentage points compared to isolated margin sharpness and texture descriptors. Results also showed that GPU increased nodule retrieval performance with a speedup of $23.7 \times$ on nodule comparison runtime.

Keywords Pulmonary nodule · Content-based image retrieval · Parallel computing · Graphics processing unit

José Raniery Ferreira Junior jrfj@ic.ufal.br

Marcelo Costa Oliveira oliveiramc@ic.ufal.br

Paulo Mazzoncini de Azevedo-Marques pmarques@fmrp.usp.br

Lab of Telemedicine and Medical Informatics, University Hospital Prof. Alberto Antunes, Institute of Computing, Federal University of Alagoas, Maceió, Alagoas, Brazil

² Center of Imaging Sciences and Medical Physics, Internal Medicine Department, Ribeirao Preto Medical School, University of São Paulo, Ribeirão Preto, São Paulo, Brazil