



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA EM TELEINFORMÁTICA
DOUTORADO EM ENGENHARIA DE TELEINFORMÁTICA

VÍTOR ALENCAR DE MESQUITA

**EMPREGO DE EQUAÇÕES BOOLEANAS E VETOR DE FILTROS NA DETECÇÃO
DE NÓDULOS PULMONARES EM TOMOGRAFIAS COMPUTADORIZADAS DO
TÓRAX**

FORTALEZA

Maio - 2022

VÍTOR ALENCAR DE MESQUITA

EMPREGO DE EQUAÇÕES BOOLEANAS E VETOR DE FILTROS NA DETECÇÃO DE
NÓDULOS PULMONARES EM TOMOGRAFIAS COMPUTADORIZADAS DO TÓRAX

Tese apresentada ao Programa de Pós-Graduação em Engenharia em Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Engenharia de Teleinformática. Área de Concentração: Processamento de Sinais e Imagens

Orientador: Prof. Dr. Paulo César Cortez

Coorientador: Prof. Dr. Victor Hugo C. de Albuquerque

FORTALEZA

Maio - 2022

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

M546e Mesquita, Vítor.
Emprego de Equações Booleanas e Vetor de Filtros na Detecção de Nódulos Pulmonares em Tomografias Computadorizadas do Tórax / Vítor Mesquita. – 2022.
78 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2022.

Orientação: Prof. Dr. Paulo César Cortez.

Coorientação: Prof. Dr. Víctor Hugo Costa de Albuquerque.

1. IoT na Medicina. 2. Sistemas de Assistência Médica. 3. Inteligência Artificial. 4. Nódulos Pulmonares. I. Título.

CDD 621.38

Dedico esta tese à memória do meu grande amigo Vinícius Bezerra Braga. Que sua alma descanse em paz.



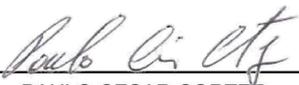
ATA DA SESSÃO DE DEFESA DE TESE DE DOUTORADO

UNIVERSIDADE FEDERAL DO CEARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

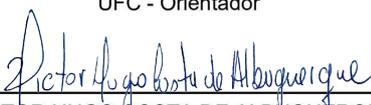
Como parte das exigências para concessão do grau de doutor, às 08:30 horas do dia 31 de Maio de 2022, realizou-se a sessão pública da defesa de tese de doutorado do aluno VÍTOR ALENCAR DE MESQUITA. O trabalho tinha como título: "Emprego de Equações Booleanas e Vetor de Filtros na Detecção de Nódulos Pulmonares em Tomografias Computadorizadas do Tórax".

Compunham a banca examinadora os professores(as) doutores(as) PAULO CESAR CORTEZ, orientador, JOSE MARQUES SOARES, VICTOR HUGO COSTA DE ALBUQUERQUE, ALEXANDRE AUGUSTO DA PENHA COELHO, AUZUIR RIPARDO DE ALEXANDRIA, JOÃO ALEXANDRE LÔBO MARQUES, PLÁCIDO ROGÉRIO PINHEIRO e WAGNER COELHO DE ALBUQUERQUE PEREIRA. O candidato expôs oralmente a tese, em seguida os membros da banca procederam à arguição, e a sessão foi finalizada com a APROVAÇÃO, por parte da banca examinadora, do trabalho sem ressalvas.

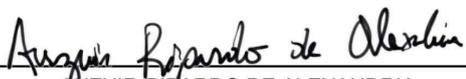
Foi lavrada a presente ata que é abaixo assinada pelos membros da referida banca:



PAULO CESAR CORTEZ
UFC - Orientador



VICTOR HUGO COSTA DE ALBUQUERQUE
UFC - Coorientador



AUZUIR RIPARDO DE ALEXANDRIA
.IFCE - Examinador Externo à Instituição



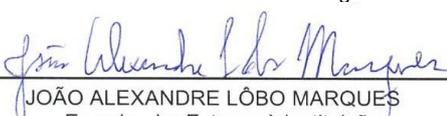
PLÁCIDO ROGÉRIO PINHEIRO
UNIFOR - Examinador Externo à Instituição



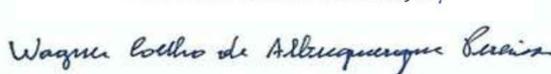
JOSE MARQUES SOARES
UFC - Examinador Interno



ALEXANDRE AUGUSTO DA PENHA COELHO
UFC - Examinador Externo ao Programa



JOÃO ALEXANDRE LÔBO MARQUES
- Examinador Externo à Instituição/



WAGNER COELHO DE ALBUQUERQUE
PEREIRA
UFRJ - Examinador Externo à Instituição

AGRADECIMENTOS

A Deus.

A meus pais, por estarem sempre presentes, por seu amor incondicional e por nunca deixarem de nos apoiar... por tudo que eu tenha conquistado ou vier um dia a conquistar, serei sempre grato a vocês.

Ao meu orientador, Paulo César Cortez pelo seu apoio e confiança mesmo diante dos momentos de dificuldade.

Ao meu co-orientador Victor Hugo Albuquerque, pelos seus encaminhamentos e sua grande boa vontade em ajudar, fundamentais para a finalização do trabalho.

Aos meus amigos de laboratório Alyson Bezerra, por sua contribuição imprescindível para a realização do trabalho, e Igor Valente, pelo exemplo de superação frente as adversidades.

Às minhas irmãs Marta e Ana Luísa, à minha comadre Vanuza, aos meus companheiros de apartamento Eriberto e Bruno, ao meu cunhado Anderson, ao meu sobrinho Nicolás, a meus familiares e amigos, por retribuírem com companheirismo e paciência meu mau humor no período de realização deste trabalho. E fora dele.

A meus avós, sempre presentes, vivos ou já falecidos.

A Angélica, por sua ajuda providencial.

A todos os Professores que tive na vida.

À Universidade Federal do Ceará.

Muito obrigado.

RESUMO

Nesta tese, é apresentado um novo método que utiliza um Vetor de Filtros de Pré-processamento, combinado com equações relacionais e Booleanas simples para detecção de nódulos pulmonares. Para isolar nódulos de outras estruturas pulmonares, é proposto um modelo com 16 filtros de janelas medianas multi-escala, limiares baseados em métricas estatísticas e operações morfológicas 3-D. Na etapa Redução Booleana de Falsos Positivos, as equações relacionais e booleanas selecionam candidatos a nódulos oriundos dos filtros de pré-processamento, por meio de um descritor com 20 atributos ora proposto. Finalmente, uma Rede Neural Convolutiva (CNN) classifica as estruturas restantes como sendo nódulos ou não. Esse método atingiu 92,75% de sensibilidade para uma média de 8 falsos positivos por exame usando todos os exames da base pública LIDC (*Lung Image Database Consortium*) com o parâmetro de espessura do *slice* menor do que 2 mm, contendo qualquer lesão maior do que 3 mm marcada como “nódulo” por pelo menos 3 radiologistas. Para uma referência menos consensual, esse método atinge o maior nível de sensibilidade entre os resultados de outros trabalhos encontrados na literatura pesquisada. Mesmo adotando uma “verdade de campo” que dificulta a detecção de nódulos e a redução de falsos positivos, os resultados estão de acordo com os padrões recentes da literatura mostrando que o método é capaz de detectar nódulos pulmonares de modo eficaz. As principais contribuições desta tese são o Vetor de Filtros de Pré-processamento, o método de Redução Booleana de Falsos Positivos e a inclusão nos métodos de lesões desconsideradas por padrões de referência que consideram apenas a opinião da maioria dos especialistas.

Keywords: IoT na Medicina. Sistemas de Assistência Médica. Inteligência Artificial. Nódulos Pulmonares.

ABSTRACT

Detection of lung cancer in the early stages significantly increases the chances of survival. This work presents a novel method that uses a Vector of Pre-processing Filters combined with simple relational and Boolean equations for pulmonary nodule detection. To isolate nodules from other lung structures, we propose a 16 filter scheme endowed with multiscale median masks, statistical-based thresholds, and 3-D morphological operations. In the Boolean False Positive Reduction stage, relational and Boolean equations select nodule candidates from pre-processing filters using our descriptor with 20 attributes. Finally, a Convolutional Neural Network (CNN) classifies the remaining structures into nodule or no nodule. The method reached 92.75% sensitivity for an average of 8 false positives per exam using all exams in the public Lung Image Database Consortium (LIDC) with a slice thickness of less than 2 mm that contains lesions larger than 3 mm marked “nodule” by at least 3 radiologists. For a less consensual reference, our method reaches the highest sensitivity levels among listed results. Even adopting a “ground truth” which hinders nodule detection and false positive reduction tasks, our results are in line with recent literature standards showing the method is able to detect lung nodules effectively. The major contributions of this thesis are the inclusion in the methods of lesions disregarded by reference standards that consider only the opinion of the majority of specialists.

Keywords: Health of things. Healthcare Systems. Artificial Intelligence. Lung Nodules.

LISTA DE FIGURAS

Figura 1 – Exemplos de elementos estruturantes em formato vertical, de cruz e horizontal. Fonte: <i>Digital Image Processing</i> (GONZALEZ; WOODS, 2018).	25
Figura 2 – Exemplo da operação de <i>Maxpooling</i> com um <i>kernel</i> de dimensões 2x2 em uma matriz de tamanho 4x4.	28
Figura 3 – Exemplo de módulos de uma hierarquia espacial para visão do mundo: detalhes, como pequenas arestas, que se combinam em objetos locais, como olhos ou orelhas, que se combinam em conceitos de alto nível, como “gato” (<i>cat</i> na fonte original). Fonte: <i>Deep Learning with Python</i> (CHOLLET, 2017)	29
Figura 4 – Panorama geral do método de detecção de nódulos.	32
Figura 5 – Fluxograma do algoritmo de segmentação.	33
Figura 6 – Escolha do limiar inicial: –500 Unidades Hounsfield (<i>HU</i>).	34
Figura 7 – Imagem ruidosa.	35
Figura 8 – Imagem filtrada.	35
Figura 9 – Quadros laterais esquerdo e direito.	38
Figura 10 – Fluxograma do preenchimento de orifícios.	39
Figura 11 – Um mesmo contorno para pulmões.	40
Figura 12 – Uma região contendo dois pulmões.	40
Figura 13 – Resultado da divisão da imagem.	40
Figura 14 – Imagem contendo duas regiões.	40
Figura 15 – Um contorno com duas regiões.	40
Figura 16 – Detalhe de um contorno com duas regiões.	40
Figura 17 – Medidas para a divisão de um pulmão na <i>faixa principal</i> mais curta.	41
Figura 18 – Medidas para divisão do pulmão na <i>faixa secundária</i> mais curta.	41
Figura 19 – Slice binário antes da remoção de traqueia.	42
Figura 20 – Slice binário depois da remoção de traqueia.	42
Figura 21 – Contorno irregular.	43
Figura 22 – Contorno suavizado.	43
Figura 23 – Estruturas presentes nos slices mais próximos ao topo do tórax.	44
Figura 24 – Estruturas inalteradas pelo preenchimento de orifícios.	44
Figura 25 – Imagem original de TC.	45

Figura 26 – Imagem após o estágio de suavização.	45
Figura 27 – Resultado depois do estágio de limiarização.	45
Figura 28 – Resultado do processamento morfológico.	45
Figura 29 – Vetor de Filtros de Pré-processamento formado por seus 3 estágios: 4 tipos de suavização, 4 tipos de limiarização e, para cada uma das 16 combinações resultantes, um tipo particular de processamento morfológico	46
Figura 30 – Configurações de janelas em que o <i>voxel</i> central é apagado pelo filtro LUT. .	50
Figura 31 – Imagem antes do pre-processamento.	52
Figura 32 – Resultado com $w = 1$ and $k = 1$	52
Figura 33 – Resultado com $w = 7$ and $k = 1$	52
Figura 34 – Imagem antes do pre-processamento.	53
Figura 35 – Resultado com $w = 1$ and $k = 2$	53
Figura 36 – Resultado com $w = 1$ and $k = 3$	53
Figura 37 – Comprimentos L_x, L_y e L_z de um candidato.	54
Figura 38 – Áreas médias A_1, A_2 e A_3 para cada grupo de seções ortogonais.	56
Figura 39 – Conjuntos N e C	57
Figura 40 – Deslocamento entre as projeções dos centros de massa \hat{c}_A e \hat{c}_V para o plano xy . .	58
Figura 41 – 6 vizinhos conectados com conectividade tridimensional 6.	59
Figura 42 – 8 vizinhos conectados com conectividade bidimensional 8 nos planos xz, yz , ou xy	59
Figura 43 – Arquitetura da CNN empregada no processo de classificação.	67
Figura 44 – Curvas FROC com e sem Redução Booleana de Falsos Positivos.	70

LISTA DE TABELAS

Tabela 1 – Parâmetros do Processamento Morfológico	48
Tabela 2 – Volumes Mínimos	61
Tabela 3 – Volumes Máximos	61
Tabela 4 – Limites Inferiores dos Intervalos Inferiores	62
Tabela 5 – Limites Superiores dos Intervalos Inferiores	62
Tabela 6 – Limites Inferiores dos Intervalos Superiores	62
Tabela 7 – Limites Superiores dos Intervalos Superiores	62
Tabela 8 – Descritor Binário	64
Tabela 9 – Primeira variável Booleana	65
Tabela 10 – Segunda Variável Booleana	65
Tabela 11 – Arquitetura da CNN	66
Tabela 12 – Performance da Redução Booleana de Falsos Positivos	69
Tabela 13 – Resultados para técnicas de detecção automática de nódulos.	71

LISTA DE ABREVIATURAS E SIGLAS

ANODE09	<i>Automatic Nodule Detection 2009</i>
CAD	<i>Computer-aided Detection</i>
CNNs	<i>Convolutional Neural Networks</i>
FNIH	<i>Foundation for the National Institutes of Health</i>
LIDC-IDRI	<i>Lung Image Database Consortium and Image Database Resource Initiative</i>
LUNA16	<i>Lung Nodule Analysis 2016</i>
MLP	<i>Multi-layer Perceptron</i>
NCI	<i>National Cancer Institute</i>
SIFT	<i>Scale-Invariant Features Transform</i>
TC	Tomografía Computadorizada
TP	<i>True Positive</i>

SUMÁRIO

1	INTRODUÇÃO E ESTADO DA ARTE	14
1.1	Objetivos	18
<i>1.1.1</i>	<i>Objetivo Geral</i>	<i>18</i>
<i>1.1.2</i>	<i>Objetivos Específicos</i>	<i>18</i>
1.2	Contribuições	19
1.3	Produção Científica	19
1.4	Organização do Trabalho	19
2	FUNDAMENTAÇÃO TEÓRICA	20
2.1	Limiarização	20
2.2	Crescimento de Regiões	21
2.3	Rastreamento de Contornos	22
2.4	Filtros Espaciais para Suavização	23
2.5	Operações Morfológicas	24
2.6	Redes Neurais Convolucionais	27
3	MATERIAIS E MÉTODOS	31
3.1	Materiais	31
3.2	Segmentação dos Pulmões	33
<i>3.2.1</i>	<i>Limiarização Global</i>	<i>33</i>
<i>3.2.2</i>	<i>Filtragem de Exames Ruidosos</i>	<i>34</i>
<i>3.2.3</i>	<i>Exclusão do Fundo e dos Tecidos de Baixa Densidade</i>	<i>36</i>
<i>3.2.4</i>	<i>Preenchimento de Orifícios</i>	<i>37</i>
<i>3.2.4.1</i>	<i>Extração dos Contornos dos Pulmões</i>	<i>37</i>
<i>3.2.4.2</i>	<i>Divisão da Imagem para Pulmões Separados</i>	<i>38</i>
<i>3.2.4.3</i>	<i>Divisão de Imagens para Pulmões Unidos</i>	<i>39</i>
<i>3.2.5</i>	<i>Remoção da Traqueia</i>	<i>42</i>
<i>3.2.6</i>	<i>Regularização do Contorno</i>	<i>43</i>
3.3	Vetor de Filtros de Pré-processamento para Detecção de Nódulos	44
<i>3.3.1</i>	<i>Suavização</i>	<i>47</i>
<i>3.3.2</i>	<i>Limiarização</i>	<i>47</i>
<i>3.3.3</i>	<i>Processamento Morfológico</i>	<i>47</i>

3.3.3.1	<i>Corrosão de Candidatos</i>	49
3.3.3.2	<i>Restauração de Candidatos</i>	51
3.3.4	<i>Equação do Modelo e seus Parâmetros</i>	52
3.4	Redução Booleana de Falsos Positivos	54
3.4.1	<i>Descritores dos Candidatos</i>	54
3.4.2	<i>Primeiro Estágio de Redução dos Falsos Positivos</i>	60
3.4.3	<i>Segundo Estágio de Redução de Falsos Positivos</i>	63
3.5	Classificação	66
4	RESULTADOS E DISCUSSÕES	69
5	CONCLUSÕES, CONTRIBUIÇÕES E TRABALHOS FUTUROS . . .	74
	REFERÊNCIAS	76

1 INTRODUÇÃO E ESTADO DA ARTE

O câncer de pulmão é a causa mais frequente de morte por câncer no mundo, considerando ambos os sexos, sendo o prognóstico consideravelmente melhor quando a detecção ocorre em suas fases iniciais e o diagnóstico precoce essencial para aumentar as chances de sobrevivência (FERLAY *et al.*, 2019). Por exemplo, nos Estados Unidos, a taxa de sobrevivência relativa em 5 anos é de 57% para pacientes no estágio inicial da doença, mas diminui para 5% para pacientes no estágio metastático. No entanto, apenas uma pequena proporção dos pacientes (16%) é diagnosticada no estágio inicial (SIEGEL *et al.*, 2020).

A Tomografia Computadorizada (TC) do tórax tem sido usada para detectar câncer de pulmão em estágios iniciais na última década, pois é capaz de mostrar pequenos nódulos pulmonares com precisão milimétrica (JACOBS *et al.*, 2014). Tais nódulos são definidos como opacidades arredondadas medindo até 3 cm de diâmetro (HANSELL *et al.*, 2008). Portanto, analisar muitos exames (cada um contendo mais de 200 imagens) é um procedimento laborioso em que erros de percepção podem levar os radiologistas à detecção imprecisa dos nódulos (RENFREW *et al.*, 1992). No intuito de auxiliar os radiologistas no diagnóstico do câncer de pulmão, muitos algoritmos *Computer-aided Detection* (CAD) foram desenvolvidos. Essas técnicas visam destacar regiões de interesse com anormalidades específicas e chamar a atenção dos médicos para tais regiões (VALENTE *et al.*, 2016). Alguns desses algoritmos para detecção de nódulos pulmonares mais recentes têm alcançado alto desempenho quando se considera como parâmetro as diferenças entre as anotações manuais dos próprios especialistas. Apesar disso, devido à heterogeneidade das estruturas anatômicas, tais técnicas se limitam a auxiliar o diagnóstico (HSU *et al.*, 2021).

A detecção automática de nódulos é uma tarefa desafiadora. Separar nódulos justapleurais e justa-vasculares dos tecidos a seu redor é particularmente complexo (ZHANG *et al.*, 2018). Além disso, atributos dos nódulo,s como nível de cinza e textura, podem ser semelhantes aos encontrados em outras estruturas pulmonares, dificultando sua identificação. Para serem considerados úteis pela maioria dos radiologistas, as técnicas de diagnóstico auxiliados por computador precisam possuir alta sensibilidade e quantidade reduzida de falsos positivos (FP) (VALENTE *et al.*, 2016).

A evolução dos algoritmos CAD para detecção de nódulos pulmonares pode ser acompanhada ao longo do tempo por meio de bancos de dados públicos de TC bem caracterizados que provêm uma referência de imagens médicas para a comunidade científica, como o *Lung*

Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) e o *Lung Nodule Analysis 2016 (LUNA16)*. A LIDC-IDRI é uma base de dados que contém 1018 casos, cada um com imagens clínicas de uma tomografia de tórax e anotações feitas por quatro radiologistas de tórax experientes (III *et al.*, 2011). Na base LUNA16, por sua vez, os exames são oriundos da base LIDC-IDRI, sendo, entretanto descartados aqueles que não estejam em conformidade com padrões específicos do LUNA16, o que resulta em 888 exames. Nesta base, nódulos que não tenham sido marcados pela maioria dos especialistas (pelo menos 3 dos 4 radiologistas) são definidos como “achados irrelevantes”, sendo desconsiderados como exemplos positivos no padrão de referência (SETIO *et al.*, 2017).

Pesquisas recentes sobre detecção de nódulos pulmonares alcançaram alta sensibilidade e número reduzido de falsos positivos usando bases de dados públicas completas e fornecendo informações detalhadas sobre o conjunto de dados, como, por exemplo, o nível de consenso entre radiologistas utilizado como referência. Assim, divide-se o estado da arte em dois tipos de trabalho, aqueles que utilizam um consenso majoritário, no qual considera-se como nódulos lesões marcadas pela maioria dos especialistas (pelo menos 3 dentre 4), e os que consideram como nódulo marcações assim anotadas por pelo menos um especialista. Neste contexto, as pesquisas a seguir utilizam padrão de referência de consenso majoritário.

Torres *et al.* (2020) utilizam atributos de diversidade funcional como descritores de textura para detectar câncer de pulmão e reduzir a quantidade de falsos positivos. Este método obtém sensibilidade de 98,4% para 3 FP/exame usando 833 exames da base LIDC e um classificador SVM.

Tan *et al.* (2020) propõem três *Convolutional Neural Networks (CNNs)*, cada uma com diferentes tipos de entrada, para redução de falsos positivos. Para fazer melhor uso da saída oriunda da rede responsável pela segmentação, o método superpõe entradas híbridas a dados não tratados. Esse modelo atinge uma sensibilidade de 97,5% e 1 FP/exame, usando a base LUNA16, com 888 exames e 1186 nódulos.

Zhang *et al.* (2018) usam primeiramente filtros de laplaciano da gaussiana e restrições de forma e tamanho para detectar candidatos a nódulo. Então, desenvolvem uma rede neural convolucional 3D profunda densamente dilatada para identificar nódulos genuínos. Os resultados obtidos para 888 e 1186 lesões de acordo com as regras da base LUNA16 são uma sensibilidade de 94,7% com 1 falso positivo por exame.

Zheng *et al.* (2019) utilizam o mesmo conjunto de dados para explorar a viabilidade

da aplicação de imagens com máxima intensidade de projeção. Neste caso, as CNNs tem a finalidade de melhorar a efetividade da detecção de nódulos. Os autores alcançam sensibilidade de 94,2% e 2 FPs/exame.

Pezeshk *et al.* (2018) apresentam uma técnica de 2 estágios para detecção de nódulos. O primeiro estágio consiste em uma rede neural 3-D completamente convolucional para triagem de candidatos. Já o segundo estágio é definido por um conjunto de CNNs 3-D treinadas, usando transformações aplicadas a *patches* positivos e negativos. Os autores obtêm 91,0% de sensibilidade e 2 falsos positivos por exame. Este método é validado por meio da base LUNA16, com 888 exames e 1186 lesões rotuladas como nódulos.

A literatura pesquisada demonstra um crescente desenvolvimento de plataformas de assistência médica em nuvem capazes de processar remotamente as informações de pacientes. Nesse sentido, a combinação de sistemas baseados em nuvem com técnicas bem estabelecidas na comunidade científica podem produzir sistemas de identificação de nódulos com resultados eficientes e precisos.

Junior *et al.* (2016) criaram um banco de dados não-relacional baseado em nuvem e público para nódulos pulmonares utilizando a base LIDC-IDRI. Tal desenvolvimento visa o provimento de uma estrutura de serviços para integração com CADs e com o ensino de diagnóstico em escolas de medicina. Esta base é alocada em uma instância MongoDB em uma infraestrutura em nuvem e conta com 379 exames, 838 nódulos e 4208 imagens segmentadas manualmente.

Vassallo *et al.* (2019) implementam uma interface baseada em nuvem para armazenar as coordenadas de nódulos maiores do que 3 mm com e sem o auxílio de um sistema CAD, comparando a latência e a sensibilidade em ambos os casos. Os resultados obtidos, usando exames de pacientes com nódulos extra-torácicos, mostram que a interpretação com o sistema CAD melhora a sensibilidade dos radiologistas em 23% .

Masood *et al.* (2020) integram computação em nuvem a uma CNN 3-D profunda para auxiliar as decisões de radiologistas. Este sistema é testado e validado usando as bases LUNA16, *Automatic Nodule Detection 2009* (ANODE09), LIDC-IDRI e conjuntos de dados fornecidos pelo hospital *Shanghai Sixth People's Hospital*. Os resultados obtidos são 98,7% de sensibilidade para 1,97 FPs por exame.

Conforme exposto, a seguir é apresentado o estado da arte que considera como verdadeiros positivos lesões com marcações por pelo menos um especialista. Masood *et al.*

(2020) sugerem uma hibridização de operações morfológicas, realce de pontos baseado na matriz Hessiana, segmentação com conectividade *fuzzy*, algoritmo de máxima densidade local, mapa de distância geodésica e classificação por árvore de regressão para detecção de nódulos pulmonares. Um conjunto de dados contendo 294 exames da base LIDC e 631 nódulos é randomicamente dividido em dois sub-conjuntos independentes: os conjuntos de treino, com 196 exames, e de teste, com 98 exames. Neste caso, os nódulos de referência são anotados por pelo menos dois radiologistas, conforme o padrão LIDC e o resultado dos autores foram 85,2% de sensibilidade e 3,13 FP/exame.

Kuruvilla e Gunavathi (2014) aplicam redes neurais multicamadas com parâmetros estatísticos como média, desvio-padrão, assimetria, curtose, quinto momento central e sexto momento central para classificação de câncer de pulmão. Obtém-se uma sensibilidade de 82,0% e uma média de 2 falsos positivos por exame são validados com 155 exames da base LIDC e tomografias hospitalares, considerando a decisão consensual de pelo menos dois radiologistas como “verdade de campo”.

Guo e Li (2012) apresentam cinco esquemas de CAD para detecção de nódulos em TCs usando informações 3D globais (esquema 3D), informações 2D locais (esquema 2D) e informações locais em conjunto com globais (esquema 2D+3D, 2D–3 esquema D e esquema 3D–2D). Esse trabalho obtém sensibilidade de 80% e 1,6 FP/exame, resultado validado com 85 exames da base LIDC com 111 nódulos na opinião de pelo menos 2 entre 4 radiologistas.

Messay *et al.* (2010) desenvolvem um sistema CAD que combina limiarização e processamento morfológico para detectar e segmentar candidatos a nódulos simultaneamente. Este método calcula um conjunto de 245 atributos para cada candidato segmentado, e um processo de seleção sequencial determina o subconjunto ótimo de atributos para dois classificadores distintos, um classificador discriminante linear de *Fisher* e um classificador quadrático. Primeiramente, o sistema CAD é ajustado e treinado usando um conjunto de dados cortesia da e avaliado usando 84 exames e 143 nódulos do banco de dados LIDC. Este CAD mescla as listas de padrão-ouro de todos os quatro radiologistas para obter a referência conforme a qual os autores avaliam seu sistema. Assim, qualquer lesão marcada por pelo menos um dos quatro radiologistas é considerada um nódulo, tendo o método alcançado sensibilidade de 82,0% para 3 FP por exame.

Muitos trabalhos na literatura pesquisada optam por considerar apenas nódulos marcados pela maioria dos radiologistas (pelo menos 3 dentre 4) como exemplos positivos no padrão de referência. No entanto, isso pode ser uma desvantagem, pois lesões anotadas por pelo

menos 1 ou 2 radiologistas também podem ser nódulos que representam risco para a saúde dos pacientes. Nesta tese, tais achados são considerados relevantes para exames mais específicos, sendo a ausência de lesões apontadas como verdadeiros positivos pela minoria dos especialistas para a identificação de nódulos. Isto também é um problema que visa-se superar nesta tese. Entretanto, tal melhoria deve ser alcançada mantendo-se a possibilidade de comparar de modo isento os achados desta tese com os de outros trabalhos que consideram padrões-ouro diferentes.

Nesse sentido, propõe-se se um novo método que considera tais nódulos de modo a propiciar uma comparação justa com a literatura pesquisada e, ao mesmo tempo, não demande cálculos tão complexos como nas técnicas existentes. Este método consiste em um pré-processamento que utiliza técnicas de processamento de imagens básicas (suavização, limiarização e processamento morfológico) para isolamento de nódulos e uma técnica de redução de falsos-positivos baseado em equações relacionais e Booleanas simples. Consideram-se ainda neste método três tipos de padrão-ouro, cujos resultados contabilizam como verdadeiros positivos anotações não consideradas em muitos trabalhos.

1.1 Objetivos

1.1.1 Objetivo Geral

O objetivo principal desta tese consiste em propor um método de detecção de nódulos pulmonares em imagens de tomografia computadorizada que considere diversos tipos de padrão de referência.

1.1.2 Objetivos Específicos

Os objetivos específicos desta tese são:

- conceber um modelo de pré-processamento capaz de isolar nódulos de outras estruturas independentemente do nível de consenso do padrão-ouro utilizado;
- desenvolver um método de formulação algébrica simples, em relação ao estado da arte, que reduza a taxa média de falsos positivos por exame sem que haja comprometimento significativo da sensibilidade;
- propor um conjunto de atributos capaz de descrever representativamente os nódulos pulmonares; e
- buscar resultados de classificação superiores aos encontrados na literatura pesquisada

considerando o padrão de referência em que a opinião de apenas um especialista é suficiente para que uma lesão seja considerada como nódulo.

1.2 Contribuições

Esta tese se concentra na identificação de nódulos, e as principais contribuições são as seguintes:

- uma técnica de segmentação dos pulmões em tomografias para posterior detecção de nódulos;
- um Vetor de Filtros de Pré-processamento que isola nódulos de vários tamanhos e consistências de outras estruturas pulmonares e
- um método de redução de falsos positivos que seleciona candidatos com base em equações relacionais e Booleanas para classificação de nódulos posterior.

1.3 Produção Científica

O desenvolvimento desta tese possibilitou a publicação do artigo intitulado *A novel method for lung nodule detection in computed tomography scans based on Boolean equations and vector of filters techniques* na revista *Computers and Electrical Engineering*, disponível na base *Science Direct* em <<https://doi.org/10.1016/j.compeleceng.2022.107911>>, classificado com A2 no Qualis CAPES.

1.4 Organização do Trabalho

No Capítulo 2 é provida a fundamentação teórica e a modelagem matemática para os métodos utilizados nesta tese de acordo com a doutrina. No Capítulo 3 é descrita a metodologia utilizada para segmentação automática dos pulmões, o isolamento dos nódulos com o Vetor de Filtros de Pré-processamento, o método de Redução Booleana de Falsos Positivos e classificação por meio de uma CNN, provendo detalhes sobre o conjunto de dados, o aparato de hardware e as ferramentas computacionais utilizadas. Já no Capítulo 4 são discutidos os principais resultados desta tese por meio de uma comparação crítica com aqueles encontrados em outros estudos relacionados. Por fim, são apresentadas no Capítulo 5 as conclusões e as propostas para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Neste Capítulo são abordados os aspectos teóricos básicos que fundamentam a metodologia ora empregada na detecção de nódulos pulmonares. Os principais tópicos abordados são: Limiarização, Crescimento de Regiões, Rastreamento de Contornos, Filtros Espaciais para Suavização, Operações Morfológicas e Redes Neurais Convolucionais.

2.1 Limiarização

Devido a suas propriedades intuitivas, simplicidade de implementação e desempenho computacional, a limiarização possui uma importância central em aplicações para processamento de imagens, nomeadamente a segmentação. Este processo é empregado na divisão de imagens em diferentes regiões, utilizando como critério a intensidade de cinza em tais regiões (GONZALEZ *et al.*, 2004).

O uso da limiarização revela-se vantajoso em imagens compostas por objetos luminosos em uma cena escura, ou, de modo mais geral, quando o nível de cinza de *pixels* diferentes regiões se distribuem ao redor de valores dominantes (ou modas) no domínio da intensidade.

Nesse sentido, seja uma imagem $f(x,y)$ composta por regiões claras e escuras de modo que a intensidade dos *pixels* dessas regiões se agrupam ao redor de duas modas. Um procedimento intuitivo para destacar objetos consiste na escolha de um valor de limiar L que separe estas modas. Desse modo, qualquer ponto da imagem $f(x,y)$ em que $f(x,y)$ seja maior que L é identificado como região de interesse, enquanto regiões nas quais esta condição não seja obedecida são identificadas como fundo. Em outras palavras, a imagem $g(x,y)$ é denotada por (GONZALEZ; WOODS, 2018)

$$g(x,y) = \begin{cases} 1, & \text{se } f(x,y) > L, \\ 0, & \text{se } f(x,y) \leq L. \end{cases} \quad (2.1)$$

Quando L é um valor constante aplicável sobre uma imagem inteira, não variando conforme as coordenadas dos *pixels* ou ao longo das regiões da imagem, o processo descrito na Equação 2.1 é chamado de limiarização global (GONZALEZ; WOODS, 2018). Nesse caso, define-se um procedimento de escolha para o valor de L conforme os passos a seguir (GONZALEZ; WOODS, 2018):

- selecionar um valor inicial estimado para o limite global L ;

- classificar todos os *pixels* em dois grupos, o fundo, com nível de cinza menor que L , e o objeto, caso contrário;
- calcular o valor médio de intensidade μ_f para *voxels* de fundo e μ_o para *voxels* de objetos;
- calcular o novo limite $L = \frac{\mu_f + \mu_o}{2}$; e
- repetir a segunda, terceira e quarta etapas enquanto a diferença entre os limites calculados sucessivamente for maior que o parâmetro de tolerância previamente definido.

2.2 Crescimento de Regiões

O crescimento de regiões é um procedimento que agrupa *voxels* ou sub-regiões em regiões maiores, definindo-se sementes e anexando-se *pixels* vizinhos a estas sementes com base em critérios de semelhança como nível de cinza, cor ou outros atributos (GONZALEZ; WOODS, 2018).

O funcionamento geral do algoritmo de crescimento de regiões consiste em identificar com diferentes rótulos regiões em imagens que possuam *voxels* de igual valor e que estejam conectadas segundo um critério de conectividade.

Um *pixel* p nas coordenadas (x, y) tem dois vizinhos horizontais e dois verticais com coordenadas $(x + 1, y)$, $(x - 1, y)$, $(x, y + 1)$ e $(x, y - 1)$. Este conjunto de *pixels*, chamados de vizinhos de p de conectividade 4, é denotado por $N_4(p)$. Os quatro vizinhos diagonais de p têm coordenadas $(x + 1, y + 1)$, $(x + 1, y - 1)$, $(x - 1, y + 1)$ e $(x - 1, y - 1)$ são denotados por $N_D(p)$. Os vizinhos diagonais, juntamente com os vizinhos de conectividade 4, são chamados de vizinhos de p de conectividade 8, denotados por $N_8(p)$ (GONZALEZ; WOODS, 2018).

No caso de uma imagem tridimensional, cada *voxel* v possui 6 vizinhos $N_6(v)$, ou seja, aqueles que se encontram à direita, à esquerda, acima, abaixo, à frente ou atrás do *voxel* em questão. Semelhantemente aos vizinhos $N_4(p)$ de um *pixel* p em um espaço bidimensional, os vizinhos $N_6(v)$ de um *voxel* v em 3 dimensões estão a uma distância euclidiana de, no máximo, 1 *voxel* do *voxel* v . Em três dimensões, os vizinhos diagonais de um *voxel* v , juntamente com o conjunto $N_6(v)$ formam os conjuntos $N_{18}(v)$ e $N_{26}(v)$ de vizinhos de v com conectividade 18 e 26, os quais se encontram a uma distância de v de, no máximo, $\sqrt{2}$ e $\sqrt{3}$ *voxels*, respectivamente.

Um dos algoritmos utilizados para realizar o crescimento de regiões é capaz de identificar com diferentes rótulos regiões conectadas com conectividade $N_6(v)$, para regiões bidimensionais, ou $N_6(v)$, $N_{18}(v)$ ou $N_{26}(v)$, para regiões tridimensionais, usando o seguinte procedimento (SEDGEWICK, 1998):

- rastrear todos os *voxels* diferentes de zero, registrando todos os rótulos equivalentes em uma tabela de união-busca;
- resolver equivalências de classes, usando um algoritmo de união-busca; e
- renomear todos os *voxels* de acordo com as novas classes equivalentes.

2.3 Rastreamento de Contornos

É possível perceber que diversas técnicas em Processamento de Imagens necessitam de uma descrição precisa dos contornos das regiões de interesse. Por exemplo, para eliminar orifícios indesejáveis no interior de um objeto em uma imagem digital, atribuímos valores iguais a 1 a todos os *pixels* localizados no interior de seu contorno. Para tanto, necessita-se de um algoritmo que extraia os contornos desses objetos de modo a representá-los por meio das coordenadas, de seus pontos ordenadas no sentido horário ou anti-horário, ou seja, um algoritmo de rastreamento de contornos.

A extração de coordenadas para cada contorno em uma imagem binária pode se processar por meio do algoritmo de rastreamento de contorno *Moore-Neighbor* modificado pelo critério de parada desenvolvido por *Jacob* (GONZALEZ *et al.*, 2004). No sentido de introduzir este algoritmo de rastreamento de contorno assume-se que (GONZALEZ; WOODS, 2018):

- as imagens binárias e os pontos do objeto e do fundo são rotulados com 1's e 0's respectivamente; e
- as imagens possuem uma borda de 0's para eliminar a possibilidade de um objeto se fundir às bordas da imagem.

Para maior clareza, limita-se a discussão a uma única região. De todo modo, a abordagem pode ser estendida para várias regiões disjuntas processando-se tais regiões individualmente.

Desse modo, a sequência de passos a seguir rastreia o contorno ordenado em sentido horário de uma região formada por *pixels* de valores iguais 1 em uma imagem binária (GONZALEZ; WOODS, 2018):

1. seja b_0 o ponto rotulado com 1 mais acima e mais à esquerda na imagem e seja c_0 o vizinho à esquerda de b_0 , em que c_0 é sempre pertencente ao fundo, o algoritmo examina os 8 vizinhos de b_0 , começando por c_0 e seguindo no sentido horário.
2. define-se b_1 como o primeiro vizinho encontrado com valor igual a 1, e c_1 como o ponto pertencente ao fundo imediatamente anterior a b_1 na sequência, armazenando-se a

localização de b_0 para uso na Etapa 5.

3. atribui-se $b = b_0$ e $c = c_0$;
4. denota-se por $n_1, n_2, n_3 \dots n_8$ os vizinhos de b de conectividade 8, começando por c e seguindo no sentido horário. Em seguida, define-se como n_k o primeiro vizinho rotulado com 1 encontrado seguindo a ordem estabelecida;
5. atribui-se $b = n_k$ e $c = n_{k-1}$;
6. os passos 4 e 5 são repetidos até que $b = b_0$, de modo que a sequência de pontos b encontrados até o atendimento do critério de parada do algoritmo define os pontos ordenados do contorno.

2.4 Filtros Espaciais para Suavização

Em processamento de imagens, a suavização de imagens por meio da utilização de filtros espaciais pode ter diversas finalidades. Pode-se elencar como aplicações comuns deste tipo de filtragem o borramento de transições bruscas de nível de cinza da imagens e a redução de ruído. A suavização também é utilizada quando se visa eliminar detalhes indesejáveis nas imagens, bastando-se para isso que se escolha uma janela de filtragem de dimensões maiores em comparação àquelas dimensões dos detalhes em questão (GONZALEZ; WOODS, 2018).

A filtragem espacial linear para suavização é um tipo de filtragem que consiste na convolução de uma imagem com uma janela ou núcleo de convolução capaz de suavizar a imagem. Neste caso, a convolução implica em um borramento da imagem, cuja intensidade depende das dimensões do núcleo de convolução e de seus coeficientes.

Um dos filtros espaciais de suavização mais comumente utilizado em processamento de imagem é o Gaussiano. Esse filtro possui algumas propriedades particulares, como apresentar uma resposta independente da orientação, o que também pode ser chamado de simetria circular ou isotropia. Além disso, o filtro Gaussiano possui uma resposta em frequência semelhante ao seu comportamento no espaço. Um núcleo de filtro Gaussiano é definido por (GONZALEZ; WOODS, 2018)

$$w(s, t) = G(s, t) = Ke^{-\frac{s^2+t^2}{2\sigma^2}}, \quad (2.2)$$

em que s e t são as distâncias horizontal e vertical de um dado *pixel* do núcleo até o centro deste núcleo, σ é o desvio padrão da janela Gaussiana e K é um coeficiente de normalização, dado

pela soma dos coeficientes da janela (GONZALEZ; WOODS, 2018). Ao aumentar-se o desvio padrão, intensifica-se o efeito de borramento produzido pela janela, ao passo que o aumento das dimensões da janela expande a área de atuação do filtro bem como a intensidade de seus efeitos.

Diferentemente dos filtros lineares, a resposta dos filtros não-lineares não é uma combinação linear dos *pixels* da imagem que se encontram sobre o núcleo de filtragem. Exemplos conhecidos de filtros espaciais não lineares são o filtro da mediana, bem como outros filtros estatísticos. A resposta de um filtro da mediana é baseada na ordenação dos *pixels* contidos no interior da janela de filtragem. Neste caso, a filtragem é obtida substituindo-se o valor do pixel central pelo valor da mediana dos níveis de cinza nas vizinhanças desse *pixel* (GONZALEZ; WOODS, 2018).

Os filtros da mediana são usados para reduzir efetivamente o ruído do tipo impulsivo, como, por exemplo, o ruído do tipo sal e pimenta e, além disso, produzem consideravelmente menos efeito de borramento quando comparados a filtros espaciais lineares. Este efeito é obtido basicamente forçando-se cada *pixel* a assumir valores semelhantes aos de seus vizinhos, dentro de uma certa região dada pela dimensão da janela de filtragem.

2.5 Operações Morfológicas

Operações morfológicas são importantes ferramentas para extrair das imagens informações úteis na descrição da forma geométrica de seus objetos. Estas operações são definidas em termos de conjuntos de modo que, em processamento de imagens, são utilizados dois tipos de conjuntos de *pixels*: objetos e elementos estruturantes. Como tais conjuntos podem assumir diversas formas e as imagens com as quais se trabalha normalmente são retangulares, tais conjuntos são inseridos em matrizes. Desse modo, os *pixels* pertencentes a estas matrizes que não sejam membros do conjunto do objeto são identificados como *pixels* de fundo (GONZALEZ; WOODS, 2018).

Elementos estruturantes são utilizados nas operações morfológicas de modo similar às janelas empregadas na operação de convolução espacial. Apesar de as operações serem diferentes, as operações de deslizamento e *padding* são essencialmente as mesmas utilizadas na convolução. Desse modo, o elemento estruturante percorre a imagem produzindo sobre o *pixel* da imagem localizado no centro do elemento estruturante um efeito que depende do tipo de operação morfológica realizada. Nesse sentido, existem diversos tipos de elementos estruturantes construídos com finalidades específicas a depender da informação que se deseja

extrair da imagem. Por exemplo, um elemento estruturante circular pode ser utilizado para suavizar as bordas de uma imagem, enquanto um elemento estruturante que seja uma linha horizontal pode ser utilizado para eliminar estreitas linhas horizontais na imagem. A Figura 1 mostra tipos de elementos estruturantes comumente utilizados.

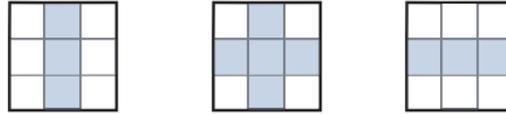


Figura 1 – Exemplos de elementos estruturantes em formato vertical, de cruz e horizontal. Fonte: *Digital Image Processing* (GONZALEZ; WOODS, 2018).

A seguir são introduzidas algumas operações morfológicas elementares usadas frequentemente, como reflexão, translação, dilatação e erosão (GONZALEZ; WOODS, 2018).

Os conceitos de reflexão e translação de conjuntos são amplamente utilizados em conjunto com elementos estruturantes em morfologia. A reflexão de um elemento estruturante B em relação a sua origem é denotada por \hat{B} e dada por

$$\hat{B} = \{w | w = -b, \text{ para } b \in B\}, \quad (2.3)$$

ou seja, se B é um conjunto de pontos em duas dimensões, então \hat{B} é o conjunto de pontos de B cujas coordenadas (x, y) foram trocadas por $(-x, -y)$.

Já a translação $z = (z_1, z_2)$ de B em cada coordenada, denotada como $(B)_z$ é dada por:

$$(B)_z = \{c | c = b + z, \text{ para } b \in B\}. \quad (2.4)$$

Em outras palavras, se B é um conjunto de *pixels* em duas dimensões, então $(B)_z$ é o conjunto de *pixels* em B cujas coordenadas (x, y) foram modificadas para $(x + z_1, y + z_2)$. Esta operação é usada para deslizar um elemento estruturante sobre uma imagem, de modo que em cada localização é realizada uma operação de conjunto entre o elemento estruturante e a área diretamente abaixo. Nesse contexto, tanto a reflexão quanto a translação são definidas em relação à origem de B .

Apresentadas as operações de reflexão e translação, define-se duas operações elementares: a erosão e a dilatação. Dados os conjuntos A e B , a erosão de A por B é dada por

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\}, \quad (2.5)$$

em que A são os *pixels* do objeto, B é o elemento estruturante e z são os valores do *pixels* do conjunto do objeto, ou seja, 1's. Em outras palavras, esta equação indica uma erosão de A por B em todos os pontos z em que a versão transladada de B está contida em A .

Similarmente, dados os conjuntos A e B , a dilatação de A por B é dada por

$$A \odot B = \{z | B_z \subseteq A\}, \quad (2.6)$$

ou seja, a dilatação de A por B é o conjunto de todos os deslocamentos de z em que os elementos de B pertencentes ao objeto interceptam pelo menos um elemento de A .

Um vez definidas as operações de dilatação e erosão, são apresentadas duas importantes operações em processamento morfológico de imagens, a abertura e o fechamento morfológico, as quais derivam das operações de erosão e dilatação ora descritas. De modo geral, a abertura suaviza os contornos de um objeto, rompe faixas estreitas que unem objetos maiores e elimina saliências finas. Por outro lado, o fechamento também tende a suavizar irregularidades em contornos, mas, ao contrário da abertura, geralmente funde objetos próximos ou conectados por faixas estreitas, fecha orifícios e preenche lacunas de contornos. Desse modo, define-se a abertura do conjunto A por um elemento estruturante B como

$$A \circ B = (A \odot B) \oplus B. \quad (2.7)$$

Assim, a abertura de A é simplesmente da erosão de A por B , seguida de uma dilatação do resultado por B . Similarmente, define-se o fechamento de um conjunto A por um elemento estruturante B como

$$A \bullet B = (A \oplus B) \odot B. \quad (2.8)$$

Em outras palavras, o fechamento de A por B consiste na dilatação de A por B , seguida pela erosão do resultado por B .

Além das técnicas mencionadas, uma transformação utilizada para diversas finalidades em Processamento de Imagens é a transformada da distância. Para compreender a transformada da distância, considere-se primeiramente que *pixels* de uma região de fundo são rotulados por 0's e *pixels* de um objeto são rotulados por 1's. Desse modo, a transformada da distância é a distância de um *pixel* até o *pixel* do objeto mais próximo, calculada para cada *pixel* da imagem (GONZALEZ; WOODS, 2018).

2.6 Redes Neurais Convolucionais

As redes neurais convolucionais (CNNs) podem ser utilizadas em problemas de classificação e regressão, produzindo desempenhos superiores às redes neurais multicamadas em grande parte dos casos. Entretanto, nas aplicações de visão computacional e processamento de imagens, este modelo, também conhecido como *convnets*, é atualmente utilizado em aplicações quase que universais (CHOLLET, 2017).

Em um resumo preliminar, pode-se entender uma rede neural convolucional como uma pilha de camadas integradas por fases de *Conv2D* e *Maxpooling*. A saída de cada uma destas camadas (*Conv2d* e *Maxpooling*) são tensores 3-D com dimensões de largura, altura e quantidade de canais. A diferença conceitual entre uma camada densamente conectada de uma rede *Multi-layer Perceptron* (MLP) e uma camada convolucional é que as camadas densas detectam padrões globais no espaço de atributos de suas entradas, enquanto camadas convolucionais são capazes de reconhecer padrões locais, no caso de imagens, padrões encontrados em pequenas janelas bidimensionais (CHOLLET, 2017).

Em uma rede neural convolucional, a saída de uma camada é uma matriz resultante da convolução entre a matriz na entrada desta camada e uma janela de convolução. Por exemplo, na primeira camada convolucional de uma CNN, uma imagem de entrada com 28×28 *pixels* pode ser submetida a convoluções com 32 filtros diferentes. Nesse caso, a saída desta camada seria um conjunto de 32 canais 26×26 (2 *pixels* a menos por dimensão devido ao *padding*), sendo os valores dos coeficientes dos filtros atualizados ao longo do desenvolvimento do aprendizado da rede.

Como é comum em muitas aplicações, quando as imagens de entrada da rede são tridimensionais, a operação da camada de convolução, denotada por *Conv3D*, é uma convolução 3-D de acordo com (HUANG *et al.*, 2017):

$$G(x, y, z) = \sum_{s=-a}^a \sum_{t=-b}^b \sum_{k=-c}^c w(s, t, k) \cdot f(x + s, y + t, z + k), \quad (2.9)$$

em que x , y e z são as coordenadas dos *voxels*, w é a máscara convolucional e f é a imagem 3-D de entrada da camada. Considerando um *kernel* convolucional de dimensões $K \times M \times N$, as variáveis a , b e c são dadas por $a = (K - 1) / 2$, $b = (M - 1) / 2$ e $c = (N - 1) / 2$.

A camada *Maxpooling*, por sua vez, reduz o tamanho da matriz de entrada, respondendo apenas com os máximos locais da imagem local na saída conforme mostrado na Figura 2

(PATTERSON; GIBSON, 2017). Analogamente, uma camada de *Maxpooling3D* é uma versão 3-D do processo ora ilustrado.



Figura 2 – Exemplo da operação de *Maxpooling* com um *kernel* de dimensões 2x2 em uma matriz de tamanho 4x4.

Por exemplo, a saída de dimensões 26x26x32 mencionada nesta secção, oriunda da primeira camada convolucional de uma CNN, pode produzir um resultado de dimensões (13,13,32) ao passar pela camada de *Maxpooling*. Em seguida, este resultado torna-se entrada para outra camada da rede composta por outro bloco de convolução e outro de *Maxpooling* e assim sucessivamente.

Conforme os blocos compostos por *Conv* e *Maxpooling* se superpõem, as dimensões de altura e de largura tendem a encolher ao passo que a rede entra em camadas mais profundas. Esta disposição construtiva torna a rede neural convolucional capaz de reconhecer padrões de diferentes dimensões em uma mesma imagem. Desta forma, uma primeira camada convolucional deve reconhecer padrões locais, como bordas, e uma segunda camada convolucional detecta padrões maiores a partir de características oriundas daquela primeira camada, e assim por diante.

Desse modo, as redes neurais convolucionais são capazes de reconhecer padrões e abstrações visuais cada vez mais complexas ao longo de suas camadas. Tal característica as torna particularmente úteis em processamento de imagens, já que o mundo visual usa padrões de reconhecimento fundamentalmente hierárquicos. Esta discussão é ilustrada na Figura 3.

Com a finalidade de compreender como a CNN realiza a classificação para a qual é projetada, retoma-se o raciocínio em que a rede recebe uma imagem de dimensões 28x28x1 como de entrada. Então, a saída da primeira camada de convolução é um tensor com dimensões 26x26x32, o qual é sub-amostrado após a primeira camada de *Maxpooling*, cuja saída é um tensor de dimensões 13x13x64, já que para cada um dos 32 canais existem 2 combinações da entrada da camada de *Maxpooling* com sua saída sub-amostrada. Assim, as dimensões das imagens quadrangulares presentes em cada canal diminuem de dimensão conforme estas imagens

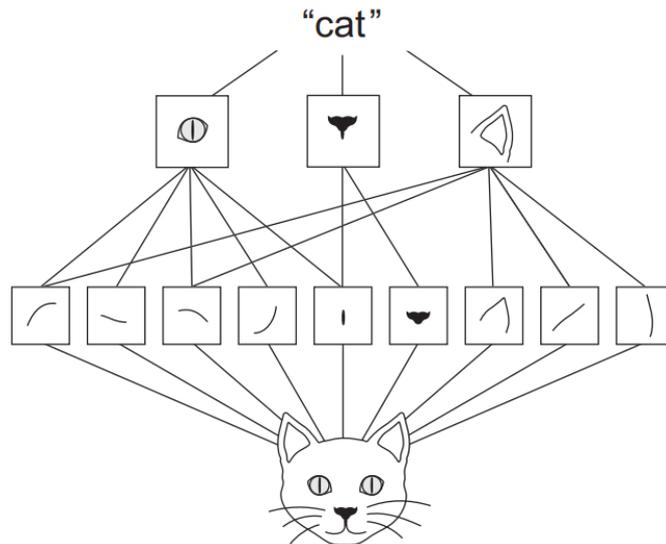


Figura 3 – Exemplo de módulos de uma hierarquia espacial para visão do mundo: detalhes, como pequenas arestas, que se combinam em objetos locais, como olhos ou orelhas, que se combinam em conceitos de alto nível, como “gato” (*cat* na fonte original).
Fonte: *Deep Learning with Python* (CHOLLET, 2017)

avançam pelas camadas da rede. Finalmente, as saídas de cada diferente camada da rede podem ter dimensões 26 (*Conv*), 13 (*Maxpooling*), 11 (*Conv*), 5 (*Maxpooling*) e 3 (*Conv*).

Caso se deseje realizar uma classificação de 10 vias, utiliza-se uma camada final com 10 saídas juntamente com uma função de ativação, por exemplo Softmax ou ReLu. Desta forma, para que o conjunto de atributos tenha sua dimensionalidade reduzida até esse ponto, uma camada *Flatten* transforma o tensor resultante das camadas *Conv* e *Softmax* em um vetor unidimensional. Este vetor passa por uma superposição de camadas densas, tais quais as de uma rede neural multicamadas. Estas camadas densas têm a finalidade de diminuir a quantidade de atributos que são considerados na entrada da função de ativação. Por exemplo, no momento em que o tensor de saída passa a ter dimensões (3, 3, 64), a camada *Flatten* transforma estes coeficientes em um vetor unidimensional com 576 elementos. Em seguida, duas camadas densas podem produzir como saídas vetores de 64 e 10 posições, respectivamente.

As funções de ativação comumente utilizadas nas últimas camadas *Softmax* e *Relu* são dadas por (SETIO *et al.*, 2016)

$$\sigma_i = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \quad (2.10)$$

$$ReLU(i) = \max(0, i). \quad (2.11)$$

Na Equação 2.10, o parâmetro i define o índice do neurônio de saída, k representa o número de neurônios em um determinado nível e z é o vetor de neurônios de saída. Em uma classificação de três grupos, a saída é um vetor de três elementos, em que cada elemento representa o grau de pertinência e sua soma é igual a 1. Já na Equação 2.11, i é a entrada da função de ativação.

Uma vez exposta toda a fundamentação teórica de que se necessita para o entendimento desta tese, são descritos e especificados os materiais e métodos empregados no desenvolvimento do novo método.

3 MATERIAIS E MÉTODOS

Definidos os conceitos fundamentais, são descritos os materiais e a sequência metodológica utilizada no desenvolvimento da técnica de detecção de nódulos pulmonares proposta.

3.1 Materiais

O conjunto de dados utilizado na validação dos métodos ora propostos é a base LIDC-IDRI. A LIDC-IDRI é uma base internacional pública acessível pela web para desenvolvimento, treinamento e avaliação de algoritmos CAD para detecção e diagnóstico de câncer de pulmão. A instituição desta base é uma colaboração entre o *National Cancer Institute* (NCI), a *Foundation for the National Institutes of Health* (FNIH), sete centros acadêmicos e oito empresas de imagens médicas.

A base LIDC é composta por 1018 exames, nos quais, para cada lesão identificada, constam marcações realizadas por quatro radiologistas de tórax. Estas marcações são feitas por meio de um processo de “leitura” de duas fases que identifica, na medida do possível, todos os nódulos pulmonares em cada TC sem exigir um consenso forçado. Em uma primeira leitura (fase de leitura cega), cada radiologista classifica independentemente as lesões como “nódulo com 3 mm ou mais”, “nódulos menores do que 3 mm” ou “não-nódulo”. Na segunda leitura (fase de leitura não-cega), os radiologistas têm a chance de revisar suas marcações de modo independente enquanto examinam as anotações anônimas dos outros três radiologistas e tomar uma decisão final (III *et al.*, 2011).

Nesta tese, todas as tomografias do conjunto de dados do LIDC com espessura de *slice* menor do que 2 mm e contendo nódulos maiores do que 3 mm são utilizadas de acordo com os vários padrões-ouro empregados, ou seja, lesões com consenso de pelo menos 3, 2 ou apenas um radiologista. No total, são 2.287 lesões em 768 exames marcadas por pelo menos um radiologista como nódulo, 1.603 lesões presentes em 694 exames consideradas nódulos por pelo menos 2 radiologistas e 607 exames contendo 1.191 lesões com pelo menos 3 anotações de radiologistas.

Para os resultados de classificação, utiliza-se um *notebook* *Asus Predator* com processador *Intel Core i7*, 7ª geração, placa de vídeo *GeForce GTX 1060* e memória RAM de 16GB DDR-4. A CNN é treinada em um computador *desktop* com processador *Intel Core i7 9700-KF*, placa de vídeo *GeForce RTX 2060 SUPER* e 64 GB de RAM DDR-4. Um *notebook*

com *Intel Core i7 3537U*, *AMD Radeon HD 7570M* e 8GB de memória RAM DDR-3 é utilizado para pré-processamento dos exames, redução de falsos positivos e implementação da classificação *Multilayer Perceptron* (MLP).

Testes de classificação e treinamento são implementados em *Python* usando *Pycharm*. Nesta etapa, a API (*Application User Interface*) *pylidc* é utilizada para a leitura das informações dos nódulos padrão-ouro oriundas da base LIDC. Além disso, o *evaluation script* do *framework* LUNA16 é utilizado para calcular as sensibilidades e produzir as curvas FROC (*Free-Response Operating Characteristic*). As etapas de pré-processamento e redução de falsos positivos são desenvolvidas com o *MATLAB* R2021b. A seguir procede-se à apresentação e discussão dos métodos desenvolvidos.

Apresentados os materiais, segue-se à metodologia adotada, descrita detalhadamente à seguir. O método de detecção de nódulos proposto é composto pelas fases de segmentação dos pulmões, Vetor de Filtros de Pré-processamento para isolamento dos nódulos, Redução Booleana de Falsos Positivos e classificação com uma CNN. A ideia geral do modelo é ilustrada na Figura 4. A seguir, procede-se segmentação para isolar a área de busca por nódulos.

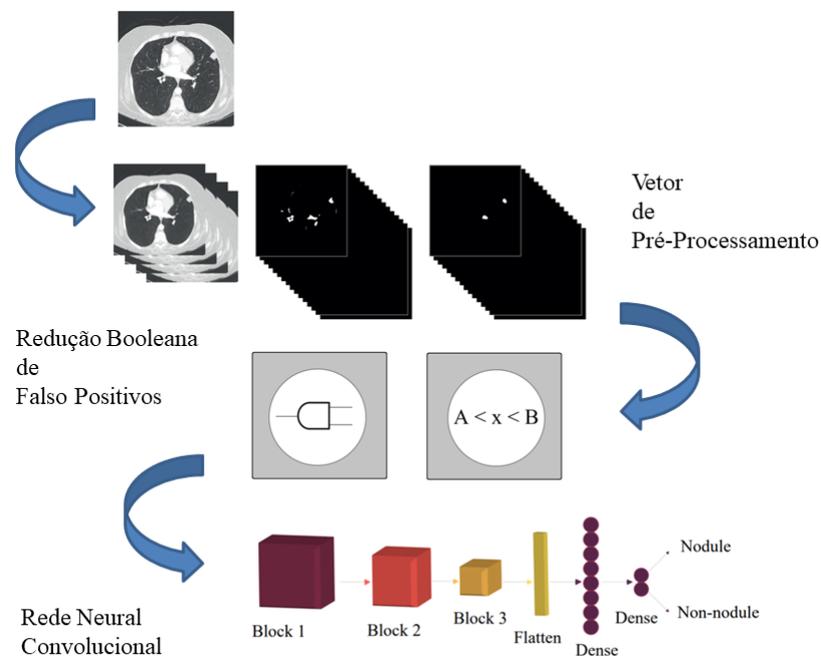


Figura 4 – Panorama geral do método de detecção de nódulos.

3.2 Segmentação dos Pulmões

A segmentação dos pulmões é um procedimento que normalmente antecede a detecção de candidatos a nódulos pulmonares e visa restringir o espaço de busca por esses nódulos às cavidades pulmonares. Nesta tese, é proposta uma técnica de segmentação que consiste na execução das seguintes tarefas: limiarização global, filtragem de exames ruidosos, exclusão de tecidos de baixa densidade e fundo, extração de contornos pulmonares, preenchimento de orifícios, remoção de traqueia e suavização de contornos. O fluxograma da Figura 5 sumariza o funcionamento geral do algoritmo de segmentação.

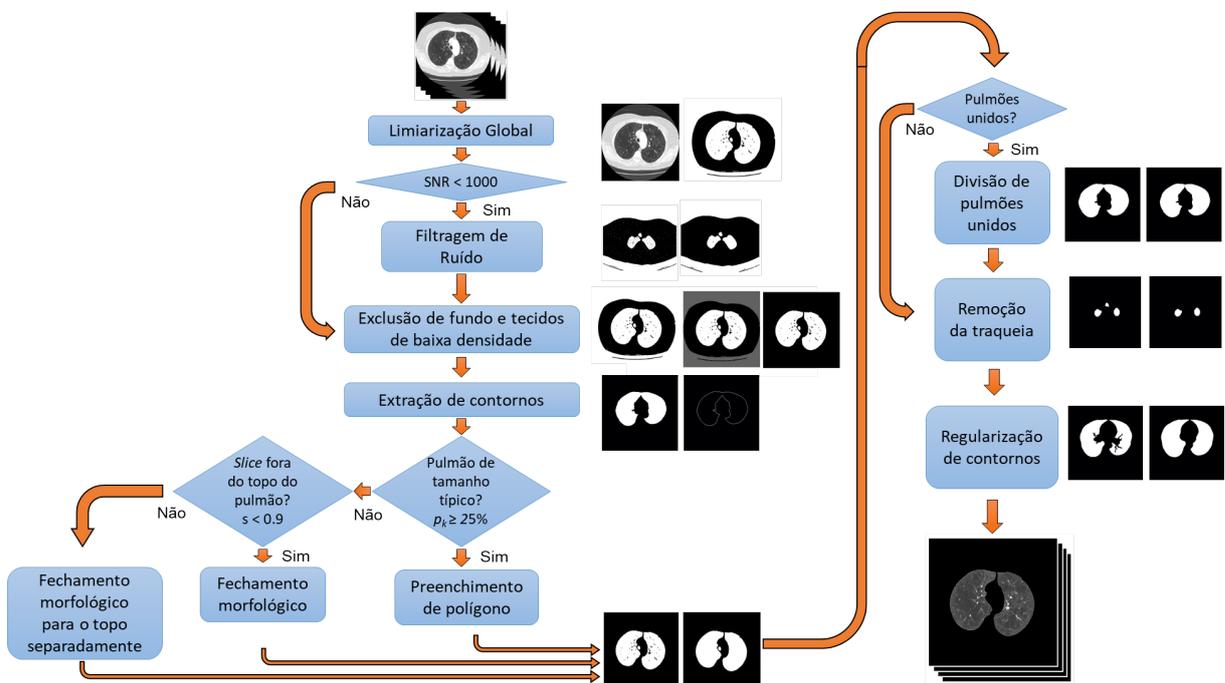


Figura 5 – Fluxograma do algoritmo de segmentação.

3.2.1 Limiarização Global

O algoritmo de limiarização global integra a etapa inicial da segmentação. Nesta tese, tal procedimento é utilizado para separar o pulmão do fundo, obtendo-se um resultado de segmentação preliminar. Em outras palavras, o objetivo da limiarização global é realizar uma classificação inicial dos *voxels* como pertencentes ou não às cavidades pulmonares. Esta técnica de limiarização global é definida na Seção 2.1, sendo empregada com o parâmetro de tolerância igual a 0. Desse modo, denota-se como T o resultado da limiarização global.

Como a atenuação do ar é de aproximadamente 1000 HU (Unidades Hounsfield), a

atenuação da maioria dos tecidos pulmonares está na faixa de $-910 HU$ a $-500 HU$, ao passo que sangue, ossos e paredes pulmonares são mais densos (acima de $500 HU$). Desse modo, considera-se a melhor escolha para o valor do limiar inicial como sendo $-500 HU$, conforme mostrado na Figura 6 (CHOI; CHOI, 2014).

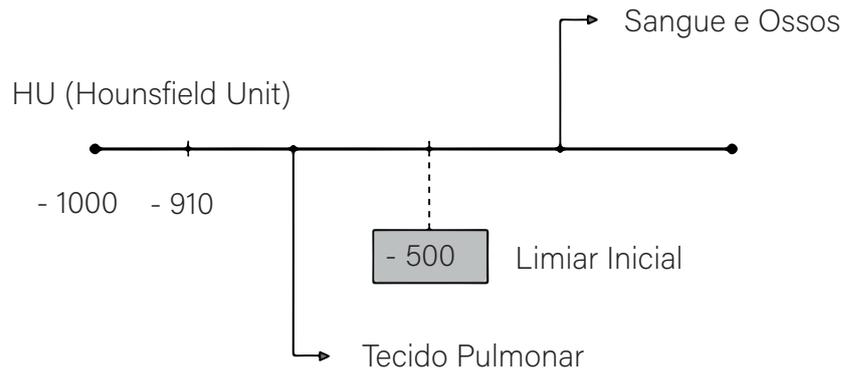


Figura 6 – Escolha do limiar inicial: -500 Unidades Hounsfield (HU).

3.2.2 Filtragem de Exames Ruidosos

Executada a primeira etapa de segmentação, observa-se que alguns exames de TC são contaminados por ruído, o qual deve ser eliminado. Para isso, é necessário identificar que exames demandam este tratamento. Conforme explica-se a seguir, esta identificação é realizada calculando-se a relação sinal-ruído (SNR) na imagem do *slice* central do exame em questão e, submetendo à filtragem de ruído os exames com SNR menor que 1000. Como a presença de diversas estruturas anatómicas no interior dos pulmões torna a imagem heterogênea nessas regiões, o valor da SNR é calculado em regiões externas ao pulmão, identificadas por meio de uma transformada da distância sobre a imagem limiarizada T . Desta forma, realiza-se uma filtragem preliminar sobre essa imagem T com um filtro da mediana e compara-se o resultado da filtragem T_f à sua versão sem filtragem, considerando-se como ruído *pixels* no exterior do pulmão, nos quais os valores dos *voxels* na imagem T são iguais a 1, mas na imagem filtrada T_f assumam valores iguais a 0. Assim, caso a SNR esteja acima do valor mínimo estabelecido, a imagem 3D T é mantida sem alterações, caso contrário, o resultado desta etapa é T_f . Portanto, a filtragem só produz efeitos nos exames com baixa relação sinal-ruído em seu *slice* central.

Desse modo, $T_f = Med_3(T)$ é definido como o resultado preliminar da filtragem, em que $Med_3(T)$ é um filtro mediana tridimensional com uma janela de tamanho 3 aplicada à imagem tridimensional limiarizada T . As regiões externas são definidas como $O = D(T_f) > 2$, em que $D(T_f)$ é a transformada da distância de uma imagem binária tridimensional T_f , a qual é calculada em todos os pontos como a distância ao *voxel* mais próximo igual a 1. O operador relacional " $>$ " produz como resultado um valor Booleano igual a 1, quando o operando à esquerda é maior do que o operando à direita, e 0, em caso contrário. Desse modo, o ruído detectado N e a relação sinal-ruído SNR são dados por

$$N = T \text{ and } \overline{T_f} \text{ and } O \quad (3.1)$$

e

$$SNR = \frac{\#O}{\sum N}. \quad (3.2)$$

Na Equação 3.1, o operador *and* diz respeito à operação lógica AND elemento a elemento, enquanto $\overline{T_f}$ é a negação lógica da matriz T_f , também elemento a elemento. Já na Equação 3.2, o termo $\#O$ é a cardinalidade do conjunto de *voxels* dentro da região " O ", ou seja, o número de 1's na matriz O , que representa o número total de *voxels* na região externa ao pulmão.

Resumidamente, uma vez conhecido o valor estimado da relação sinal-ruído, mantém-se o resultado da filtragem sempre que $SNR > 1000$ e descarta-se a filtragem em caso contrário. Um exemplo de resultado do procedimento descrito nesta seção é mostrado nas Figuras 7 e 8.



Figura 7 – Imagem ruidosa.

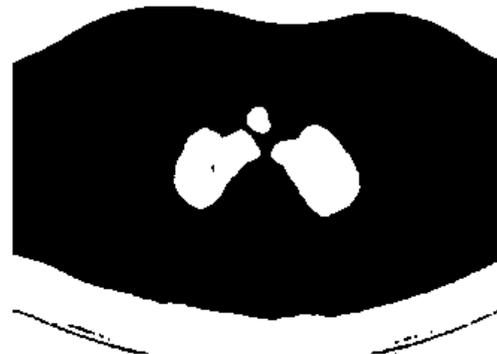


Figura 8 – Imagem filtrada.

3.2.3 Exclusão do Fundo e dos Tecidos de Baixa Densidade

O processo de segmentação deve fornecer uma imagem binária tridimensional com *voxels* pertencentes à cavidade pulmonar identificados por 1's, e os demais marcados com 0's. No entanto, os dois estágios de segmentação anteriores podem marcar com 1 *voxels* de outras regiões, como plano de fundo ou estruturas de baixa densidade. Portanto, faz-se necessário excluir tais partes da imagem do resultante da segmentação preliminar.

Desse modo, para cada exame, o algoritmo descrito na Seção 2.2 identifica com diferentes rótulos regiões conectadas com conectividade 18, em decorrência dos melhores resultados visuais obtidos por este tipo de conectividade.

Uma vez que as regiões conectadas tridimensionais são rotuladas, aplica-se um algoritmo desenvolvido nesta tese para exclusão do fundo e de estruturas de baixa densidade, o qual executa as seguintes tarefas:

- atribuir valor 1 para *voxels* que pertençam às 4 maiores regiões;
- considerando-se apenas a primeira linha e a primeira coluna do primeiro *slice*, definir como fundo o conjunto de *voxels* marcados com o rótulo que ocorre com maior frequência, ou seja, o valor da moda;
- atribuir 0 a *voxels* que pertençam ao fundo;
- atribuir 0 a regiões conectadas a qualquer *voxel* que esteja mais próximo aos limites esquerdo ou direito da imagem do que ao fundo em qualquer *slice xy*;
- apagar regiões que não ocupem as duas metades verticais da imagem para qualquer *slice xy*; e
- apagar regiões menores que 10% da região remanescente imediatamente maior desde que a deleção não inclua os dois maiores volumes.

As últimas três etapas do algoritmo descrito baseiam-se em três premissas sobre as imagens de TC dos pulmões, respectivamente. A primeira é que os pulmões se encontram na região central da maioria das imagens do exame 3-D. Em segundo lugar, os pulmões geralmente ocupam a maior parte da dimensão vertical da maioria dos *slices xy*. Por fim, um ou dois pulmões apresentam volume significativo em relação a qualquer outra região da imagem 3-D na maioria dos casos. Em outras palavras, normalmente um pulmão não corresponde a menos do que 10% do volume de outra estrutura presente na imagem.

O resultado do processamento descrito é a coleção de *slices* $\lambda = F(s)$, em que F é a versão binária do exame no *slice s* e λ é a imagem tridimensional formada por esta sequência de

slices, $s \in N | 1 \leq s \leq n$, e n é o número de *slices* do exame.

3.2.4 Preenchimento de Orifícios

Os volumes selecionados como candidatos a pulmão na Seção 3.2.3 podem envolver regiões vazias não marcadas com 1's, nas quais é possível encontrar nódulos. Por isso, define-se um processo de preenchimento de orifícios, o qual é um esforço para incluir tais volumes no espaço de busca por nódulos. No entanto, a proximidade entre as duas paredes pulmonares pode conduzir ao preenchimento de regiões que estejam entre os pulmões, porém fora de ambos. Por essa razão, o processo de preenchimento dos orifícios dos pulmões esquerdo e direito ocorre separadamente de acordo com as seguintes etapas, executadas *slice* por *slice*: extração dos contornos dos pulmões, divisão da imagem para pulmões separados e para pulmões unidos.

3.2.4.1 Extração dos Contornos dos Pulmões

O primeiro passo para preencher os orifícios presentes na segmentação dos pulmões é extrair os contornos das regiões candidatas a pulmão, em cada *slice* da imagem binária tridimensional λ definida na secção anterior. Essa extração ocorre por meio da repetição das seguintes tarefas para cada plano xy :

- extração de coordenadas C para cada contorno na imagem binária, usando o algoritmo de rastreamento de contorno descrito na Secção 2.3;
- ordenação de contornos por ordem decrescente de valores de perímetro;
- seleção dos contornos mais à esquerda e mais à direita como aqueles contendo os *pixels* mais à esquerda e mais à direita do *slice*, respectivamente;
- atribuição do contorno mais à esquerda como sendo candidato a contorno do pulmão esquerdo, desde que este possua um dos dois maiores perímetros; e
- atribuição do contorno mais à direita como sendo candidato a contorno do pulmão direito desde que este possua um dos dois maiores perímetros.

Após este processo, pode ocorrer de os pulmões se encontrarem unidos, dificultando os procedimentos seguintes para a detecção de seus nódulos, em especial, os demais passos da segmentação. Este fato se deve à tendência que algumas operações morfológicas, como, por exemplo, o fechamento, possuem de anexar aos pulmões regiões que estejam entre os mesmos. Portanto, com o objetivo de evitar este problema, há a necessidade de separar os pulmões.

3.2.4.2 Divisão da Imagem para Pulmões Separados

Para dividir os pulmões presentes em um *slice* s 2-D, define-se um quadro Q como o retângulo de menor área que envolve todos os *pixels* de ambos os pulmões e k como um índice lateral, sendo $k = 1$ o índice que identifica o lado esquerdo e $k = 2$ o índice do lado direito. Neste caso, o lado esquerdo Q_1 do quadro é definido como a região localizada da borda esquerda do quadro Q até a coluna central da imagem do *slice*. Da mesma forma, o lado direito Q_2 do quadro é a região localizada da coluna central até a borda direita do quadro Q , conforme mostrado na Figura 9. Desse modo, a fração p_k de área do quadro Q_k não ocupada pelo fundo é dada por

$$p_k = \frac{\sum Q_k}{\#Q_k}. \quad (3.3)$$

Quando os candidatos a pulmão esquerdo ou direito ocupam uma fração de área de seus respectivos quadros laterais compatível com as observadas em um pulmão típico, definida nesta tese como sendo $p_k \geq 25\%$, são preenchidas as regiões dentro dos contornos poligonais obtidos na Secção 3.2.4.1.

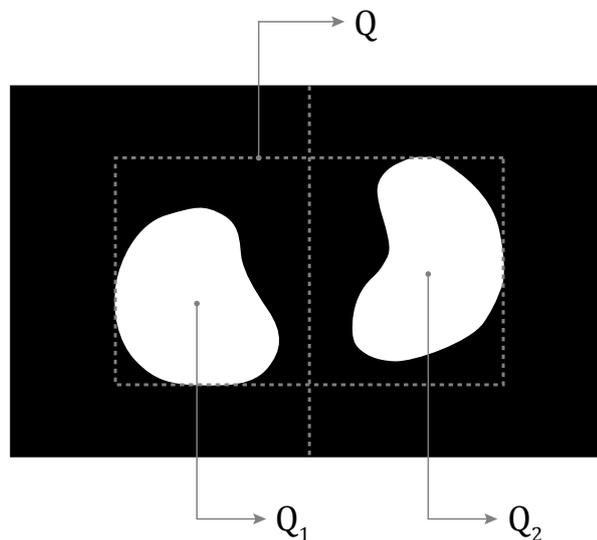


Figura 9 – Quadros laterais esquerdo e direito.

No entanto, se p_k não satisfizer o valor percentual mínimo, pode ocorrer de múltiplos contornos representarem um único pulmão. Nesse caso, uma técnica de reconstrução morfológica (SOILLE, 1999) preenche as áreas vazias, cujos parâmetros de entrada do algoritmo são as regiões selecionadas na Secção 3.2.3. No entanto, se um determinado candidato a pulmão estiver

localizado nos *slices* mais próximos à parte superior do tórax, o que se dá quando $s \geq 0,9n$, em que n é o número de *slices* do exame, e também não satisfizer o percentual mínimo de 25% da área de um quadro lateral, este algoritmo morfológico realiza o processo de preenchimento, porém o tratamento posterior de cada candidato ocorre separadamente. Essa diferenciação ocorre porque, tanto o topo do pulmão quanto outras estruturas anatômicas, podem ser encontrados nessa região.

No intuito de descrever o algoritmo de divisão e preenchimento dos pulmões *slice* por *slice*, define-se F_k como o lado k (direito ou esquerdo) de um dado *slice* binário F , C_k como o contorno de um objeto presente no lado k , $H_g(C)$ como função de preenchimento geométrico de um polígono C , $H_m(F)$ como função de preenchimento morfológico de uma imagem binária F , e finalmente T_k e B_k como versões do exame λ contendo, respectivamente, apenas a parte superior e inferior do pulmão no lado k . As imagens binárias tridimensionais T_k e B_k são entradas para processamento posterior. O fluxograma na Figura 10 representa o algoritmo descrito.

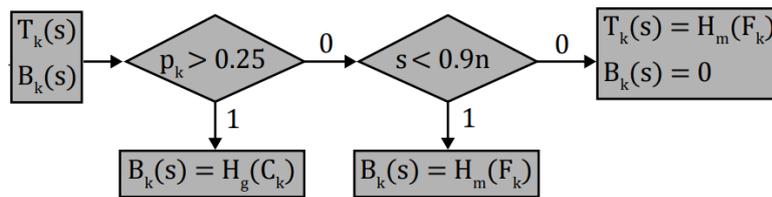


Figura 10 – Fluxograma do preenchimento de orifícios.

3.2.4.3 Divisão de Imagens para Pulmões Unidos

O método de segmentação proposto necessita que ocorra a separação de cada um dos dois pulmões individualmente. Em *slices* nos quais os pulmões se encontram muito próximos, é possível que o mesmo contorno envolva ambos em apenas uma região, conforme mostrado nas Figuras 11 e 12. No sentido de evitar divisões desnecessárias de *slices* que não contenham pulmões, o processo de divisão aplica-se apenas à parte inferior dos pulmões B_k , sendo a própria matriz B_k atualizada após este processamento e não havendo necessidade da criação de novas estruturas de dados. A divisão da imagem para pulmões unidos é mostrada na Figura 13.

Ainda assim, existem casos em que o mesmo contorno envolve duas regiões diferentes. Esta situação é detectada por meio da comparação entre a quantidade de regiões conectadas e a quantidade de contornos fechados em um *slice* da imagem de TC. A diferença entre estas quantidades pode causar a falsa impressão de que os pulmões foram segmentados separadamente,

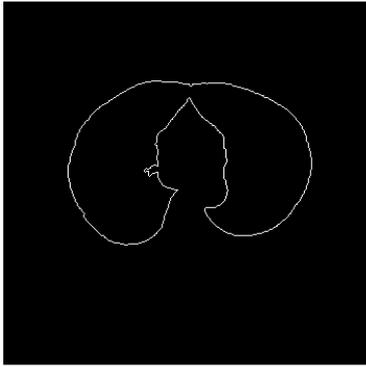


Figura 11 – Um mesmo contorno para pulmões.

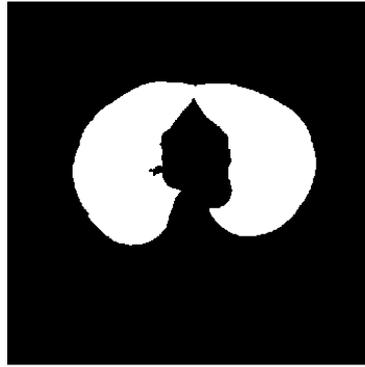


Figura 12 – Uma região contendo dois pulmões.

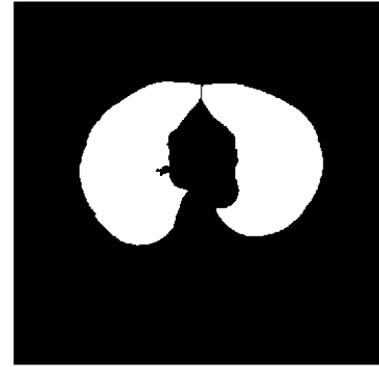


Figura 13 – Resultado da divisão da imagem.

como mostrado nas Figuras 14, 15 e 16. Assim, quando um *slice* contém um único contorno e duas regiões, uma operação de fechamento morfológico funde os pulmões. Nesta operação, com o intuito de produzir as alterações necessárias com o mínimo de impacto sobre a imagem, usa-se um elemento estruturante no formato de disco com raio igual a 1, amplamente usado na literatura.

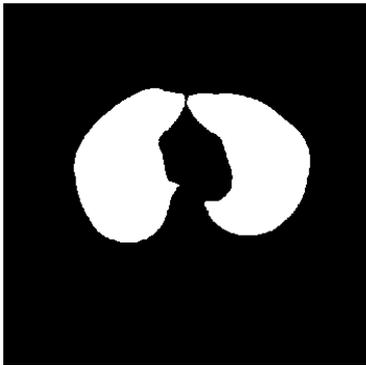


Figura 14 – Imagem contendo duas regiões.

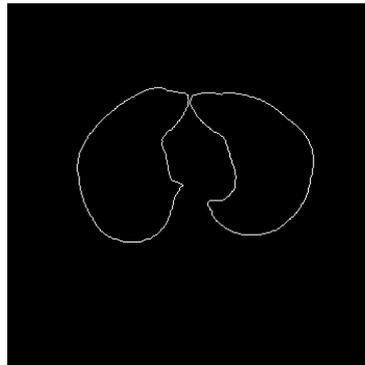


Figura 15 – Um contorno com duas regiões.

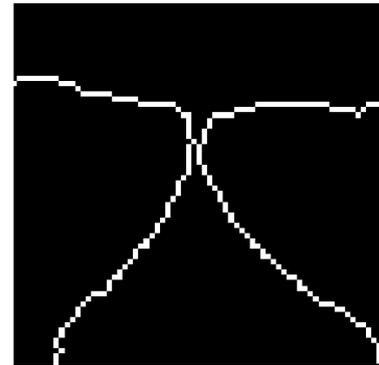


Figura 16 – Detalhe de um contorno com duas regiões.

Uma vez que os pulmões estejam inequivocamente unidos, o processo de divisão pode ser realizado definindo-se uma coluna em cada imagem para divisão vertical. Existem duas opções para isto. A primeira opção consiste em dividir a imagem na coluna com a *faixa principal* mais curta, conforme mostrado na Figura 17 e considerando as seguintes definições.

Uma *faixa* é um segmento contíguo de 1's em uma coluna. Já um *segmento expresso* é um segmento horizontal contíguo de 1's que perpassa ambos os pulmões de um extremo a outro. Um *segmento central* é definido como um tipo de *segmento expresso* cujos lados esquerdo e direito correspondem individualmente a pelo menos $1/4$ do comprimento total do segmento

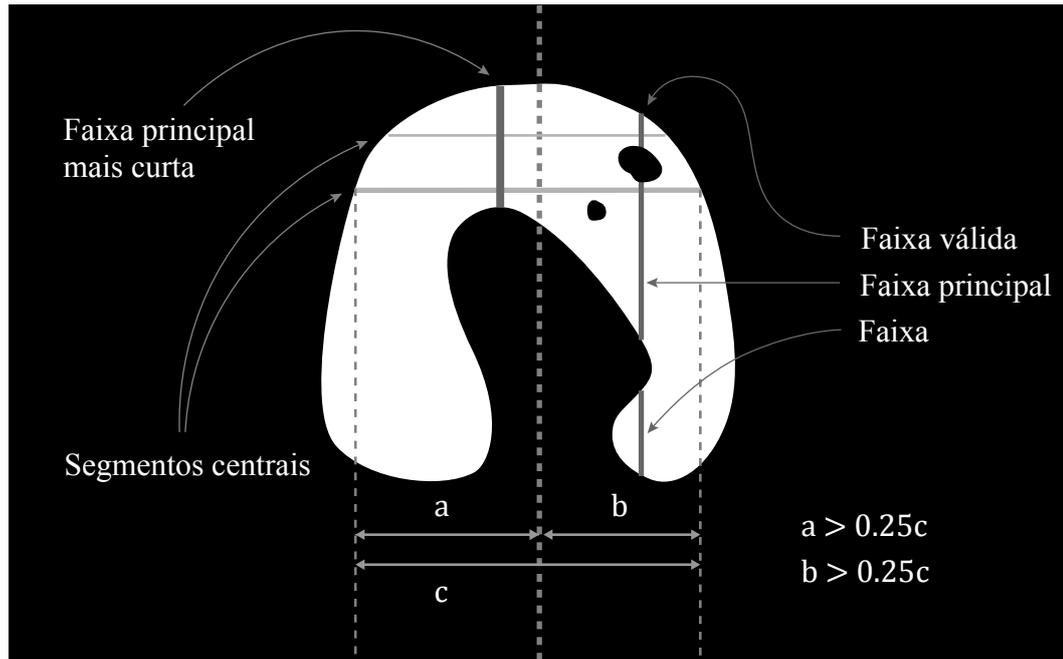


Figura 17 – Medidas para a divisão de um pulmão na *faixa principal* mais curta.

expresso em questão. Os lados esquerdo e direito são resultados da divisão do segmento em questão a partir do centro da imagem. A seguir, define-se uma *faixa central* como uma *faixa* que intercepta pelo menos um *segmento central*. Assim, uma *faixa válida* é uma *faixa central* pela qual passa pelo menos um *segmento exposto*. Finalmente, a *faixa principal* é a mais longa *faixa válida* de uma coluna da imagem.

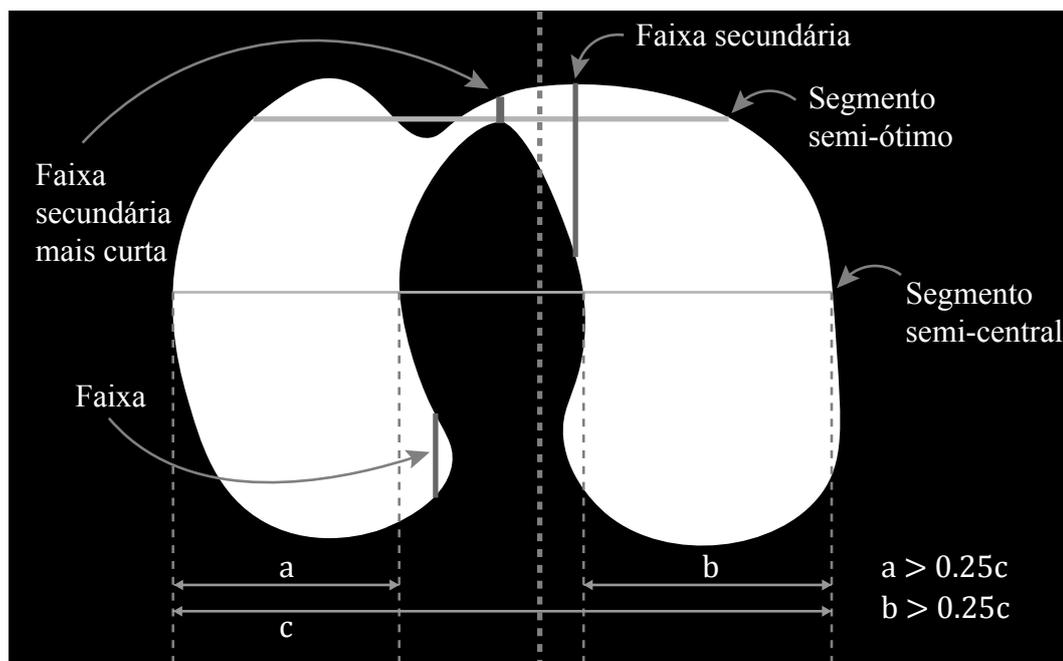


Figura 18 – Medidas para divisão do pulmão na *faixa secundária* mais curta.

Caso não haja uma *faixa principal*, a segunda opção é dividir a imagem na coluna

com a *faixa secundária* mais curta de acordo com as definições a seguir, as quais são ilustradas na Figura 18. Um *segmento semi-central* é um segmento horizontal contíguo de 1's que perpassa ambos os pulmões de um extremo a outro e cuja quantidade de 1's em cada lado (esquerdo e direito) corresponde a pelo menos 1/4 do total. Os lados esquerdo e direitos são resultados da divisão do segmento a partir do centro da imagem. Já um *segmento semi-ótimo* é o *segmento semi-central* com a menor quantidade de *pixels* pertencentes ao fundo que intercepta pelo menos a *faixa* de uma dada coluna. Finalmente, a *faixa secundária* é a faixa de uma coluna que intercepta o *segmento semi-ótimo* desta coluna.

3.2.5 Remoção da Traqueia

O método descrito nas etapas anteriores pode considerar a traqueia como uma região pertencente aos pulmões, sendo conveniente removê-la. No intuito de remover a traqueia, é definida uma representação preliminar do pulmão no lado k (direito ou esquerdo), como a matriz tridimensional binária L_k . Portanto, considerando $T_k + B_k$ como a união das imagens tridimensionais do topo e da base do pulmão do lado k (resultado da etapa descrita na Seção 3.2.4), L_k é uma matriz, cujos *voxels* são 1's quando correspondem à maior região conectada de $T_k + B_k$ e 0's caso contrário. Tais regiões resultam do algoritmo mencionado na Seção 3.2.3 (SEDFEWICK, 1998). No sentido de acelerar os cálculos, o processo de remoção considera a maior região conectada como sendo aquela cujo rótulo possui o maior número de ocorrências em 100 sorteios entre todos os *voxels* não pertencentes ao fundo. Este processo resulta da remoção da traqueia que é mostrado nas Figuras 19 e 20.

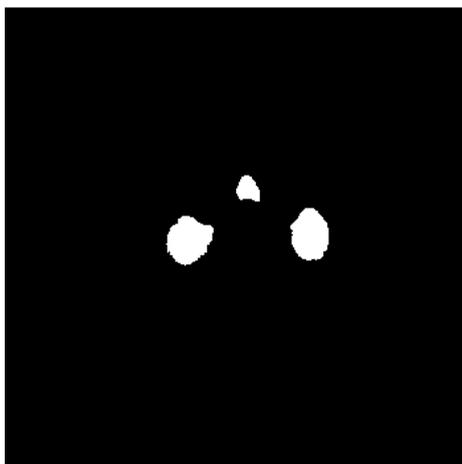


Figura 19 – Slice binário antes da remoção de traqueia.

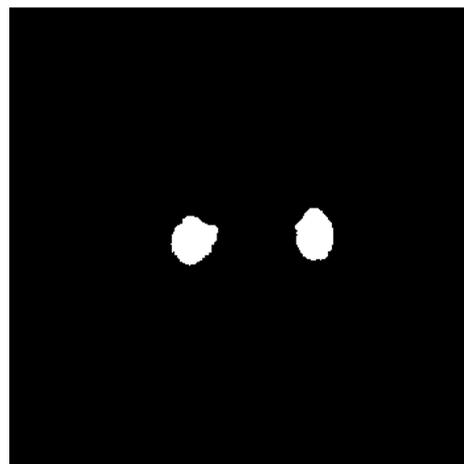


Figura 20 – Slice binário depois da remoção de traqueia.

3.2.6 Regularização do Contorno

O processo de segmentação ora desenvolvido pode excluir vasos pulmonares, vias aéreas e áreas próximas às aberturas das paredes pulmonares. Como é possível que haja nódulos nessas regiões, a técnica de preenchimento exposta a seguir suaviza os contornos pulmonares e o processo de segmentação geral passa a considerar tais regiões como parte do pulmão.

Desse modo, seja t o tamanho máximo tolerável para uma irregularidade no contorno, limitado a 20 *pixels*, $D(\cdot)$ a transformada de distância aplicada à imagem binária 3-D “.”, e L_k e T_k matrizes definidas de acordo com a Secção 3.2.5. Então, a matriz tridimensional, que representa o pulmão com bordas suavizadas *slice* a *slice* do lado k (esquerdo ou direito) é dada por

$$R_k = \overline{D[D(L_k - \overline{T_k}) \leq t]} \geq t. \quad (3.4)$$

O processo de suavização dos contornos é ilustrado nas Figuras 21 e 22 nos lados $k = 1$ e $k = 2$ de R_k , sendo o tamanho das irregularidades suavizadas regulado pelo parâmetro $l = 20$.

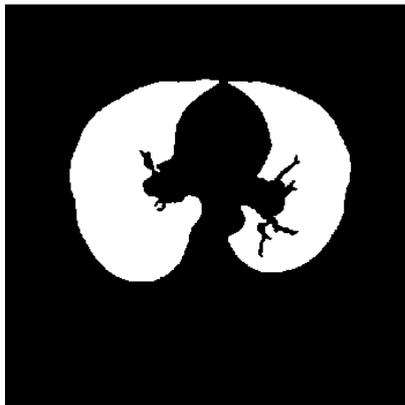


Figura 21 – Contorno irregular.

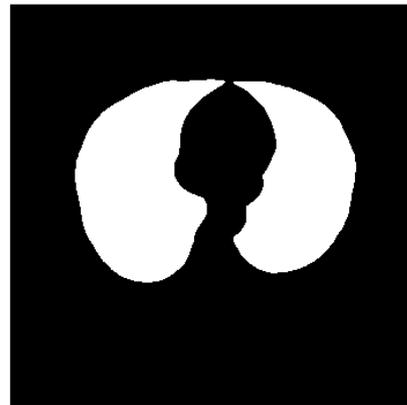


Figura 22 – Contorno suavizado.

Como os *slices* mais altos dos exames podem conter tanto o topo do pulmão quanto outras estruturas anatômicas, conforme mostram as Figuras 23 e 24, o preenchimento referente à fase de suavização de contorno se aplica apenas aos *slices* localizados abaixo de 90% da altura total do exame.

Após a suavização dos contornos, a imagem auxiliar de segmentação, resultado do processo de segmentação como um todo, é obtida unindo-se o topo dos pulmões T_k à matriz R_k , que contém as demais regiões do pulmão que passaram pelo processo de suavização de bordas.



Figura 23 – Estruturas presentes nos slices mais próximos ao topo do tórax.

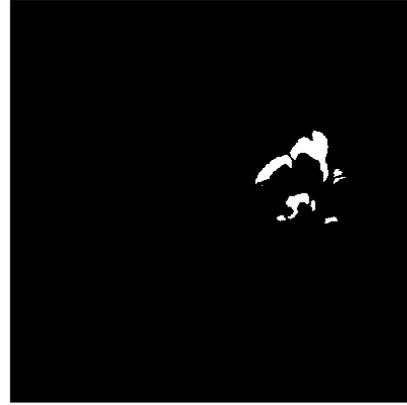


Figura 24 – Estruturas inalteradas pelo preenchimento de orifícios.

Assim, a imagem auxiliar de segmentação M assume valor igual a 1 em *voxels* que correspondem à região do pulmão e 0 nas demais regiões e é dada por

$$M = \sum_{k=1}^2 R_k \text{ ou } \sum_{k=1}^2 T_k. \quad (3.5)$$

Em exames em que fatores excepcionais como presença de marca-passos, ruído de radiação ou outras condições anormais comprometam a continuidade da imagem da parede pulmonar, as bordas ausentes são preenchidas manualmente para completar o processo de segmentação e permitir detecção de nódulos posteriormente.

3.3 Vetor de Filtros de Pré-processamento para Detecção de Nódulos

A geração de imagens que evidenciem todos os nódulos existentes em um exame de TC para sua posterior detecção é uma tarefa complexa. Além de evidenciados, para sua correta identificação, tais nódulos devem ser isolados de outras estruturas, como vias aéreas e vasos pulmonares. Dessa forma, um pré-processamento permissivo tende a manter os nódulos conectados a outras estruturas anatômicas. Por outro lado, um pré-processamento rígido pode suprimir nódulos pequenos ou com baixa intensidade de nível de cinza. Esta tese propõe um Vetor de Filtros de Pré-processamento para superar tais dificuldades.

O Vetor de Filtros de Pré-processamento é um conjunto de filtros, cuja finalidade é maximizar a quantidade de nódulos isolados. Neste método, são executadas múltiplas etapas de pré-processamento com valores predefinidos para cada parâmetro ao invés de realizar-se uma busca por um conjunto sub-ótimo de parâmetros para uma etapa única. Tais argumentos

são entradas para um algoritmo com três estágios: suavização, limiarização e processamento morfológico. Desse modo, todos os volumes resultantes dos múltiplos pré-processamentos são candidatos a nódulo, sendo submetidos à etapa posterior de classificação. Os efeitos dos estágios de suavização, limiarização e processamento morfológico são ilustrados nas Figuras 25, 26, 27 e 28. Ao A Figura 28 mostra duas regiões obtidas ao final do processo, sendo a região mais acima um nódulo unido às paredes do pulmão remanescente das etapas anteriores (protuberância presente na parte superior direita das imagens 25, 26, 27, na borda da região cinza escura).



Figura 25 – Imagem original de TC.



Figura 26 – Imagem após o estágio de suavização.



Figura 27 – Resultado depois do estágio de limiarização.

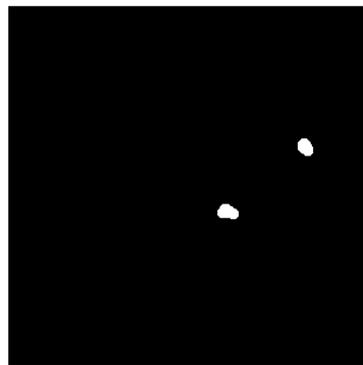


Figura 28 – Resultado do processamento morfológico.

O Vetor de Filtros de Pré-processamento assemelha-se ao algoritmo de reconhecimento de objetos *Scale-Invariant Features Transform* (SIFT) no sentido de que o SIFT filtra a imagem original com janelas gaussianas de diferentes dimensões e desvios padrão antes de proceder à extração de atributos (LOWE, 1999). Semelhantemente, o método proposto pré-processa uma TC com filtros mediana multi-escala e realiza várias limiarizações com diferentes valores de limiar para cada escala. Esta propriedade confere ao Vetor de Filtros de Pré-processamento a capacidade de detectar nódulos de modo mais independente em relação ao tamanho ou ao nível

de cinza destes nódulos. De modo similar, o SIFT é capaz, segundo sua formulação, de extrair atributos descritivos invariantes à escala ou à iluminação de um objeto.

As diferentes versões do mesmo exame, cada uma suavizada por um filtro da mediana com janela de tamanho w (assumindo valores 1, 3, 5 e 7, indexados por $i = 1, 2, 3$ ou 4) e limiarizada com valores de limiar z (0,5, 1, 2 ou 3 desvios padrão acima da moda, indexados por $j = 1, 2, 3$ ou 4) são mostrados na Figura 29. O resultado são 4 versões suavizadas para um exame, e, para cada versão suavizada, 4 versões limiarizadas, totalizando 16 versões tridimensionais de um mesmo exame, sendo cada um submetida a um tipo particular de processamento morfológico. Desse modo, todas as 16 imagens tridimensionais resultantes das etapas de suavização e limiarização são submetidas a um processamento morfológico que depende dos valores de w e de z . Tal medida visa evidenciar os nódulos presentes e diminuir a quantidade prevalente de regiões que não são nódulos. Os detalhes de cada um dos três estágios de pré-processamento e a definição do modelo geral de Vetor de Filtros de Pré-processamento são mostrados a seguir.

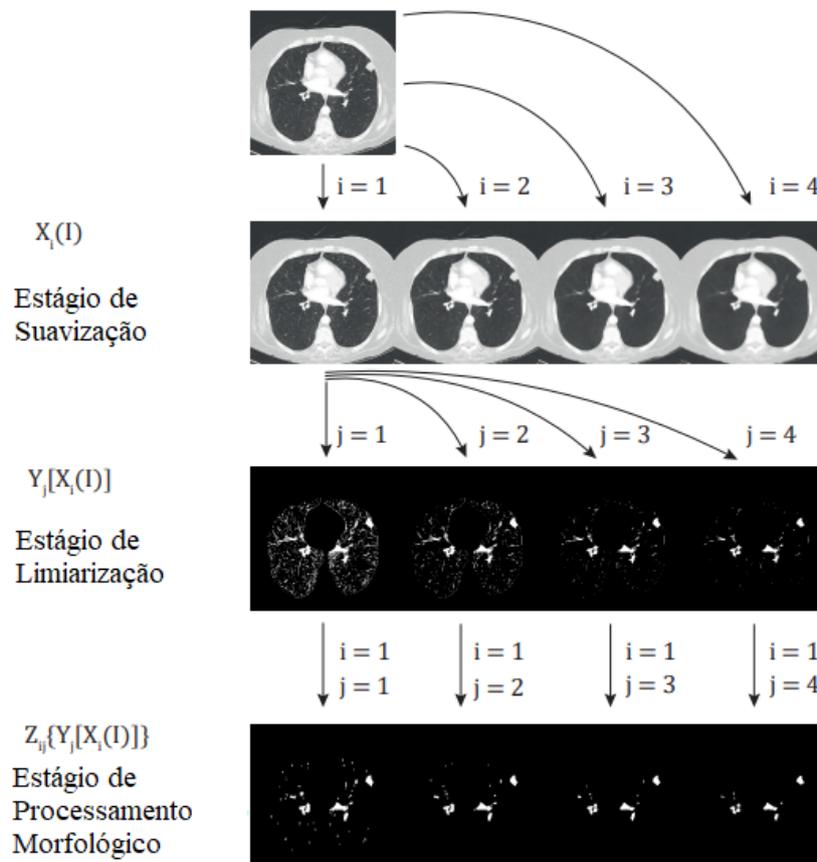


Figura 29 – Vetor de Filtros de Pré-processamento formado por seus 3 estágios: 4 tipos de suavização, 4 tipos de limiarização e, para cada uma das 16 combinações resultantes, um tipo particular de processamento morfológico

3.3.1 Suavização

Seja $w_i = 2i - 1$ a dimensão da janela de suavização, em que $i = 1, 2, 3$ ou 4 . A operação de suavização X_i consiste na convolução do filtro da mediana tridimensional com janela de tamanho w_i com a imagem de entrada “.” dada pela equação

$$X_i(\cdot) = Med_{w_i}(\cdot). \quad (3.6)$$

Portanto, este estágio produz quatro versões do exame original suavizadas por filtros da mediana com janelas de tamanho 1, 3, 5 e 7, respectivamente. Esse procedimento visa evidenciar nódulos de diferentes dimensões.

3.3.2 Limiarização

Seja k_j a componente escalar índice j do vetor $k = [0.5, 1, 2, 3]$, onde $j = 1, 2, 3$, ou 4 , e seja M seja a máscara de segmentação obtida conforme descrito no Capítulo 3.2. A operação de limiarização Y_j aplicada aos níveis de cinza dos *voxels* no exame “.” é dada por

$$Y_j(\cdot) = [(\cdot) > T_j] \text{ and } M, \quad (3.7)$$

em que $T_j(\cdot) = Mo(\cdot) + k_j\sigma(\cdot)$ é o limiar, $\sigma(\cdot)$ é o desvio padrão, e $Mo(\cdot)$ é a moda do nível de cinza dos *voxel* no exame, desconsiderando-se os zeros.

Em resumo, o processo de limiarização seleciona os níveis de cinza que estão a $0,5\sigma$, 1σ , 2σ ou 3σ acima da moda. Esta técnica permite a detecção de nódulos de vários níveis de cinza. Como a fase de limiarização produz quatro versões limiarizadas da entrada e recebe quatro entradas oriundas da etapa de suavização, são produzidas 16 versões do exame inicial.

3.3.3 Processamento Morfológico

O principal objetivo da fase de processamento morfológico é desvincular os nódulos justa-pleurais e justa-vasculares de outras estruturas às quais se encontrem fisicamente conectados. No entanto, cada uma das 16 imagens 3-D provenientes das fases anteriores de pré-processamento contêm nódulos com propriedades distintas. Por esta razão, cada versão suavizada e limiarizada de exame é submetida a um tipo particular de processamento morfológico.

O comportamento dos múltiplos processamentos morfológicos é regulado por seis parâmetros em comum: *tamMin*, *espParede*, *tamCorte*, *minRest*, *maxRest* e *tolVol*. Ao longo desta seção são fornecidas explicações adicionais acerca desses parâmetros, os quais são definidos preliminarmente conforme a seguir.

O argumento *tamMin* é o volume mínimo para que uma região seja considerada grande para os propósitos do algoritmo. Já o parâmetro, *espParede* consiste na espessura das paredes do pulmão considerada nos cálculos. O argumento *tamCorte* é o comprimento do elemento estruturante utilizado para separar os nódulos de outras estruturas. O volume mínimo requerido para que um candidato possa ser restaurado é dado pelo parâmetro *minRest*. Semelhantemente, *maxRest* regula o volume máximo requerido para que um candidato possa ser restaurado. Por último, o parâmetro *tolVol* limita até que volume um candidato que sofreu corrosão pode recuperar na etapa de restauração.

A definição dos valores para os parâmetros utilizados nos múltiplos processamentos morfológicos é mostrada na Tabela 1. Estes valores são empregados para processar morfológica-mente os 16 resultados das etapas de suavização e limiarização, sendo cada resultado indexado por *p* nas linhas da Tabela. A escolha dos parâmetros mostrados na Tabela 1 é realizada de forma experimental e visa maximizar o número de candidatos a nódulo isolados. O processamento morfológico é dividido em duas fases: corrosão de candidatos e restauração de candidatos.

Tabela 1 – Parâmetros do Processamento Morfológico

<i>p</i>	<i>tamMin</i>	<i>espParede</i>	<i>tamCorte</i>	<i>minRest</i>	<i>maxRest</i>	<i>tolVol</i>
1	105	2	0	10	100	2
2	110	2	0	10	100	2
3	120	2	0	10	100	2
4	130	2	0	10	100	2
5	305	2	4	30	400	2
6	310	2	4	30	400	2
7	320	2	5	30	400	2
8	330	2	5	30	400	2
9	505	3	8	50	1600	2
10	510	3	8	50	1600	2
11	520	3	11	50	1600	3
12	530	3	11	50	1600	3
13	705	5	8	70	6400	4
14	710	5	8	70	6400	4
15	720	5	12	70	6400	4
16	730	5	15	70	6400	4

3.3.3.1 Corrosão de Candidatos

A fase de corrosão dos candidatos visa eliminar os *voxels* que conectam os candidatos a nódulos a outras estruturas anatômicas. Os critérios de eliminação dos *voxels* são os seguintes:

1. *Voxels* marcados como candidatos a vasos sofrem corrosão; e
2. *Voxels* próximos às paredes do pulmão são eliminados caso pertençam a regiões consideradas grandes para os propósitos do algoritmo.

O primeiro passo visa isolar nódulos justa vasculares sem eliminar pequenos nódulos em outras partes do pulmão. Já o segundo passo, visa desconsiderar como candidatos a nódulos estruturas remanescentes do processo de segmentação que não pertençam ao pulmão. Desse modo, nas vizinhanças das paredes pulmonares, apenas estruturas pequenas ou estruturas grandes com alguma porção dentro do pulmão permanecem após este estágio. Então, a matriz resultante da operação de corrosão M_C é dada pelo processamento elemento a elemento

$$M_C = [(\overline{M_V} \text{ and } Y) \text{ or } (M_V \text{ and } M_{cut})] \text{ and } M_B. \quad (3.8)$$

Com a finalidade de definir as matrizes binárias 3-D M_V , M_B e M_{cut} na Equação 3.8, considera-se uma imagem 3-D com regiões conectadas com conectividade $N_{18}(v)$, ou conectividade 18, obtidas por meio da aplicação de um algoritmo de crescimento de regiões sobre a matriz Y definida na Seção 3.3.2 (SEDGEWICK, 1998). Esses termos são tais que M_V é uma matriz que contém candidatos a vasos, definidos como a maior região conectada exceto pelo fundo. Desse modo, M_V é uma imagem binária 3-D contendo 1's nesta região. Já Y é o resultado das fases de suavização e limiarização. Por sua vez, a matriz M_B representa regiões grandes próximas às paredes pulmonares. Para definir M_B , considera-se M como sendo a máscara de segmentação e $D(\cdot)$ a transformada da distância bidimensional aplicada sobre cada *slice* da imagem 3-D “.” individualmente. Neste caso, a região que envolve o parênquima pulmonar resulta da operação $\overline{D(M)} \geq espParede \text{ and } M$, em que $espParede$ é a espessura das paredes pulmonares considerada nos cálculos mostrada na Tabela 1. Em outras palavras, as bordas do parênquima são um invólucro 3-D com uma espessura dada por $espParede$ que cresce a partir da interface da máscara de segmentação. Portanto, a matriz 3-D M_B representa regiões que pertencem à superfície dessas bordas e também pertencem a qualquer estrutura considerada grande. Então, os *voxels* de M_B possuem valor igual 1 em tais regiões. Nesse contexto, uma região conectada é considerada grande caso seu volume seja maior que o parâmetro $tamMin$. Finalmente, define-se M_{cut} como a

versão recortada da imagem tridimensional Y .

O método de corte que produz a matriz M_{cut} depende da dimensão da janela utilizada na etapa de suavização prévia, na qual cada uma dessas janelas corresponde a um tipo de recorte. Caso a janela possua dimensão igual 1, esse método aplica um filtro não linear 3×3 *slice* a *slice*, o filtro *look-up table* (filtro LUT). Este filtro elimina o *voxel* central de janelas da imagem que sejam iguais a qualquer um dos conjuntos mostrados na Figura 30. Esta medida tem por objetivo apagar apenas pequenos objetos que se assemelhem às conexões entre nódulos e outras estruturas.

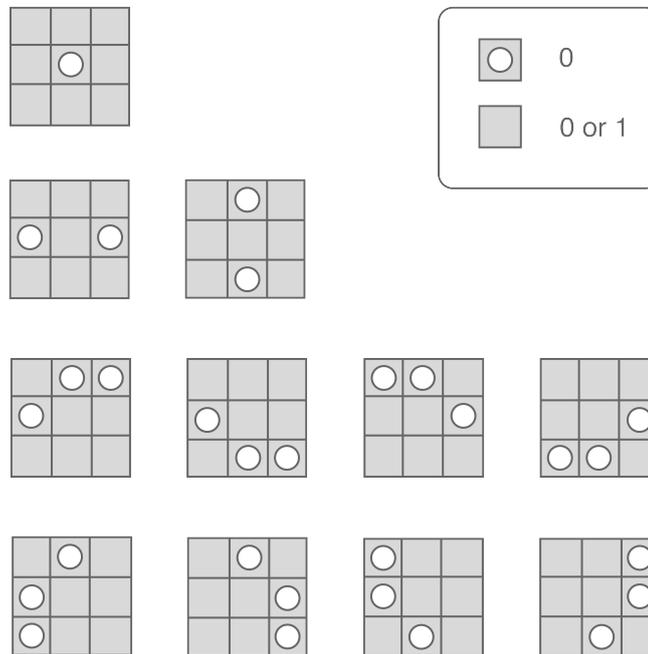


Figura 30 – Configurações de janelas em que o *voxel* central é apagado pelo filtro LUT.

A janela LUT 3×3 filtra os *slices* da imagem Y ao longo dos planos ortogonais xy , yz e xz em 3 etapas de filtragem, uma para cada plano. Estas etapas são seriais e respeitam a sequência xy , yz e xz , resultando na imagem de corte M_{cut} .

Para janelas com dimensões maiores que 1, que evidenciam nódulos maiores, os filtros LUT não são severos o suficiente para quebrar a maioria das rígidas conexões entre esses nódulos e outras estruturas pulmonares. Neste caso, são aplicadas duas operações morfológicas de abertura 2-D a cada *slice*, usando elementos estruturantes com apenas uma dimensão dominante, horizontal ou vertical. A ideia é cortar conexões resistentes com elementos estruturantes geometricamente semelhantes a um objeto cortante. Neste caso, o resultado apenas considera

as estruturas remanescentes após os dois processos de abertura. Desse modo, a operação de abertura no *slice* s de uma imagem Y 3-D é dada por $O[Y(s), \varepsilon]$, com um elemento estruturante ε . Então, cada *slice* da imagem cortada M_{cut} é dado por

$$M_{cut}(s) = O[Y(s), \varepsilon] \text{ and } O[Y(s), \varepsilon^T], \quad (3.9)$$

em que o elemento estruturante ε é um vetor linha de comprimento $tamCorte$, e ε^T é sua versão transposta.

3.3.3.2 Restauração de Candidatos

O estágio de corrosão de candidatos pode degradar nódulos provenientes do pré-processamento. Na fase de restauração de candidatos, no entanto, é possível rotular novamente os *voxels* que foram perdidos. Por outro lado, esses *voxels* devem respeitar condições que evitem que conexões previamente quebradas sejam reestabelecidas. Desse modo, o resultado desta etapa é tal que determinados candidatos podem restaurar sua forma original após serem degradados.

Assim, considerando-se que \hat{x} , \hat{y} e \hat{z} sejam as coordenadas espaciais dos candidatos a nódulo. Existem dois cenários possíveis para rotular novamente os *voxels*: 1) um determinado voxel está dentro da região limitada pelo candidato $[\min(\hat{x}) \leq \hat{x} \leq \max(\hat{x}), \min(\hat{y}) \leq \hat{y} \leq \max(\hat{y}), \min(\hat{z}) \leq \hat{z} \leq \max(\hat{z})]$, ou, 2) em caso negativo, o voxel está dentro do volume definido pelo parâmetro $tolVol$, ou seja, $[\min(\hat{x}) - tolVol \leq \hat{x} \leq \max(\hat{x}) + tolVol, \min(\hat{y}) - tolVol \leq \hat{y} \leq \max(\hat{y}) + tolVol, \min(\hat{z}) - tolVol \leq \hat{z} \leq \max(\hat{z}) + tolVol]$. Nesse processo, os *voxels* de rótulo l dentro da região limitada pelo candidato a nódulo recebem um novo rótulo se:

- mudaram de rótulo depois da fase de corrosão;
- possuíam rótulo igual a l antes da fase de corrosão; e
- não eram parte do fundo antes da fase de corrosão.

Desse modo, os *voxels* próximos ao candidato de rótulo l que não se encaixem nas condições anteriores, mas se localizam dentro da região limitada pelo parâmetro $tolVol$ recebem um novo rótulo se estão conectados ao candidato de rótulo l e obedecem às três regras elencadas antes. Utiliza-se uma vizinhança de conectividade 26 por seus melhores resultados visuais.

Portanto, sintetiza-se o processamento morfológico como $Z_{ij}(\cdot)$, em que i e j são os respectivos índices de suavização e limiarização utilizadas, e “.” é a entrada binária 3-D oriunda do estágio de limiarização.

3.3.4 Equação do Modelo e seus Parâmetros

Uma vez conhecidas as equações dos estágios de suavização, limiarização e processamento morfológico, pode-se sumarizar todo o modelo através de uma única equação. Seja I o conjunto de imagens de TC originais, i o índice do processo de suavização e j o índice da operação de limiarização, conforme descrito nas Seções 3.3.1 e 3.3.2, respectivamente. Cada uma das 16 versões pré-processadas de um exame é dada por

$$P_{ij} = Z_{ij}\{Y_j[X_i(I)].\} \quad (3.10)$$

Os efeitos que o tamanho da janela de suavização produzem no resultado do pré-processamento para um limiar fixo são mostrados nas Figuras 31, 32 e 33. É possível observar na Figura 32 que uma janela de dimensão $w = 1$ não é suficiente para separar o nódulo de outras estruturas adjacentes. Por exemplo, no caso mostrado na Figura 33, uma dimensão de janela $w = 7$ é necessária para o cumprimento desta tarefa.

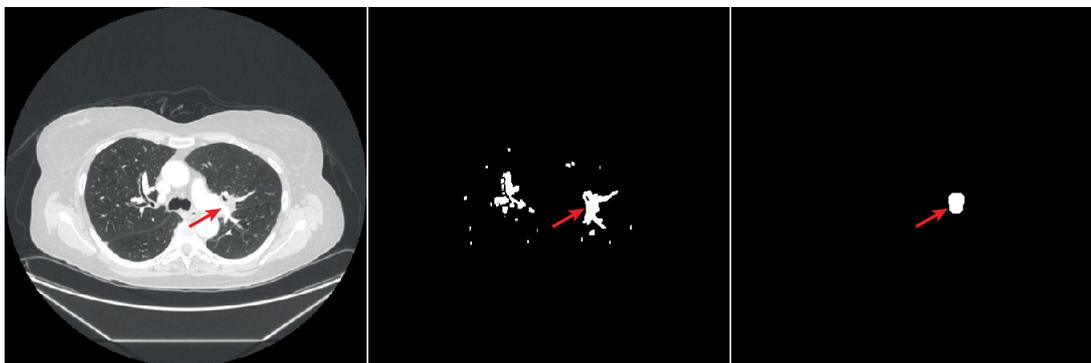


Figura 31 – Imagem antes do pré-processamento.

Figura 32 – Resultado com $w = 1$ and $k = 1$.

Figura 33 – Resultado com $w = 7$ and $k = 1$.

Similarmente, as Figuras 34, 35 e 36 mostram como uma variação no valor do limiar causa impactos no resultado de pré-processamento para uma dada dimensão de janela de suavização. Como se verifica na Figura 35, um valor de limiar de $k = 2$ pode separar de outras estruturas 3 nódulos presentes na imagem. No entanto, de modo semelhante aos exemplos anteriores, o nódulo na parte superior da imagem é perdido quando o valor de k aumenta para 3, conforme mostrado na Figura 36. Ressalta-se que cada etapa de pré-processamento produz seus próprios candidatos a nódulos, sendo considerados para a fase posterior candidatos oriundos de todas as etapas. Desse modo, candidatos progressivamente identificados ao longo de uma dentre

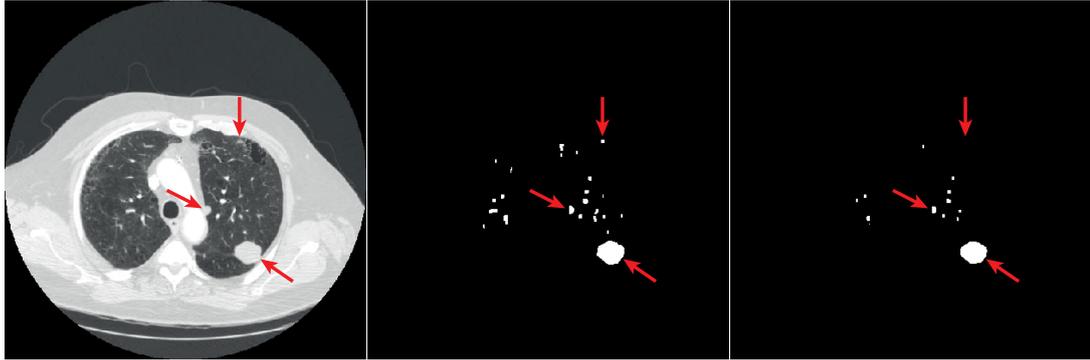


Figura 34 – Imagem antes do pré-processamento.

Figura 35 – Resultado com $w = 1$ and $k = 2$.

Figura 36 – Resultado com $w = 1$ and $k = 3$.

as 16 sequencias possíveis de suavização, limiarização e processamento morfológico não podem ser excluídos pelas outras 15 configurações.

Os parâmetros *tamCorte* e *tolVol* calibram os níveis de corrosão e restauração, respectivamente. Por exemplo, escolher um valor alto para o parâmetro *tamCorte* pode tornar os nódulos indistinguíveis, ao passo que um valor baixo pode falhar em corroer conexões entre nódulos e vasos ou vias aéreas. Isto também ocorre para o parâmetro *tolVol*, em que a escolha de valores altos pode reconstruir conexões indesejáveis, e valores baixos podem manter os nódulos degradados.

O volume dos candidatos com maior chance de serem nódulos se concentra dentro de uma faixa de valores diferente para cada dimensão de janela. Também os parâmetros *minRest* e *maxRest* são estimativas desses valores limites. Assim, a escolha por esses dois parâmetros visa reduzir o tempo de processamento, evitando a reconstrução de candidatos improváveis.

A finalidade do parâmetro *tamMin* é selecionar quais estruturas anatômicas na imagem binária devem sofrer o processo de corrosão que separa os nódulos de outras regiões. Definir um valor baixo permite que pequenos nódulos sofram uma corrosão violenta e, dessa forma, desapareçam. Por outro lado, a escolha de um valor alto pode inviabilizar o processo de corrosão em estruturas como vasos e vias aéreas, mantendo conexões que dificultam a identificação dos nódulos.

Uma vez que a fase de pré-processamento isola os candidatos a nódulos de outras estruturas anatômicas, procede-se à redução falsos positivos e às etapas de classificação.

3.4 Redução Booleana de Falsos Positivos

A fase de pré-processamento consiste em rotular estruturas dentro dos pulmões. Como cada uma das 16 versões dos exames produz um conjunto de candidatos a nódulos, a quantidade total de candidatos é elevada. Tal fato aumenta o custo computacional e dificulta a classificação correta dos nódulos por produzir falsos positivos.

A fase de redução de falsos positivos precede a classificação e é dividida em duas etapas. Na primeira etapa elimina-se o excesso de candidatos e na segunda etapa é realizada uma triagem mais refinada destes candidatos.

Portanto, no intuito de diminuir o volume de dados espúrios não classificados, a técnica de Redução Booleana de Falsos Positivos é aplicada usando-se 20 atributos que descrevem os candidatos conforme mostrado nas Secções 3.4.3 e 3.4.3.

3.4.1 Descritores dos Candidatos

Com a finalidade de eliminar antecipadamente os candidatos a nódulo não elegíveis, é necessário listar os atributos descritivos desses candidatos. Tais atributos são armazenados juntamente com os respectivos rótulos de identificação atribuídos a cada candidato e são descritos a seguir.

O volume de um candidato é dado pelo seu número de *voxels*. Já o comprimento consiste no número de *slices* do candidato nas dimensões x , y e z , dado por L_x , L_y e L_z , respectivamente, como mostra a Figura 37.

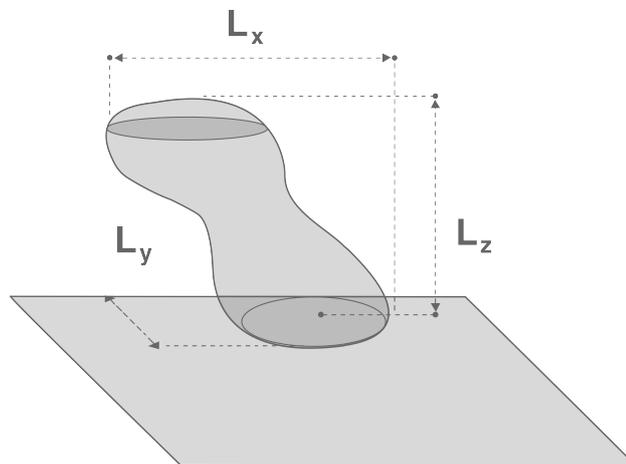


Figura 37 – Comprimentos L_x, L_y e L_z de um candidato.

Na primeira fase da redução de falsos positivos, um candidato é elegível se todas as três medidas estiverem dentro do intervalo definido na Tabela de redução de falsos positivos, definida nesta etapa posterior de processamento. Na segunda etapa de redução de falsos positivos, o atributo L é o mínimo $\min(L_x, L_y, L_z)$ dos três valores de comprimento descritos. Outro atributo descritivo é a proporção de um candidato, dada pela razão entre os comprimentos de um candidato e calculada como $P_{xy} = \frac{L_x}{L_y}$, $P_{yz} = \frac{L_y}{L_z}$ e $P_{xz} = \frac{L_x}{L_z}$. No primeiro estágio da redução de falsos positivos, um candidato alcança a próxima fase se todos os três valores obedecerem aos limites P_{xy} , P_{xz} e P_{yz} da Tabela de redução de falsos positivos. Já no segundo estágio de redução de falsos positivos, o atributo P é dado pelo mínimo $\min(P_{xy}, P_{xz}, P_{yz})$, entre os três valores calculados. Já o parâmetro *Box Count* é a quantidade Q_{box} de *voxels* dentro do menor paralelepípedo que contém um candidato e o parâmetro *Box Count* para este candidato é dado por $P_{box} = \frac{V}{Q_{box}}$.

Ainda no contexto da descrição de nódulos pulmonares, uma observação comum durante o exame *slice a slice* de um nódulo é a existência de um pico no valor da área bidimensional do nódulo, o que, em uma visualização breve da sequência de imagens, se assemelha ao piscar de uma luz. Essas variações, denominadas como parâmetros *blink*, são quantificadas ao longo dos planos xy , yz e xz por meio das seguintes etapas:

- contar o número de seções ortogonais que interceptam o candidato em cada perspectiva xy , yz e xz ;
- calcular a área bidimensional do candidato como o número de *voxels* candidatos que pertencem a cada interseção ortogonal para todas as três perspectivas;
- dividir cada conjunto de seções ortogonais em três grupos com a mesma quantidade de *slices* e calcular as áreas médias (A_1 , A_2 e A_3) para cada grupo, conforme mostrado na Figura 38;
- para candidatos com menos de 3 *slices* em uma determinada perspectiva xy , yz ou xz , os parâmetros B_0 e B_F são 0; e
- caso contrário, calcular os parâmetros B_0 e B_F para as perspectivas xy , yz e xz como

$$B_0 = \frac{A_2}{A_1} \quad (3.11)$$

e

$$B_F = \frac{A_3}{A_2}. \quad (3.12)$$

O resultado desse atributo consiste nas 6 medidas B_{xy_0} , B_{xy_F} , B_{yz_0} , B_{yz_F} , B_{xz_0} e B_{xz_F} . No entanto, como o valor dessas seis medidas individuais pode variar significativamente, os três valores escolhidos para os parâmetros *Blink* são: o maior entre os seis valores (B_3), o segundo maior valor entre as medidas (B_2) e o maior valor entre os 3 menores valores para cada perspectiva, dado por:

$$B_1 = \max \left(\begin{bmatrix} \min(B_{xy_0}, B_{xy_F}) \\ \min(B_{yz_0}, B_{yz_F}) \\ \min(B_{xz_0}, B_{xz_F}) \end{bmatrix} \right). \quad (3.13)$$

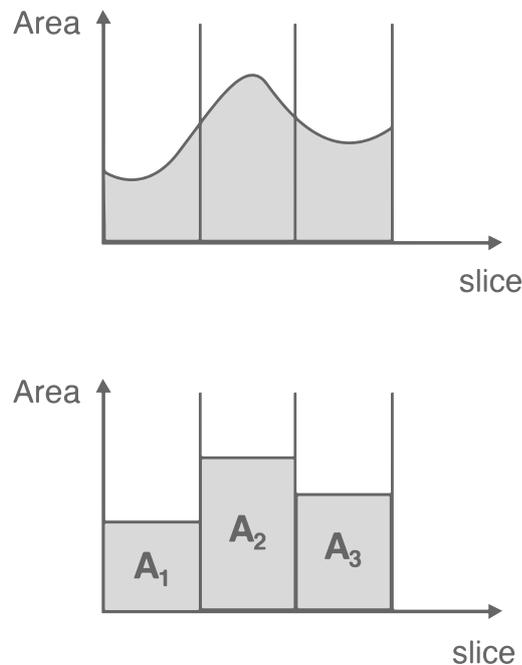


Figura 38 – Áreas médias A_1 , A_2 e A_3 para cada grupo de seções ortogonais.

Um atributo que define o quão a forma geométrica de um candidato a nódulo se assemelha à de um círculo em cada *slice* é chamado de circularidade. O valor deste atributo é uma média de medidas tomadas nos *slices* do candidato, conforme define-se a seguir. Seja \hat{N} o vetor de coordenadas \hat{x} e \hat{y} que localizam um candidato em um *slice* s_{xy} , e seja \hat{S} o círculo de centro (μ_X, μ_Y) e raio $R = \mu_d + \sigma_d$, em que μ_u e σ_u são as respectivas média e desvio padrão de u e $d = \sqrt{(\hat{x} - \mu_X)^2 + (\hat{y} - \mu_Y)^2}$. Já a circularidade de um candidato no *slice* s_{xy} é dada por $\left| \frac{\hat{N} \cap \hat{S}}{\hat{N} \cup \hat{S}} \right|$. A Figura 39 mostra os conjuntos de *pixels* \hat{N} e \hat{C} usados no cálculo da circularidade.

De uma forma similar, este procedimento é realizado para as perspectivas yz e xz , produzindo os parâmetros C_{xy} , C_{yz} e C_{xz} . No sentido de refletir com maior fidelidade a seme-

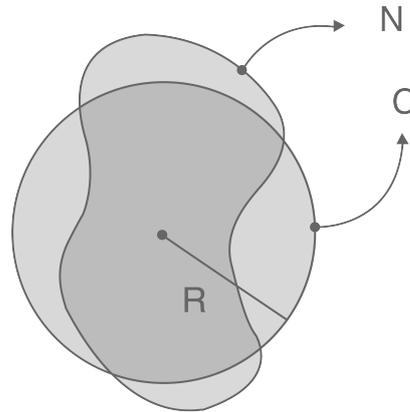


Figura 39 – Conjuntos N e C .

lhança com um círculo, os valores escolhidos para estes parâmetros são $C_1 = \min(C_{xy}, C_{yz}, C_{xz})$ e $C_2 = \text{mediana}(C_{xy}, C_{yz}, C_{xz})$. Adicionalmente, o terceiro atributo consiste na média de uma nova medida de circularidade tomada ao longo dos *slices* xy que interceptam o candidato, sendo cada medida ponderada pela área do candidato. No entanto, agora a circularidade de cada *slice* é dada por $\frac{4\pi A}{P^2}$, em que A e P são a área e o perímetro do candidato um dado plano.

No exame detalhado de uma sequência de imagens de TC, é possível perceber que o centro de massa dos nódulos pulmonares tende a se deslocar pouco, o que não ocorre para outras estruturas, como vasos e vias aéreas. Desse modo, seja um nódulo com centro de massa 3-D dado por \hat{c}_V e centro de massa 2-D no primeiro *slice* xy dado por \hat{c}_A . Na análise *slice* a *slice* de um nódulo pulmonar em uma tomografia computadorizada do tórax, o deslocamento axial D_{xy} entre \hat{c}_V e \hat{c}_A deste nódulo tende a possuir valores pequenos. Então, o parâmetro de deslocamento do centro de massa é o desvio do centro de massa em relação a sua localização inicial considerando a projeção do candidato em cada plano ortogonal. Nesse sentido, seja \hat{N} como o vetor de coordenadas \hat{x} , \hat{y} e \hat{z} de um candidato. Desse modo, o centro de massa 2-D do candidato é dado por $\hat{c}_A = \frac{1}{A} \int_A \hat{N} dA$ e o centro de massa 3-D é definido por $\hat{c}_V = \frac{1}{V} \int_V \hat{N} dV$, em que A é a área do candidato projetada em seu primeiro *slice* xy , yz ou xz e V é o espaço volumétrico do candidato. Na Figura 40 é mostrado o deslocamento do centro de massa para o plano xy , cujos parâmetros para o deslocamento do centro de massa para cada perspectiva são dados por

$$D_{xy} = | \hat{c}_V \cdot \hat{u}_{xy} - \hat{c}_A |, \quad (3.14)$$

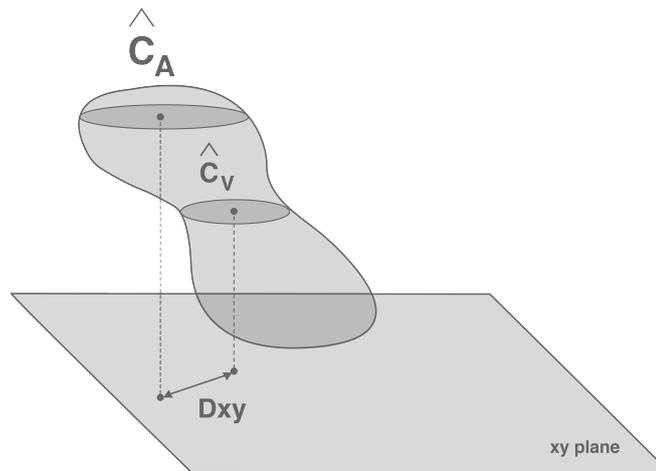


Figura 40 – Deslocamento entre as projeções dos centros de massa \hat{c}_A e \hat{c}_V para o plano xy .

$$D_{yz} = | \hat{c}_V \cdot \hat{u}_{yz} - \hat{c}_A | \quad (3.15)$$

e

$$D_{xz} = | \hat{c}_V \cdot \hat{u}_{xz} - \hat{c}_A | . \quad (3.16)$$

em que $\hat{u}_{xy} = [1,1,0]$, $\hat{u}_{yz} = [0,1,1]$, e $\hat{u}_{xz} = [1,0,1]$ são os vetores unitários dos espaços xy , yz , and xz , respectivamente. Então, os parâmetros do deslocamento do centro de massa são definidos como $D_1 = \text{Min}(D_{xy}, D_{yz}, D_{xz})$, $D_2 = \text{Mediana}(D_{xy}, D_{yz}, D_{xz})$ e $D_3 = \text{Max}(D_{xy}, D_{yz}, D_{xz})$. Na etapa de redução de falsos positivos, os requisitos mais rígidos são alcançados para os parâmetros de deslocamento para os menores valores.

Há ainda atributos estatísticos que são importantes para a descrição de candidatos a nódulos, tais como a média e o desvio padrão. Então, considerando-se que todos os *voxels* candidatos são amostras da imagem, os parâmetros μ e σ são a média e o desvio padrão dos níveis de cinza desses *voxels*. Como ocorre no parâmetro deslocamento do centro de massa, os menores valores de desvio padrão são aqueles que alcançam os requisitos mais rígidos para este parâmetro no estágio de redução falso positivo.

Alguns atributos do descritor fornecem uma representação quantitativa para a consistência do candidato, são eles os parâmetros de conectividade KV e KP . O primeiro, KV , representa quantidade de *voxels* de um candidato que estão tridimensionalmente conectados com conectividade 6, ou $N_6(v)$, a exatamente 6 vizinhos que também pertencem ao candidato, conforme mostrado na Figura 41. O parâmetro KP é o número de *voxels* do candidato que

possuem exatamente 8 conexões bidimensionais de conectividade 8, ou $N_8(p)$, em pelo menos um dos planos xy , yz ou xz , com vizinhos que também pertencem ao candidato. Para ilustrar o cálculo deste parâmetro, na Figura 42 são mostradas três estruturas nas quais o valor de KP é igual a 1.

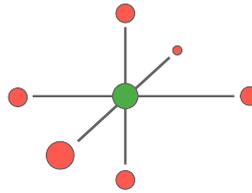


Figura 41 – 6 vizinhos conectados com conectividade tridimensional 6.

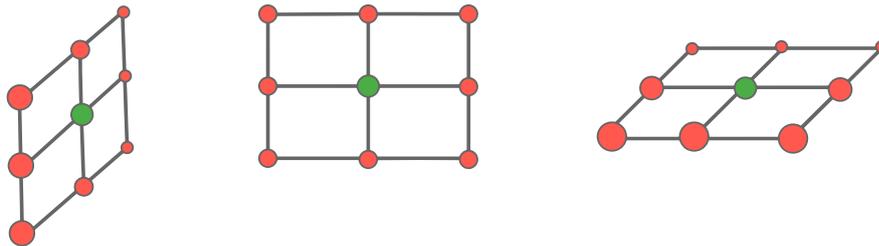


Figura 42 – 8 vizinhos conectados com conectividade bidimensional 8 nos planos xz , yz , ou xy .

Finalmente, o atributo *wall Flag* consiste em uma variável binária que identifica os candidatos que tocam as paredes do pulmão. Nesse sentido, um procedimento para definir a região da parede expande a imagem auxiliar de segmentação M obtida no Capítulo 3.2 e compara sua versão expandida com a original. A parede pulmonar é definida como a região onde ambas as imagens valem 1. Para realizar a expansão da imagem auxiliar de segmentação, aplica-se uma operação de erosão *slice a slice* a esta imagem. O elemento estruturante utilizado é uma matriz 3×3 em que os 1's são dispostos em formato de cruz, sendo os outros *pixels* iguais a 0. Em seguida, compara-se o resultado da operação anterior com a imagem auxiliar de segmentação original para evidenciar as paredes do pulmão. Tal comparação é realizada por meio de uma operação lógica *and bit a bit* a imagem auxiliar de segmentação e o resultado da operação de erosão. Desse modo, o valor 1 é atribuído ao parâmetro *Wall Flag* para candidatos com qualquer *voxel* que pertença à região definida como parede pulmonar. Nas seções de redução de falsos positivos, o algoritmo proposto utiliza este atributo para dividir os candidatos em subgrupos de

acordo com sua proximidade às paredes pulmonares.

3.4.2 Primeiro Estágio de Redução dos Falsos Positivos

Para o primeiro estágio de redução falsos positivos, os candidatos a nódulos são divididos em 16 grupos, sendo cada grupo proveniente de um resultado de pré-processamento diferente. Em seguida, são eliminados todos os candidatos a nódulo com valores de atributo fora de intervalo para cada grupo. Desse modo, cada um desses grupos se divide novamente em 16 subgrupos de acordo com o volume dos candidatos, sendo os intervalos permitidos para cada atributo definidos por subgrupo. Portanto, em um conjunto de candidatos oriundo de um dado tipo de pré-processamento de candidatos, existem 16 subgrupos, os 12 primeiros contêm candidatos localizados dentro da cavidade pulmonar e os 4 últimos contêm candidatos apontados como pertencentes às paredes do pulmão pelo atributo *Wall Flag*.

Os valores de volume mínimo e máximo que delimitam cada subgrupo são mostrados nas Tabelas 2 e 3. Nestas Tabelas, os números 1, 2, 3 e 4 nas colunas referem-se aos tipos de pré-processamento que usam janelas de tamanho 1 e fatores de limiar k iguais a 0,5, 1, 2 e 3, respectivamente. Da mesma forma, os tipos de pré-processamento 5, 6, 7 e 8 usam os mesmos fatores k , mas janelas de dimensão 3. Aplica-se o mesmo raciocínio aos tipos de pré-processamento de dimensão de janela 5.

Como muitos nódulos, por serem volumosos, apenas podem ser isolados utilizando-se janelas de dimensão 7, uma parametrização criteriosa posterior a estes tipos de pré-processamento é uma estratégia para conservar os verdadeiros positivos. Essa estratégia é realizada separando-se duas colunas nas Tabelas 2 e 3, especificamente para nódulos oriundos de pré-processamentos com janelas de dimensão $w = 7$, uma coluna a mais que se utiliza para outras dimensões de janela. No intuito de tratar adequadamente essa quantidade elevada de candidatos, os conjuntos de candidatos provenientes dos pré-processamentos de tipos 13 e 14 possuem faixas de volume diferentes dos tipos 15 e 16, todos filtrados com de janelas de tamanho 7. A escolha do intervalo de valores é obtida empiricamente e visa balancear o número de candidatos nos diferentes subgrupos provenientes de um mesmo tipo de pré-processamento.

Uma vez estabelecido o intervalo de volume de cada subgrupo, é definida a matriz que regula o primeiro estágio da redução falso positivo. A ideia principal é estabelecer limites no domínio de atributos para cada subgrupo e eliminar candidatos fora desses domínios. Desse modo, os valores preliminares limites para os domínios dos atributos são dados por I_{MIN} , I_{MAX} ,

Tabela 2 – Volumes Mínimos

Sub-grupo	Tipos 1 a 4 ($w = 1$)	Tipos 5 a 8 ($w = 3$)	Tipos 9 a 12 ($w = 5$)	Tipos 13 a 14 ($w = 7$)	Tipos 15 a 16 ($w = 7$)
1	10	15	35	60	60
2	15	50	150	100	1200
3	20	100	200	250	1500
4	25	115	250	450	2000
5	30	130	300	800	2500
6	40	145	350	1000	3500
7	50	160	400	1250	4000
8	75	175	450	1500	4500
9	100	200	500	1750	5000
10	125	220	750	2000	6000
11	150	300	1000	2500	9500
12	200	500	2000	5000	10000
13	10	15	35	60	60
14	15	35	60	150	150
15	20	50	75	300	300
16	25	100	100	600	600

Tabela 3 – Volumes Máximos

Sub-grupo	Tipos 1 a 4 ($w = 1$)	Tipos 5 a 8 ($w = 3$)	Tipos 9 a 12 ($w = 5$)	Tipos 13 a 14 ($w = 7$)	Tipos 15 a 16 ($w = 7$)
1	15	50	150	100	1200
2	20	100	200	250	1500
3	25	115	250	450	2000
4	30	130	300	800	2500
5	40	145	350	1000	3500
6	50	160	400	1250	4000
7	75	175	450	1500	4500
8	100	200	500	1750	5000
9	125	220	750	2000	6000
10	150	300	1000	2500	9500
11	200	500	2000	5000	10000
12	1001	4000	4000	20000	20000
13	15	35	60	150	150
14	20	50	75	300	300
15	25	100	100	600	600
16	41	300	500	3000	3000

S_{MIN} e S_{MAX} . Portanto, dois intervalos são definidos inicialmente para cada atributo, o inferior $[I_{MIN}, I_{MAX}]$ e o superior $[S_{MIN}, S_{MAX}]$. Estes valores são posteriormente refinados e utilizados para gerar a matriz de redução de falsos positivos. A escolha dos limites é experimental e visa garantir que a maioria dos nódulos isolados atinja a fase de classificação. Os intervalos

preliminares inferiores e superiores para subgrupos de todos os tipos de pré-processamento são mostrados nas Tabelas 4, 5, 6 e 7.

Tabela 4 – Limites Inferiores dos Intervalos Inferiores

	<i>L</i>	<i>P</i>	<i>P_{box}</i>	<i>B₁</i>	<i>B₂</i>	<i>B₃</i>	<i>C₁</i>	<i>C₂</i>	<i>C₃</i>	<i>D₁</i>	<i>D₂</i>	<i>D₃</i>	<i>M</i>	<i>S</i>	<i>K_P</i>	<i>K_V</i>
Subgrupos 1 a 4	0	0,4	0	0	1,5	1	0,3	0,3	0,3	15	30	30	0,3	0,2	0	0
Subgrupos 5 a 8	0	0,4	0	0	1,5	1	0,3	0,3	0,3	15	30	30	0,3	0,2	0	0
Subgrupos 9 a 12	0	0,4	0	0	1,5	1	0,3	0,3	0,3	15	30	30	0,3	0,2	0	0
Subgrupos 13 a 16	0	0,4	0	0	1,5	1	0,3	0,3	0,3	15	30	30	0,3	0,2	0	0

Tabela 5 – Limites Superiores dos Intervalos Inferiores

	<i>L</i>	<i>P</i>	<i>P_{box}</i>	<i>B₁</i>	<i>B₂</i>	<i>B₃</i>	<i>C₁</i>	<i>C₂</i>	<i>C₃</i>	<i>D₁</i>	<i>D₂</i>	<i>D₃</i>	<i>M</i>	<i>S</i>	<i>K_P</i>	<i>K_V</i>
Subgrupos 1 a 4	6	0,5	0,3	1,25	2,25	1,5	0,8	0,9	0,6	5	20	20	0,6	0,1	10 ²	10
Subgrupos 5 a 8	6	1	0,3	1,25	2,25	1,5	0,8	0,9	0,6	5	20	20	0,6	0,1	10 ²	10
Subgrupos 9 a 12	6	1	0,3	1,25	2,25	1,5	0,8	0,9	0,6	5	20	20	0,6	0,1	10 ²	10
Subgrupos 13 a 16	6	1	0,3	1,25	5	5	0,75	0,9	0,6	5	20	20	0,6	0,1	10 ²	10

Tabela 6 – Limites Inferiores dos Intervalos Superiores

	<i>L</i>	<i>P</i>	<i>P_{box}</i>	<i>B₁</i>	<i>B₂</i>	<i>B₃</i>	<i>C₁</i>	<i>C₂</i>	<i>C₃</i>	<i>D₁</i>	<i>D₂</i>	<i>D₃</i>	<i>M</i>	<i>S</i>	<i>K_P</i>	<i>K_V</i>
Subgrupos 1 a 4	10	1,5	0,7	2,25	4,5	7,5	0,9	0,95	0,8	2,5	10	10	0,8	0,05	10 ³	10 ²
Subgrupos 5 a 8	10	1,5	0,4	1,5	3	2,5	0,9	0,95	0,8	2,5	10	10	0,8	0,05	10 ³	10 ²
Subgrupos 9 a 12	10	1,5	0,4	1,5	3	2,5	0,9	0,95	0,8	2,5	10	10	0,8	0,05	10 ³	10 ²
Subgrupos 13 a 16	10	1,5	0,4	1,5	6	2,5	0,9	0,95	0,8	2,5	10	10	0,8	0,05	10 ³	10 ²

Tabela 7 – Limites Superiores dos Intervalos Superiores

	<i>L</i>	<i>P</i>	<i>P_{box}</i>	<i>B₁</i>	<i>B₂</i>	<i>B₃</i>	<i>C₁</i>	<i>C₂</i>	<i>C₃</i>	<i>D₁</i>	<i>D₂</i>	<i>D₃</i>	<i>M</i>	<i>S</i>	<i>K_P</i>	<i>K_V</i>
Subgrupos 1 a 4	40	2	0,8	2,5	7	30	1	1	1,5	0	0	0	1	0	10 ³	10 ⁴
Subgrupos 5 a 8	40	2	0,8	2,5	7	30	1	1	1,5	0	0	0	1	0	10 ³	10 ⁴
Subgrupos 9 a 12	40	2	0,8	2,5	7	30	1	1	1,5	0	0	0	1	0	10 ³	10 ⁴
Subgrupos 13 a 16	40	2	0,8	2,5	7	30	1	1	1,5	0	0	0	1	0	10 ³	10 ⁴

Após a definição inicial das faixas de atributos inferior e superior, uma fração dos nódulos que se encontram em tais intervalos é utilizada para construir a matriz que regula o primeiro estágio de redução falsos positivos conforme detalhado a seguir. Nesse sentido, os limites externos dos valores dos atributos definem um paralelepípedo de 16 dimensões no domínio dos atributos para cada subgrupo. Desse modo, o lugar geométrico em que pelo menos 15 dos 16 atributos assumem valores limite corresponde às arestas que partem de cada vértice do paralelepípedo. Em outras palavras, regiões próximas às arestas do paralelepípedo possuem pelo menos 15 dos 16 atributos pertencentes ao intervalo inferior ou ao intervalo superior mostrados

nas Tabelas 4, 5, 6 e 7. No caso dos tipos de pré-processamento que utilizam janelas de dimensão 1, em que a quantidade de nódulos detectados é normalmente menor, a condição de contorno ora descrita é flexibilizada e pelo menos 12 atributos devem possuir valores limítrofes para que tal condição seja obedecida. Após definir as condições de contorno iniciais, refina-se o ajuste das faixas permitidas para cada atributo usando os valores de atributos mínimos e máximos apenas dos nódulos que se encontram próximos às arestas do paralelepípedo, os quais correspondem a menos de 25% do total de nódulos. Esses valores limites são armazenados na matriz usada no primeiro estágio da redução de falsos positivos e definem a faixa de atributos tolerável para os 16 subgrupos de cada tipo de pré-processamento.

Como o presente algoritmo contém uma sucessão de regras heurísticas, é desejável considerar possíveis variações de desempenho face sua aplicação a outros bancos de dados ou a imagens oriundas de diferentes equipamentos. Com a finalidade de tornar o processo de redução de falsos positivos mais independente de tais fatores, procede-se ao relaxamento dos intervalos dos atributos. Este relaxamento é obtido realizando-se uma suavização e um arredondamento dos valores limites de cada atributo na matriz de redução de falsos positivos. Uma janela mediana móvel de dimensão 3 é aplicada sobre os valores limites de um mesmo atributo presentes em cada coluna da matriz ao longo dos diferentes subgrupos enumerados nas linhas. Esta operação torna a redução de falsos positivos menos rígida e menos dependente do conjunto de dados.

Já o processo de arredondamento que ocorre após a suavização dos valores dos atributos da Tabela também torna os limites dos atributos mais tolerantes em todos os casos. Como cada atributo possui suas próprias faixas típicas de valores, a precisão de arredondamento depende do tipo de atributo. Nomeadamente, a precisão do arredondamento é 0,5 para o parâmetro M , 0,25 para B_1, B_2, B_3, D_1, D_2 e D_3 , 0,1 para C_1, C_2 e C_3 , 0,05 para P e S e 0,005 para P_{BOX} . Para KV e KP , a precisão do arredondamento é 1, 2, 5, 5, 10 e 25 para volumes nos intervalos $[10, 20]$, $[20, 50]$, $[50, 100]$, $[50, 100]$ e superior a 100, respectivamente.

Após a utilização de matriz de redução de falsos positivos, reduz-se apenas o excesso de falsos positivos. Desse modo, ainda há a necessidade de um segundo estágio de redução de falsos positivos para melhorar o processo de detecção de nódulos.

3.4.3 Segundo Estágio de Redução de Falsos Positivos

Como após a primeira etapa de redução de falsos positivos ainda há uma grande quantidade de candidatos, estes candidatos são submetidos a uma nova seleção para a fase de

classificação por meio de operações Booleanas. Desse modo, as variáveis dos atributos devem ser transformadas em variáveis Booleanas, utilizando-se um descritor binário. A exemplo da etapa anterior, a seleção dos candidatos ocorre dividindo-se estes candidatos em subgrupos e processando-se diferentes operações para cada subgrupo.

No segundo estágio da redução falsos positivos, os candidatos a nódulos são divididos de acordo com seu tipo de pré-processamento de origem, sendo sete grupos ao todo. Os grupos 1, 2 e 3 contêm candidatos provenientes de tipos de pré-processamento que usam janelas de dimensões 1, 3 e 5, respectivamente. Os grupos 4, 5, 6 e 7 incluem candidatos oriundos do pré-processamento com janela de dimensão $w = 7$ e fator de limiar k de 0,5, 1, 2 e 3, respectivamente. Similarmente à etapa anterior, existem 16 subgrupos para cada tipo de pré-processamento divididos de acordo com as dimensões dos candidatos, sendo os valores limite de volume para cada subgrupo os mesmos mostrados nas Tabelas 2 e 3. As operações Booleanas que realizam a seleção dos nódulos são aplicadas a cada um desses subgrupos. Ainda como na primeira etapa, os primeiros 12 dos 16 subgrupos contêm candidatos que se localizam dentro da cavidade pulmonar e os 4 últimos contêm candidatos que tocam as paredes pulmonares.

Então, para cada conjunto de 4 subgrupos, uma função Booleana simples define o critério de eliminação dos candidatos. Desse modo, um novo descritor de 16 *bits* representa um candidato, em que cada *bit* X_i se refere a um dos atributos A_i definidos na Seção 3.4.1. A correspondência entre os atributos e seus respectivos *bits* no descritor é mostrada na Tabela 8.

Tabela 8 – Descritor Binário

i	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
A_i	L	P	P_{box}	B_1	B_2	B_3	C_1	C_2	C_3	D_1	D_2	D_3	μ	σ	KV	KP
X_i	l	p	$pbox$	b_1	b_2	b_3	c_1	c_2	c_3	d_1	d_2	d_3	m	s	kv	kp

Uma vez conhecidos os valores dos atributos de cada candidato, a operação relacional a seguir define o valor do *bit* X_i relacionado ao atributo A_i como $X_i = |A_i - A_{iMAX}| \leq |A_i - A_{iMIN}|$, em que A_i é a medida do i -ésimo atributo e A_{iMIN} e A_{iMAX} são os valores limites definidos para este atributo. Ou seja, $X_i = 1$, caso a comparação $|A_i - A_{iMAX}| \leq |A_i - A_{iMIN}|$ produza um valor lógico verdadeiro e 0, caso contrário. Tais limites são os mesmos armazenados na matriz de redução de falsos positivos na Seção e dependem do subgrupo de origem do candidato. Por exemplo, para $i = 15$, o atributo A_i corresponde a L , sendo este valor comparado ao valor mínimo e ao valor máximo para L definido na matriz de redução de falsos positivos.

Tabela 9 – Primeira variável Booleana

M_{ij}	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$
$i = 1$	$\overline{b_2}$	$\overline{d_1}$	$\overline{d_3}$	\overline{l}	$\overline{c_2}$	$\overline{c_2}$	$\overline{c_1}$
$i = 2$	$\overline{b_2}$	$\overline{d_1}$	$\overline{d_3}$	\overline{l}	$\overline{c_2}$	$\overline{c_2}$	$\overline{c_1}$
$i = 3$	$\overline{b_2}$	$\overline{d_1}$	$\overline{d_3}$	\overline{l}	$\overline{c_2}$	$\overline{c_2}$	$\overline{c_1}$
$i = 4$	$\overline{b_2}$	$\overline{d_1}$	$\overline{d_3}$	\overline{l}	$\overline{c_2}$	$\overline{c_2}$	$\overline{c_1}$
$i = 5$	$\overline{b_2}$	$\overline{d_3}$	\overline{p}	$\overline{d_3}$	$\overline{d_1}$	\overline{p}	\overline{l}
$i = 6$	$\overline{b_2}$	$\overline{d_3}$	\overline{p}	$\overline{d_3}$	$\overline{d_1}$	\overline{p}	\overline{l}
$i = 7$	$\overline{b_2}$	$\overline{d_3}$	\overline{p}	$\overline{d_3}$	$\overline{d_1}$	\overline{p}	\overline{l}
$i = 8$	$\overline{b_2}$	$\overline{d_3}$	\overline{p}	$\overline{d_3}$	$\overline{d_1}$	\overline{p}	\overline{l}
$i = 9$	\overline{l}	\overline{pbox}	\overline{l}	$\overline{b_2}$	\overline{pbox}	\overline{l}	\overline{l}
$i = 10$	\overline{l}	\overline{pbox}	\overline{l}	$\overline{b_2}$	\overline{pbox}	\overline{l}	\overline{l}
$i = 11$	\overline{l}	\overline{pbox}	\overline{l}	$\overline{b_2}$	\overline{pbox}	\overline{l}	\overline{l}
$i = 12$	\overline{l}	\overline{pbox}	\overline{l}	$\overline{b_2}$	\overline{pbox}	\overline{l}	\overline{l}
$i = 13$	\overline{l}	\overline{m}	$\overline{d_1}$	$\overline{d_1}$	\overline{m}	$\overline{d_1}$	\overline{s}
$i = 14$	\overline{l}	\overline{m}	$\overline{d_1}$	$\overline{d_1}$	\overline{m}	$\overline{d_1}$	\overline{s}
$i = 15$	\overline{l}	\overline{m}	$\overline{d_1}$	$\overline{d_1}$	\overline{m}	$\overline{d_1}$	\overline{s}
$i = 16$	\overline{l}	\overline{m}	$\overline{d_1}$	$\overline{d_1}$	\overline{m}	$\overline{d_1}$	\overline{s}

Tabela 10 – Segunda Variável Booleana

N_{ij}	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$
$i = 1$	\overline{m}	$\overline{d_3}$	\overline{s}	$\overline{d_2}$	$\overline{d_3}$	$\overline{d_3}$	\overline{m}
$i = 2$	\overline{m}	$\overline{d_3}$	\overline{s}	$\overline{d_2}$	$\overline{d_3}$	$\overline{d_3}$	\overline{m}
$i = 3$	\overline{m}	$\overline{d_3}$	\overline{s}	$\overline{d_2}$	$\overline{d_3}$	$\overline{d_3}$	\overline{m}
$i = 4$	\overline{m}	$\overline{d_3}$	\overline{s}	$\overline{d_2}$	$\overline{d_3}$	$\overline{d_3}$	\overline{m}
$i = 5$	$\overline{d_2}$	$\overline{c_2}$	$\overline{d_1}$	$\overline{d_1}$	\overline{m}	$\overline{d_3}$	\overline{l}
$i = 6$	$\overline{d_2}$	$\overline{c_2}$	$\overline{d_1}$	$\overline{d_1}$	\overline{m}	$\overline{d_3}$	\overline{l}
$i = 7$	$\overline{d_2}$	$\overline{c_2}$	$\overline{d_1}$	$\overline{d_1}$	\overline{m}	$\overline{d_3}$	\overline{l}
$i = 8$	$\overline{d_2}$	$\overline{c_2}$	$\overline{d_1}$	$\overline{d_1}$	\overline{m}	$\overline{d_3}$	\overline{l}
$i = 9$	\overline{pbox}	\overline{s}	$\overline{d_3}$	$\overline{b_3}$	$\overline{d_3}$	$\overline{d_3}$	$\overline{d_3}$
$i = 10$	\overline{pbox}	\overline{s}	$\overline{d_3}$	$\overline{b_3}$	$\overline{d_3}$	$\overline{d_3}$	$\overline{d_3}$
$i = 11$	\overline{pbox}	\overline{s}	$\overline{d_3}$	$\overline{b_3}$	$\overline{d_3}$	$\overline{d_3}$	$\overline{d_3}$
$i = 12$	\overline{pbox}	\overline{s}	$\overline{d_3}$	$\overline{b_3}$	$\overline{d_3}$	$\overline{d_3}$	$\overline{d_3}$
$i = 13$	$\overline{d_1}$	$\overline{d_2}$	\overline{m}	$\overline{d_3}$	$\overline{d_1}$	\overline{m}	$\overline{d_1}$
$i = 14$	$\overline{d_1}$	$\overline{d_2}$	\overline{m}	$\overline{d_3}$	$\overline{d_1}$	\overline{m}	$\overline{d_1}$
$i = 15$	$\overline{d_1}$	$\overline{d_2}$	\overline{m}	$\overline{d_3}$	$\overline{d_1}$	\overline{m}	$\overline{d_1}$
$i = 16$	$\overline{d_1}$	$\overline{d_2}$	\overline{m}	$\overline{d_3}$	$\overline{d_1}$	\overline{m}	$\overline{d_1}$

Finalmente, os candidatos provenientes do tipo de filtro j e pertencentes ao sub-grupo i são eliminados quando os atributos M_{ij} e N_{ij} são ambos iguais a zero, em que M e N são as matrizes mostradas nas Tabelas 9 e 10. Em outras palavras, as regras de eliminação para os candidatos são funções lógicas AND de 2 variáveis, em que a primeira variável é mostrada na Tabela 9, a segunda variável figura na Tabela 10 e estas variáveis estão localizadas em linhas e

colunas definidas pelo tamanho do subgrupo e pelo tipo de pré-processamento, respectivamente.

Por exemplo, se um candidato é oriundo de um pré-processamento com janela de dimensão 5, o valor de j nas Tabelas 9 e 10 é 3. Caso este candidato possua volume igual a 425 *voxels* e esteja no interior da cavidade pulmonar, o mesmo pertence ao subgrupo $i = 7$, cujos valores limites estão entre 400 e 450 *voxels*, de acordo com as Tabelas 2 e 3. Desse modo, agora de acordo com as Tabelas 9 e 10 em $i = 7$ e $j = 3$ a operação lógica \bar{p} AND \bar{d}_1 seleciona este candidato para a fase de classificação resultando em 1 ou 0, respectivamente.

3.5 Classificação

Uma vez que o estágio de redução falsos positivos seleciona os candidatos a nódulos, a resposta de uma Rede Neural Convolutiva (CNN) é utilizada para estimar a probabilidade de um candidato ser nódulo ou não. De acordo com essa resposta, é possível estabelecer limites de probabilidade que podem alterar a sensibilidade e a taxa média de falsos positivos. Geralmente, um ajuste de limiar de probabilidade que aumenta a sensibilidade também aumenta a taxa média de falsos positivos (RIBEIRO, 2021).

Os candidatos são dispostos como entradas da rede CNN no formato de volumes 32x32x32 tomados a partir de seus centróides. A arquitetura da CNN utilizada nesta tese, a qual é explicada a seguir, é mostrada na Tabela 11 (RIBEIRO, 2021).

Tabela 11 – Arquitetura da CNN

Camada	Quantidade de filtros	Tamanho do Filtro	Função de Ativação	Taxa de Descarte	Número de Neurônios
Conv3D	16	3x3x3	-	-	-
Conv3D	8	3x3x3	Relu	-	-
MaxPooling3D	-	2x2x2	-	-	-
Dropout	-	-	-	0.1	-
Conv3D	8	3x3x3	Relu	-	-
MaxPooling3D	-	2x2x2	-	-	-
Dropout	-	-	-	0.1	-
Conv3D	8	3x3x3	Relu	-	-
MaxPooling3D	-	2x2x2	-	-	-
Dropout	-	-	-	0.25	-
Faltn	-	-	-	-	-
Dense	-	-	Relu	-	8
Dense	-	-	Softmax	-	2

Nesta CNN, a camada *Conv3D* atualiza os pesos do *kernel* via retropropagação para realçar o objeto de interesse. Por sua vez, a camada *Maxpooling* reduz a quantidade de

parâmetros para as camadas seguintes e, desse modo, diminui as chances da rede apenas aprender as características do conjunto de treinamento (BROWNLEE, 2019). Nesta tese, as camadas de *Maxpooling* são tridimensionais devido a natureza das imagens de entrada e obedecem às dimensões das janelas ora definidas.

No sentido de reduzir a ocorrência de *overfitting*, a camada *Dropout* elimina uma fração das amostras de treinamento em cada época após as camadas de *Maxpooling3D* e *Conv3D*. Nesse sentido, o parâmetro da taxa de descarte na Tabela 11 é a fração de *voxels* descartados na camada *Dropout*. Portanto, um dado volume de entrada de $32 \times 32 \times 32$ *voxels* passa por um bloco de extração de características sequencial de 3 etapas com uma ou mais camadas *Conv3D*, *Maxpooling* e *Dropout*, conforme mostrado na Figura 43. Após a execução dos blocos de extração de características mostrados na Tabela 11, uma camada *Flatten* redimensiona o volume alinhando os *voxels* resultantes e em um vetor unidimensional. Este vetor é a entrada da camada densa, a qual consiste em um *perceptron* multicamadas com 8 neurônios e uma função de ativação *ReLu*, cuja vantagem é a simplicidade de otimização devido à sua semelhança com a função identidade (CHOLLET, 2017). Por fim, a última camada densa define o resultado da classificação com 2 neurônios e uma função de ativação *Softmax*.

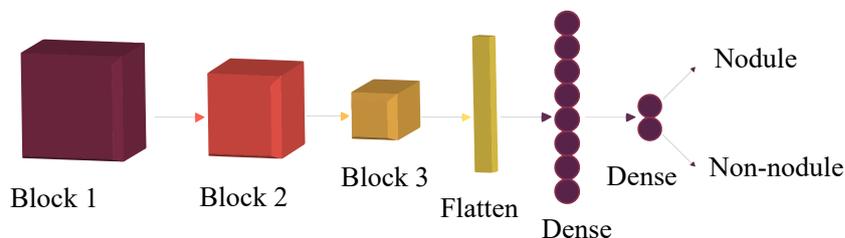


Figura 43 – Arquitetura da CNN empregada no processo de classificação.

O treinamento dessa rede usa o método de validação cruzada $k - fold$ com $k = 10$ para 100 épocas de treinamento. No entanto, na hipótese de o erro não diminuir em pelo menos 0,05 em até 25 épocas, o processo reinicia com outros valores de peso randômicos. Este processo de treinamento aplica a otimização estocástica Adam, com 80% das imagens utilizadas para treinamento e 20% para teste (KINGMA; BA, 2015).

Com a finalidade de gerar mais dados de qualidade admissível para treinamento da rede, são aplicadas técnicas de *data augmentation* de rotação e translação. Assim, mais imagens são geradas a partir de rotações de 90, 180 e 270 graus no plano axial e de operações de espelhamento. Além disso, translações com passos de 4 *voxels*, combinando variações de x ,

y e z, geram candidatos para cada volume que haja sofrido rotação e espelhamento. Para fins de treinamento, são considerados verdadeiros negativos 400 candidatos por exame amostrados aleatoriamente dentre aqueles que estejam a uma distância euclidiana superior a 40 *voxels* de qualquer verdadeiro positivo segundo o padrão de referência. Para fins de contabilização do resultado da classificação, considera-se que um candidato a nódulo produz um verdadeiro positivo *True Positive* (TP) quando seu centro de massa está a uma distância euclidiana inferior a 10 *voxels* do centróide de uma lesão considerada como nódulo no padrão de referência (RIBEIRO, 2021).

Esse método não reutiliza atributos da rede ou aplica qualquer técnica de transferência de aprendizado, como adaptação de domínio, para compensar uma possível escassez de dados de treinamento.

Uma vez especificados a arquitetura e as propriedades da rede utilizada como ferramenta para a classificação de candidatos em nódulos ou não-nódulos, procede-se à apresentação dos principais resultados e às discussões sobre os mesmos.

4 RESULTADOS E DISCUSSÕES

Neste Capítulo, resumem-se os resultados e as discussões para evidenciar a aplicabilidade do novo método proposto nesta tese. Para tanto, utiliza-se a definição de verdadeiro positivo (TP) fornecida na Secção 3.5. Por sua vez, um falso positivo consiste em qualquer lesão apontada como nódulo pelo método que não seja um verdadeiro positivo. De modo similar, um nódulo de referência produz um falso negativo (FN) quando o centro de massa de nenhum candidato está menos de 10 *voxels* de distância de seu centróide. Desse modo, a sensibilidade do método é dada por $SENS = \frac{TP}{TP+FN}$.

No sentido de avaliar o desempenho da técnica de redução de falsos positivos somente, o número de candidatos e a sensibilidade antes e depois da aplicação do método de Redução Booleana de Falsos Positivos são mostrados na Tabela 12. Esse método produz uma redução de 90,68% no número de candidatos, seguido por uma consequente redução na sensibilidade de 3,88%, 2,84% e 2,33%, considerando-se as anotações de pelo menos 1, 2 e 3 radiologistas respectivamente. Dessa forma, reduz-se acentuadamente o número de candidatos sem uma perda significativa de sensibilidade.

Tabela 12 – Performance da Redução Booleana de Falsos Positivos

	Antes da Redução de FP	Depois da Redução de FP
Quantidade de candidatos	6.062.878	565.059
Sensibilidade (3 anotações)	0,984	0,961
Sensibilidade (2 anotações)	0,983	0,955
Sensibilidade (1 anotação)	0,979	0,941

O fato de o Vetor de Filtros de Pré-processamento isolar 98,4% e 97,9% de todos os nódulos para as anotações de pelo menos um especialista ou da maioria dos mesmos, respectivamente, também é mostrado na Tabela 12. Nesse caso, são utilizados para assumir que um nódulo está isolado os mesmos critérios adotados na etapa de classificação para contabilizar um verdadeiro positivo.

A curva FROC *Free-Response Operating Characteristic* para resultados de classificação com e sem o estágio prévio de redução falsos positivos é mostrado na Figura 44 utilizando-se o padrão-ouro que considera as anotações de pelo menos 3 especialistas. A generalização e estabilização dos resultados é realizada pelo método de validação cruzada *k-fold* com $k = 10$ e 100 iterações. A curva FROC sempre mostra valores de sensibilidade mais altos para classifica-

ções que utilizam a Redução Booleana de Falsos Positivos como uma etapa prévia em relação àquelas que não utilizam esta técnica dada qualquer taxa de falsos positivos. Por exemplo, a CNN produz 92,75% de sensibilidade com média de 8 falsos positivos por exame quando o método de redução de falsos positivos proposto é aplicado antes da classificação. Por outro lado, a sensibilidade diminui para 80,19%, dada a mesma taxa média de falsos positivos, quando não há etapa de redução de FP. Tais resultados mostram que o método de redução de falsos positivos proposto melhora efetivamente o desempenho da classificação.

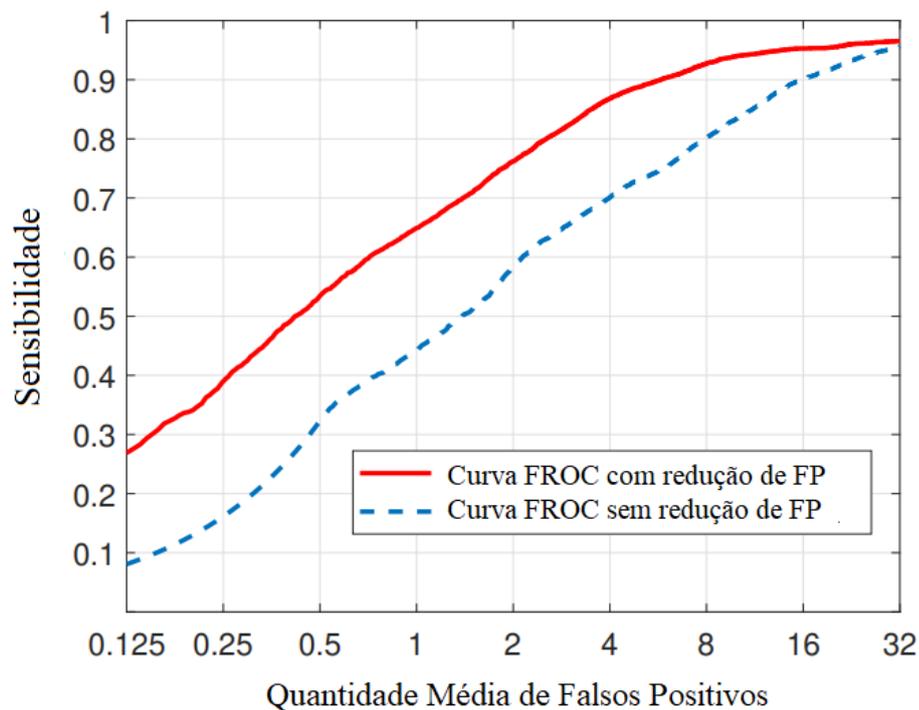


Figura 44 – Curvas FROC com e sem Redução Booleana de Falsos Positivos.

Para situar esta tese no contexto da literatura pesquisada, a Tabela 13 resume seus resultados juntamente com os resultados de outras pesquisas relevantes na detecção de nódulos pulmonares. Esta Tabela está dividida em 3 conjuntos de linhas. O primeiro conjunto de linhas da tabela lista os resultados que apenas contabilizam lesões consideradas nódulos pela maioria dos radiologistas (pelo menos 3 anotações dentre um total de 4 especialistas). Por outro lado, os resultados apresentados nos 2 últimos conjuntos de linhas consideram como verdadeiros positivos lesões marcadas como "nódulo" por pelo menos um radiologista ou pelo menos dois radiologistas, respectivamente.

Para resultados de detecção, baixos níveis de consenso tendem a dificultar o aumento da sensibilidade para uma taxa média de falsos positivos determinada. Isso ocorre porque

Tabela 13 – Resultados para técnicas de detecção automática de nódulos.

	Base	Número de nódulos	Número de exames	Consenso	SENS $\frac{FP}{exam} = 1$	SENS $\frac{FP}{exam} = 2$	SENS $\frac{FP}{exam} = 3$	SENS $\frac{FP}{exam} = 4$	SENS $\frac{FP}{exam} = 8$	SENS $\frac{FP}{exam} = 16$
(TAN <i>et al.</i> , 2020)	LUNA16	888	1186	3/4	0,975	0,982	–	0,988	0,990	–
(ZHANG <i>et al.</i> , 2018)	LUNA16	888	1186	3/4	0,949	0,965	–	0,972	0,976	–
(ZHENG <i>et al.</i> , 2019)	LUNA16	888	1186	3/4	0,927	0,942	–	0,948	0,953	–
(PEZESHK <i>et al.</i> , 2018)	LUNA16	888	1186	3/4	0,865	0,907	–	0,938	0,952	–
Método proposto (3 anotações)	LIDC-IDRI	607	1191	3/4	0,649	0,761	0,824	0,868	0,927	0,952
(GUO; LI, 2012)	LIDC	85	111	2/4	–	0,800 ¹	0,850 ¹	–	–	–
(LU <i>et al.</i> , 2015)	LIDC	294	631	2/4	–	–	0,852 ¹	–	–	–
(KURUVILLA; GUNAVATHI, 2014)	LIDC/ hospital scans	155	–	2/4	–	0,820	–	–	–	0,900 ¹
Método proposto (2 anotações)	LIDC-IDRI	684	1603	2/4	0,578	0,696	0,758	0,802	0,878	0,916²
(MESSAY <i>et al.</i> , 2010)	LIDC/ base da UTBM	84	143	1/4	–	–	0,804	–	–	–
Método proposto (1 anotação)	LIDC-IDRI	768	2287	1/4	0,482	0,592	0,661	0,706	0,791	0,845

[1] As taxas médias de falsos positivos para sensibilidades marcadas com “1” são aproximações para os valores constantes de FP/exam em cada coluna as quais visam encaixar os resultados na tabela adequadamente. Em (GUO; LI, 2012), SENS = 0,800 corresponde a FP/exam = 1,9. Em (GUO; LI, 2012), SENS = 0,850 corresponde a FP/exam = 2,7. Em (LU *et al.*, 2015), SENS = 0,852 corresponde a FP/exam = 3,13. Em (KURUVILLA; GUNAVATHI, 2014), SENS = 0,900 corresponde a FP/exam = 15.

[2] Para nosso método, SENS = 0,916 corresponde a FP/exam = 15 para uma comparação razoável com o resultado abaixo.

algoritmos que classificam lesões não consensuais como sendo nódulos tendem a ser permissivos também com falsos positivos, dificultando a detecção dos nódulos. Além do nível de consenso, outras informações são necessárias para avaliar o desempenho de um sistema CAD. Desse modo, a Tabela 13 mostra o número de nódulos, de exames e os conjuntos de dados usados em cada pesquisa para prover uma comparação mais clara entre os métodos. Nos resultados, cada medida de sensibilidade corresponde a um valor de taxa média de falsos positivos de acordo com o que é reportado nas pesquisas listadas.

De acordo com essa tabela, o método proposto atinge uma sensibilidade de 0,927 para uma média de FP/exame = 8 e referência de verdadeiros positivos com anotações de pelo menos radiologistas. Este resultado é comparável aos encontrados em outros trabalhos que utilizam padrões-ouro com o mesmo nível de consenso (TAN *et al.*, 2020; ZHANG *et al.*, 2018; ZHENG *et al.*, 2019; PEZESHK *et al.*, 2018). Considerando os padrões de referência que contabilizam como nódulos lesões que tenham sido marcadas por pelo menos 1 ou 2 radiologistas, a referida tabela mostra que o método proposto obtém sensibilidades de 0,845 e 0,916 para

16 PF/exame, respectivamente. Estes são os maiores valores de sensibilidade presentes nesta Tabela para padrões de referência com consensos não-majoritários (GUO; LI, 2012; LU *et al.*, 2015; KURUVILLA; GUNAVATHI, 2014). Dessa forma, estes resultados mostram que o método proposto é comparável a outros métodos de detecção de nódulos encontrados na literatura pesquisada com a vantagem de utilizar uma álgebra simples e vários tipos de padrão de referência, sugerindo a validação da metodologia empregada nesta tese.

Como o volume de dados é maior na fase de redução de falsos positivos, então a técnica com menor latência dentre as apresentadas, ou seja, a Redução Booleana de Falso Positivo, é escolhida para cumprir esta tarefa. Essa técnica é menos complexa e utiliza atributos descritivos em vez de amostras de imagens sem tratamento. Por outro lado, para testes com sensibilidade acima de 94%, tal técnica apresenta valores médios de FP/exame superiores a 500, tornando-se impraticável para aplicações médicas. Considerando o desempenho preditivo das CNNs e o número de candidatos selecionados, escolhe-se uma CNN para a etapa final de classificação.

Então, para uma comparação mais clara entre as latências das duas técnicas empregadas nas etapas de redução falso positivo e classificação, nas quais há grande disparidade no número de candidatos na entrada, utiliza-se para unidade de medida de tempo o segundo calculado por candidato. Os valores médios de latência para Redução Booleana de Falso Positivo e Classificação CNN são 0,285 e 0,379 segundos por candidato, respectivamente, o que corresponde a uma diferença de 25%.

Em geral, a CNN é uma técnica de alto custo computacional. Considerando todo o conjunto de dados, a rede leva aproximadamente 24 horas para gerar um modelo, sendo o tempo de teste por exame de 5 minutos em média (RIBEIRO, 2021). A fim de fornecer respostas mais rápidas para os especialistas e possivelmente integrar um sistema CAD em uma plataforma baseada em nuvem, outro método de classificação é testado. Esses testes são executados usando um Perceptron Multicamadas (MLP) dados os 20 atributos propostos como entrada, duas camadas, 30 neurônios na camada oculta e utilizando a ferramenta *Graphic User Interface (GUI)* do *MATLAB R2021b Neural Network Toolbox*.

Os resultados desse experimentos mostram que a latência média diminui para 6,99 segundos para todo o conjunto de dados, considerando anotações de pelo menos um radiologista, quando se utiliza uma MLP. No entanto, a taxa de falsos positivos aumenta sensivelmente para o mesmo cenário. Por exemplo, para uma sensibilidade de 84,5%, a taxa média de falsos positivos

aumenta de 16 para 103 ao se optar por uma MLP em vez de uma CNN.

Uma vez expostos os resultados, finalmente são apresentadas as conclusões e sugestões para trabalhos futuros.

5 CONCLUSÕES, CONTRIBUIÇÕES E TRABALHOS FUTUROS

Esta tese apresenta um novo algoritmo para detecção de nódulos pulmonares em exames de TC baseado em um vetor de 16 filtros de pré-processamento combinado com equações relacionais e Booleanas. Nesse método, os filtros de pré-processamento isolam os nódulos usando diferentes janelas de filtragem medianas, limiares baseados em métricas estatísticas e operações morfológicas tridimensionais. Em seguida, a técnica de Redução Booleana de Falsos Positivos carrega os dados de cada filtro de pré-processamento e seleciona os candidatos por meio das equações relacionais e Booleanas, as quais utilizam o descritor de 20 atributos proposto. Uma vez que os candidatos são selecionados é realizada a classificação.

De acordo com os resultados, o Vetor de Filtros de Pré-processamento provê altas taxas de isolamento de nódulos independentemente do nível de consenso do padrão de referência. Os resultados também mostram que a Redução Booleana de Falsos Positivos atinge sua finalidade usando equações relacionais e Booleanas simples sem redução significativa da sensibilidade. Esse fato também sugere que os atributos escolhidos são capazes de descrever representativamente os verdadeiros positivos do conjunto de dados. Além disso, a técnica de redução de falsos positivos proposta melhora efetivamente os resultados da classificação final, mostrando seu potencial para integração em sistemas CAD.

O método proposto atinge alta sensibilidade para uma taxa média de falsos positivos por exame baixa considerando como padrão-ouro um consenso majoritário entre especialistas, sendo tal resultado comparável aos relatados na literatura pesquisada. Para os padrões de referência não majoritários, o método proposto alcança os valores mais altos de sensibilidade entre os reportados nas pesquisas mencionadas com o mesmo nível de consenso. Tais resultados evidenciam que a técnica de redução de falsos positivos proposta contribui efetivamente para a detecção e classificação de nódulos pulmonares por meio de equações lógicas e relacionais simples. Os resultados do método concordam com os padrões do estado da arte mesmo para um vasto conjunto de dados e um padrão de referência desafiador com um baixo nível de consenso.

Desse modo, destacam-se como contribuições desta tese:

- o desenvolvimento de um novo algoritmo de segmentação automática dos pulmões para separação da área de interesse e posterior detecção de nódulos;
- a elaboração de um novo método de isolamento de nódulos que os separa de outras estruturas adjacentes, facilitando sua classificação posterior: o Vetor de Filtros de Pré-processamento; e

- o desenvolvimento de uma técnica de seleção de candidatos e redução de falsos positivos por meio de equações Booleanas e relacionais.

Ainda assim, o método proposto possui pontos potenciais para evolução. Desse modo, para sua melhoria, as estratégias a seguir são sugeridas como trabalhos futuros:

- aprimoramento do Vetor de Filtros de Pré-processamento por meio do aumento da quantidade de filtros no vetor, ou seja, do número de janelas de diferentes dimensões e da quantidade de valores de limiar;
- utilização de novos atributos descritivos no método de Redução Booleana de Falsos Positivos;
- modificação da arquitetura da CNN utilizada no estágio de classificação por meio da busca heurística por seus parâmetros;
- emprego de outras técnicas conhecidas na literatura no estágio de classificação de nódulos, como *Support Vector Machines* (SVMs), MLP; e
- redução da latência para melhor integração do método com outros sistema por meio da reescrita dos algoritmos em linguagem C, ou similar.

REFERÊNCIAS

- BROWNLEE, J. **Deep learning for computer vision: image classification, object detection, and face recognition in python**. San Juan: Machine Learning Mastery, 2019.
- CHOI, W.-J.; CHOI, T.-S. Automated pulmonary nodule detection based on three-dimensional shape-based feature descriptor. **Computer methods and programs in biomedicine**, v. 113, n. 1, p. 37–84, 2014.
- CHOLLET, F. **Deep learning with Python**. New York City: Manning Publications Company, 2017.
- FERLAY, J.; COLOMBET, M.; SOERJOMATARAM, I.; MATHERS, C.; PARKIN, D. M.; PIÑEROS, M.; ZNAOR, A.; BRAY, F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. **International journal of cancer**, v. 144, n. 8, p. 1941–1953, 2019.
- GONZALEZ, R.; WOODS, R. **Digital Image Processing, Global Edition**. New York City: Pearson, 2018.
- GONZALEZ, R.; WOODS, R.; EDDINS, S. **Digital Image Processing Using MATLAB**. Upper Saddle River: Prentice Hall, 2004.
- GUO, W.; LI, Q. High performance lung nodule detection schemes in ct using local and global information. **Medical Physics**, v. 39, n. 8, p. 5157–5168, 2012.
- HANSELL, D. M.; BANKIER, A. A.; MACMAHON, H.; MCLOUD, T. C.; MULLER, N. L.; REMY, J. Fleischner society: glossary of terms for thoracic imaging. **Radiology**, v. 246, n. 3, p. 697–722, 2008.
- HSU, H.-H.; KO, K.-H.; CHOU, Y.-C.; WU, Y.-C.; CHIU, S.-H.; CHANG, C.-K.; CHANG, W.-C. Performance and reading time of lung nodule identification on multidetector ct with or without an artificial intelligence-powered computer-aided detection system. **Clinical Radiology**, Elsevier, 2021.
- HUANG, X.; SHAN, J.; VAIDYA, V. Lung nodule detection in CT using 3D convolutional neural networks. In: **Proceedings - International Symposium on Biomedical Imaging**. Melbourne: [s.n.], 2017. p. 379–383.
- III, S. G. A.; MCLENNAN, G.; BIDAUT, L.; MCNITT-GRAY, M. F.; MEYER, C. R.; REEVES, A. P.; ZHAO, B.; ABERLE, D. R.; HENSCHKE, C. I.; HOFFMAN, E. A.; KAZEROONI, E. A.; MACMAHON, H.; BEEK, E. J. R. van; YANKELEVITZ, D.; BIANCARDI, A. M.; BLAND, P. H.; BROWN, M. S. 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, n. 2, p. 915–931, 2011.
- JACOBS, C.; RIKXOORT, E. M. van; TWELLMANN, T.; SCHOLTEN, E. T.; JONG, P. A. de; KUHNIGK, J.-M.; OUDKERK, M.; KONING, H. J. de; PROKOP, M.; SCHAEFER-PROKOP, C. *et al.* Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images. **Medical image analysis**, v. 18, n. 2, p. 374–384, 2014.

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

KINGMA, D. P.; BA, J. L. Adam: A method for stochastic optimization. **3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings**, 2015.

KURUVILLA, J.; GUNAVATHI, K. Lung cancer classification using neural networks for CT images. **Computer methods and programs in biomedicine**, v. 113, n. 1, p. 202–209, 2014.

LOWE, D. G. Object recognition from local scale-invariant features. In: **Proceedings of the seventh IEEE international conference on computer vision**. Corfu: [s.n.], 1999. v. 2, p. 1150–1157.

LU, L.; TAN, Y.; SCHWARTZ, L. H.; ZHAO, B. Hybrid detection of lung nodules on CT scan images. **Medical physics**, v. 42, n. 9, p. 5042–5054, 2015.

MASOOD, A.; YANG, P.; SHENG, B.; LI, H.; LI, P.; QIN, J.; LANFRANCHI, V.; KIM, J.; FENG, D. D. Cloud-based automated clinical decision support system for detection and diagnosis of lung cancer in chest CT. **IEEE Journal of Translational Engineering in Health and Medicine**, v. 8, p. 1–13, 2020.

MESSAY, T.; HARDIE, R. C.; ROGERS, S. K. A new computationally efficient CAD system for pulmonary nodule detection in CT imagery. **Medical image analysis**, v. 14, n. 3, p. 390–406, 2010.

PATTERSON, J.; GIBSON, A. **Deep learning: A practitioner's approach**. Beijing: O'Reilly, 2017.

PEZESHK, A.; HAMIDIAN, S.; PETRICK, N.; SAHINER, B. 3-D convolutional neural networks for automatic detection of pulmonary nodules in chest CT. **IEEE journal of biomedical and health informatics**, v. 23, n. 5, p. 2080–2090, 2018.

RENFREW, D. L.; JR., E. A. F.; BERBAUM, K. S.; WEIGELT, F. H.; ABU-YOUSEF, M. M. Error in radiology: classification and lessons in 182 cases presented at a problem case conference. **Radiology**, v. 183, n. 1, p. 145–150, 1992.

RIBEIRO, A. B. N. **A new automatic system for detection and classification of pulmonary nodules in thoracic CT images using a single CNN**. Tese (Doutorado) — Federal University of Ceara, 2021.

SEDGEWICK, R. **Algorithms in C, 3rd Ed**. Pennsylvania State University: Addison-Wesley, 1998.

SETIO, A. A. A.; CIOMPI, F.; LITJENS, G.; GERKE, P.; JACOBS, C.; Van Riel, S. J.; WILLE, M. M. W.; NAQIBULLAH, M.; SANCHEZ, C. I.; Van Ginneken, B. Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks. **IEEE Transactions on Medical Imaging**, v. 35, n. 5, p. 1160–1169, 2016.

SETIO, A. A. A.; TRAVERSO, A.; BEL, T. D.; BERENS, M. S.; BOGAARD, C. van den; CERELLO, P.; CHEN, H.; DOU, Q.; FANTACCI, M. E.; GEURTS, B.; GUGTEN, R.; HENG, P.; JANSEN, B.; KASTE, M. M. de; KOTOV, V.; LIN, J. Y.-H.; MANDERS, J. T.;

SÉÑORA-MENGANA, A.; GARCÍA-NARANJO, J. C.; PAPAVALASILEIOU, E.; PROKOP, M.; SALETTA, M.; SCHAEFER-PROKOP, C. M.; SCHOLTEN, E. T.; SCHOLTEN, L.; SNOEREN, M. M.; TORRES, E. L.; VANDEMEULEBROUCKE, J.; WALASEK, N.; ZUIDHOF, G. C.; GINNEKEN, B. v.; JACOBS, C. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. **Medical image analysis**, v. 42, p. 1–13, 2017.

SIEGEL, R. L.; MILLER, K. D.; JEMAL, A. Cancer statistics, 2020. **CA: A Cancer Journal for Clinicians**, v. 70, n. 1, p. 7–30, 2020.

SOILLE, P. **Morphological Image Analysis: Principles and Applications**. University of California: Springer-Verlag, 1999.

TAN, M.; WU, F.; YANG, B.; MA, J.; KONG, D.; CHEN, Z.; LONG, D. Pulmonary nodule detection using hybrid two-stage 3D CNNs. **Medical Physics**, v. 47, n. 8, p. 3376–3388, 2020.

TORRES, W. de O.; FILHO, A. O. de C.; RABÊLO, R. d. A. L.; SILVA, R. R. V. e. Texture analysis of lung nodules in computerized tomography images using functional diversity. **Computers & Electrical Engineering**, Elsevier, v. 84, p. 106618, 2020.

VALENTE, I. R. S.; CORTEZ, P. C.; NETO, E. C.; SOARES, J. M.; ALBUQUERQUE, V. H. C. de; TAVARES, J. M. R. Automatic 3D pulmonary nodule detection in CT images: A survey. **Computer methods and programs in biomedicine**, v. 124, n. 91–107, p. 173–196, 2016.

VASSALLO, L.; TRAVERSO, A.; AGNELLO, M.; BRACCO, C.; CAMPANELLA, D.; CHIARA, G.; FANTACCI, M. E.; TORRES, E. L.; MANCA, A.; SALETTA, M.; GIANNINI, V.; MAZZETTI, S.; STASI, M.; CERELLO, P.; REGGE, D. A cloud-based computer-aided detection system improves identification of lung nodules on computed tomography scans of patients with extra-thoracic malignancies. **European radiology**, Springer, v. 29, n. 1, p. 144–152, 2019.

ZHANG, J.; XIA, Y.; ZENG, H.; ZHANG, Y. NODULE: Combining constrained multi-scale LoG filters with densely dilated 3D deep convolutional neural network for pulmonary nodule detection. **Neurocomputing**, v. 317, p. 159–167, 2018.

ZHENG, S.; GUO, J.; CUI, X.; VELDHUIS, R. N. J.; OUDKERK, M.; OOIJEN, P. M. A. V. Automatic pulmonary nodule detection in CT scans using convolutional neural networks based on maximum intensity projection. **IEEE transactions on medical imaging**, v. 39, n. 3, p. 797–805, 2019.