



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
E DE COMPUTAÇÃO

THIAGO FELLIPE ORTIZ DE CAMARGO

**Inteligência artificial aplicada na quantificação de
nódulos pulmonares em imagens de tomografia
computadorizada**

[GOIÂNIA - GOIÁS - BRASIL]

2025



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a [Lei 9.610/98](#), o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

Dissertação Tese Outro*: _____

*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

2. Nome completo do autor

THIAGO FELLIPE ORTIZ DE CAMARGO

3. Título do trabalho

“Inteligência artificial aplicada na quantificação de nódulos pulmonares em imagens de tomografia computadorizada”

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:

a) consulta ao(à) autor(a) e ao(à) orientador(a);

b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação. O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Thiago Fellipe Ortiz De Camargo, Discente**, em 21/07/2025, às 10:41, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Wesley Pacheco Calixto, Usuário Externo**, em 21/07/2025, às 15:57, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5515543** e o código CRC **7A372C04**.

THIAGO FELLIPE ORTIZ DE CAMARGO

**Inteligência artificial aplicada na quantificação de
nódulos pulmonares em imagens de tomografia
computadorizada**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Computação, da Faculdade Engenharia Elétrica, Mecânica e Computação, da Universidade Federal de Goiás (UFG), como requisito para obtenção do título de Mestre em Engenharia Elétrica e de Computação.

Área de concentração: Engenharia de computação.

Orientador: Professor Doutor Wesley Pacheco Calixto.

Coorientadora: Professora Doutora Marcella Scoczynski Ribeiro Martins.

[GOIÂNIA - GOIÁS - BRASIL]

2025

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Camargo, Thiago Fellipe Ortiz de
INTELIGÊNCIA ARTIFICIAL APLICADA NA QUANTIFICAÇÃO DE
NÓDULOS PULMONARES EM IMAGENS DE TOMOGRAFIA
COMPUTADORIZADA [manuscrito] / Thiago Fellipe Ortiz de
Camargo. - 2025.
XCVI, 96 f.: il.

Orientador: Prof. Dr. Wesley Pacheco Calixto; co-orientadora Dra.
Marcella Scoczynski Ribeiro Martins.

Dissertação (Mestrado) - Universidade Federal de Goiás, Escola
de Engenharia Elétrica, Mecânica e de Computação (EMC), Programa
de Pós-Graduação em Engenharia Elétrica e de Computação, Goiânia,
2025.

Bibliografia.

Inclui siglas, abreviaturas, símbolos, gráfico, tabelas, lista de
figuras, lista de tabelas.

1. Nódulo Pulmonar. 2. Aprendizado Profundo. 3. Tomografia
Computadorizada. 4. Quantificação. 5. Segmentação Semântica. I. Calixto,
Wesley Pacheco, orient. II. Título.

CDU 621.3



UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA, MECÂNICA E DE COMPUTAÇÃO

ATA EMC 12/2025

Ata nº 12 da sessão de Defesa de Dissertação de **Thiago Felliipe Ortiz de Camargo**, que confere o título de Mestre em **Engenharia Elétrica e de Computação**, na área de concentração em **Engenharia de Computação**.

Aos **onze dias do mês de junho de dois mil e vinte e cinco**, a partir das **14:00h**, realizou-se a sessão pública de Defesa de Dissertação intitulada “**Inteligência artificial aplicada na quantificação de nódulos pulmonares em imagens de tomografia computadorizada**”. Os trabalhos foram instalados pelo Orientador, Professor Doutor **Wesley Pacheco Calixto - (EMC/UFG)** com a participação dos demais membros da Banca Examinadora: Professora Doutora **Marcella Scoczynski Ribeiro Martins - (ELE/UTFPR)** Coorientadora, Professor Doutor **Hugo Valadares Siqueira (ELE/UTFPR)** Membro Titular Externo, Professor Doutor **Guilherme Alberto Sousa Ribeiro (HIAE/COMP)** Membro Titular Externo, Professor Doutor **Gelson da Cruz Junior (EMC/UFG)** Membro Titular Interno, Professora Doutora **Solange Amorim Nogueira (HIAE/BIOMED)**: **cujas participações ocorreram através de videoconferência pelo link: <https://meet.google.com/uux-wmms-yn>**. Durante a arguição os membros da banca **não fizeram** sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Dissertação, tendo sido o candidato **aprovado** pelos seus membros. Proclamados os resultados pelo Professor Doutor **Wesley Pacheco Calixto**, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos **onze dias do mês de junho de dois mil e vinte e cinco**.

TÍTULO SUGERIDO PELA BANCA



Documento assinado eletronicamente por **Wesley Pacheco Calixto, Usuário Externo**, em 24/07/2025, às 10:18, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Thiago Felliipe Ortiz De Camargo, Discente**, em 24/07/2025, às 10:25, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marcella Scoczynski Ribeiro Martins, Usuário Externo**, em 24/07/2025, às 21:11, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Guilherme Alberto Sousa Ribeiro, Usuário Externo**, em 26/07/2025, às 16:29, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **SOLANGE AMORIM NOGUEIRA**, **Usuário Externo**, em 28/07/2025, às 19:43, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Hugo V Siqueira registrado(a) civilmente como Hugo Valadares Siqueira**, **Usuário Externo**, em 29/07/2025, às 10:05, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gelson Da Cruz Junior**, **Professor do Magistério Superior**, em 30/07/2025, às 15:54, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5524068** e o código CRC **A542BAFA**.

”Tal foi ou deve ter sido a origem da sociedade e das leis, que deram novos entraves ao fraco e novas forças ao rico, destruíram irreparavelmente a liberdade natural, fixaram para sempre a lei da propriedade e da desigualdade, fizeram uma hábil usurpação um direito irrevogável e, em proveito de alguns ambiciosos, sujeitaram , a partir de então, todo o gênero de humano ao trabalho, à servidão e à miséria”.

JEAN-JAQUES ROUSSEAU

em “Discours sur l’origine et les fondements de l’inégalité parmi les hommes”, 1754.

Dedico esta obra, antes de tudo, a todos quantos, privados de iguais oportunidades, mantêm viva a chama dos seus sonhos. Dedico-a, ainda, àqueles de quem me ausentei para erigir estas páginas e que, mesmo assim, me ampararam com paciência e afeto. E dedico-a, por fim, aos mestres que me enriqueceram com saber e humanidade, descortinando-me horizontes e ensinando-me a transformar o mundo pelo conhecimento. A educação liberta.

AGRADECIMENTOS

Agradeço a todas as pessoas que reconheceram em mim um potencial que, por vezes, eu mesmo não percebia; que me motivaram, orientaram e ajudaram a moldar uma personalidade questionadora, cética, curiosa e humana. A essas pessoas, expresso meus mais sinceros agradecimentos.

Agradeço à minha família, pelo suporte incondicional, por me conceder o privilégio de dedicar-me integralmente aos estudos durante a graduação e pelos afagos recebidos mesmo nas minhas ausências em virtude da elaboração deste trabalho.

Agradeço aos meus amigos, que compartilharam suas visões de mundo e trouxeram alegria em momentos desafiadores. Em especial, registro minha gratidão a Sthefanie Mônica Premebida, cuja extraordinária capacidade de criar conexões abriu caminhos decisivos para que eu chegasse até aqui — e continua gerando frutos.

Agradeço à minha coorientadora, Prof.^a Dra. Marcella Scoczynski Ribeiro Martins, por ter despertado meu interesse pela pesquisa, proporcionando-me experiências acadêmicas e profissionais fundamentais para minha maturidade, além de acolhimento e suporte genuinamente maternais.

Agradeço ao meu orientador, Prof. Dr. Wesley Pacheco Calixto, por partilhar conhecimento e visão com generosidade. Seu brilho intelectual ímpar enriquece-me diariamente. Sou grato pelo verdadeiro interesse que dedica à pesquisa de seus orientandos e por sua disponibilidade constante.

Agradeço às profissionais da Psicologia, Grazieli Gonçalves Vilarinho da Rocha e Vanessa Mascareli Piedade Casaro, que me ampararam e me guiaram na jornada de autoconhecimento, amadurecimento e cura. Seus cuidados transformaram meu recorte da realidade — feliz, em vez de apático; leve, em vez de desproporcionalmente pesado; colorido, em vez de monocromático. Meu mais sincero obrigado por cada gesto que contribuiu para a construção da minha maturidade emocional.

Por fim, agradeço às instituições públicas brasileiras, que me garantiram acesso a educação e saúde de qualidade. Sem elas, eu não teria alcançado este ponto nem poderia almejar um futuro ainda melhor.

RESUMO

A quantificação precisa de nódulos pulmonares em tomografias computadorizadas enfrenta desafios relacionados à variabilidade interobservador e à ausência de métodos escaláveis capazes de generalizar entre bases de dados heterogêneas. Este estudo propõe uma solução automatizada que integra aprendizado profundo para gerar máscaras de segmentação de nódulos e calcular volumes de maneira precisa, utilizando informações espaciais das imagens. A metodologia é estruturada em três etapas principais: preparação e processamento dos dados, treinamento e validação do modelo de segmentação, e avaliação da eficiência da solução. A utilização da arquitetura nnU-Net, que automatiza o pré-processamento, a segmentação e o pós-processamento, confere escalabilidade e adaptação dinâmica ao fluxo de trabalho, potencializando a aplicabilidade clínica da proposta. Os resultados indicam estabilidade nas medições de volume e diâmetro em exames sucessivos e alta aproximação às máscaras de consenso, mesmo em nódulos de anatomia complexa. A arquitetura U-Net 3D F destaca-se, alcançando $\bar{D}_c = 0,7846$ com desvio padrão $\sigma = 0,18$, superando o índice interobservador de $D_c = 0,5218$ e demonstrando baixo desvio volumétrico entre aquisições. A metodologia desenvolvida promove avanços na quantificação automatizada de nódulos pulmonares, oferecendo solução resiliente e adaptável para suporte ao diagnóstico médico em cenários clínicos reais.

Palavras-chave: Nódulo Pulmonar; Aprendizado Profundo; Tomografia Computadorizada; Quantificação; Segmentação Semântica.

ARTIFICIAL INTELLIGENCE APPLIED TO THE QUANTIFICATION OF PULMONARY NODULES IN COMPUTED TOMOGRAPHY IMAGES

ABSTRACT

Accurate quantification of pulmonary nodules in computed tomography remains challenging due to interobserver variability and the lack of scalable methods capable of generalizing across heterogeneous datasets. This study proposes an automated solution that integrates deep learning to generate nodule segmentation masks and compute volumetric measurements using spatial information from the images. The methodology is structured into three main stages: data preparation and pre-processing, model training and validation, and performance evaluation. The use of the nnU-Net architecture, which automates pre-processing, segmentation, and post-processing, provides scalability and dynamic adaptation to the workflow, enhancing the clinical applicability of the solution. The results indicate consistent volume and diameter measurements across successive scans and strong agreement with consensus masks, even for anatomically complex nodules. The 3D U-Net architecture achieved a mean Dice coefficient of $\bar{D}_c = 0.7846$ with a standard deviation of $\sigma = 0.18$, outperforming the interobserver Dice index of $D_c = 0.5218$ and exhibiting a low volumetric deviation between acquisitions. The proposed methodology advances automated quantification of pulmonary nodules, offering a resilient and adaptable solution to support medical diagnosis in real-world clinical scenarios.

Keywords: Lung Nodule; Deep Learning; Computed Tomography; Quantification; Semantic Segmentation.

SUMÁRIO

Pág.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SÍMBOLOS

LISTA DE ABREVIATURAS E SIGLAS

CAPÍTULO 1 INTRODUÇÃO	29
CAPÍTULO 2 IMAGIOLOGIA DE NÓDULOS PULMONARES .	35
2.1 Problemas relacionados a imagens de nódulos	35
2.2 Segmentação de imagens médicas	36
2.3 Tomografia computadorizada	38
2.4 Considerações finais	39
CAPÍTULO 3 INTELIGÊNCIA ARTIFICIAL	41
3.1 Conceituação	41
3.2 Avaliação de modelos de inteligência artificial	42
3.3 nnU-Net	45
3.3.1 Pré-processamento	46
3.3.2 Definição de Arquitetura e Hiperparâmetros	48
3.3.3 Conjunto de modelos, pós-processamento e residual encoder	49
3.4 Considerações finais	50
CAPÍTULO 4 METODOLOGIA	51
4.1 Contextualização	51
4.2 Preparação dos dados	52
4.2.1 Processamento de máscaras	52
4.2.2 Recorte de volume	53
4.3 Modelo de segmentação	53
4.4 Validação da solução	55
4.4.1 Cálculo do volume e diâmetro	56

4.5	Considerações finais	56
CAPÍTULO 5 RESULTADOS		57
5.1	Escolha das bases e pré-processamento dos dados	57
5.1.1	Avaliação da concordância interobservador e desempenho do modelo de segmentação	59
5.1.2	Avaliação qualitativa e análise de erros	63
5.1.3	Avaliação e validação da solução	71
5.2	Discussão	75
CAPÍTULO 6 CONCLUSÃO		79
6.1	Contribuições do Trabalho	80
REFERÊNCIAS BIBLIOGRÁFICAS		81

LISTA DE FIGURAS

	<u>Pág.</u>
4.1 Fluxograma da metodologia proposta.	52
5.1 Imagens de TC representativas das bases de dados: (a) BD_1 e (b) BD_2 . .	58
5.2 Função de perda de treino T_L e de validação V_L , pseudo-DICE P_{DC} , e média móvel do pseudo-DICE $MA_{P_{DC}}$ para fold de número 0 para os modelos: a) 2D, b) 2D REM, c) 2D REL, d) 3D L, e) 3D LREM, f) 3D F, g) 3D FREM, h) 3D CF e i) 3D CFREM.	61
5.3 Diagrama do modelo final U-Net 2D.	61
5.4 Diagrama do modelo final U-Net 3D F.	62
5.5 <i>Boxplot</i> da métrica D_c para cada modelo em nódulos individuais do subconjunto de teste BD_1	62
5.6 VP, FN e FP para cada modelo no subconjunto de teste BD_1	63
5.7 Comparação entre os modelos U-Net 2D, U-Net 3D F e U-Net 3D FREM: (a), (c), (e) e (g) exibem cortes axiais dos volumes LIDC-0181 e LIDC- 0773, enquanto (b), (d), (f) e (h) apresentam as máscaras de segmentação sobrepostas correspondentes.	65
5.8 Análise de verdadeiros positivos para o volume LIDC-0235 do conjunto de dados BD_1 : (a), (c) e (e) apresentam cortes das fatias, enquanto (b), (d) e (f) apresentam a segmentação dos nódulos geradas pelo modelo U-Net 3D F.	66
5.9 Análise de verdadeiros e falsos positivos para a imagem LIDC-0235 do conjunto de dados BD_1 : (a), (c), (e) e (g) apresentam cortes das fatias, enquanto (b), (d), (f) e (h) mostram as segmentações geradas pelo modelo U-Net 3D F.	66
5.10 Segmentações produzidas pelo modelo U-Net 3D F avaliados como não satisfatórias pelo radiologista torácico no volume LIDC-0061 de BD_1 . (a) fatia axial, (b) fatia axial ampliada, (c) eixo coronal, e (d) eixo sagital.	69
5.11 Segmentações produzidas pelo modelo 3D F contendo falsos positivos, verificados pelo radiologista torácico, nos volumes LIDC-0080 e LIDC- 0925 de BD_1 : (a) falso positivo em região de estria atelectásica, (b) eixo axial, (c) eixo coronal e (d) eixo sagital da mesma região.	70

5.12	Segmentações produzidas pelo modelo U-Net 3D F em que contêm falsos positivos que foram classificados como verdadeiros positivos pelo radiologista torácico: (a) LIDC-0879, (b) LIDC-0925 e (c) LIDC-1008.	70
5.13	Segmentações geradas pelo modelo U-Net 3D F de falsos positivos, verificado pelo radiologista torácico: (a) LIDC-0080, (b) segmentação por limiar da região de falso negativo e (c) LIDC-1008.	71
5.14	Gráfico de Bland–Altman das diferenças de volume dos nódulos em escala logarítmica.	72
5.15	Gráfico de Bland–Altman das diferenças de diâmetro dos nódulos em escala logarítmica.	73

LISTA DE TABELAS

	<u>Pág.</u>
1.1 Síntese dos trabalhos que abordam cálculo volumétrico de nódulos pulmonares em imagens de tomografia computadorizada de tórax.	32
5.1 Arquiteturas representativas configuradas pela nnU-Net.	60
5.2 Média e desvio-padrão de D_c , J_c e HD_{3D} no subconjunto de teste de BD_1	62
5.3 Precisão e sensibilidade para cada modelo no subconjunto de teste BD_1	63
5.4 Imagens com maior quantidade de FN recorrentes entre as variações de U-Net, subconjunto de teste de BD_1	64
5.5 Imagens com maior quantidade de FP recorrentes entre as variações de U-Net. Subconjunto de teste de BD_1	65
5.6 Avaliação de delineamento e justificativas.	67
5.7 Falsos positivos e nível de dificuldade de interpretação.	68
5.8 Falsos negativos e nível de dificuldade.	69
5.9 Resumo estatístico das diferenças de volume e diâmetro para os modelos no conjunto de dados BD_2	72
5.10 Viés, limites de concordância, e intervalos de confiança no gráfico de Bland-Altman das diferenças de volume de nódulos.	72
5.11 Viés, limites de concordância, e intervalos de confiança no gráfico de Bland-Altman das diferenças de diâmetro de nódulos.	73
5.12 Coeficientes da regressão do modelo $\Delta V = \beta_0 + \beta_1 \bar{V}$ por modelo.	73
5.13 Comparação de volume e diâmetro entre TC1 e TC2.	74

LISTA DE SÍMBOLOS

μ	– Média
σ	– Desvio padrão
\log_{10}	– Logaritmo na base 10
BD_1	– Banco de dados 1
BD_2	– Banco de dados 2
D_c	– Coeficiente DICE
\bar{D}_c	– Coeficiente DICE médio
DH_{3D}	– Distância Hausdorff tridimensional
J_c	– Coeficiente Jaccard
N	– Quantitativo total de elementos
T_m	– Taxa de malignidade
p	– p -valor
ΔV	– Diferença de volume ($V_{CT2} - V_{CT1}$)
\bar{V}	– Volume médio ($(V_{CT1} + V_{CT2})/2$)
β_0	– Intercepto da regressão (viés fixo)
β_1	– Coeficiente angular da regressão (viés proporcional)
R_{aj}^2	– Coeficiente de determinação ajustado da regressão

LISTA DE ABREVIATURAS E SIGLAS

2D REL	–	U-Net 2D Residual Encoder L
2D REM	–	U-Net 2D Residual Encoder M
3D CF	–	U-Net 3D Cascada Fullres
3D CFREM	–	U-Net 3D Cascade Fullres Residual Encoder M
3D F	–	U-Net 3D Fullres
3D FREM	–	U-Net 3D Fullres Residual Encoder M
3D L	–	U-Net 3D Lowres
3D LREM	–	U-Net 3D Lowres Residual Encoder M
AGI	–	Artificial General Intelligence
CNN	–	Convolutional Neural Network
DICOM	–	Digital Imaging and Communications in Medicine
FN	–	Falsos Negativos
FP	–	Falsos Positivos
GPU	–	Graphics Processing Unit
IA	–	Inteligência Artificial
JSON	–	JavaScript Object Notation
LIDC-IDRI	–	Image Database Consortium and Image Database Resource Initiative
NAI	–	Narrow Artificial Intelligence
ND	–	Não Detectado
NIfTI	–	Neuroimaging Informatics Technology Initiative
nnU-Net	–	no-new-U-Net
RIDER	–	Reference Image Database to Evaluate Response to Therapy in Lung Cancer Project
TC	–	Tomografia Computadorizada
UH	–	Unidades Hounsfield
OOD	–	Out-of-Distribution
ResEnc	–	Residual Encoder
RM	–	Ressonância Magnética
ROC-AUC	–	Receiver Operating Characteristic - Area Under the Curve
VP	–	Verdadeiros Positivos
VRAM	–	Video Random Access Memory

CAPÍTULO 1

INTRODUÇÃO

O câncer de pulmão é a principal causa de morte decorrente de câncer e tem apresentado alta incidência de diagnósticos em todo o mundo desde 2018, sendo um dos tipos mais frequentemente identificados (BRAY *et al.*, 2018; SUNG *et al.*, 2021). A detecção precoce do câncer de pulmão é fundamental na obtenção do prognóstico favorável. Uma das abordagens para identificar a doença em estágios iniciais é por meio da utilização de tomografia computadorizada (TC), permitindo a observação de possíveis nódulos pulmonares (TEAM, 2011; KONING *et al.*, 2020).

A interpretação de nódulos encontrados em imagens de tomografia computadorizada do tórax requer a aplicação de diretrizes clínicas para auxiliar na tomada de decisões. Diretrizes como Lung-RADS, PanCan e Fleischner Society fornecem recomendações específicas com base em características como tipo de nódulo, tamanho, crescimento, formato e fatores de risco do paciente (American College of Radiology Committee on Lung-RADS, 2022; MCWILLIAMS *et al.*, 2013; NAIDICH *et al.*, 2013; MACMAHON *et al.*, 2017). A categorização adequada do nódulo é necessária, pois diferentes tipos têm probabilidades de malignidade distintas (HENSCHKE *et al.*, 2002). Conhecer o tamanho do nódulo pulmonar é importante, pois é a principal característica utilizada para guiar a estratégia e o procedimento a serem adotados. No entanto, a medição do diâmetro e volume dos nódulos pode variar entre observadores (STEENBAKKERS *et al.*, 2005; BOGOT *et al.*, 2005). Para superar este problema, a utilização de ferramentas semiautomáticas e automáticas com base no diâmetro do nódulo ou na contagem de voxels¹ pode ser uma das soluções (NAIDICH *et al.*, 2013).

A análise de imagens de tomografia computadorizada envolve diversos métodos de segmentação para extrair informações relevantes. Os principais tipos de segmentação utilizados têm base em: i) intensidade, ii) regiões, iii) contorno, iv) modelos e v) na semântica. A segmentação semântica é importante na classificação dos pixels da imagem, criando máscara que fornece informações estruturadas sobre o objeto de interesse (LONG *et al.*, 2015). Esta abordagem é amplamente aplicada em várias áreas,

¹O termo voxel é a abreviação de elemento volumétrico e representa valor de intensidade em espaço tridimensional. A contagem de voxels da máscara de segmentação é o processo de quantificar o número de unidades tridimensionais de volume dentro da região de interesse. É abordagem objetiva e automatizada que auxilia os profissionais de saúde na análise e no acompanhamento de lesões, como os nódulos pulmonares, proporcionando informações valiosas para o diagnóstico e tratamento (GLASSNER, 1984; ASHBURNER; FRISTON, 2000; HUGHES *et al.*, 2014).

incluindo a médica, na qual auxilia no acompanhamento de doenças e na tomada de decisão clínica (WORTH et al., 1997; DALE et al., 1999; FISCHL et al., 2002).

Na área médica, a obtenção manual da máscara de segmentação é processo oneroso em termos de tempo e recursos financeiros (WORTH et al., 1997; YEPES-CALDERON; MCCOMB, 2019). Embora existam metodologias clássicas com base em técnicas de visão computacional, como valores de limiar, clusterização e crescimento de regiões, estas abordagens exigem intervenção manual devido à sua acurácia limitada (PHAM et al., 2000b; SHARMA; AGGARWAL, 2010). Diante disto, métodos com base em aprendizado profundo, especialmente a arquitetura U-Net (FALK et al., 2019), têm se apresentado como opção viável e promissora para a tarefa de segmentação semântica (RONNEBERGER et al., 2015; HELLER et al., 2019).

Estas abordagens utilizam algoritmos de aprendizado de máquina que são treinados em grandes conjuntos de dados para reconhecer padrões complexos e realizar segmentações precisas e automatizadas (LIN et al., 2014; HE et al., 2017). Com a arquitetura U-Net, por exemplo, é possível obter resultados precisos e confiáveis, reduzindo a dependência de intervenção manual e agilizando o processo de segmentação semântica. Isto representa avanço significativo na área médica, possibilitando análise mais rápida e eficiente das imagens de TC e fornecendo informações valiosas para diagnóstico, acompanhamento de doenças e tomada de decisão clínica (FALK et al., 2019; HOLLON et al., 2020; CUI et al., 2022).

Apesar das arquiteturas com base na U-Net para segmentação semântica alcançarem resultados satisfatórios em métricas de desempenho, a transferência destes fluxos de processamento entre conjuntos de dados de imagens médicas para segmentar achados radiológicos geralmente não é viável (HELLER et al., 2021). Isto ocorre porque cada conjunto de dados possui características únicas que exigem adaptações no pré-processamento das imagens, na arquitetura do modelo e nas técnicas de aumento de dados, por exemplo, produzindo impacto nos resultados e ocasionando análise manual, o que requer a avaliação de especialistas (LITJENS et al., 2017).

Esta limitação na transferência do fluxo de processamento afeta o cálculo volumétrico dos nódulos pulmonares, devido ao fato de que algumas técnicas dependem da qualidade da máscara produzida (YOON et al., 2022). Portanto, a obtenção de resultados confiáveis na volumetria dos nódulos pulmonares requer adaptação cuidadosa do fluxo de processamento para cada conjunto de dados específico. A superação

desta limitação é essencial para garantir a acurácia e a consistência na medição volumétrica dos nódulos pulmonares, contribuindo assim para análise precisa e nas práticas clínicas de diagnóstico e monitoramento de doenças pulmonares (DEVARAJ et al., 2017).

Tan e Le (2019) propõem técnica de dimensionamento de modelos para gerar arquiteturas eficientes, considerando o equilíbrio entre a profundidade, largura e resolução das arquiteturas. Esta abordagem possibilita o desenvolvimento de sete variações, denominadas EfficientNet², em que suas características são escaladas utilizando coeficiente composto³. Em testes, a metodologia alcança resultados satisfatórios no conjunto de dados ImageNet (RUSSAKOVSKY et al., 2015), especialmente no contexto da classificação de imagens.

Dosovitskiy et al. (2020) adaptam a metodologia dos Transformers⁴ de Vaswani et al. (2017) para o contexto da visão computacional, resultando na arquitetura Vision Transformer, que mantém as capacidades de paralelização de cálculos computacionais e mecanismos de autoatenção. Apesar da dependência do pré-treinamento em elevado conjunto de imagens, esta abordagem alcança resultados comparáveis ou superiores aos melhores resultados nos conjuntos de dados ImageNet, CIFAR-10, CIFAR-100 e Oxford-IIIT Pets (DOSOVITSKIY et al., 2020; KRIZHEVSKY; HINTON, 2009; OLIVA; TORRALBA, 2001; PARKHI et al., 2012).

Ronneberger et al. (2015), com o objetivo de reduzir a quantidade de imagens necessárias para o treinamento de modelos de segmentação semântica aplicados a imagens médicas, propõem a U-Net, arquitetura com base na Fully Convolutional Neural Network⁵ de Long et al. (2015), que inclui camadas de convolução de aumento de resolução, camadas de agregação e conexões residuais. Com base nestas arquiteturas, vários autores propuseram variações, como V-Net, U-Net++ e Attention U-Net, que alcançam resultados equivalentes ou superiores (MILLETARI et al., 2016; OKTAY

²Técnica que permite gerar arquiteturas de redes neurais profundas automaticamente, ajustando o tamanho da arquitetura com base no fator de escala (TAN; LE, 2019).

³Parâmetro que controla o dimensionamento proporcional das características da arquitetura, permitindo ajustar a capacidade de representação dos modelos de acordo com as necessidades específicas (TAN; LE, 2019).

⁴Abordagem para processamento de linguagem natural e aprendizado de máquina que permite capturar relações de longo alcance em sequências de texto, superando limitações de abordagens como Redes Neurais Recorrentes e Redes Neurais Convolucionais (VASWANI et al., 2017).

⁵Arquitetura de rede neural que permite segmentação semântica em alta resolução, preservando informações espaciais finas, através do uso de camadas de convolução de aumento de resolução (LONG et al., 2015).

et al., 2018; ZHOU et al., 2018).

O trabalho de maior destaque com base na U-Net é a nnU-Net proposta por Fabian et al. (2020), que é arquitetura de rede neural desenvolvida para segmentação semântica de imagens médicas. A nnU-Net utiliza treinamento em cascata, técnicas de aumento de dados e ajuste de hiperparâmetros para obter resultados precisos e adaptáveis a diferentes conjuntos de dados. A abordagem da nnU-Net supera outros métodos em 33 dos 53 desafios de segmentação semântica em imagens médicas, apresentando sua eficácia e flexibilidade para lidar com diversas necessidades de processamento de dados.

Diversas investigações têm sido conduzidas para avaliar o cálculo de volume de nódulos pulmonares, como disposto na Tabela 1.1, que sintetiza alguns destes estudos em ordem cronológica. No entanto, não há metodologia que se adapte às demandas dos dados em todas as etapas do fluxo de processamento, incluindo a medição de parâmetros volumétricos, como diâmetro e volume, de nódulos pulmonares em imagens de tomografia computadorizada de tórax, por meio do emprego de técnicas de aprendizado profundo, apresentando assim a originalidade desta pesquisa. Portanto, este estudo inova ao preencher a lacuna de desenvolvimento de metodologia que aborde as variadas necessidades dos conjuntos de dados na tarefa de quantificação de nódulos pulmonares, justificando o trabalho.

Tabela 1.1 - Síntese dos trabalhos que abordam cálculo volumétrico de nódulos pulmonares em imagens de tomografia computadorizada de tórax.

Proposta/Abordagem	Referência	Contexto/Descrição
Validação Externa de ferramenta de mercado	Murchison et al. (2022)	Avalia impacto de ferramenta CAD no desempenho de médicos na tarefa de classificação, quantificação e segmentação.
Comparação entre métodos de cálculo de volume	Yoon et al. (2022)	Avalia a discrepância entre três métodos de cálculo de volume de nódulos pulmonares se baseando em máscaras de segmentação geradas por software de mercado.
Modelo estatístico para predição de malignidade	Goo (2011)	Aplica modelo de regressão para, através de volumetria, avaliar o crescimento de nódulos pulmonares.
Algoritmo para cálculo de volume	Heckel et al. (2013)	Aplica algoritmo matemático para calcular volume de nódulos pulmonares visando reduzir variabilidade inerente da máscara de segmentação.
Comparação entre ferramentas de quantificação	Scholten et al. (2013)	Avalia a variação na quantificação de nódulos pulmonares entre método manual e ferramenta semiautomática.
Validação Externa de ferramenta de mercado	Lancaster et al. (2022)	Avalia o desempenho de ferramenta autônomo na tarefa de avaliação de nódulos pulmonares em imagens de TC com dose ultrabaixa de radiação.

O objetivo principal deste trabalho é a implementação de metodologia que atenda as necessidades heterogênicas de diferentes conjuntos de dados para o cálculo volumétrico de nódulos pulmonares através de segmentação semântica, seguindo as

recomendações de medição da Fleischner Society (MACMAHON et al., 2017; BANKIER et al., 2017). Como objetivos específicos têm-se: i) construção do modelo que resolva o problema das necessidades heterogênicas, ii) treino e validação interna⁶ do modelo de segmentação semântica e iii) avaliação da robustez da solução⁷ que será medida por meio da análise de variabilidade de volume e diâmetro dos nódulos pulmonares durante o tratamento dos pacientes.

A relevância deste trabalho reside no aprimoramento do processo de tomada de decisão clínica por médicos, ao proporcionar maior agilidade através da redução do tempo necessário para a análise de imagens de TC e ao diminuir a variabilidade na quantificação de nódulos pulmonares. Com relação a razoabilidade, observa-se que a efetiva utilização desta ferramenta requer a realização de série de etapas de validação clínica, integração ao sistema do usuário e adaptação dos profissionais ao uso da tecnologia. Quanto à aplicabilidade, a ferramenta proposta é destinada à utilização na área médica hospitalar.

O presente trabalho contém a estrutura: o Capítulo 2 apresenta os conceitos fundamentais de imagiologia aplicados à análise de nódulos pulmonares, enquanto o Capítulo 3 aborda os princípios básicos de inteligência artificial. Estes capítulos fornecem o embasamento teórico necessário para a compreensão da metodologia e dos resultados deste trabalho. O Capítulo 4 descreve a metodologia proposta para o cálculo do volume nodular, incluindo os procedimentos de teste e validação. O Capítulo 5 apresenta os resultados alcançados e os discute criticamente, ao passo que o Capítulo 6 expõe as conclusões do estudo.

⁶Validação interna testa o desempenho de modelos de IA no mesmo conjunto de dados utilizado para treinamento, frequentemente utilizando técnicas como validação cruzada para verificar a generalização interna do modelo. Em contraste, a validação externa avalia modelos em conjunto de dados de origem distinta, proporcionando medida de sua eficácia em condições reais e variadas (YU et al., 2022)

⁷Neste trabalho, o termo modelo refere-se ao algoritmo que gera máscaras dos nódulos pulmonares, enquanto o termo solução indica o modelo treinado, testado e pronto para uso prático.

CAPÍTULO 2

IMAGIOLOGIA DE NÓDULOS PULMONARES

Neste capítulo são descritas as causas potenciais dos nódulos pulmonares, sua caracterização e os fatores que influenciam a probabilidade de malignidade, fundamentando a conduta clínica. Também fornece o conceito de segmentação de imagens, destacando técnicas clássicas e aquelas com base em aprendizado de máquina. Por fim, este capítulo apresenta características das imagens de tomografia computadorizada e o padrão de formatação de imagens médicas.

2.1 Problemas relacionados a imagens de nódulos

Os nódulos pulmonares representam pequenas formações anormais nos pulmões, cujas origens podem estar associadas a variedade de processos patológicos, como granulomas decorrentes de infecções, proliferação celular anormal não neoplásica (neurofibromas), infecções fúngicas (histoplasmose), doenças autoimunes e processos neoplásicos. A definição de nódulo pulmonar está na natureza da sua dimensão, sendo considerado nódulo se apresentar até $3cm$ de diâmetro, enquanto formações maiores são classificadas como massas pulmonares. Estes nódulos são categorizados em três classes distintas: i) sólidos pequenos, com diâmetro $< 8mm$; ii) sólidos grandes, com diâmetro $\geq 8mm$ e iii) sub-sólidos, subdivididos em nódulos em vidro fosco na ausência de componentes sólidos e nódulos parcialmente sólidos, quando contêm parcela em vidro fosco e parcela sólida (NATHAN et al., 1962; MURTHY; RICE, 2002; BUENO et al., 2018; MAZZONE; LAM, 2022).

A avaliação da probabilidade de malignidade de nódulos pulmonares leva em consideração vários fatores, incluindo tamanho, taxa de crescimento, formato, localização e fatores de risco do paciente. Observa-se que a maioria, $\approx 95\%$, dos nódulos pulmonares é benigna. Nódulos com diâmetro de até $8mm$ apresenta baixa probabilidade de malignidade, com taxas de malignidade $T_m \approx 1\%$ para nódulos menores que $6mm$ e $1\% \leq T_m \leq 2\%$ para nódulos com diâmetro de $6mm$ a $8mm$. No entanto, a probabilidade de malignidade para nódulos com diâmetro de $8mm$ a $30mm$ varia consideravelmente, indo de $1\% \leq T_m \leq 70\%$, dependendo dos fatores de risco do paciente, formato, localização e presença de calcificação (MAZZONE; LAM, 2022).

A conduta a ser adotada quando o nódulo pulmonar é identificado na tomografia computadorizada de tórax depende das diretrizes clínicas estabelecidas. Por exem-

plô, a Fleischner Society propõe abordagens distintas com base no diâmetro e volume dos nódulos, levando também em consideração o perfil de risco do paciente. Em contraste, as diretrizes do Lung-RADS consideram principalmente o diâmetro do nódulo como critério para determinar a conduta clínica ([American College of Radiology Committee on Lung-RADS, 2022](#); [NAIDICH et al., 2013](#); [BUENO et al., 2018](#)).

Portanto, o mapeamento do crescimento do nódulo é fundamental para sua caracterização. Sendo assim, a medição destas propriedades deve seguir o padrão para garantir consistência e confiabilidade, reduzindo as incertezas. Conforme as recomendações da Fleischner Society, ao realizar a medição manual com paquímetros eletrônicos, o diâmetro dos nódulos pulmonares deve ser aferido no eixo axial, com base na média do diâmetro máximo e do respectivo diâmetro perpendicular máximo. No caso de medição volumétrica automática ou semiautomática, é aconselhável empregar o mesmo software para a monitoramento do crescimento dos nódulos ([BANKIER et al., 2017](#)).

2.2 Segmentação de imagens médicas

A tarefa de segmentação de imagem está intrinsecamente ligada à identificação e diferenciação de estruturas com semelhanças em suas características, como intensidade de pixel, coloração, profundidade e textura. O intuito é realizar esta segmentação de maneira homogênea, evitando qualquer forma de sobreposição. A metodologia adotada para a reformulação da imagem envolve a classificação de todos os pixels presentes na imagem original, agrupando-os em conjuntos que representam entidades discretas e distintas. O resultado desejado consiste na obtenção da máscara de segmentação, que delinea o conjunto de regiões na imagem por meio da aplicação de rótulos de interesse ([SETAREHDAN; SINGH, 2001](#); [MOORTHY; GANDHI, 2022](#); [DESPOTOVIĆ et al., 2015](#); [PHAM et al., 2000a](#)). Esta máscara pode ser gerada por meio da aplicação de técnicas tradicionais da área de visão computacional ou adotando abordagens de aprendizado de máquina.

As técnicas clássicas de segmentação de imagem abarcam operações de pré-processamento, como a aplicação de filtros, detecção de bordas, limiarização (thresholding) e clusterização. A aplicação de filtros visa reduzir ruídos presentes na imagem e realçar estruturas de interesse específicos, enquanto a detecção de bordas tem o intuito de identificar transições abruptas de intensidade de pixel na imagem. A limiarização realiza a divisão dos pixels em diferentes regiões com base no valor

de limiar predefinido, atribuindo-lhes rótulos específicos. Por último, a clusterização agrupa pixels com características semelhantes. No entanto, é importante destacar que técnicas de segmentação clássica podem apresentar sensibilidade a variações na intensidade dos pixels, exigindo ajustes específicos para diferentes tipologias de imagens (SHARMA; AGGARWAL, 2010; DESPOTOVIĆ et al., 2015; PHAM et al., 2000a).

Frente a estas limitações, os métodos de aprendizado de máquina se destacam pela robustez e eficiência. Estes métodos utilizam algoritmos de aprendizado, podendo ser supervisionados ou não supervisionados, com o intuito de automatizar a identificação de padrões e características nas imagens analisadas. A obtenção da máscara por meio do aprendizado profundo de máquina ganha relevância com o desenvolvimento da Fully Convolutional Network. Sua singularidade reside na substituição das camadas totalmente conectadas por camadas convolucionais, possibilitando que a rede processe imagens de tamanhos diversos e gere mapas de segmentação com alta resolução (LONG et al., 2015).

Na área da medicina, foi concebida a rede com arquitetura U-Net, especialmente voltada para a tarefa de segmentação de imagens biomédicas, embora sua aplicabilidade se estenda a diversas outras áreas. Esta rede é projetada com base em estrutura em forma de U, composta por caminho de contração para capturar o contexto e caminho de expansão simétrico para permitir a localização precisa. Esta configuração ajuda a rede a aprender representações em várias escalas, sendo particularmente eficiente na segmentação de imagens, na qual é fundamental distinguir entre diferentes estruturas anatômicas ou patológicas. A U-Net destaca-se em vários desafios de segmentação devido à sua capacidade de trabalhar com quantidade limitada de dados e à sua habilidade de generalização a partir de amostras de treinamento anotadas (RONNEBERGER et al., 2015).

Para treinar qualquer abordagem supervisionada de IA, é necessário realizar a segmentação da imagem. Assim, utiliza-se a segmentação manual, criada por especialistas no contexto do problema, para avaliar artefatos nas imagens (LIU et al., 2021). Na área da medicina, na qual imagens volumétricas como TC são comuns, uma das ferramentas utilizadas para visualização, processamento e segmentação manual e semiautomática é a Plataforma 3D Slicer, software de código aberto para computação de imagens médicas (EGGER et al., 2013). Esta ferramenta permite manipular detalhadamente os dados de imagem, possibilitando a segmentação de estruturas específicas, como tecidos ou tumores. Utilizando técnicas como marcação manual e

ajustes finos da segmentação, os usuários podem alcançar alto grau de precisão (FEDOROV et al., 2012).

2.3 Tomografia computadorizada

A tomografia computadorizada (TC) é a ferramenta diagnóstica e terapêutica importante, proporcionando imagens volumétricas de alta resolução espacial e contraste aprimorado. Estas características oferecem visão detalhada das estruturas internas do tórax. As Unidades Hounsfield [HU] são fundamentais na interpretação destas imagens, fornecendo medida quantitativa da radiodensidade para distinguir entre diferentes tipos de tecidos e patologias (HOUNSFIELD, 1973; KALENDER, 2006). Tanto o tamanho quanto o espaçamento dos vóxeis são determinantes para a qualidade da imagem tridimensional, influenciando diretamente a resolução espacial, que é a capacidade de discernir pequenas estruturas anatômicas. A resolução de contraste, por sua vez, define a habilidade do sistema de TC em diferenciar tecidos com densidades semelhantes, sendo necessária para a detecção de alterações sutis nos tecidos (KALENDER, 2006; GOLDMAN, 2007).

A TC se destaca na área da imagem médica devido à sua capacidade de proporcionar representação volumétrica detalhada do tórax. Esta representação é composta por várias fatias, que são imagens transversais do corpo adquiridas em planos paralelos durante o exame. A espessura destas fatias é parâmetro importante que influencia a resolução espacial da imagem. Fatias mais finas melhoram a resolução, possibilitando a detecção de pequenas estruturas e detalhes anatômicos, enquanto fatias mais espessas podem reduzir a resolução, ao mesmo tempo em que aumentam a velocidade de aquisição da imagem e reduzem a dose de radiação para o paciente (KALENDER, 2006; GOLDMAN, 2007).

A isotropia da imagem refere-se à uniformidade das resoluções espaciais nas três dimensões: axial, coronal e sagital. Imagens isotrópicas possibilitam a reconstrução em múltiplos planos sem perda de detalhes, sendo necessária para a segmentação e diagnóstico preciso. Em contraste, imagens anisotrópicas apresentam resoluções diferentes nas três dimensões, o que pode limitar algumas análises e aplicações clínicas (KALENDER, 2006). Entretanto, as imagens de TC do tórax estão suscetíveis a artefatos decorrentes de diversos fatores, como limitações do equipamento, movimentos do paciente e presença de materiais metálicos que interferem na passagem dos raios-X. Estes artefatos podem mimetizar patologias ou comprometer a qua-

lidade da imagem a ponto de tornar o diagnóstico desafiador (KALENDER, 2006; GOLDMAN, 2007; BOAS; FLEISCHMANN, 2012).

O padrão atualmente utilizado para o formato de imagens médicas é o Digital Imaging and Communications in Medicine (DICOM). Ele define como as imagens médicas e informações associadas são trocadas entre diferentes sistemas de tecnologia da informação em saúde, estabelecendo o padrão para armazenamento, compartilhamento e gerenciamento de imagens médicas. O DICOM suporta ampla gama de dados médicos, incluindo imagens radiológicas, ultrassonografias, tomografias e ressonâncias magnéticas (RM), além de conter informações do paciente, detalhes do exame e parâmetros de imagem, como características espaciais (PIANYKH; PIANYKH, 2012). Este padrão facilita a interoperabilidade entre sistemas distintos, permitindo a transição suave de imagens e dados associados através de instituições de saúde, independentemente do fabricante do equipamento. Além disto, o DICOM incorpora medidas de segurança para proteger a confidencialidade e a integridade dos dados do paciente (PIANYKH; PIANYKH, 2012; HUANG, 2011).

2.4 Considerações finais

Este capítulo abordou as características dos nódulos pulmonares, assim como diretrizes para a tomada de decisões clínicas com base no tamanho e no crescimento destes nódulos. Foram apresentados conceitos e técnicas de segmentação de imagem, além da utilização da tomografia computadorizada como ferramenta necessária para a visualização de estruturas anatômicas e patologias. Estes conceitos serão utilizados para aplicação de modelo de segmentação automática que será abordado no Capítulo 3.

CAPÍTULO 3

INTELIGÊNCIA ARTIFICIAL

Este capítulo discorre sobre as características fundamentais do modelo de segmentação automática nnU-Net, com base em aprendizado profundo. É detalhado a metodologia adotada para a configuração automática do pré-processamento, da arquitetura da rede neural e do pós-processamento. No que concerne à arquitetura da rede, são incluídas as quatro variações possíveis, destacando-se as particularidades de cada uma.

3.1 Conceituação

Inteligência artificial pode ser definida como a capacidade de uma máquina replicar a habilidade humana de resolver problemas (RUSSELL; NORVIG, 2020). Contudo, algoritmos que empregam técnicas de IA buscam resolver problemas específicos em contextos específicos, e não configuram inteligência geral. Neste contexto, surge a diferenciação entre a tecnologia atual denominada inteligência artificial estreita (*Narrow Artificial Intelligence* – NAI) e a vindoura tecnologia de inteligência artificial geral (*Artificial General Intelligence* – AGI) (GOERTZEL, 2014). A capacidade da NAI de replicar a habilidade humana de resolver problemas específicos pode ser adquirida por meio do processo de aprendizado de máquina, em que a capacidade de resolver problemas é aprendida em conjunto de dados e generalizada para novos dados (VALIANT, 1984; CORTES; VAPNIK, 1995; SHALEV-SHWARTZ; BEN-DAVID, 2014).

Uma forma de se conduzir o aprendizado de máquina é mediante a adoção do paradigma de aprendizado supervisionado, em que o modelo é treinado utilizando dados de interesse e seus respectivos rótulos. Desta forma, é possível calcular erros e aplicar ajustes aos parâmetros do modelo (RUMELHART et al., 1986; LECUN et al., 2015). A forma dos dados, rótulos e modelo dependem da natureza do problema. Em visão computacional, os dados são imagens, e os modelos de aprendizado de máquina tem base na técnica de convolução, resultando nas Redes Neurais Convolucionais (*Convolutional Neural Network* – CNN) (KRIZHEVSKY et al., 2012). Por outro lado, os rótulos podem assumir diferentes formatos dependendo da tarefa a ser executada. Na tarefa de classificação de imagens, os rótulos são as classes verdadeiras do conteúdo das imagens. De forma mais elaborada, na tarefa de segmentação, o rótulo

consiste em imagem em que os píxeis assumem valores que representam as classes dos objetos de interesse.

A arquitetura de redes neurais totalmente convolucionais (*Fully Convolutional Neural Networks* – FCNN) marca avanço significativo na segmentação de imagens, possibilitando aprendizado profundo do modelo de ponta-a-ponta e processamento da imagem completa (LONG et al., 2015). A FCNN alavanca o desenvolvimento da U-Net, arquitetura de rede projetada especificamente para a segmentação de imagens médicas, que se destaca pela sua eficiência em trabalhar com quantidades limitadas de dados e proporcionar segmentação precisa ao incorporar conexões residuais entre os módulos de extração de características e convolução de aumento de resolução para capturar contexto, garantindo localização precisa (RONNEBERGER et al., 2015).

A relevância da U-Net se estende por várias aplicações no campo da medicina, modernizando a análise de imagens médicas ao facilitar a segmentação detalhada de estruturas anatômicas e patológicas complexas (DONG et al., 2017; ISENSEE et al., 2018; FALK et al., 2019). No entanto, a U-Net apresenta limitações operacionais quanto à sua flexibilidade e otimização, o que demanda intervenção de especialista na segmentação. A nnU-Net, variação da U-Net, foi desenvolvida para superar estes problemas por meio de abordagem automatizada que ajusta sua configuração para maximizar o desempenho em variedade de problemas de segmentação médica (ISENSEE et al., 2024).

3.2 Avaliação de modelos de inteligência artificial

A metodologia para avaliar o desempenho do modelo de inteligência artificial varia conforme a tarefa executada. Modelos de regressão, por exemplo, são comumente avaliados por métricas como erro quadrático médio e sua raiz, erro absoluto médio e coeficiente de determinação R^2 (BOTCHKAREV, 2018). Para classificação, utiliza-se acurácia, sensibilidade, especificidade, precisão, e área sob a curva de características operacionais do receptor (*Receiver Operating Characteristic - Area Under the Curve* – ROC-AUC) (STEHMAN, 1997; POWERS, 2020; DAVIS; GOADRICH, 2006). Na segmentação, os coeficientes de Jaccard J_c e Dice D_c são padrões (TANIMOTO, 1958; DICE, 1945; SUDRE et al., 2017).

A segmentação de imagens médicas é avaliada por diversas métricas, sendo os coeficientes de Jaccard J_c e Dice D_c os mais utilizados (TANIMOTO, 1958; DICE, 1945;

SUDRE et al., 2017). O coeficiente J_c , definido na equação (3.1), expressa a razão entre a interseção e a união das máscaras original e segmentada, enquanto D_c , dado pela equação (3.2), atribui peso dobrado à interseção. Ambos variam no intervalo $[0,1]$, indicando maior precisão à medida que se aproximam de 1. O coeficiente D_c é particularmente relevante como função de perda no treinamento supervisionado de modelos de segmentação, sendo minimizado conforme seu valor aumenta, dado pela equação (3.3). Este comportamento reflete a melhoria da precisão da segmentação por meio do algoritmo Backpropagation (RUMELHART et al., 1986). Esta abordagem permite lidar de forma eficiente com condições variáveis, desbalanceamento de classes e dados esparsos, uma vez que cada píxel corretamente classificado impacta significativamente na função de perda (SUDRE et al., 2017).

$$J_c(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (3.1)$$

$$D_c(A, B) = \frac{2 \cdot |A \cap B|}{|A| + |B|} \quad (3.2)$$

$$Dice\ Loss(A, B) = 1 - D_c(A, B) \quad (3.3)$$

Além da sobreposição entre máscaras segmentadas, a métrica 3D Hausdorff Distance (DH) é utilizada para quantificar discrepâncias nas fronteiras volumétricas segmentadas (AGARWAL et al., 2010). A DH mede a maior distância entre quaisquer pontos de dois conjuntos, sendo formalmente definida pela equação (3.4), onde $d_{3D}(a, b)$ representa a distância Euclidiana entre os pontos a e b , dado pela equação (3.5). A primeira parte da equação (3.4) identifica o ponto de A mais distante de qualquer ponto B , enquanto a segunda parte realiza a operação inversa para B . O maior destes valores representa o pior alinhamento entre as fronteiras volumétricas (AGARWAL et al., 2010). Na segmentação 3D, valores baixos de $DH_{3D}(A, B)$ indicam que as fronteiras segmentadas estão próximas das reais, enquanto valores elevados sugerem discrepâncias significativas. Esta métrica é empregada na avaliação do erro máximo da segmentação, embora apresente sensibilidade a valores atípicos (T. et al., 2018; V.; I., 2017).

$$DH_{3D}(A, B) = \max \left\{ \max_{a \in A} \min_{b \in B} d_{3D}(a, b), \max_{b \in B} \min_{a \in A} d_{3D}(b, a) \right\}, \quad (3.4)$$

$$d_{3D}(a, b) = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2 + (a_z - b_z)^2}. \quad (3.5)$$

Embora amplamente utilizada, a métrica DH apresenta limitações em segmentações classificatórias, o que motivou o desenvolvimento de versões balanceadas para minimizar erros (AYDIN et al., 2021). Tanto a DH quanto o coeficiente D_c possuem baixa relevância clínica, podendo não refletir com precisão as implicações médicas (KIM et al., 2015). Para superar esta limitação, são propostas métricas mais alinhadas ao contexto clínico, como o Índice de Similaridade Médica, que estabelece a correlação entre a qualidade da segmentação e a precisão volumétrica (MORAIN-NICOLIER et al., 2007; KIM et al., 2015). Além disto, estudos apontam variações nas implementações métricas entre diferentes ferramentas, ressaltando a necessidade de padronização. Métodos fundamentados na DH têm sido aplicados para quantificar a progressão tumoral cerebral e avaliar a localização do câncer de próstata (MORAIN-NICOLIER et al., 2007), ampliando sua aplicabilidade em contextos clínicos.

Dado o impacto das métricas na avaliação de segmentações médicas, torna-se fundamental adotar estratégias de validação que garantam a confiabilidade dos modelos. A validação cruzada é importante neste processo, pois contribui para a generalização dos resultados (GEISSER, 1975; STONE, 1974; ARLOT; CELISSE, 2010). No método hold-out, os dados são divididos aleatoriamente, enquanto na validação cruzada k -fold, o conjunto é particionado em k subconjuntos, sendo treinado em $k - 1$ deles e validado no subconjunto restante (HASTIE et al., 2009). Este procedimento possibilita ajustes de hiperparâmetros e reduz o risco de sobreajuste (HAWKINS, 2004). No caso da nnU-Net, a validação cruzada k -fold é amplamente adotada, combinando a função de perda de Dice com entropia cruzada para otimização (FABIAN et al., 2020). A avaliação de desempenho é conduzida por meio de coeficientes como D_c e J_c , fornecendo indicadores quantitativos da qualidade da segmentação.

A segmentação pulmonar constitui um dos principais focos de pesquisa na aplicação da visão computacional à medicina. Estudos recentes indicam a eficiência das CNN, especialmente das arquiteturas U-Net, na segmentação de pulmões em imagens de TC, com coeficientes D_c variando entre 82,61% e 97,75% (QINHUA et al., 2020). O de-

sempenho da U-Net decorre de sua capacidade de preservar informações espaciais e extrair características em múltiplas escalas. Para aprimorar a segmentação, técnicas de pré-processamento, como o aprimoramento com base em histograma, têm sido investigadas (AHMET *et al.*, 2023). Além disto, estudos comparativos demonstram que modelos CNN pré-treinados podem proporcionar resultados superiores (QINHUA *et al.*, 2020). Além da U-Net, outras abordagens vêm sendo exploradas na segmentação pulmonar. A Mask R-CNN apresenta alta precisão, atingindo 97,68% (QINHUA *et al.*, 2020; SOUZA *et al.*, 2021). Técnicas de segmentação automatizada têm se apresentado promissoras no diagnóstico e na detecção de doenças pulmonares, proporcionando maior rapidez e precisão em comparação à segmentação manual (OULEFKI *et al.*, 2021; FRANSISKUS *et al.*, 2022). Assim, ferramentas especializadas são desenvolvidas para a segmentação automática de pulmões em imagens de TC, destacando-se a Lungmask, amplamente empregada no processamento destas imagens (GANG *et al.*, 2019; BONDARENKO; SYRYH, 2024).

3.3 nnU-Net

A U-Net, apesar de sua eficiência, apresenta limitações quando aplicada a ampla variedade de tarefas de segmentação sem otimizações específicas para cada conjunto de dados. Este desafio é aumentado pela exigência de conhecimento especializado em aprendizado profundo de máquina, no qual ajustes mínimos no processamento podem resultar em perdas significativas de desempenho (LITJENS *et al.*, 2017). Além disto, intervenções manuais frequentemente levam a otimizações subótimas (BERGSTRA; BENGIO, 2012). Neste contexto, destaca-se a no-new-U-Net (nnU-Net), que avança além da U-Net original, oferecendo arquitetura capaz de se adaptar automaticamente a ampla gama de configurações de dados de segmentação, eliminando a necessidade de ajustes manuais ou aprimoramento por especialistas. Esta flexibilidade é necessária diante da heterogeneidade e complexidade das imagens médicas e dos desafios associados à segmentação (FABIAN *et al.*, 2020).

A nnU-Net é fundamentada em três pilares essenciais: i) adaptação de pré-processamento, ii) personalização da arquitetura da rede e iii) refinamento de pós-processamento. No pré-processamento, o conjunto de características é extraído do conjunto de dados (data fingerprint), permitindo à nnU-Net ajustar automaticamente a normalização das imagens, incluindo redimensionamento, alterações nas características espaciais e transformações na intensidade dos pixels. Com base nestas características, a configuração da U-Net é otimizada para preservar as informa-

ções contextuais necessárias ao longo das camadas do modelo e definir a estratégia de aumento de dados que maximiza o desempenho em segmentações específicas. No pós-processamento, a nnU-Net utiliza abordagens empíricas para aprimorar as segmentações, como a seleção de componentes conectados orientada por heurísticas específicas do conjunto de dados (FABIAN et al., 2020).

Existem três variantes da U-Net que a nnU-Net pode implementar: (i) 2D U-Net, (ii) 3D full-resolution U-Net (3D F), (iii) 3D low-resolution (3D L), and (iv) 3D cascade U-Net (3D C) (PUNN; AGARWAL, 2022). A U-Net 2D é ideal para o processamento de imagens fatia a fatia, sendo adequada para dados anisotrópicos, como conjuntos de imagens com variação no espaçamento entre fatias. A U-Net 3D F opera com imagens em tamanho original e frequentemente produz os melhores resultados. No entanto, ao processar imagens de alta resolução, esta variante pode enfrentar desafios devido ao tamanho da dimensão espacial da região retangular, resultando em perda de contexto e segmentações imprecisas. Na U-Net 3D C, ocorre a abordagem em duas fases projetada para imagens de alta resolução. Na primeira fase, a rede aprende a gerar segmentações rudimentares em baixa resolução, e na segunda fase, aprimora estas segmentações em alta resolução. Esta abordagem é particularmente útil para imagens com alta resolução, nas quais o tamanho da dimensão espacial pode se tornar excessivamente pequena (LI et al., 2022). A abordagem adaptativa da nnU-Net oferece arquitetura de segmentação de imagem versátil e eficiente. Além de preservar as vantagens da U-Net original em termos de eficiência e precisão, ela amplia sua aplicabilidade, reduzindo a necessidade de intervenção manual (FABIAN et al., 2020).

3.3.1 Pré-processamento

A nnU-Net automatiza a segmentação de imagens médicas por meio de técnicas de pré-processamento, como normalização de intensidade, recorte e reamostragem, que são ajustadas dinamicamente com base no *data fingerprint* e no *pipeline fingerprint* (FABIAN et al., 2020). O *data fingerprint* armazena informações sobre o conjunto de dados, incluindo tamanho, espaçamento e distribuição das intensidades, permitindo análises estatísticas detalhadas (K. et al., 2021; D. et al., 2023), enquanto o *pipeline fingerprint* registra as configurações ideais para pré-processamento, arquitetura e pós-processamento. O recorte reduz o tamanho das imagens, concentrando-se nas regiões de interesse para otimizar o custo computacional, ao mesmo tempo em que preserva a estrutura original para reconstrução posterior. Eliminando a necessidade

de ajustes manuais, a nnU-Net emprega parâmetros fixos, regras interdependentes e decisões empíricas para maximizar o desempenho da segmentação (ISENSEE et al., 2019). Superando soluções especializadas em diversos conjuntos de dados públicos, sem intervenção humana (FABIAN et al., 2020; RICKY, 2021), a nnU-Net demonstra alta adaptabilidade a diferentes tarefas, tornando a segmentação automatizada mais acessível para aplicações clínicas (CHAYMAE; NAJIBA, 2023).

A normalização da intensidade é necessária para otimizar o desempenho da segmentação de imagens médicas, especialmente em RM e TC (SHINOHARA et al., 2014; REINHOLD et al., 2019). Em TC, os valores de intensidade são limitados a percentis predefinidos e padronizados por média e desvio padrão, enquanto em RM, a normalização z -score é aplicada para garantir consistência nos dados (ONOFREY et al., 2019). Técnicas de normalização, como equalização de histograma e z -score, demonstram aprimorar a precisão da segmentação e a qualidade da síntese de imagens (JACOBSEN et al., 2019; CARRÉ et al., 2020). A reamostragem padroniza o espaçamento das imagens considerando a anisotropia dos dados (JUCEVICIUS et al., 2022). Para imagens isotrópicas, aplica-se interpolação por *spline* de terceira ordem, enquanto imagens anisotrópicas utilizam esta técnica nos eixos de altura e largura, combinada à interpolação por vizinho mais próximo no eixo de profundidade (HU et al., 2021). O espaçamento final é definido como o modelo adotado: na U-Net 3D F, utiliza-se a média dos espaçamentos das imagens de treinamento, enquanto na U-Net 2D, a mediana do espaçamento determina os eixos empregados na extração de fatias, reduzindo variações estruturais das imagens (JESUS et al., 2020; CARRÉ et al., 2020).

A nnU-Net ajusta automaticamente sua arquitetura e hiperparâmetros para otimizar a segmentação, considerando as características do conjunto de dados e as restrições computacionais (FABIAN et al., 2020; ISENSEE et al., 2019). Sua estrutura tem base no modelo U-Net padrão, sendo modificada iterativamente para definir o número de estágios, filtros e tamanhos dos *kernels* de convolução (ISENSEE et al., 2018). O *downsampling* é aplicado de forma independente em cada eixo, reduzindo a resolução até que o mapa de características atinja quatro ou menos voxels ou se torne anisotrópico. O tamanho padrão dos *kernels* é $3 \times 3 \times 3$ para U-Net 3D e 3×3 para U-Net 2D, sendo ajustado caso a razão de espaçamento entre eixos ultrapasse dois. O *patch size* utilizado no treinamento é definido com base na mediana das dimensões das imagens reamostradas, sendo adaptado à capacidade da memória da GPU (*Graphics Processing Unit*). Caso a VRAM (*Video Random Access Memory*)

não suporte o tamanho inicial, o processo iterativo reduz o *patch size*, reconfigurando a arquitetura até que o modelo possa ser treinado sem ultrapassar os limites computacionais (FABIAN et al., 2020).

Sem a necessidade de intervenção manual, a nnU-Net supera soluções especializadas em diversos conjuntos de dados e competições (FABIAN et al., 2020; ISENSEE et al., 2019). Estudos recentes investigam a incorporação de blocos avançados, como residual, dense e inception, demonstrando ganhos de desempenho dependentes do conjunto de dados (MCCONNELL et al., 2022). Versões aprimoradas, como a nnU-Net de fluxo duplo, apresentam resultados promissores para a segmentação de tumores (CHAYMAE; NAJIBA, 2023). Apesar do surgimento de novas arquiteturas, pesquisas indicam que variantes configuradas da U-Net, como ResNet e ConvNeXt, continuam apresentando alto desempenho quando integradas ao framework nnU-Net (ISENSEE et al., 2024).

3.3.2 Definição de Arquitetura e Hiperparâmetros

A definição da arquitetura e a escolha dos hiperparâmetros são cruciais para o desempenho do modelo nnU-Net. Este processo envolve a adaptação da topologia da rede e a seleção de parâmetros específicos para treinamento e inferência.

A partir de um template U-Net, a topologia é ajustada iterativamente, configurando-se a quantidade de estágios, filtros e tamanhos dos *kernels* de convolução conforme as características do conjunto de dados, visando otimizar a utilização de recursos computacionais e garantir bom desempenho na tarefa de segmentação. A quantidade de operações de *downsampling* é determinada pelo número de aplicações necessárias para reduzir o mapa de características a quatro ou menos voxels, ou até que o mapa se torne anisotrópico. O *downsampling* ocorre de maneira independente para cada eixo. Inicialmente, caso existam eixos de maior resolução, estes são reduzidos até atingirem um fator de 2 do eixo de menor resolução, ou seja, os eixos maiores são reduzidos até serem, no mínimo, duas vezes maiores que o eixo menor; em seguida, todos os eixos são reduzidos simultaneamente. Uma vez que um eixo atinja a condição de interrupção, o *downsampling* para esse eixo é encerrado, e continua-se para os outros eixos, caso seja possível.

O tamanho padrão do *kernel* para convoluções mantém-se em $(3 \times 3 \times 3)$ para U-Net 3D e (3×3) para U-Net 2D. Caso exista diferença inicial de resolução entre os eixos,

ou seja, uma razão de espaçamento maior que dois, o tamanho do *kernel* é ajustado para reduzir a razão dos espaçamentos. Após este ajuste, o tamanho do *kernel* de convolução permanece três para todos os eixos.

O tamanho da porção da imagem utilizada durante o treinamento (*Patch Size*) é ajustado de acordo com a capacidade de memória da GPU. Inicialmente, o tamanho do *patch* é definido como a mediana do tamanho das imagens após a reamostragem, e verifica-se se a quantidade de VRAM é suficiente. Caso a VRAM não suporte o tamanho do *patch*, inicia-se um processo iterativo de redução do tamanho do *patch*, atualização da arquitetura conforme o novo tamanho e nova verificação. Este processo é repetido até que a quantidade de VRAM seja suficiente. Finalmente, o tamanho do lote é definido para 2 caso o tamanho do *patch* tenha sido alterado; caso contrário, o tamanho do lote é aumentado até esgotar a capacidade da GPU.

3.3.3 Conjunto de modelos, pós-processamento e residual encoder

A nnU-Net otimiza automaticamente as etapas de pré-processamento, definição da arquitetura, treinamento e pós-processamento, ajustando-se às características do conjunto de dados (ISENSEE et al., 2018; FABIAN et al., 2020). Durante a inferência, seleciona a configuração mais adequada ou combina múltiplas configurações com base no coeficiente D_c médio das classes-alvo, calculado na validação cruzada. Quando múltiplas configurações são utilizadas, os resultados são combinados por meio da fusão das probabilidades Softmax, aprimorando a precisão da segmentação (RONNEBERGER et al., 2015; ISENSEE et al., 2019). No pós-processamento, a nnU-Net emprega a técnica de componentes conectados, avaliando iterativamente o impacto da supressão de pequenas regiões segmentadas (RICKY, 2021). Inicialmente, todas as classes de primeiro plano são tratadas como única classe, removendo-se componentes menores caso isto melhore o coeficiente D_c médio sem comprometer outras classes. Se esta estratégia for benéfica, ela é aplicada individualmente a cada classe, refinando ainda mais a segmentação.

A nnU-Net tem demonstrado desempenho superior em diversas aplicações médicas, como segmentação de tumores cerebrais (FABIAN et al., 2020) e detecção de pólipos (HUANG et al., 2024), superando soluções especializadas em múltiplas competições biomédicas (FABIAN et al., 2020). Seu êxito decorre da capacidade de adaptação a diferentes bases de dados sem necessidade de intervenção manual, tornando a segmentação automatizada mais acessível e eficiente para aplicações clínicas. Como

aprimoramento desta abordagem, a nnU-Net com Residual Encoder (nnU-Net ResEnc) introduz conexões residuais no codificador, facilitando o fluxo de gradientes e reduzindo dificuldades no treinamento de redes profundas (ISENSEE et al., 2022). Esta variante ajusta automaticamente seus parâmetros de configuração com base na capacidade do hardware disponível, utilizando predefinições escaláveis M, L e XL para otimizar tamanho do lote e do patch de acordo com a memória da GPU (ISENSEE et al., 2022).

Mantendo a estrutura de auto-configuração, a nnU-Net ResEnc preserva a automação do processo, enquanto as conexões residuais aprimoram a resiliência em segmentações volumétricas complexas, como demonstrado em benchmarks com os conjuntos de dados KiTS e AMOS (ISENSEE et al., 2019; FABIAN et al., 2020). Extensões recentes da nnU-Net incorporam blocos avançados, como residual, dense e inception, além de mecanismos de atenção, para aprimorar a segmentação (MCCONNELL et al., 2022; MCCONNELL et al., 2023). Modelos otimizados, como a Two-Stream nnU-Net, combinam informações estruturais e funcionais para maior precisão em aplicações como segmentação de tumores em cânceres de cabeça e pescoço (RICKY, 2021; CHAYMAE; NAJIBA, 2023).

3.4 Considerações finais

A nnU-Net oferece ferramentas para a segmentação de imagens médicas em ampla gama de contextos, com elevada capacidade de adaptação automática. Este modelo é capacitado para normalizar imagens anisotrópicas e de variados tamanhos. As diversas arquiteturas disponíveis conferem versatilidade ao modelo, permitindo a abordagem eficiente de desafios de segmentação em que o objeto de interesse varia significativamente em tamanho em relação às dimensões totais da imagem. Deste modo, a metodologia que será apresentada no próximo capítulo tem como elemento central a aplicação da nnU-Net.

CAPÍTULO 4

METODOLOGIA

Este capítulo fornece o contexto e os detalhes dos procedimentos empregados na construção da solução destinada à quantificação de nódulos pulmonares em imagens de tomografia computadorizada. Após a obtenção do conjunto de dados, são realizadas operações de formatação e processamento. Com os dados pré-processados, o modelo de segmentação é treinado, avaliado e validado internamente. Para validar a eficiência do sistema, a solução é implementada e seu desempenho é analisado.

4.1 Contextualização

Este estudo objetiva desenvolver método de cálculo volumétrico e de diâmetro do maior eixo em nódulos pulmonares de diferentes conjuntos de dados, utilizando segmentação semântica em conformidade com as diretrizes da Fleischner Society (MACMAHON et al., 2017; BANKIER et al., 2017). O objetivo é criar solução eficiente e versátil que supere os desafios impostos pela heterogeneidade dos dados, através do treinamento, validação interna e avaliação do modelo de segmentação. Este trabalho não pretende substituir a avaliação realizada por radiologistas torácicos na análise de imagens de tomografia computadorizada, ele busca auxiliar estes profissionais ao fornecer quantificação precisa, tanto em termos de diâmetro do maior eixo quanto de volume dos nódulos pulmonares. Para atingir este objetivo, o estudo se concentra no desenvolvimento de modelo de segmentação semântica com base em aprendizado profundo, capaz de identificar e quantificar nódulos pulmonares, facilitando assim a estimativa de seus volumes e diâmetros.

A solução proposta é estruturada em duas partes fundamentais: i) modelo de segmentação, responsável pela identificação dos nódulos pulmonares, e ii) rotina de quantificação, que calcula o volume e diâmetro dos nódulos a partir das máscaras geradas pela segmentação. Para a implementação e avaliação desta solução, são utilizados dois conjuntos de dados distintos: o primeiro serve para construir e realizar a validação interna do modelo de segmentação, enquanto o segundo é aplicado para testar a eficiência da rotina de quantificação. O processo integral da metodologia adotada é detalhadamente apresentado na Figura 4.1.

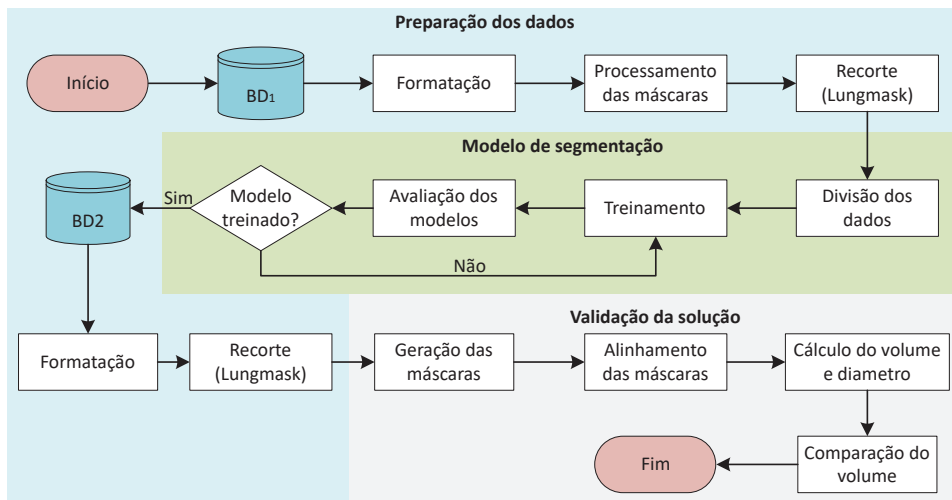


Figura 4.1 - Fluxograma da metodologia proposta.

4.2 Preparação dos dados

A metodologia se inicia com a preparação do conjunto de dados do BD_1 , designado para o treinamento e validação do modelo de segmentação. Este processo começa com a agregação das imagens, na qual cada fatia do exame, originalmente em arquivo DICOM individual, é compilada em único arquivo. Esta etapa é necessária, já que o modelo de segmentação necessita da análise da imagem completa para operar. Após a agregação das imagens, é realizada a conversão do formato DICOM para o Neuroimaging Informatics Technology Initiative (NIfTI). Esta mudança é justificada pelo fato de que o formato NIfTI é mais leve, devido à ausência de metadados, e proporciona maior facilidade de manipulação para operações de processamento de imagem, elementos importantes para a eficiência do processamento subsequente. Durante esta conversão, o novo arquivo é salvo com nome apropriado para ser processado pelo modelo nnU-Net. Posteriormente, as máscaras de segmentação originais presentes no BD_1 são processadas para se ajustarem ao desenvolvimento do modelo de segmentação.

4.2.1 Processamento de máscaras

As máscaras originais do conjunto de dados BD_1 são formatadas para consolidar múltiplas segmentações de único nódulo em única imagem e em único arquivo. Este processo se inicia com o estabelecimento de consenso entre as diversas segmentações realizadas por especialistas, disponíveis para cada nódulo. Para formar a máscara

de consenso, as várias segmentações são sobrepostas, e as áreas com sobreposição em pelo menos 60% das segmentações são incorporadas na nova máscara de consenso. Este procedimento é replicado para cada nódulo presente na imagem. Cada máscara de nódulo é inicialmente armazenada separadamente. Portanto, uma vez que as máscaras de consenso são criadas, é necessário compilar todas estas máscaras individuais em único arquivo para cada imagem de TC.

A unificação destas máscaras em volume único envolve não apenas a agregação das mesmas máscaras, mas também a incorporação de informações espaciais críticas da imagem de TC original. Estas informações incluem a origem, o espaçamento, o tamanho e a direção da imagem, que são importantes para garantir a correta correspondência e alinhamento espacial entre as máscaras e a respectiva imagem de TC. Este alinhamento é fundamental para garantir o correto treinamento do modelo de segmentação, assegurando que as máscaras reflitam adequadamente a localização e a extensão dos nódulos nas imagens de TC.

4.2.2 Recorte de volume

Para reduzir o tamanho da imagem e o tempo de processamento, além de restringir a região de busca, a imagem de TC é recortada para conter apenas o parênquima pulmonar. Com base na máscara segmentada do parênquima pulmonar, calculam-se as coordenadas da caixa limitadora que engloba esta região, permitindo o recorte da TC e da máscara dos nódulos pulmonares. Devido à presença de artefatos variados na imagem de TC, como vestimentas, objetos pessoais do paciente e a cama de exame, além de tecidos e estruturas externas ao parênquima pulmonar, há risco de que o modelo interprete equivocadamente estes elementos como nódulos pulmonares, problema associado ao fenômeno de Out-of-Distribution (OOD) ([HENDRYCKS et al., 2019](#)). Para reduzir este efeito, objetos segmentados fora do parênquima pulmonar são considerados amostras fora da distribuição e descartados.

4.3 Modelo de segmentação

Após a formatação das imagens de TC, o processamento das máscaras e o recorte subsequente das imagens e máscaras do BD_1 , inicia-se o treinamento do modelo de segmentação semântica nnU-Net. Os dados são divididos em subconjunto de treinamento e de teste, garantindo que não existam imagens do mesmo paciente entre subconjuntos. Posteriormente, é necessário criar o arquivo JavaScript Object Nota-

tion (JSON). Este arquivo descreve o conjunto de dados, permitindo que o fluxo de processamento da nnU-Net se ajuste à solução proposta. As informações detalhadas no arquivo JSON incluem a modalidade de imagem, as classes de segmentação e seus rótulos correspondentes.

Em seguida, a ferramenta `TaskCreator` é utilizada para criar a estrutura de diretórios necessária, pré-processar as imagens de entrada e gerar os arquivos de configuração para o treinamento do modelo de segmentação dentro da nnU-Net. Esta ferramenta extrai a impressão digital dos dados do subconjunto de treino e define as estratégias de adequação do fluxo de processamento, incluindo reamostragem das segmentações e das imagens, ajuste do espaçamento e normalização de intensidade. Além disso, a ferramenta `TaskCreator` realiza a divisão dos dados em subconjuntos, seguindo a metodologia de validação cruzada k -fold.

Este estudo utiliza diferentes configurações da nnU-Net para explorar as capacidades destas arquiteturas, incluindo U-Net 2D, U-Net 2D Residual Encoder M (2D REM), U-Net 2D Residual Encoder L (2D REL), U-Net 3D Fullres (3D F), U-Net 3D Fullres Residual Encoder M (3D FREM), U-Net 3D Lowres (3D L), U-Net 3D Lowres Residual Encoder M (3D LREM), U-Net 3D Cascada Fullres (3D CF), U-Net 3D Cascade Fullres Residual Encoder M (3D CFREM) (FABIAN et al., 2020). A otimização dos hiperparâmetros é realizada por meio da busca em grade, testando e avaliando diferentes combinações no conjunto de validação para identificar a configuração que maximiza o desempenho do modelo, utilizando o coeficiente D_c como métrica de avaliação. Este processo é iterativo, repetindo-se até que a combinação ótima de hiperparâmetros seja encontrada. A validação dos modelos de segmentação nnU-Net é realizada no subconjunto de teste, onde as máscaras geradas pelos modelos são comparadas com as máscaras originais para calcular métricas de avaliação, incluindo coeficiente D_c , J_c e HD_{3D} . Estas métricas permitem análise quantitativa da qualidade da segmentação. Além da avaliação da segmentação, a capacidade de detecção de nódulos é examinada por meio da quantificação de verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos, permitindo a mensuração da sensibilidade e especificidade do modelo.

A avaliação da capacidade de detecção do modelo de segmentação de nódulos pulmonares amplia a análise além das métricas D_c , J_c e HD_{3D} , permitindo a interpretação mais abrangente do desempenho do modelo. Verdadeiros positivos indicam que o modelo segmentou o nódulo previamente anotado por especialistas, independentemente

da precisão da segmentação. Falsos positivos ocorrem quando o modelo segmenta uma região sem anotação original, o que pode resultar de diversos fatores, como a presença de tecidos patológicos atípicos, dificuldades na diferenciação entre estruturas torácicas e nódulos ou limitações na qualidade das anotações. Por outro lado, falsos negativos indicam falhas na segmentação, em que o modelo não identificou o nódulo existente, ou casos em que a segmentação foi realizada corretamente, mas a anotação não atingiu consenso entre os especialistas. Este tipo de erro pode refletir desafios inerentes à heterogeneidade dos nódulos pulmonares, variabilidade na anotação manual ou limitações no conjunto de treinamento.

4.4 Validação da solução

Para validar a solução, inicialmente realiza-se a preparação dos dados, incluindo a formatação das imagens do banco de dados BD_2 . Neste processo, os arquivos DICOM individuais de cada fatia do exame são compilados em único arquivo e convertidos para o formato NIfTI, facilitando manipulações subsequentes e otimizando o processamento. Após o treinamento e validação do modelo de segmentação, bem como a formatação e recorte dos dados do BD_2 , aplica-se um passo adicional de registro entre as máscaras geradas em exames distintos de um mesmo paciente. A partir da máscara do parênquima pulmonar gerada pelo LungMask (HOFMANNINGER et al., 2020). Este procedimento consiste no alinhamento das máscaras por meio da transformação estimada, utilizando registro de imagens em três dimensões. Esta abordagem aprimora a coincidência espacial entre os nódulos pulmonares segmentados, minimizando desalinhamentos que poderiam inflar artificialmente a diferença de quantificação calculada.

Para avaliar o desempenho da solução, computa-se a diferença absoluta de volume e diâmetro de cada nódulo pulmonar segmentado nos exames espacialmente registrados. A análise da média μ e do desvio padrão σ destas diferenças permite quantificar a consistência da segmentação tridimensional, sendo desejáveis valores reduzidos de μ e σ . As diferenças também são analisadas por meio do gráfico de Bland-Altman e de testes estatísticos para verificar se existe diferença estatisticamente significativa nas medidas de volume e diâmetro (BLAND; ALTMAN, 1986). Além disto, a concordância entre múltiplos anotadores do conjunto de dados BD_1 pode ser avaliada por meio do coeficiente D_c , possibilitando a mensuração da variabilidade de interpretação quanto à presença, extensão e fronteiras dos nódulos pulmonares. Esta análise é necessária, visto que o processo de anotação em imagens volumétricas apresenta

elevada complexidade (BOGOT et al., 2005).

4.4.1 Cálculo do volume e diâmetro

O cálculo do volume dos nódulos pulmonares inicia-se com a geração das máscaras de segmentação utilizando os modelos nnU-Net. Em seguida, o volume é determinado com o auxílio da biblioteca Python SimpleITK (YANIV et al., 2018), que permite recuperar os dados de espaçamento da imagem original a partir do arquivo DICOM, correspondendo ao tamanho espacial de cada voxel. O volume total é então obtido pela multiplicação do número de voxels da máscara segmentada pelo espaçamento de cada voxel.

Para medir o diâmetro do maior eixo de cada nódulo, conforme as recomendações da Fleischner Society, primeiramente, os nódulos são segmentados e as máscaras resultantes são decompostas em componentes conectados. Para cada nódulo identificado, percorrem-se todas as fatias no plano axial z e, em cada uma, determina-se a caixa delimitadora bidimensional do objeto, ou seja, o menor retângulo alinhado aos eixos x e y que contém todos os voxels pertencentes ao nódulo naquela fatia. A extensão da caixa delimitadora é convertida para milímetros considerando os metadados espaciais do DICOM. Em cada corte axial, o diâmetro é definido como o maior valor entre a largura em $x \times \Delta_x$ e a altura em $y \times \Delta_y$. Finalmente, o diâmetro do maior eixo do nódulo é determinado pelo valor máximo obtido entre todas as fatias axiais.

4.5 Considerações finais

A metodologia proposta aborda o processamento do conjunto de dados em BD_1 com o objetivo de treinar e validar o modelo de segmentação semântica nnU-Net. A etapa importante deste processamento é a geração da máscara de consenso. Com base neste conjunto preparado, dois modelos de segmentação são desenvolvidos: a arquitetura U-Net 2D e a arquitetura U-Net 3D Fullres. Adicionalmente, o conjunto de dados em BD_2 é utilizado para testar a eficiência da solução desenvolvida. Este teste envolve a avaliação das diferenças dos volumes dos nódulos do mesmo paciente, porém identificados em diferentes imagens de TC, proporcionando medida da consistência e precisão da solução proposta em condições variadas. No próximo capítulo, serão apresentados os resultados obtidos com a aplicação da metodologia proposta.

CAPÍTULO 5

RESULTADOS

O capítulo de resultados aborda a implementação da metodologia proposta para preparar os dados de treinamento dos modelos de segmentação semântica nnU-Net, bem como a preparação dos dados necessários para validar a solução de quantificação do volume e diâmetro de nódulos pulmonares. A avaliação do desempenho dos modelos de segmentação tem base no D_c . Por outro lado, a avaliação da solução de quantificação concentra-se na análise das diferenças na μ e no σ dos volumes e diâmetros dos nódulos pulmonares medidos em imagens sucessivas do mesmo paciente em momentos temporais distintos. Após a apresentação de todos os resultados desta proposta, conduz-se a discussão que aborda de forma concisa os resultados observados na lacuna identificada.

5.1 Escolha das bases e pré-processamento dos dados

Para o treinamento e validação do modelo de segmentação de nódulos pulmonares, utilizou-se o conjunto de dados do *Image Database Consortium and Image Database Resource Initiative* (LIDC-IDRI) (ARMATO et al., 2011), armazenado em BD_1 . Este banco de dados consiste em imagens de tomografia computadorizada do tórax, acompanhadas de máscaras de segmentação para nódulos com diâmetro $\geq 3mm$, elaboradas por especialistas em radiologia torácica. No caso de nódulos $< 3mm$, o conjunto de dados fornece apenas as coordenadas espaciais do centroide, limitando a utilidade a nódulos maiores. O BD_1 contém 1018 imagens de TC, com um total de 2669 lesões segmentadas por pelo menos um radiologista, e 928 lesões avaliadas por quatro radiologistas. Não são disponibilizadas informações adicionais, como gênero, idade e origem das imagens, para esta base de dados. Durante a fase de preparação do BD_1 , das 1018 imagens de TC disponíveis, 1011 foram processadas para criar as máscaras de segmentação. As imagens restantes foram desconsideradas devido a problemas de disponibilidade na base de dados de origem, ou então problemas nos metadados dos arquivos DICOM, que impediram a geração das máscaras.

Para validar a solução, empregou-se o conjunto de dados do *Reference Image Database to Evaluate Response to Therapy in Lung Cancer Project* (RIDER) (III et al., 2008; ZHAO et al., 2009), na categoria Lung CT, armazenado em BD_2 . Este conjunto foi concebido para fornecer recursos destinados ao desenvolvimento e validação de

ferramentas voltadas para avaliar a resposta ao tratamento do câncer de pulmão. O BD_2 é caracterizado por incluir duas imagens de tomografia computadorizada por paciente, obtidas no mesmo dia, porém sem máscaras de segmentação. Composto por 64 imagens de 32 pacientes, a base de dados também não oferece informações detalhadas sobre gênero, idade ou origem das imagens. Durante o pré-processamento deste conjunto de dados, identificou-se apenas um exame de TC para um dos indivíduos, resultando em total de 31 pares de imagens. As imagens foram processadas de forma análoga as da base de dados BD_1 . A Figura 5.1 apresenta uma imagem da base de dados BD_1 e uma imagem da base de dados BD_2 .

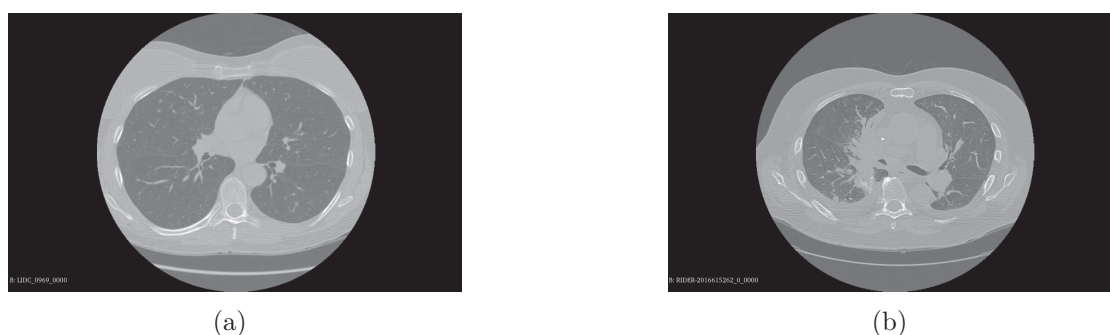


Figura 5.1 - Imagens de TC representativas das bases de dados: (a) BD_1 e (b) BD_2 .

Na construção dos subconjuntos, buscou-se uma distribuição equitativa de imagens contendo diferentes quantidades de nódulos. O objetivo foi assegurar equilíbrio entre a diversidade dos casos empregados no processo de aprendizado e daqueles utilizados para avaliação do desempenho. Como resultado, o subconjunto de treinamento foi composto por 822 imagens, enquanto o subconjunto de teste incluiu 189 imagens.

Na construção dos subconjuntos, buscou-se uma distribuição equitativa de imagens contendo diferentes quantidades de nódulos. O objetivo foi assegurar equilíbrio entre a diversidade dos casos empregados no processo de aprendizado e daqueles utilizados para avaliação do desempenho. Como resultado, o subconjunto de treinamento foi composto por 822 imagens, enquanto o subconjunto de teste incluiu 189 imagens.

5.1.1 Avaliação da concordância interobservador e desempenho do modelo de segmentação

Com o objetivo de avaliar o nível de concordância entre os anotadores do conjunto de dados BD_1 , foi calculado o D_c interobservador. Devido à ausência de informações que permitam associar cada máscara de nódulo a um anotador específico, o cálculo foi conduzido segundo as premissas: i) o D_c é calculado por nódulo, e não por paciente, ii) para determinar o número de anotadores a ser considerado em cada paciente, identifica-se o nódulo com o maior número de anotações, sendo esta quantidade adotada como referência para todos os nódulos do mesmo paciente, iii) caso um nódulo possua menos anotações que o número máximo definido para o respectivo paciente, as máscaras ausentes são tratadas como *background*, ou seja, regiões sem nódulo e iv) o valor final de D_c corresponde à média dos índices calculados para todos os nódulos.

Com base neste procedimento, obteve-se $D_c = 0,5218$ para os interobservadores no conjunto de dados BD_1 , valor que indica considerável variabilidade entre os especialistas na segmentação de nódulos pulmonares. Este resultado destaca a complexidade intrínseca da tarefa de segmentação, reforçando a necessidade de abordagens automatizadas robustas que possam, ao menos, se aproximar do nível de concordância observado entre anotadores humanos. Neste cenário, a proposta desenvolvida destaca-se no campo da segmentação de imagens médicas, principalmente devido à sua abordagem prática, à cuidadosa seleção dos conjuntos de dados, à otimização dos hiperparâmetros e à aplicação da metodologia de validação na base de dados BD_2 . A otimização dos hiperparâmetros é necessária para melhor controlar o comportamento e o desempenho da rede neural U-Net, tanto durante o treinamento quanto na fase de inferência. Os hiperparâmetros otimizados incluem a taxa de aprendizado, o número de camadas, o tamanho do lote, a função de ativação e a arquitetura específica da rede. Estes ajustes resultam em um modelo mais eficiente, capaz de fornecer resultados mais precisos na tarefa de segmentação de imagens médicas.

O treinamento dos modelos foi realizado utilizando o *framework* PyTorch v2.1.0a0+gitef05c5f, em Python 3.9.17 (PSF) (PASZKE et al., 2019). O processo de treinamento ocorreu ao longo de 1000 épocas para cada um dos cinco *folds* de validação cruzada, com aplicação de técnicas avançadas de aumento de dados, incluindo rotação, escalonamento, adição de ruído gaussiano, desfoque gaussiano, ajustes de brilho e contraste, simulação de baixa resolução e transformações gama, visando

promover a resiliência e a generalização dos modelos. O algoritmo de otimização utilizado foi o gradiente descendente estocástico com momento de Nesterov de 0,99 e taxa de aprendizado inicial de 0,01. A taxa de aprendizagem foi modulada utilizando a técnica Polynomial Learning Rate Scheduler (PLRS), que promove a convergência suave e precisa para o mínimo da função de perda (ROBBINS; MONRO, 1951; NESTEROV, 1983).

A Figura 5.2 apresenta a evolução da função de perda Dice durante o treinamento T_L , em azul e a validação V_L , em vermelho. A métrica pseudo-Dice P_{Dc} é exibida em verde, acompanhada de sua média móvel MM_{pdc} , representada por linha tracejada. O cálculo de P_{Dc} utiliza amostras do conjunto de validação, buscando otimizar o processo de treinamento, uma vez que a avaliação da métrica em todas as imagens ao final de cada época implicaria em alto custo computacional. Em decorrência do procedimento de validação cruzada k -fold, são gerados 45 modelos. Assim, para cada arquitetura, apresenta-se apenas um dos cinco gráficos resultantes, selecionado de forma arbitrária.

Com o objetivo de exemplificar a arquitetura final ajustada pela nnU-Net, a Tabela 5.1 apresenta os resultados do processo de configuração automática das arquiteturas U-Net 2D, U-Net 3D F e U-Net 3D CFREM, considerando as restrições impostas pelos recursos computacionais disponíveis. A arquitetura resultante da U-Net 2D é apresentada na Figura 5.3, enquanto a arquiteturas da U-Net 3D F está apresentada na Figura 5.4. Os diagramas das arquiteturas foram geradas com o auxílio da biblioteca PlotNeuralNet (IQBAL, 2018).

Tabela 5.1 - Arquiteturas representativas configuradas pela nnU-Net.

Características	Modelo 2D	Modelo 3D F	U-Net 3D CFREM
Número de Estágios	7	6	6
Características por Estágio	32, 64, 128, 256, 512, 512, 512	32, 64, 128, 256, 320, 320	32, 64, 128, 256, 320, 320
Tipo de Convolução	2D	3D	3D
Tamanhos de Kernel	[3 × 3]	[3 × 3 × 3]	[3 × 3 × 3]
Número de Convs por Estágio	[2, 2, 2, 2, 2, 2, 2]	[2, 2, 2, 2, 2, 2]	[1, 3, 4, 6, 6, 6]
Número de Convs no Decoder	[2, 2, 2, 2, 2, 2]	[2, 2, 2, 2, 2]	[1, 1, 1, 1, 1]
Função de Ativação	LeakyReLU	LeakyReLU	LeakyReLU
Input Channels	1	1	2
Patch Size	[320, 448]	[96, 128, 192]	[96, 128, 192]
Batch Size	23	2	2
Tipo de Arquitetura	U-Net	U-Net	Residual Encoder U-Net

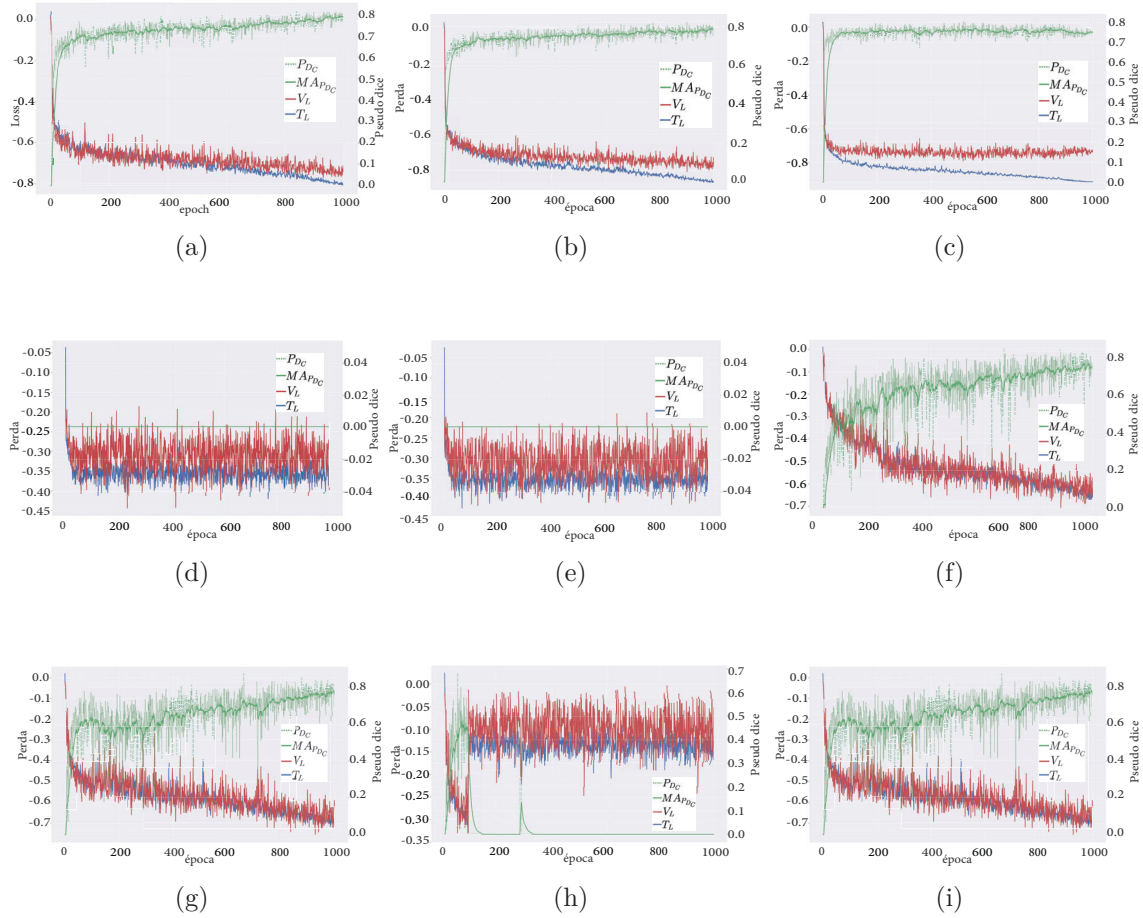


Figura 5.2 - Função de perda de treino T_L e de validação V_L , pseudo-DICE P_{DC} , e média móvel do pseudo-DICE $MA_{P_{DC}}$ para fold de número 0 para os modelos: a) 2D, b) 2D REM, c) 2D REL, d) 3D L, e) 3D LREM, f) 3D F, g) 3D FREM, h) 3D CF e i) 3D CFREM.

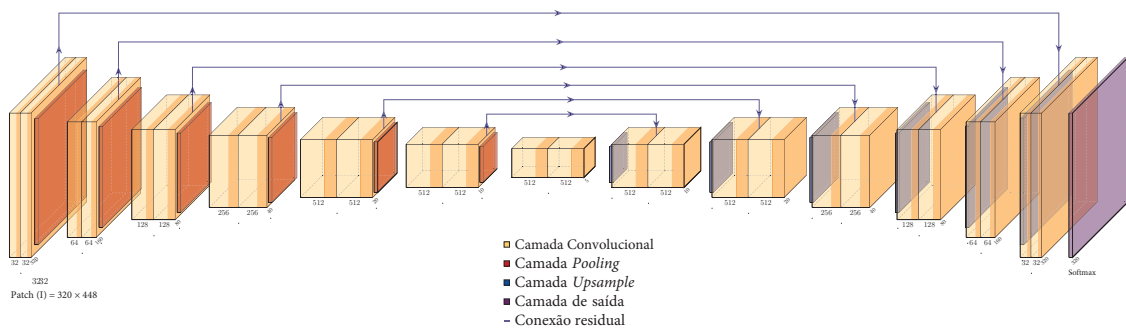


Figura 5.3 - Diagrama do modelo final U-Net 2D.

A Tabela 5.2 apresenta as métricas D_c , J_c e HD_{3D} utilizadas para avaliar o desempenho de cada variação da U-Net na tarefa de segmentação, considerando o

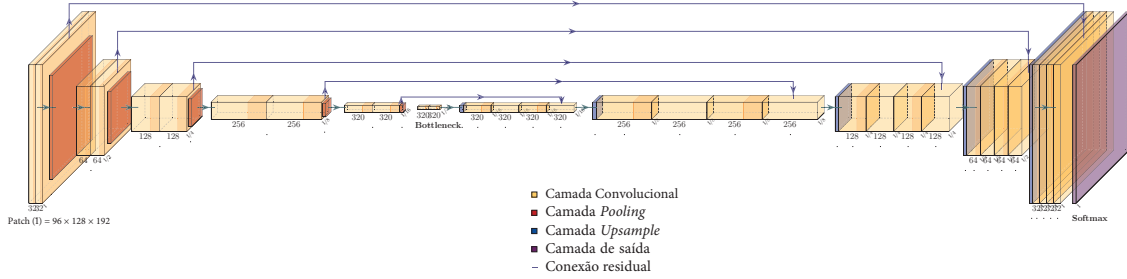


Figura 5.4 - Diagrama do modelo final U-Net 3D F.

subconjunto de teste do banco de dados BD_1 . A Figura 5.5 apresenta o *boxplot* da distribuição da métrica D_c por nódulo, considerando apenas aqueles cuja segmentação apresentou intersecção com a máscara de consenso. Os valores de verdadeiros positivos (VP), falsos negativos (FN) e falsos positivos (FP) estão apresentados na Figura 5.6, enquanto as métricas de precisão e sensibilidade estão presentes na Tabela 5.3.

Tabela 5.2 - Média e desvio-padrão de D_c , J_c e HD_{3D} no subconjunto de teste de BD_1 .

Modelo	μD_c [σ]	μJ_c [σ]	μHD_{3D} [σ]
2D	0.6416 [0.2446]	0.5155 [0.2441]	5.2232 [6.0996]
2D REM	0.6753 [0.2162]	0.5457 [0.2243]	4.3480 [4.7456]
2D REL	0.6584 [0.2320]	0.5301 [0.2323]	4.8629 [5.4173]
3D F	0.7846 [0.1502]	0.6656 [0.1659]	3.5160 [3.7703]
3D FREM	0.7943 [0.1423]	0.6769 [0.1571]	3.5932 [3.9692]
3D L	nan	nan	nan
3D LREM	nan	nan	nan
3D CF	0.6470 [0.1687]	0.4988 [0.1679]	4.3480 [4.0733]
3D CFREM	0.6508 [0.1956]	0.5114 [0.2042]	4.5498 [4.0926]

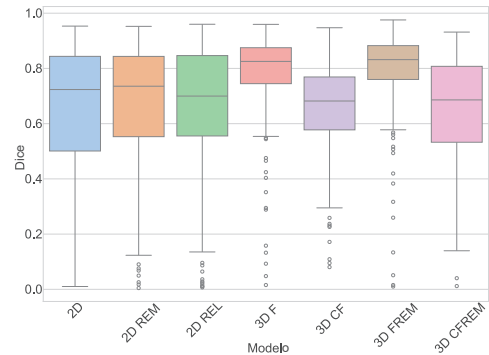


Figura 5.5 - *Boxplot* da métrica D_c para cada modelo em nódulos individuais do subconjunto de teste BD_1 .

Os resultados obtidos apresentam a eficiência do processo de configuração automática das arquiteturas U-Net, conduzido pela abordagem nnU-Net, na adaptação das redes às particularidades do conjunto de dados analisado. A evolução das curvas de perda e a estabilidade observada durante as fases de treinamento e validação confirmam a consistência dos modelos gerados, refletindo o efeito positivo da otimização dos hiperparâmetros e das técnicas de aumento de dados na resiliência e capacidade de generalização das redes. A análise comparativa das arquiteturas resultantes, asso-

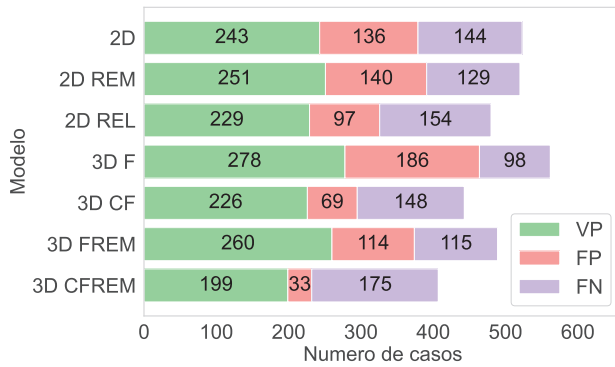


Figura 5.6 - VP, FN e FP para cada modelo no subconjunto de teste BD_1

Tabela 5.3 - Precisão e sensibilidade para cada modelo no subconjunto de teste BD_1

Modelo	Precisão	Sensibilidade
2D	0.641161	0.627907
2D REM	0.641944	0.660526
2D REL	0.702454	0.597911
3D F	0.599138	0.739362
3D CF	0.766102	0.604278
3D FREM	0.695187	0.693333
3D CFREM	0.857759	0.532086

ciada ao desempenho quantitativo avaliado por múltiplas métricas, oferece uma base para a validação do modelo proposto, estabelecendo equilíbrio adequado entre complexidade computacional e acurácia na segmentação de nódulos pulmonares. Estes achados fundamentam a continuidade da avaliação qualitativa e da análise de erros, com o objetivo de aprofundar a compreensão das limitações e das potencialidades do modelo no contexto clínico.

5.1.2 Avaliação qualitativa e análise de erros

Com o objetivo de avaliar o desempenho quantitativo das diferentes variações das arquiteturas U-Net, foram aplicadas múltiplas métricas de segmentação ao subconjunto de teste do banco de dados BD_1 . As métricas D_c , J_c e HD_{3D} foram selecionadas por quantificarem, de forma complementar, tanto a sobreposição entre as máscaras geradas pelos modelos e as máscaras de consenso quanto a distância entre as bordas das segmentações. Esta abordagem permite uma análise abrangente da acurácia, da precisão espacial e da consistência das segmentações, viabilizando a comparação direta entre arquiteturas 2D e 3D, incluindo variantes com diferentes estratégias de codificação residual. Os modelos U-Net 3D L e U-Net 3D LREM foram suprimidos devido ao desempenho insuficiente na tarefa de segmentação. A Tabela 5.4 apresenta os casos com maior frequência de FN, nos quais, de forma recorrente, as variações da U-Net não segmentaram os nódulos. De modo semelhante, a Tabela 5.5 apresenta os casos de FP recorrentes entre as diferentes variações da U-Net, sendo que ND representa casos não detectados.

A Figura 5.7 apresenta comparação entre os modelos U-Net 2D, U-Net 3D F e U-Net

Tabela 5.4 - Imagens com maior quantidade de FN recorrentes entre as variações de U-Net, subconjunto de teste de BD_1 .

Paciente - Nódulo	2D	2D REM	2D REL	3D F	3D CF	3D FREM	3D CFREM
LIDC_0061 - Nodule_1	VP	VP	VP	VP	VP	VP	VP
LIDC_0061 - Nodule_2	FN	FN	FN	FN	FN	FN	FN
LIDC_0061 - Nodule_3	FN	FN	FN	FN	FN	FN	FN
LIDC_0061 - Nodule_4	VP	VP	FN	VP	FN	VP	VP
LIDC_0061 - Nodule_5	FN	FN	FN	FN	FN	FN	FN
LIDC_0061 - Nodule_6	FN	FN	FN	FN	FN	FN	FN
LIDC_0061 - Nodule_7	VP	VP	VP	VP	VP	VP	VP
LIDC_0061 - Nodule_8	VP	FN	FN	VP	VP	VP	VP
LIDC_0061 - Nodule_9	VP	VP	VP	VP	VP	VP	VP
LIDC_0080 - Nodule_1	FN	FN	FN	FN	FN	FN	FN
LIDC_0080 - Nodule_2	FN	FN	FN	VP	FN	VP	VP
LIDC_0620 - Nodule_1	FN	FN	FN	FN	FN	FN	FN
LIDC_0620 - Nodule_2	FN	VP	FN	FN	FN	FN	FN
LIDC_0620 - Nodule_3	VP	VP	VP	VP	VP	VP	FN
LIDC_0620 - Nodule_4	VP	VP	VP	VP	VP	VP	VP
LIDC_1008 - Nodule_5	FN	FN	FN	FN	FN	FN	FN
LIDC_1008 - Nodule_1	VP	VP	VP	VP	VP	VP	FN
LIDC_1008 - Nodule_2	VP	VP	VP	VP	VP	VP	VP
LIDC_1008 - Nodule_3	VP	VP	VP	FN	FN	FN	FN
LIDC_1008 - Nodule_4	FN	FN	FN	VP	FN	VP	FN
LIDC_1008 - Nodule_5	VP	VP	FN	VP	FN	VP	VP
LIDC_1008 - Nodule_6	FN	FN	FN	FN	FN	FN	FN

3D FREM com a máscara de consenso, no eixo axial, para os volumes LIDC-0181 e LIDC-0773. A máscara de consenso está representada em vermelho, a gerada pelo modelo U-Net 2D em amarelo, pela U-Net 3D F em verde e pela U-Net 3D FREM em azul. A Figura 5.7(a) e a Figura 5.7(c) apresentam cortes axiais do volume LIDC-0181, enquanto a Figura 5.7(e) e a Figura 5.7(g) apresentam cortes do volume LIDC-0773. A Figura 5.7(b), a Figura 5.7(d), a Figura 5.7(f) e a Figura 5.7(h) apresentam as máscaras sobrepostas às imagens originais dos volumes LIDC-0181 e LIDC-0773.

A Figura 5.8(a) até a Figura 5.8(c) apresentam recortes de diferentes fatias do volume analisado. A Figura 5.8(d) até a Figura 5.8(f) apresentam as segmentações geradas pelo modelo U-Net 3D F, comparadas às máscaras originais. Nestas imagens, os contornos produzidos pelo modelo estão na cor rosa, enquanto as máscaras dos radiologistas aparecem em vermelho, amarelo, verde e azul. Na Figura 5.8(e), para melhor visualização, apresentam-se apenas os contornos das máscaras dos radiologistas, devido à alta sobreposição das anotações. De forma semelhante à Figura 5.8, a Figura 5.9 apresenta recortes da imagem original e a sobreposição entre a máscara gerada pelo modelo U-Net 3D F e as máscaras originais no eixo axial para o volume LIDC-IDRI-0235, destacando casos de falsos positivos. A Figura 5.9(a) até a Figura 5.9(d) apresentam os recortes de diferentes fatias deste volume, enquanto

Tabela 5.5 - Imagens com maior quantidade de FP recorrentes entre as variações de U-Net. Subconjunto de teste de BD_1 .

Paciente - FP	2D	2D REM	2D REL	3D F	3D CF	3D FREM	3D CFREM
LIDC_0925 - FP_1	ND	ND	ND	FP	ND	FP	ND
LIDC_0925 - FP_2	FP	ND	FP	ND	ND	ND	ND
LIDC_0925 - FP_3	ND	FP	FP	ND	ND	ND	ND
LIDC_0925 - FP_4	FP	FP	FP	ND	ND	ND	ND
LIDC_0925 - FP_5	FP	FP	FP	FP	FP	FP	FP
LIDC_0925 - FP_6	FP	FP	ND	ND	ND	ND	ND
LIDC_0925 - FP_7	FP	FP	ND	ND	ND	FP	ND
LIDC_0925 - FP_8	FP	FP	ND	ND	ND	ND	ND
LIDC_0922 - FP_1	FP	FP	FP	ND	ND	ND	ND
LIDC_0922 - FP_2	FP	FP	ND	ND	ND	ND	ND
LIDC_0922 - FP_3	ND	FP	FP	ND	ND	ND	ND
LIDC_0922 - FP_4	FP	FP	ND	ND	ND	ND	ND
LIDC_0922 - FP_5	FP	FP	ND	ND	ND	ND	ND
LIDC_0922 - FP_6	FP	FP	ND	FP	ND	ND	ND
LIDC_1008 - FP_1	ND	ND	ND	FP	FP	ND	ND
LIDC_1008 - FP_2	ND	ND	ND	FP	ND	FP	FP
LIDC_1008 - FP_3	ND	FP	FP	ND	ND	ND	ND
LIDC_1008 - FP_4	FP	FP	ND	ND	ND	ND	ND
LIDC_1008 - FP_5	FP	FP	FP	FP	FP	FP	FP
LIDC_1008 - FP_6	ND	ND	FP	FP	ND	ND	ND
LIDC_0235 - FP_1	ND	ND	FP	FP	ND	FP	ND
LIDC_0235 - FP_2	FP	FP	ND	FP	ND	ND	ND
LIDC_0235 - FP_3	FP	FP	FP	FP	FP	FP	ND
LIDC_0235 - FP_4	FP	FP	FP	FP	ND	ND	ND
LIDC_0235 - FP_5	FP	FP	ND	ND	ND	ND	ND
LIDC_0097 - FP_1	FP	FP	ND	ND	ND	ND	ND
LIDC_0097 - FP_2	FP	FP	ND	ND	ND	ND	ND
LIDC_0097 - FP_3	FP	FP	ND	ND	ND	ND	ND
LIDC_0097 - FP_4	ND	ND	ND	FP	ND	FP	ND
LIDC_0097 - FP_5	ND	ND	ND	FP	FP	FP	FP

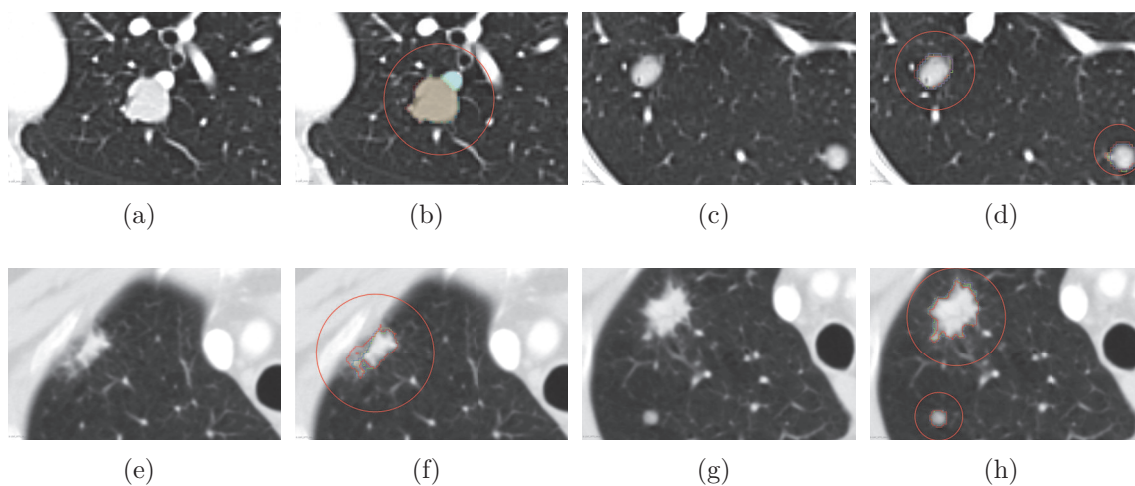


Figura 5.7 - Comparação entre os modelos U-Net 2D, U-Net 3D F e U-Net 3D FREM: (a), (c), (e) e (g) exibem cortes axiais dos volumes LIDC-0181 e LIDC-0773, enquanto (b), (d), (f) e (h) apresentam as máscaras de segmentação sobrepostas correspondentes.

a Figura 5.9(e) até a Figura 5.9(h) apresentam os exemplos de falsos positivos.

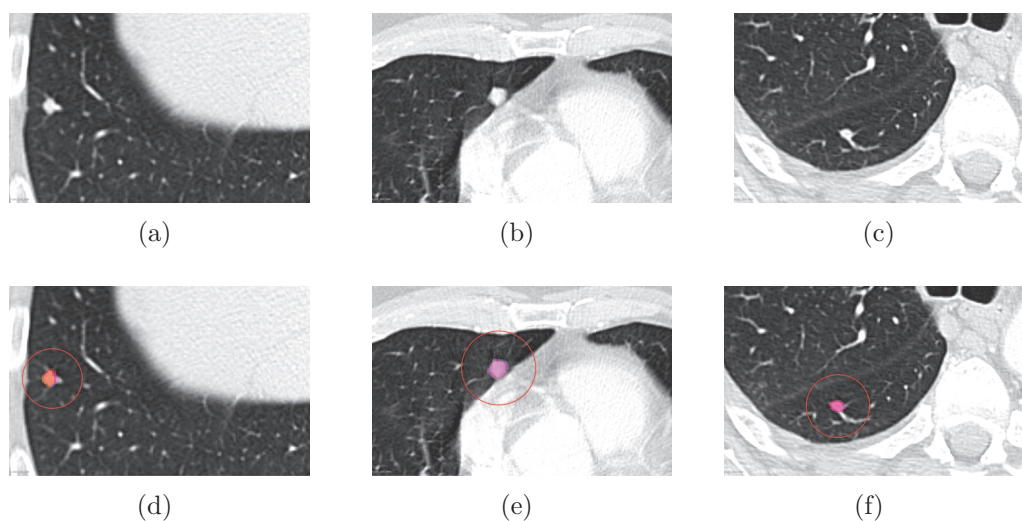


Figura 5.8 - Análise de verdadeiros positivos para o volume LIDC-0235 do conjunto de dados BD_1 : (a), (c) e (e) apresentam cortes das fatias, enquanto (b), (d) e (f) apresentam a segmentação dos nódulos geradas pelo modelo U-Net 3D F.

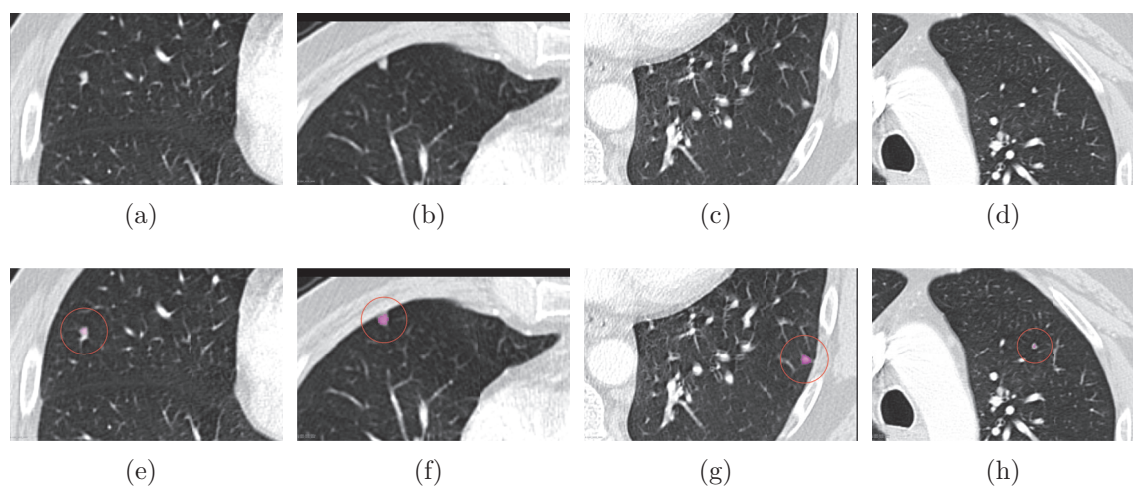


Figura 5.9 - Análise de verdadeiros e falsos positivos para a imagem LIDC-0235 do conjunto de dados BD_1 : (a), (c), (e) e (g) apresentam cortes das fatias, enquanto (b), (d), (f) e (h) mostram as segmentações geradas pelo modelo U-Net 3D F.

Com o objetivo de investigar as possíveis causas associadas aos casos recorrentes de FP e FN, foi conduzida uma análise qualitativa com um radiologista com seis anos de experiência em radiologia cardiotorácica. A análise considerou os cinco volumes

com maior número de FP e os cinco volumes com maior número de FN. O procedimento consistiu na análise visual dos volumes juntamente com as máscaras geradas pelo modelo U-Net 3D F e com a máscara de consenso. A Tabela 5.6 apresenta a avaliação da qualidade do delineamento das máscaras, quantificada em escala de 0 a 10, com justificativas para notas ≤ 5 . A Tabela 5.7 apresenta a análise dos FP e o respectivo nível subjetivo de dificuldade de interpretação, enquanto a Tabela 5.8 dispõe a análise dos FN e seu nível subjetivo de dificuldade de interpretação.

Tabela 5.6 - Avaliação de delineamento e justificativas.

ID	Avaliação*	ID	Avaliação*
LIDC-0061	Nódulo 1 - Localizado na língua. Nota 8. Nódulo 2 - Localizado no segmento anterior do LSE. Nota 10. Nódulo 3 - Paramediastinal à esquerda. Nota 10. Nódulo 4 - Lingular inferior. Nota 10. Nódulo 5 - Paramediastinal direita. Nota 1. Delimitou apenas uma porção. Justificativa: sem justificativa evidente. Nódulo 6 - Região justadiafragmática esquerda. Nota 8. Nódulo 7 - Localizado no LID. Nota 10. Nódulo 8 - Localizado no lobo médio. Nota 5. Justificativa: junto a um vaso.	LIDC-0097	Nódulo 1 - Justafissural à direita. Nota 5. Segmentou apenas porção central. Justificativa: Justafissural.
LIDC-0080	Nódulo 1 - Lobo inferior direito, espículado. Nota 7. O modelo concentrou-se nas porções mais centrais, não marcou as espículas. Caso difícil, mas bom desempenho geral.	LIDC-0235	Nódulo 1 - Segmento anterior do LSD. Nota 10. Nódulo 2 - Periferia do LID. Nota 10.
LIDC-0620	Nódulo 1 - Campo pulmonar médio direito, peribroncovascular. Nota 10. Nódulo 2 - Segmento anterior do LSD. Nota 10.	LIDC-0879	Nódulo 1 - Ápice pulmonar direito. Nota 10. Nódulo 2 - Ápice pulmonar esquerdo. Nota 10. Nódulo 3 - Lobo superior esquerdo, justapleural. Nota 10. Nódulo 4 - Língua. Nota 10. Nódulo 5 - Lobo médio, maior nódulo do exame. Nota 10.
LIDC-0922	Nódulo 1 - Campo médio esquerdo. Nota 10.	LIDC-0925	Nódulo 1 - Nota 10.
LIDC-1008	Nódulo 1 - Segmento posterior do LSD. Nota 10. Nódulo 2 - Lobo médio. Nota 10. Nódulo 3 - Lobo inferior esquerdo. Nota 10.		

* Em uma escala de 0 a 10, como você classifica as capacidades do modelo em delinear as fronteiras do nódulo? Caso a nota seja ≤ 5 , justifique.

No total, foram avaliados 24 casos de VP, com nota média de 8,92 para a qualidade do delineamento das máscaras, sendo que apenas três nódulos receberam nota ≤ 5 . A Figura 5.10 apresenta recortes destes nódulos juntamente com as respectivas justi-

Tabela 5.7 - Falsos positivos e nível de dificuldade de interpretação.

ID	Avaliação*	ID	Avaliação*
LIDC-0061	Caso 1 - FP no pulmão direito. Não existem nódulos. Escala: 1. Justificativa: podem mimetizar vasos. Caso 2 - FP justapleural direita. Não existe nódulo. Escala: 1. Justificativa: pode mimetizar vaso.	LIDC-0097	Caso 1 - FP diminutos esparsos pelos pulmões. Não existem nódulos. Escala: 8. Justificativa: podem mimetizar vias aéreas distais.
LIDC-0080	Caso 1 - FP ao longo de cicatriz pulmonar no campo superior esquerdo. Não existe nódulo. Escala: 5. Justificativa: estrias atelectásicas. Caso 2 - FP justapleural à direita. Escala: 1. Justificativa: atelectasia. Caso 3 - FP na periferia do campo superior esquerdo, justapleural. Escala: 10. Justificativa: pode mimetizar vaso ou preenchimento de via aérea. Caso 4 - FP na periferia do pulmão direito. Escala: 10. Justificativa: podem mimetizar vasos ou preenchimento de via aérea distal.	LIDC-0925	Caso 1 - Dois FP micronódulos no LSD. Escala: 7. Justificativa: cicatriz (atelectásica). Caso 2 - FP. Justificativa: em outro eixo percebe-se que não é nódulo, mas opacidades adjacentes a cicatriz. Caso 3 - Micronódulo no LSE (inferior). Este nódulo é real. Nota 10. Caso 4 - Micronódulo justafissural no lobo médio. Nódulo real (provável linfonodo). Caso 5 - Nódulo justadiafragmático à direita. Nódulo real. Escala: 1. Justificativa: atelectasia. Caso 6 - FP ínfimo no lobo médio. Escala: 1. Justificativa: vaso pulmonar.
LIDC-0620	Caso 1 - Três FP diminutos no LID. Escala: 8. Justificativa: preenchimento de via aérea distal ou vaso.	LIDC-0879	Caso 1 - FP no lobo superior direito. Escala: 1. Justificativa: vaso. Caso 2 - Marcou nódulo no lobo inferior direito que não existe na máscara, mas este nódulo existe de fato. Segmentação: Nota 10.
LIDC-1008	Caso 1 - Nódulo semissólido peribroncovascular à direita (real). Nota 10. Caso 2 - FP próximo ao caso 1. Escala: 1. Justificativa: vaso pulmonar. Caso 3 - FP micronódulo justapleural no lobo médio (provável linfonodo). Não existe nódulo. Caso 4 - Micronódulo no lobo inferior direito (real). Nota 10.	LIDC-0235	Caso 1 - FP no campo superior esquerdo. Escala: 1. Justificativa: vaso. Caso 2 - FP justapleural direita. Escala: 1. Justificativa: linfonodo (morfologia triangular). Caso 3 - FP esparsos pelos pulmões. Escala: 10. Justificativa: podem mimetizar vasos ou via aérea distal.
LIDC-0922	Caso 1 - FP no campo médio esquerdo. Escala: 1. Justificativa: vaso. Caso 2 - FP mais anterior no campo médio direito. Escala: 1. Justificativa: vaso.		

* Existe algum falso positivo? Se sim, qual patologia/tecido foi confundido com um nódulo? Em uma escala de 0 a 10 qual o nível de dificuldade de interpretação do caso? Justifique sua resposta.

ficativas. Dos 25 casos identificados como FP, a partir do confronto entre a máscara de consenso e a máscara gerada pelo modelo U-Net 3D F, treze correspondem à segmentação incorreta de vasos pulmonares ou vias aéreas distais, com dificuldade de interpretação classificada como 1 pelo radiologista. Em seis casos, o nódulo estava presente na imagem, mas não constava na máscara de consenso. Os demais casos apresentam níveis variáveis de dificuldade de interpretação e justificativas distintas, como estrias atelectásicas, linfonodos e opacidades. A Figura 5.11 apresenta exemplos de FP com suas respectivas justificativas, enquanto a Figura 5.12 apresenta

Tabela 5.8 - Falsos negativos e nível de dificuldade.

ID	Avaliação*	ID	Avaliação*
LIDC-0061	Não.	LIDC-0925	Não.
LIDC-0097	Não.	LIDC-0235	Não.
LIDC-0620	Caso 1 - FN no segmento superior do LSE, justapleural. Existe nódulo. Escala: 1. Justificativa: pode mimetizar linfonodo ou via aérea. Caso 2 - FN de nódulo espiculado no segmento basal do lobo inferior esquerdo.	LIDC-0879	Caso 1 - FN micronódulo no ápice esquerdo. Existe nódulo. Escala: 7. Justificativa: em meio a irregularidades cicatriciais. Caso 2 - FN micronódulo superior esquerdo. Existe nódulo. Escala: 7. Justificativa: em meio a irregularidades cicatriciais.
LIDC-0922	Caso 1 - FN no ápice pulmonar esquerdo.	LIDC-0080	Caso 1 - FN na região justadiafragmática direita. Existe nódulo. Escala: 1. Justificativa: sem justificativa evidente.
LIDC-1008	Caso 1 - FN micronódulo justafissural à direita. Escala: 1. Justificativa: provavelmente linfonodo. Caso 2 - FN opacidade nodular no lobo médio, junto a atelectasia. Escala: 8. Justificativa: difícil por estar junto à atelectasia. Caso 3 - FN opacidade no LIE, composta por vasos pulmonares. Escala: 8.		

* Existe algum falso negativo? Se sim, em uma escala de 0 a 10 qual o nível de dificuldade de interpretação do caso?

casos em que a máscara de consenso não continha segmentação, mas o radiologista identificou o nódulo como positivo. Todos os nove casos de FN foram validados pelo radiologista, sendo que a Figura 5.13 apresenta exemplos destes casos.

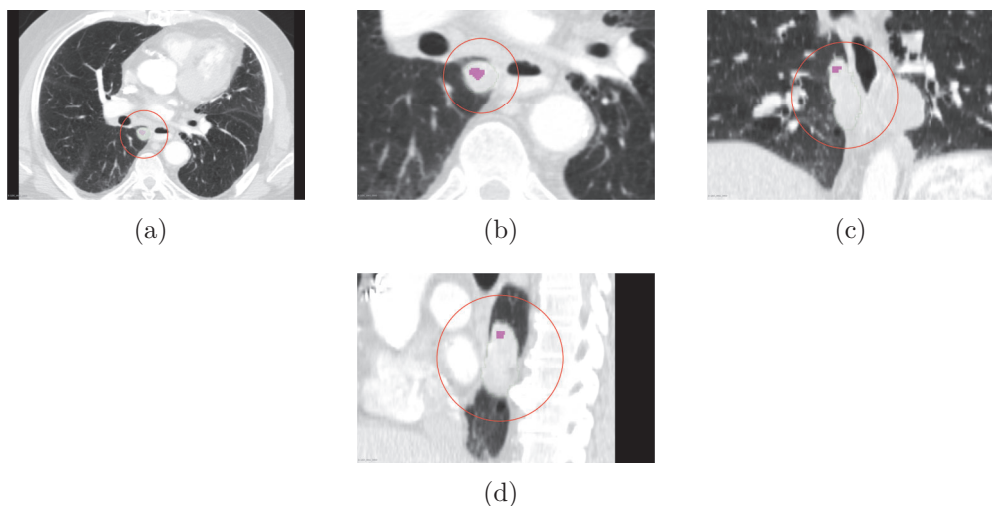


Figura 5.10 - Segmentações produzidas pelo modelo U-Net 3D F avaliados como não satisfatórias pelo radiologista torácico no volume LIDC-0061 de BD_1 . (a) fatia axial, (b) fatia axial ampliada, (c) eixo coronal, e (d) eixo sagital.

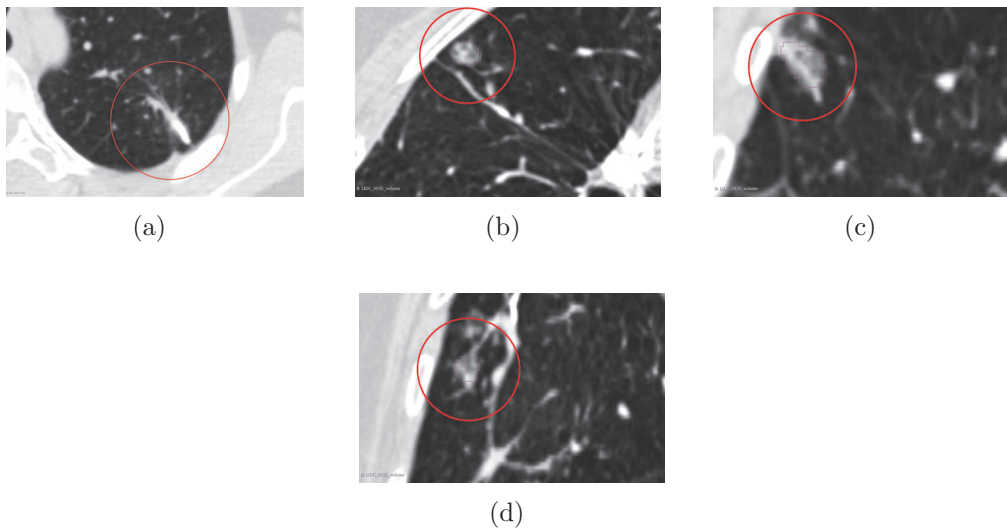


Figura 5.11 - Segmentações produzidas pelo modelo 3D F contendo falsos positivos, verificados pelo radiologista torácico, nos volumes LIDC-0080 e LIDC-0925 de BD_1 : (a) falso positivo em região de estria atelectásica, (b) eixo axial, (c) eixo coronal e (d) eixo sagital da mesma região.

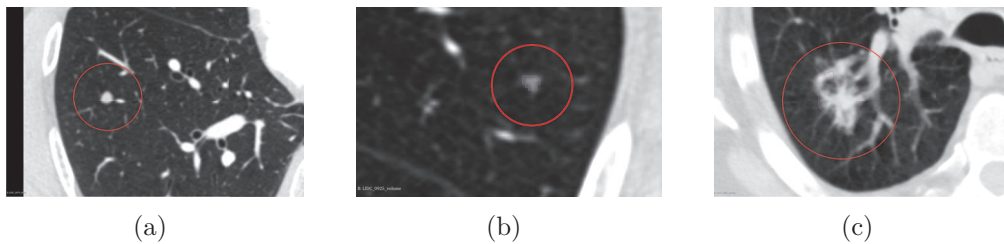


Figura 5.12 - Segmentações produzidas pelo modelo U-Net 3D F em que contêm falsos positivos que foram classificados como verdadeiros positivos pelo radiologista torácico: (a) LIDC-0879, (b) LIDC-0925 e (c) LIDC-1008.

Os resultados apresentam o potencial das arquiteturas U-Net 3D para a segmentação de nódulos pulmonares em tomografias computadorizadas. O aumento significativo nos índices D_c e J_c , observado na comparação entre as arquiteturas 2D e 3D, destaca a relevância da preservação da estrutura volumétrica das imagens médicas para aprimorar o desempenho da segmentação. Contudo, a elevada variabilidade observada nos boxplots por nódulo, Figura 5.5, associada à recorrência de FP em regiões de vasos pulmonares e vias aéreas distais, indicam limitações específicas das arquiteturas avaliadas. Tais limitações podem ser reduzidas por meio da incorporação de informações anatômicas.

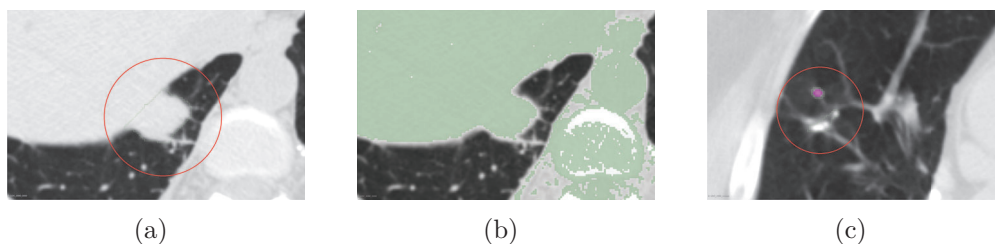


Figura 5.13 - Segmentações geradas pelo modelo U-Net 3D F de falsos positivos, verificado pelo radiologista torácico: (a) LIDC-0080, (b) segmentação por limiar da região de falso negativo e (c) LIDC-1008.

Portanto, a análise qualitativa conduzida com o radiologista indicou que parte dos casos classificados como FP pelo modelo decorre de inconsistências nas máscaras de consenso, sugerindo que o algoritmo pode identificar nódulos omitidos nas anotações originais. Este achado reforça não apenas o potencial dos modelos de segmentação, mas também a necessidade de aprimorar os processos de anotação manual, incorporando revisões por especialistas ou métodos semiautomáticos, de modo a fortalecer a confiabilidade dos conjuntos de dados utilizados na validação de soluções com base em inteligência artificial.

5.1.3 Avaliação e validação da solução

Para avaliar a consistência da solução no conjunto de dados BD_2 , foram analisadas as diferenças de volume e diâmetro entre exames sucessivos do mesmo paciente. A Tabela 5.9 apresenta síntese das médias μ e os desvios padrão σ das diferenças de volume e diâmetro para todos os modelos. Para análise visual, são apresentados os gráficos de Bland–Altman referentes ao volume e ao diâmetro. A Figura 5.14 apresenta os volumes em escala logarítmica (\log_{10}), na qual a linha contínua representa o viés estimado e as linhas tracejadas indicam os limites de concordância, calculados por $\pm 1,98 \cdot \sigma$. Diferenças de volume superiores a 25% estão assinaladas com o marcador \times . Os valores absolutos de viés, limites de concordância e respectivos intervalos de confiança encontram-se na Tabela 5.10. A Figura 5.15 apresenta as diferenças de diâmetro, em que Valores inferiores a 2mm foram representados em escala linear, e não logarítmica (\log_{10}), o que os aproxima do eixo das abscissas. As estatísticas correspondentes estão dispostas na Tabela 5.11.

A Figura 5.14 apresenta um padrão de dispersão em formato cônico, sugerindo possível relação entre o volume dos nódulos e as diferenças observadas. Para investigar

Tabela 5.9 - Resumo estatístico das diferenças de volume e diâmetro para os modelos no conjunto de dados BD_2 .

Modelo	Volume (mm^3)		Diâmetro (mm)	
	μ	σ	μ	σ
2D	828.5835	2508.7333	0.9740	1.8330
2D REM	530.4377	1857.2101	1.3377	2.8476
2D REL	606.3745	1593.4779	1.5540	3.1827
3D F	220.1032	702.6879	1.0401	2.0412
3D FREM	207.6049	576.3138	1.0295	1.6104
3D CF	501.0527	1259.8241	1.2797	1.9189
3D CFREM	503.6713	1212.2626	1.2634	1.8962

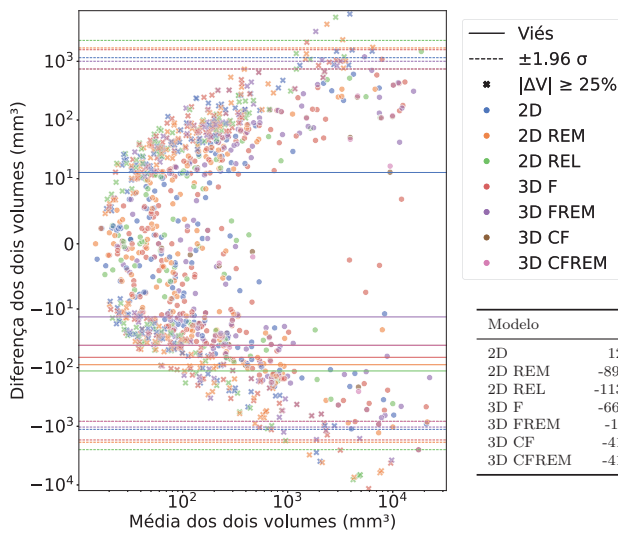


Tabela 5.10 - Viés, limites de concordância, e intervalos de confiança no gráfico de Bland-Altman das diferenças de volume de nódulos.

Modelo	Bias[CI]	-1.96σ[CI]	+1.96σ[CI]
2D	12.74 [-12.42, 37.89]	-378.56[-422.12, -334.99]	404.03[360.46, 447.59]
2D REM	-89.26 [-128.94, -49.59]	-700.05[-768.77, -631.33]	521.53[452.81, 590.25]
2D REL	-113.93 [-176.36, -51.49]	-932.71[-1040.85, -824.58]	704.86[596.73, 813.00]
3D F	-66.66 [-102.10, -31.21]	-630.41[-691.80, -569.02]	497.10[435.71, 558.49]
3D FREM	-13.66 [-42.27, 14.95]	-361.75[-411.31, -312.19]	334.43[284.87, 383.99]
3D CF	-41.67 [-70.22, -13.12]	-307.99[-357.44, -258.53]	224.65[175.20, 274.10]
3D CFREM	-41.64 [-70.27, -13.01]	-308.70[-358.29, -259.11]	225.41[175.82, 275.00]

Figura 5.14 - Gráfico de Bland-Altman das diferenças de volume dos nódulos em escala logarítmica.

esta hipótese, aplicou-se a técnica de regressão linear, cujos resultados estão dispostos na Tabela 5.12. Observa-se que nenhum coeficiente de interceptação β_0 difere significativamente de zero, indicando ausência de viés fixo. Por outro lado, foi identificado viés proporcional em cinco modelos, nos quais $\beta_1 < 0$ e $p_{\beta_1} < 0,05$, o que sugere tendência de subestimação em volumes maiores. Ainda assim, os valores de R^2 ajustado R_{ad}^2 , indicam que a variabilidade das diferenças não é plenamente explicada pela regressão em nenhum dos modelos. Como exemplo, considera-se um nódulo esférico com diâmetro de $30mm$ e volume estimado de $V = 14140mm^3$ avaliado pelo modelo 2D REM. Neste cenário de pior caso, a variação estimada seria de $-0,136 \cdot 14140 = -1920mm^3$, correspondente a subestimação de $-13,6\%$. É importante destacar que o viés observado para o diâmetro do maior eixo jamais

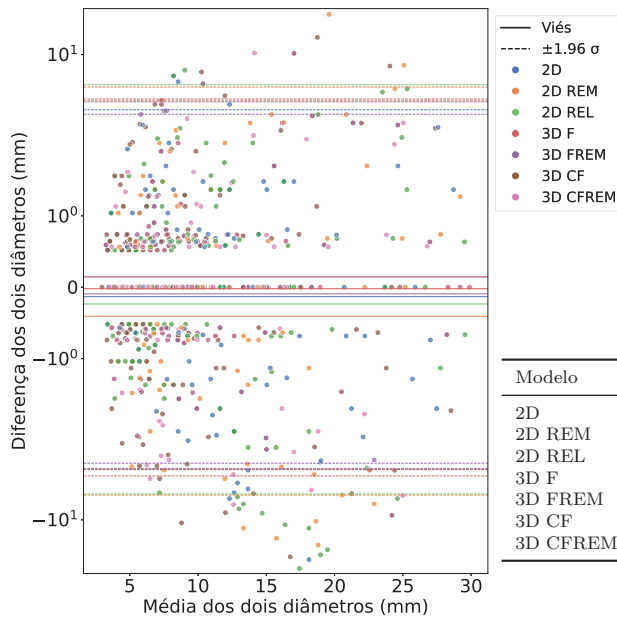


Tabela 5.11 - Viés, limites de concordância, e intervalos de confiança no gráfico de Bland-Altman das diferenças de diâmetro de nódulos.

Modelo	Bias[CI]	-1.96σ[CI]	+1.96σ[CI]
2D	-0.13[-0.39, 0.12]	-3.98[-4.43, -3.54]	3.72[3.28, 4.16]
2D REM	-0.41[-0.80, -0.01]	-6.41[-7.10, -5.72]	5.60[4.91, 6.28]
2D REL	-0.23[-0.71, 0.24]	-6.30[-7.12, -5.47]	5.83[5.00, 6.65]
3D F	-0.02[-0.31, 0.27]	-4.55[-5.06, -4.05]	4.51[4.01, 5.02]
3D FREM	-0.09[-0.39, 0.21]	-3.62[-4.14, -3.11]	3.44[2.92, 3.96]
3D CF	0.15[-0.32, 0.62]	-4.06[-4.88, -3.25]	4.35[3.54, 5.17]
3D CFREM	0.15[-0.32, 0.61]	-4.00[-4.80, -3.19]	4.29[3.48, 5.09]

Figura 5.15 - Gráfico de Bland-Altman das diferenças de diâmetro dos nódulos em escala logarítmica.

ultrapassou a margem de erro aleatório de 2 mm descrita na literatura, e que o pior desvio proporcional daí derivado para o volume ($\approx 13\%$) permanece muito aquém do limiar de variabilidade de 25% considerado clinicamente aceitável (LARICI et al., 2017; CALLISTER, 2017).

Tabela 5.12 - Coeficientes da regressão do modelo $\Delta V = \beta_0 + \beta_1 \bar{V}$ por modelo.

Modelo	N	β_0 (IC 95%)	p_{β_0}	β_1 (IC 95%)	p_{β_1}	R_{aj}^2
2D	242	9.64 [-67.82; 87.11]	0.81	0.0049 [-0.0337; 0.0435]	0.804	-0.004
2D REL	172	-40.56 [-230.48; 149.37]	0.674	-0.114 [-0.208; -0.019]	0.018	0.027
2D REM	237	-0.90 [-117.60; 115.81]	0.988	-0.136 [-0.192; -0.079]	$< 10^{-5}$	0.083
3D CF	87	43.26 [-40.14; 126.67]	0.305	-0.062 [-0.089; -0.036]	$< 10^{-4}$	0.195
3D CFREM	87	43.45 [-40.25; 127.15]	0.305	-0.062 [-0.089; -0.036]	$< 10^{-4}$	0.194
3D F	253	-11.84 [-120.71; 97.02]	0.831	-0.0578 [-0.0982; -0.0175]	0.0051	0.027
3D FREM	148	13.77 [-77.14; 104.68]	0.77	-0.0190 [-0.0443; 0.0062]	0.139	0.008

Para verificar se as diferenças de volume e diâmetro observadas no subconjunto BD_2 são estatisticamente significativas, aplicou-se o teste de hipótese pareado, cujos resultados de seus p -valores estão apresentados na Tabela 5.13. Esta abordagem é apropriada porque ambas as aquisições tomográficas pertencem ao mesmo paciente,

foram realizadas no mesmo dia e os nódulos foram segmentados pelo mesmo modelo em ambas as imagens, permitindo comparação direta nódulo a nódulo. Inicialmente, o teste de Shapiro–Wilk (SW) indicou ausência de normalidade em todas as distribuições de diferenças, com valores de $p < 0,05$, o que levou à rejeição da hipótese nula de normalidade. Em decorrência, utilizou-se o teste não paramétrico de Wilcoxon (Wlx) (SHAPIRO; WILK, 1965; REY; NEUHÄUSER, 2011), cujos valores de p foram superiores a 0,05, não permitindo rejeitar a hipótese nula de que a mediana das diferenças seja igual a zero. Assim, não se identificaram diferenças estatisticamente significativas entre as medições de volume e diâmetro nas duas aquisições.

Tabela 5.13 - Comparação de volume e diâmetro entre TC1 e TC2.

Modelo	N	Volume		Diâmetro	
		p(SW)	p(Wlx)	p(SW)	p(Wlx)
2D	242	1.604327e-28	0.910164	1.488039e-22	0.677984
2D REL	172	5.741843e-16	0.659059	1.628509e-19	0.253559
2D REM	237	5.889903e-18	0.662973	7.521694e-24	0.316592
3D CF	87	1.559395e-06	0.823639	1.022571e-09	0.643906
3D CFREM	87	2.704957e-06	0.756013	7.340807e-10	0.643906
3D F	253	9.943687e-22	0.668144	8.973661e-23	0.464491
3D FREM	148	2.485966e-13	0.799455	1.272923e-13	0.784075

A análise global dos resultados permite identificar aspectos adicionais relevantes para a interpretação do desempenho do modelo de segmentação proposto. O índice $D_c = 0,52$ obtido na avaliação interobservador indica variabilidade significativa entre os radiologistas no delineamento das fronteiras dos nódulos, reforçando a complexidade inerente à segmentação manual. Este valor também sugere possível limite superior para o desempenho de modelos automatizados, considerando que mesmo os melhores algoritmos tendem a refletir a inconsistência presente nas anotações humanas. Além disto, a melhoria observada com a transição de arquiteturas 2D para 3D, com valores médios de $D_c \approx 0,79$, especialmente nas configurações com codificadores residuais (FREM), destaca a relevância da preservação da estrutura volumétrica das imagens para a detecção e segmentação de nódulos pulmonares. As características tridimensionais destas lesões exigem modelos capazes de capturar adequadamente as relações espaciais entre as fatias.

Outro aspecto relevante é a diferença observada entre os índices globais de D_c e os resultados individuais por nódulo, apresentados nos boxplots. Embora os valores médios sejam satisfatórios, a elevada dispersão em casos específicos indica que o modelo ainda enfrenta dificuldades em situações particulares, como nódulos de

pequenas dimensões ou localizados em regiões anatômicas complexas. Além disto, a recorrência de falsos positivos em estruturas como vasos pulmonares e vias aéreas distais apresenta limitações da arquitetura empregada na diferenciação entre nódulos e tecidos com morfologia semelhante. Este resultado aponta direções futuras para o aprimoramento do modelo, como a incorporação de informações anatômicas adicionais.

Por fim, a validação dos casos classificados como falsos positivos, em que o modelo identificou nódulos ausentes nas máscaras de consenso, sugere que o algoritmo pode detectar lesões omitidas originalmente pelos anotadores. Este achado reabre a discussão sobre a confiabilidade das máscaras de consenso, destacando a necessidade de revisões especializadas para aprimorar a qualidade dos dados utilizados na avaliação de modelos de inteligência artificial.

5.2 Discussão

A proposta deste trabalho buscou desenvolver uma solução automatizada para segmentação e quantificação de nódulos pulmonares, integrando redes neurais profundas à análise volumétrica em tomografia computadorizada. Os resultados obtidos demonstraram que a arquitetura U-Net 3D F, em especial, apresentou desempenho superior às demais variantes, com média de $\bar{D}_c = 0,7846$, menor distância de Hausdorff HD_{3D} e maior proporção de verdadeiros positivos na base BD_1 . A baixa variabilidade nas diferenças de volume e diâmetro entre exames consecutivos no conjunto BD_2 , confirmada por testes estatísticos e análise de Bland–Altman, reforça a estabilidade da solução em medições repetidas, propriedade desejável em contextos de rastreamento clínico longitudinal. Embora a arquitetura U-Net 3D F tenha enfrentado dificuldades em segmentar nódulos acompanhados de outras patologias, suas máscaras de segmentação apresentaram correspondência com as máscaras de consenso, indicando capacidade de representação anatômica adequada mesmo em casos desafiadores.

Além da consistência nas métricas quantitativas, a U-Net 3D F demonstrou capacidade de identificar nódulos não segmentados nas anotações originais, sugerindo que alguns falsos positivos podem, na verdade, representar detecções válidas não anotadas. Esta constatação é reforçada pela avaliação qualitativa do radiologista, que apontou que parte dos falsos negativos atribuídos ao modelo decorrem de omissões nas máscaras de referência. Por outro lado, as variantes bidimensionais da U-Net

apresentaram desempenho limitado, especialmente em nódulos localizados junto à pleura ou em regiões com complexidade anatômica, com maior incidência de falsos positivos e negativos, indicando a limitação destas arquiteturas na preservação do contexto espacial completo (FALK et al., 2019; FABIAN et al., 2020).

Comparando-se com a literatura, observa-se consonância entre os achados do presente estudo e trabalhos prévios que destacam a importância do contexto tridimensional na segmentação de nódulos pulmonares (FALK et al., 2019). A variabilidade interobservador observada com $D_c = 0,52$ é compatível com dados reportados em estudos anteriores (ARMATO et al., 2011) e reforça o desafio da anotação manual em imagens de TC. A presença recorrente de falsos positivos em regiões vasculares e estruturas adjacentes também converge com indicação de que tais áreas são suscetíveis a interpretação ambígua de estruturas adjacentes (KUBOTA et al., 2011), tanto por algoritmos quanto por especialistas (TANG et al., 2019). O desempenho da U-Net 3D F confirma o potencial de arquiteturas tridimensionais para reduzir estas ambiguidades e amplia os avanços recentes da literatura ao integrar uma abordagem autoadaptativa com capacidade de generalização entre diferentes conjuntos de dados (FABIAN et al., 2020).

Embora existam pontos de convergência, o presente trabalho também traz contribuições que enfrentam limitações recorrentes da literatura. A proposta desenvolvida é capaz de adaptar seu fluxo de pré-processamento, segmentação e pós-processamento automaticamente por meio da arquitetura nnU-Net (FABIAN et al., 2020), o que representa um avanço relevante frente a modelos que requerem ajuste manual ou validação restrita a bases únicas. A capacidade de quantificar de forma estável o volume e o diâmetro de nódulos, mesmo em bases distintas como BD_1 e BD_2 , reforça a contribuição metodológica do estudo no enfrentamento da heterogeneidade dos dados médicos. Esta abordagem contribui para reduzir o impacto da variabilidade inerente aos protocolos de aquisição, ampliando o potencial de uso da solução em contextos clínicos reais (NAYYAR et al., 2021).

Apesar dos resultados promissores, o estudo apresenta limitações importantes. A ausência de segmentações manuais no conjunto BD_2 impede a avaliação direta da qualidade da segmentação nesta base de dados, limitando a análise à quantificação. Além disto, o desempenho de variantes como U-Net 3D CF e 3D CFREM foi comprometida pela dependência da etapa inicial de segmentação da variante U-Net 3D L, que não conseguiu segmentar nódulos de pequena dimensão. Outro fator limitante

foi a restrição de hardware, que impediu a experimentação de todas as variantes da nnU-Net. Do ponto de vista metodológico, a unificação dos tipos de nódulos (sólidos, sub-sólidos e em vidro fosco) em única classe pode ter introduzido variações que influenciaram a acurácia dos resultados. Não foi possível considerar informações sobre patologias associadas, o que reduz a capacidade de análise qualitativa em cenários mais complexos. Como alternativas, recomenda-se a validação externa em bases clínicas diversas, a separação dos nódulos por classe e a incorporação de avaliações qualitativas conduzidas por especialistas.

As perspectivas futuras envolvem a incorporação de técnicas emergentes de aprendizado, como autoaprendizado e aprendizado contínuo (KIM et al., 2024), que permitiriam ao modelo adaptar-se a novas informações sem necessidade de reanotação extensiva. Arquiteturas com mecanismos de atenção, como TransUNet e Swin UNETR, também devem ser consideradas para capturar relações espaciais mais complexas entre os nódulos e estruturas adjacentes. Outras possibilidades incluem o desenvolvimento de modelos longitudinais capazes de acompanhar a evolução dos nódulos ao longo do tempo, auxiliando diretamente em estratégias de rastreamento e avaliação de resposta ao tratamento. Além disto, sugere-se manter o uso da arquitetura nnU-Net por sua escalabilidade e adaptabilidade, complementando as métricas tradicionais com medidas como Hausdorff para melhor avaliação das bordas. A experimentação em múltiplas arquiteturas e conjuntos de dados continua sendo recomendada, a fim de evitar sobreajustes e assegurar a generalização da solução em ambientes clínicos heterogêneos.

CAPÍTULO 6

CONCLUSÃO

O presente estudo desenvolveu uma metodologia para segmentação e quantificação volumétrica de nódulos pulmonares com diâmetro superior a $3mm$ em imagens de tomografia computadorizada do tórax, utilizando redes neurais profundas. A abordagem adotada, fundamentada na arquitetura nnU-Net, permitiu a criação automática de máscaras de segmentação, viabilizando a medição precisa dos volumes por contagem de voxels associados às informações espaciais das imagens originais. O treinamento foi realizado sobre o conjunto de dados LIDC-IDRI, consolidado na literatura médica, e validado de forma interna e em bases externas, visando ampliar a aplicabilidade em cenários clínicos reais da solução proposta.

A hipótese de que seria possível desenvolver uma solução automatizada capaz de quantificar com precisão o volume e o diâmetro de nódulos pulmonares foi corroborada. Esta confirmação decorre dos resultados obtidos, destacando-se o desempenho da arquitetura U-Net 3D F, que alcançou média Dice $\bar{D}_c = 0,7846$ com desvio padrão $\sigma = 0,18$, superando o índice interobservador de $D_c = 0,5218$ entre radiologistas. No conjunto RIDER, testes estatísticos indicaram estabilidade das medições entre exames consecutivos, sem diferenças estatisticamente significativas, apresentando a consistência da solução em diferentes condições de aquisição de imagem.

Os objetivos estabelecidos foram atingidos em sua totalidade. A proposta não apenas desenvolveu um processo automatizado de segmentação e quantificação, como também demonstrou sua generalização em diferentes bases de dados. Entre os resultados mais relevantes, destaca-se a baixa variabilidade nas diferenças de volume e diâmetro entre aquisições, a alta capacidade de detecção de nódulos e a aproximação dos resultados gerados automaticamente às máscaras de consenso. A utilização da abordagem tridimensional adaptativa, com validação cruzada interna e análise crítica de resultados, diferenciou o trabalho em termos de rigor metodológico e aplicabilidade clínica futura.

Portanto, conclui-se que a solução proposta apresenta potencial para aplicação prática na quantificação automatizada de nódulos pulmonares, contribuindo para maior precisão no monitoramento longitudinal de pacientes e para o aprimoramento dos processos diagnósticos. A estabilidade das medições, a capacidade de generaliza-

ção entre bases heterogêneas e a proximidade com padrões de anotação humana reforçam sua relevância para aplicações clínicas e para futuras pesquisas na área de inteligência artificial aplicada à saúde.

6.1 Contribuições do Trabalho

As contribuições do trabalho podem ser desta forma descritas:

Artigos publicados

NOGUEIRA, S. A.; LUZ, F. A. B.; **CAMARGO, T. F. O.**; OLIVEIRA, J. C. S.; NETO, G. C. C.; CARVALHAES, F. B. F., REIS, M. R. C.; SANTOS, P. V.; MENDES, G. S.; LOUREURO, R. M.; TORNIERI, D.; PACHECO, V. M. G.; COIMBRAS; A. P., CALIXTO, W. P.. Artificial intelligence applied in identifying left ventricular walls in myocardial perfusion scintigraphy images: a pilot study. Plos one, 20(1), e0312257. 2025.

Artigos submetidos em revista

CAMARGO, T. F. O.; MARTINS, M. S. R.; REIS, M. R. C.; PACHECO, V. M. G.; V. M. G.; COIMBRAS; CALIXTO, W. P.. Quantification of pulmonary nodules in computed tomography using automated 3D segmentation.

REFERÊNCIAS BIBLIOGRÁFICAS

AGARWAL, P. K.; HAR-PELED, S.; SHARIR, M.; WANG, Y. Hausdorff distance under translation for points and balls. **ACM Transactions on Algorithms**, Association for Computing Machinery (ACM), v. 6, n. 4, p. 1–26, 8 2010. ISSN 1549-6325. [43](#)

AHMET, I.; K., A.; B., Ş.; R., A. Covid-19 Lung CT image segmentation using localization and enhancement methods with U-Net. **Procedia Computer Science**, 2023. [45](#)

American College of Radiology Committee on Lung-RADS. **Lung-RADS Assessment Categories 2022**. 2022. [29](#), [36](#)

ARLOT, S.; CELISSE, A. A survey of cross-validation procedures for model selection. **Project Euclid**, 2010. [44](#)

ARMATO, S. G. r.; 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.; BEEKE, E. J. R. V.; YANKELEVITZ, D.; BIANCARDI, A. M.; BLAND, P. H.; BROWN, M. S.; ENGELMANN, R. M.; LADERACH, G. E.; MAX, D.; PAIS, R. C.; QING, D. P. Y.; ROBERTS, R. Y.; SMITH, A. R.; STARKEY, A.; BATRAH, P.; CALIGIURI, P.; FAROOQI, A.; GLADISH, G. W.; JUDE, C. M.; MUNDEN, R. F.; PETKOVSKA, I.; QUINT, L. E.; SCHWARTZ, L. H.; SUNDARAM, B.; DODD, L. E.; FENIMORE, C.; GUR, D.; PETRICK, N.; FREYMAN, J.; KIRBY, J.; HUGHES, B.; CASTEELE, A. V.; GUPTA, S.; SALLAM, M.; HEATH, M. D.; KUHN, M. H.; DHARAIYA, E.; BURNS, R.; FRYD, D. S.; SALGANICOFF, M.; ANAND, V.; SHRETER, U.; VASTAGH, S.; CROFT, B. Y. The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. **Medical physics**, Wiley Online Library, v. 38, n. 2, p. 915–931, 2011. [57](#), [76](#)

ASHBURNER, J.; FRISTON, K. J. Voxel-based morphometry—the methods. **Neuroimage**, Elsevier, v. 11, n. 6, p. 805–821, 2000. [29](#)

AYDIN, O. U.; TAHA, A. A.; HILBERT, A.; KHALIL, A. A.; GALINOVIC, I.; FIEBACH, J. B.; FREY, D.; MADAI, V. I. On the usage of average Hausdorff

distance for segmentation performance assessment: hidden error when used for ranking. **European Radiology Experimental**, Springer Science and Business Media LLC, v. 5, n. 1, jan 21 2021. ISSN 2509-9280. [44](#)

BANKIER, A. A.; MACMAHON, H.; GOO, J. M.; RUBIN, G. D.; SCHAEFER-PROKOP, C. M.; NAIDICH, D. P. Recommendations for measuring pulmonary nodules at ct: a statement from the fleischner society. **Radiology**, v. 285, n. 2, p. 584–600, 2017. [33](#), [36](#), [51](#)

BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. **Journal of machine learning research**, v. 13, n. 2, 2012. [45](#)

BLAND, J. M.; ALTMAN, D. Statistical methods for assessing agreement between two methods of clinical measurement. **The lancet**, Elsevier, v. 327, n. 8476, p. 307–310, 1986. [55](#)

BOAS, F. E.; FLEISCHMANN, D. Ct artifacts: Causes and reduction techniques. **Imaging in Medicine**, Future Medicine Ltd London, UK, v. 4, n. 2, p. 229–240, 2012. [39](#)

BOGOT, N. R.; KAZEROONI, E. A.; KELLY, A. M.; QUINT, L. E.; DESJARDINS, B.; NAN, B. Interobserver and intraobserver variability in the assessment of pulmonary nodule size on ct using film and computer display methods1. **Academic Radiology**, Elsevier, v. 12, n. 8, p. 948–956, 2005. [29](#), [56](#)

BONDARENKO, G. O.; SYRYH, A. S. Predicting pneumothorax progression: a methodology using lung mask comparison. In: IEEE. **2024 XXVII International Conference on Soft Computing and Measurements (SCM)**. Saint Petersburg, Russian Federation, 2024. p. 440–443. [45](#)

BOTCHKAREV, A. Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. **arXiv preprint arXiv:1809.03006**, 2018. [42](#)

BRAY, F.; FERLAY, J.; SOERJOMATARAM, I.; SIEGEL, R. L.; TORRE, L. A.; JEMAL, A. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: a cancer journal for clinicians**, Wiley Online Library, v. 68, n. 6, p. 394–424, 2018. [29](#)

BUENO, J.; LANDERAS, L.; CHUNG, J. H. Updated fleischner society guidelines for managing incidental pulmonary nodules: common questions and challenging scenarios. **Radiographics**, Radiological Society of North America, v. 38, n. 5, p. 1337–1350, 2018. [35](#), [36](#)

CALLISTER, M. Es 02.02 the fleischner guideline / lung-rads. **Journal of Thoracic Oncology**, Elsevier, v. 12, n. 11, p. S1612–S1613, 2017. ISSN 1556-0864. [73](#)

CARRÉ, A.; KLAUSNER, G.; EDJLALI, M.; LEROUSSEAU, M.; BRIEND-DIOP, J.; SUN, R.; AMMARI, S.; REUZÉ, S.; ANDRES, E. A.; ESTIENNE, T.; NIYOTEKA, S.; BATTISTELLA, E.; VAKALOPOULOU, M.; DHERMAIN, F.; PARAGIOS, N.; DEUTSCH, E.; OPPENHEIM, C.; PALLUD, J.; ROBERT, C. Standardization of brain MR images across machines and protocols: bridging the gap for MRI-based radiomics. **Scientific Reports**, Springer Science and Business Media LLC, v. 10, n. 1, jul 23 2020. ISSN 2045-2322. [47](#)

CHAYMAE, E. M.; NAJIBA, E. A. E. I. Two-Stream nnU-Net: A Novel Architecture for Precise Tumor Segmentation in Medical Imaging. **International Conference on Wireless Networks and Mobile Communications**, 2023. [47](#), [48](#), [50](#)

CORTES, C.; VAPNIK, V. Support-vector networks. **Machine learning**, Springer, v. 20, p. 273–297, 1995. [41](#)

CUI, Z.; FANG, Y.; MEI, L.; ZHANG, B.; YU, B.; LIU, J.; JIANG, C.; SUN, Y.; MA, L.; HUANG, J.; LIU, Y.; ZHAO, Y.; LIAN, C.; DING, Z.; ZHU, M.; SHEN, D. A fully automatic ai system for tooth and alveolar bone segmentation from cone-beam ct images. **Nature communications**, Nature Publishing Group, v. 13, n. 1, p. 1–11, 2022. [30](#)

D., V.; A., L.; SVETLANA, S. T.; ALEKSANDR, M. S. Data Preprocessing Influence on the Medical Images Segmentation Quality. **IEEE International Conference on Photonics**, 2023. [46](#)

DALE, A. M.; FISCHL, B.; SERENO, M. I. Cortical surface-based analysis: I. segmentation and surface reconstruction. **Neuroimage**, Elsevier, v. 9, n. 2, p. 179–194, 1999. [30](#)

DAVIS, J.; GOADRICH, M. The relationship between precision-recall and roc curves. In: **Proceedings of the 23rd international conference on Machine learning**. New York, NY, United States: Association for Computing Machinery, 2006. p. 233–240. [42](#)

DESPOTOVIĆ, I.; GOOSSENS, B.; PHILIPS, W. Mri segmentation of the human brain: challenges, methods, and applications. **Computational and mathematical methods in medicine**, Hindawi, v. 2015, 2015. [36](#), [37](#)

DEVARAJ, A.; GINNEKEN, B. van; NAIR, A.; BALDWIN, D. Use of volumetry for lung nodule management: theory and practice. **Radiology**, Radiological Society of North America, v. 284, n. 3, p. 630–644, 2017. [31](#)

DICE, L. R. Measures of the amount of ecologic association between species. **Ecology**, JSTOR, v. 26, n. 3, p. 297–302, 1945. [42](#), [43](#)

DONG, H.; YANG, G.; LIU, F.; MO, Y.; GUO, Y. Automatic brain tumor detection and segmentation using u-net based fully convolutional networks. In: SPRINGER. **Medical Image Understanding and Analysis: 21st Annual Conference, MIUA 2017, Edinburgh, UK, July 11–13, 2017, Proceedings 21**. Edinburgh, UK,, 2017. p. 506–517. [42](#)

DOSOVITSKIY, A.; BEYER, L.; KOLESNIKOV, A.; WEISSENBORN, D.; ZHAI, X.; UNTERTHINER, T.; DEHGhani, M.; MINDERER, M.; HEIGOLD, G.; GELLY, S.; USZKOREIT, J.; HOULSBY, N. An image is worth 16x16 words: Transformers for image recognition at scale. **arXiv preprint arXiv:2010.11929**, 2020. [31](#)

EGGER, J.; KAPUR, T.; FEDOROV, A.; PIEPER, S.; MILLER, J. V.; VEERARAGHAVAN, H.; FREISLEBEN, B.; GOLBY, A. J.; NIMSKY, C.; KIKINIS, R. Gbm volumetry using the 3d slicer medical image computing platform. **Scientific reports**, Nature Publishing Group UK London, v. 3, n. 1, p. 1364, 2013. [37](#)

FABIAN, I.; P., J.; SIMON, A. A. K.; JENS, P.; KLAUS, M.-H. nnu-Net: a self-configuring method for deep learning-based biomedical image segmentation. **Nature Methods**, 2020. [32](#), [44](#), [45](#), [46](#), [47](#), [48](#), [49](#), [50](#), [54](#), [76](#)

FALK, T.; MAI, D.; BENSCH, R.; ÇIÇEK, O.; ABDULKADIR, A.; MARRAKCHI, Y.; BÖHM, A.; DEUBNER, J.; JÄCKEL, Z.; SEIWALD, K.;

DOVZHENKO, A.; TIETZ, O.; BOSCO, C. D.; WALSH, S.; SALTUKOGLU, D.; TAY, T. L.; PRINZ, M.; PALME, K.; SIMONS, M.; DIESTER, I.; BROX, T.; RONNEBERGER, O. U-net: deep learning for cell counting, detection, and morphometry. **Nature methods**, Nature Publishing Group, v. 16, n. 1, p. 67–70, 2019. [30](#), [42](#), [76](#)

FEDOROV, A.; BEICHEL, R.; KALPATHY-CRAMER, J.; FINET, J.; FILLION-ROBIN, J.-C.; PUJOL, S.; BAUER, C.; JENNINGS, D.; FENNESSY, F.; SONKA, M.; BUATTI, J.; AYLWARD, S.; MILLER, J.; PIEPER, S.; KIKINIS, R. 3d slicer as an image computing platform for the quantitative imaging network. **Magnetic resonance imaging**, Elsevier, v. 30, n. 9, p. 1323–1341, 2012. [38](#)

FISCHL, B.; SALAT, D. H.; BUSA, E.; ALBERT, M.; DIETERICH, M.; HASELGROVE, C.; KOUWE, A. van der; KILLIANY, R.; KENNEDY, D.; KLAVENESS, S.; MONTILLO, A.; MAKRIS, N.; ROSEN, B.; DALE, A. M. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. **Neuron**, Elsevier, v. 33, n. 3, p. 341–355, 2002. [30](#)

FRANSISKUS, X. F.; E., M. Y.; I., K. E. P.; M., P. Covid-19 Lung Segmentation using U-Net CNN based on Computed Tomography Image. **CIVEMSA**, 2022. [45](#)

GANG, P.; ZENG, W.; GORDIENKO, Y.; KOCHURA, Y.; ALIENIN, O.; ROKOVYI, O.; STIRENKO, S. Effect of data augmentation and lung mask segmentation for automated chest radiograph interpretation of some lung diseases. In: SPRINGER. **International conference on neural information processing**. Sydney, Australia, 2019. p. 333–340. [45](#)

GEISSER, S. The predictive sample reuse method with applications. **Journal of the American statistical Association**, Taylor & Francis, v. 70, n. 350, p. 320–328, 1975. [44](#)

GLASSNER, A. S. Space subdivision for fast ray tracing. **IEEE Computer Graphics and applications**, IEEE, v. 4, n. 10, p. 15–24, 1984. [29](#)

GOERTZEL, B. Artificial general intelligence: concept, state of the art, and future prospects. **Journal of Artificial General Intelligence**, v. 5, n. 1, p. 1–48, 2014. [41](#)

GOLDMAN, L. W. Principles of ct and ct technology. **Journal of Nuclear Medicine Technology**, Society of Nuclear Medicine, v. 35, n. 3, p. 115–128, 2007. [38](#), [39](#)

GOO, J. M. A computer-aided diagnosis for evaluating lung nodules on chest ct: the current status and perspective. **Korean journal of radiology**, The Korean Society of Radiology, v. 12, n. 2, p. 145–155, 2011. [32](#)

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. H.; FRIEDMAN, J. H. **The elements of statistical learning: data mining, inference, and prediction**. New York, NY, USA: Springer, 2009. v. 2. [44](#)

HAWKINS, D. M. The problem of overfitting. **Journal of chemical information and computer sciences**, ACS Publications, v. 44, n. 1, p. 1–12, 2004. [44](#)

HE, K.; GKIOXARI, G.; DOLLÁR, P.; GIRSHICK, R. Mask r-cnn. In: **Proceedings of the IEEE international conference on computer vision**. Venice, Italy.: IEEE, 2017. p. 2961–2969. [30](#)

HECKEL, F.; MEINE, H.; MOLTZ, J. H.; KUHNIGK, J.-M.; HEVERHAGEN, J. T.; KIESSLING, A.; BUERKE, B.; HAHN, H. K. Segmentation-based partial volume correction for volume estimation of solid lesions in ct. **IEEE Transactions on Medical Imaging**, IEEE, v. 33, n. 2, p. 462–480, 2013. [32](#)

HELLER, N.; ISENSEE, F.; MAIER-HEIN, K. H.; HOU, X.; XIE, C.; LI, F.; NAN, Y.; MU, G.; LIN, Z.; HAN, M.; YAO, G.; ZHANG, Y.; WANG, Y.; HOU, F.; YANG, J.; XIONG, G.; TIAN, J.; ZHONG, C.; MA, J.; RICKMAN, J.; DEAN, J.; STAI, B.; TEJPAUL, R.; OESTREICH, M.; BLAKE, P.; KALUZNIAK, H.; RAZA, S.; ROSENBERG, J.; MOORE, K.; WALCZAK, E.; RENGEL, Z.; EDGERTON, Z.; VASDEV, R.; PETERSON, M.; MCSWEENEY, S.; PETERSON, S.; KALAPARA, A.; SATHIANATHEN, N.; PAPANIKOLOPOULOS, N.; WEIGHT, C. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced ct imaging: Results of the kits19 challenge. **Medical image analysis**, Elsevier, v. 67, p. 101821, 2021. [30](#)

HELLER, N.; SATHIANATHEN, N.; KALAPARA, A.; WALCZAK, E.; MOORE, K.; KALUZNIAK, H.; ROSENBERG, J.; BLAKE, P.; RENGEL, Z.; OESTREICH, M.; DEAN, J.; TRADEWELL, M.; SHAH, A.; TEJPAUL, R.; EDGERTON, Z.; PETERSON, M.; RAZA, S.; REGMI, S.;

PAPANIKOLOPOULOS, N.; WEIGHT, C. The kits19 challenge data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes. **arXiv preprint arXiv:1904.00445**, 2019. 30

HENDRYCKS, D.; BASART, S.; MAZEIKA, M.; ZOU, A.; KWON, J.; MOSTAJABI, M.; STEINHARDT, J.; SONG, D. Scaling out-of-distribution detection for real-world settings. **arXiv preprint arXiv:1911.11132**, 2019. 53

HENSCHKE, C. I.; YANKELEVITZ, D. F.; MIRTICHEVA, R.; MCGUINNESS, G.; MCCAULEY, D.; MIETTINEN, O. S. Ct screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. **American Journal of Roentgenology**, Am Roentgen Ray Soc, v. 178, n. 5, p. 1053–1057, 2002. 29

HOFMANNINGER, J.; PRAYER, F.; PAN, J.; RÖHRICH, S.; PROSCH, H.; LANGS, G. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. **European Radiology Experimental**, Springer, v. 4, p. 1–13, 2020. 55

HOLLON, T. C.; PANDIAN, B.; ADAPA, A. R.; URIAS, E.; SAVE, A. V.; KHALSA, S. S. S.; EICHBERG, D. G.; D'AMICO, R. S.; FAROOQ, Z. U.; LEWIS, S.; PETRIDIS, P. D.; MARIE, T.; SHAH, A. H.; GARTON, H. J. L.; MAHER, C. O.; HETH, J. A.; MCKEAN, E. L.; SULLIVAN, S. E.; HERVEY-JUMPER, S. L.; PATIL, P. G.; THOMPSON, B. G.; SAGHER, O.; 2ND, G. M. M.; KOMOTAR, R. J.; IVAN, M. E.; SNUDERL, M.; OTTEN, M. L.; JOHNSON, T. D.; SISTI, M. B.; BRUCE, J. N.; MURASZKO, K. M.; TRAUTMAN, J.; FREUDIGER, C. W.; CANOLL, P.; LEE, H.; CAMELO-PIRAGUA, S.; ORRINGER, D. A. Near real-time intraoperative brain tumor diagnosis using stimulated raman histology and deep neural networks. **Nature medicine**, Nature Publishing Group, v. 26, n. 1, p. 52–58, 2020. 30

HOUNSFIELD, G. N. Computerized transverse axial scanning (tomography): Part 1. description of system. **The British journal of radiology**, The British Institute of Radiology, v. 46, n. 552, p. 1016–1022, 1973. 38

HU, Z.; ZHUANG, Q.; XIAO, Y.; WU, G.; SHI, Z.; CHEN, L.; WANG, Y.; YU, J. Mil normalization — prerequisites for accurate MRI radiomics analysis. **Computers in Biology and Medicine**, Elsevier BV, v. 133, p. 104403, 6 2021. ISSN 0010-4825. 47

HUANG, H. K. *Pacs and imaging informatics: basic principles and applications*. John Wiley & Sons, 2011. [39](#)

HUANG, L.; MIRON, A.; HONE, K.; LI, Y. Segmenting Medical Images: From UNet to Res-UNet and nnUNet. In: IEEE. **2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)**. Guadalajara, Mexico, 2024. p. 483–489. [49](#)

HUGHES, J. F.; DAM, A. van; MCGUIRE, M.; SKLAR, D. F.; FOLEY, J. D.; FEINER, S. K.; AKELEY, K. **Computer Graphics: Principles and Practice**. Third edition. Upper Saddle River, New Jersey: Addison-Wesley, 2014. Revised edition of: *Computer Graphics / James D. Foley. . . [et al.]*.—2nd ed., Reading, Mass.: Addison-Wesley, 1995. ISBN 978-0-321-39952-6. [29](#)

III, S. G. A.; MEYER, C. R.; MCNITT-GRAY, M. F.; MCLENNAN, G.; REEVES, A.; CROFT, B. Y.; CLARKE, L. P.; GROUP, R. R. The reference image database to evaluate response to therapy in lung cancer (rider) project: A resource for the development of change-analysis software. **Clinical Pharmacology & Therapeutics**, Wiley Online Library, v. 84, n. 4, p. 448–456, 2008. [57](#)

IQBAL, H. **HarisIqbal88/PlotNeuralNet v1.0.0**. Europe: Zenodo, 2018. [60](#)

ISENSEE, F.; JAEGER, P. F.; FULL, P. M.; WOLF, I.; ENGELHARDT, S.; MAIER-HEIN, K. H. Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features. In: SPRINGER. **Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers 8**. Quebec City, Canada, 2018. p. 120–129. [42](#), [47](#), [49](#)

ISENSEE, F.; PETERSEN, J.; KLEIN, A.; ZIMMERER, D.; JAEGER, P. F.; KOHL, S.; WASSERTHAL, J.; KOEHLER, G.; NORAJITRA, T.; WIRKERT, S.; MAIER-HEIN, K. H. Abstract: nnu-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation. In: _____. **Informatik aktuell**. Wiesbaden, Germany: Springer Fachmedien Wiesbaden, 2019. p. 22–22. ISBN 9783658253257. [47](#), [48](#), [49](#), [50](#)

ISENSEE, F.; ULRICH, C.; WALD, T.; MAIER-HEIN, K. H. Extending nnU-Net is all you need. **Bildverarbeitung für die Medizin**, arXiv, 2022. [50](#)

ISENSEE, F.; WALD, T.; ULRICH, C.; BAUMGARTNER, M.; ROY, S.; MAIER-HEIN, K.; JAEGER, P. F. nnu-net revisited: A call for rigorous validation in 3d medical image segmentation. In: SPRINGER. **International Conference on Medical Image Computing and Computer-Assisted Intervention**. MARRAKESH, MOROCCO, 2024. p. 488–498. [42](#), [48](#)

JACOBSEN, N.; DEISTUNG, A.; TIMMANN, D.; GOERICKE, S. L.; REICHENBACH, J. R.; GÜLLMAR, D. Analysis of intensity normalization for optimal segmentation performance of a fully convolutional neural network. **Zeitschrift für Medizinische Physik**, Elsevier BV, v. 29, n. 2, p. 128–138, 5 2019. ISSN 0939-3889. [47](#)

JESUS, T.; MAGALHÃES, R.; ALVES, V. Spatial Normalization of MRI Brain Studies Using a U-Net Based Neural Network. In: _____. **Advances in Intelligent Systems and Computing**. Budva, Montenegro: Springer International Publishing, 2020. p. 493–502. ISBN 9783030456962. [47](#)

JUCEVICIUS, J.; TREIGYS, P.; BERNATAVICIENE, J.; TRAKYMAS, M.; NARUSEVICIUTE, I.; BRIEDIENE, R. Segmentation Mask Resampling for MRI Prostate Localization Improvement. In: IEEE. **2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)**. Prague, Czech Republic, 2022. p. 1–6. [47](#)

K., B. D. R.; KARIN, A. v. G.; M., S.; S., V. D. V.; FATIH, I.; E., O.; J., H.; S., K.; M., S. The Effect of Preprocessing on Convolutional Neural Networks for Medical Image Segmentation. **IEEE International Symposium on Biomedical Imaging**, 2021. [46](#)

KALENDER, W. A. Computed tomography fundamentals, system technology, image quality, applications. **Publicis**, Erlangen, 2006. [38](#), [39](#)

KIM, H.; MONROE, J. I.; LO, S.; YAO, M.; HARARI, P. M.; MACHTAY, M.; SOHN, J. W. Quantitative evaluation of image segmentation incorporating medical consideration functions. **Medical Physics**, Wiley, v. 42, n. 6Part1, p. 3013–3023, may 26 2015. ISSN 0094-2405. [44](#)

KIM, J.; BROWN, M. S.; RUAN, D. Self-supervised learning without annotations to improve lung chest x-ray segmentation. In: SPIE. **Medical Imaging 2024: Computer-Aided Diagnosis**. Manchester, UK, 2024. v. 12927, p. 774–781. [77](#)

KONING, H. J. de; AALST, C. M. van der; JONG, P. A. de; SCHOLTEN, E. T.; NACKAERTS, K.; HEUVELMANS, M. A.; LAMMERS, J.-W. J.; WEENINK, C.; YOUSAF-KHAN, U.; HOREWEG, N.; WESTEINDE, S. van 't; PROKOP, M.; MALI, W. P.; HOESEIN, F. A. A. M.; OOIJEN, P. M. A. van; AERTS, J. G. J. V.; BAKKER, M. A. den; THUNNISSEN, E.; VERSCHAKELEN, J.; Vliegenthart, R.; Walter, J. E.; HAAF, K. T.; Groen, H. J. M.; Oudkerk, M. Reduced lung-cancer mortality with volume ct screening in a randomized trial. **New England Journal of Medicine**, Massachusetts Medical Society, v. 382, n. 6, p. 503–513, 2020. [29](#)

KRIZHEVSKY, A.; HINTON, G. Learning multiple layers of features from tiny images. Toronto, ON, Canada, 2009. [31](#)

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Advances in neural information processing systems**, v. 25, 2012. [41](#)

KUBOTA, T.; JEREBKO, A. K.; DEWAN, M.; SALGANICOFF, M.; KRISHNAN, A. Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. **Medical Image Analysis**, Elsevier, v. 15, n. 1, p. 133–154, 2011. [76](#)

LANCASTER, H. L.; ZHENG, S.; ALESHINA, O. O.; YU, D.; CHERNINA, V. Y.; HEUVELMANS, M. A.; BOCK, G. H. de; DORRIUS, M. D.; GRATAMA, J. W.; MOROZOV, S. P.; GOMBOLEVSKIY, V. A.; SILVA, M.; YI, J.; Oudkerk, M. Outstanding negative prediction performance of solid pulmonary nodule volume ai for ultra-ldct baseline lung cancer screening risk stratification. **Lung Cancer**, Elsevier, v. 165, p. 133–140, 2022. [32](#)

LARICI, A. R.; FARCHIONE, A.; FRANCHI, P.; CILIBERTO, M.; CICCETTI, G.; CALANDRIELLO, L.; CIELLO, A. D.; BONOMO, L. Lung nodules: size still matters. **European respiratory review**, European Respiratory Society, v. 26, n. 146, 2017. [73](#)

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. [41](#)

LI, Z.; ZHU, Q.; ZHANG, L.; YANG, X.; LI, Z.; FU, J. A deep learning-based self-adapting ensemble method for segmentation in gynecological brachytherapy. **Radiation Oncology**, BioMed Central, v. 17, n. 1, p. 1–10, 2022. [46](#)

- LIN, T.-Y.; MAIRE, M.; BELONGIE, S.; HAYS, J.; PERONA, P.; RAMANAN, D.; DOLLÁR, P.; ZITNICK, C. L. Microsoft coco: Common objects in context. In: SPRINGER. **Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13**. Zurich, Switzerland, 2014. p. 740–755. [30](#)
- LITJENS, G.; KOOL, T.; BEJNORDI, B. E.; SETIO, A. A. A.; CIOMPI, F.; GHAFLOORIAN, M.; LAAK, J. A. V. D.; GINNEKEN, B. V.; SÁNCHEZ, C. I. A survey on deep learning in medical image analysis. **Medical image analysis**, Elsevier, v. 42, p. 60–88, 2017. [30](#), [45](#)
- LIU, X.; SONG, L.; LIU, S.; ZHANG, Y. A review of deep-learning-based medical image segmentation methods. **Sustainability**, MDPI, v. 13, n. 3, p. 1224, 2021. [37](#)
- LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. p. 3431–3440, 2015. [29](#), [31](#), [37](#), [42](#)
- MACMAHON, H.; NAIDICH, D. P.; GOO, J. M.; LEE, K. S.; LEUNG, A. N. C.; MAYO, J. R.; MEHTA, A. C.; OHNO, Y.; POWELL, C. A.; PROKOP, M.; RUBIN, G. D.; SCHAEFER-PROKOP, C. M.; TRAVIS, W. D.; SCHIL, P. E. V.; BANKIER, A. A. Guidelines for management of incidental pulmonary nodules detected on ct images: from the fleischner society 2017. **Radiology**, Radiological Society of North America, v. 284, n. 1, p. 228–243, 2017. [29](#), [33](#), [51](#)
- MAZZONE, P. J.; LAM, L. Evaluating the patient with a pulmonary nodule: a review. **Jama**, American Medical Association, v. 327, n. 3, p. 264–273, 2022. [35](#)
- MCCONNELL, N.; MIRON, A.; WANG, Z.; LI, Y. Integrating Residual, Dense, and Inception Blocks into the nnUNet. In: IEEE. **2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS)**. Shenzhen, China, 2022. p. 217–222. [48](#), [50](#)
- MCCONNELL, N.; NDIPENOCH, N.; CAO, Y.; MIRON, A.; LI, Y. Exploring advanced architectural variations of nnUNet. **Neurocomputing**, Elsevier BV, v. 560, p. 126837, 12 2023. ISSN 0925-2312. [50](#)
- MCWILLIAMS, A.; TAMMEMAGI, M. C.; MAYO, J. R.; ROBERTS, H.; LIU, G.; SOGHRATI, K.; YASUFUKU, K.; MARTEL, S.; LABERGE, F.; GINGRAS, M.; ATKAR-KHATTRA, S.; BERG, C. D.; EVANS, K.; FINLEY, R.; YEE, J.; ENGLISH, J.; NASUTE, P.; GOFFIN, J.; PUKSA, S.; STEWART, L.; TSAI, S.;

JOHNSTON, M. R.; MANOS, D.; NICHOLAS, G.; GOSS, G. D.; SEELY, J. M.; AMJADI, K.; TREMBLAY, A.; BURROWES, P.; MACEACHERN, P.; BHATIA, R.; TSAO, M.-S.; LAM, S. Probability of cancer in pulmonary nodules detected on first screening ct. **New England Journal of Medicine**, Massachusetts Medical Society, v. 369, n. 10, p. 910–919, 2013. [29](#)

MILLETARI, F.; NAVAB, N.; AHMADI, S.-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. p. 565–571, 2016. [31](#), [32](#)

MOORTHY, J.; GANDHI, U. D. A survey on medical image segmentation based on deep learning techniques. **Big Data and Cognitive Computing**, MDPI, v. 6, n. 4, p. 117, 2022. [36](#)

MORAIN-NICOLIER, F.; LEBONVALLET, S.; BAUDRIER, E.; RUAN, S. Hausdorff Distance based 3d Quantification of Brain Tumor Evolution from MRI Images. In: IEEE. **2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society**. Lyon, France, 2007. p. 5597–5600. ISSN 1557-170X. [44](#)

MURCHISON, J. T.; RITCHIE, G.; SENYSZAK, D.; NIJWENING, J. H.; VEENENDAAL, G. van; WAKKIE, J.; BEEK, E. J. van. Validation of a deep learning computer aided system for ct based lung nodule detection, classification, and growth rate estimation in a routine clinical population. **Plos one**, Public Library of Science San Francisco, CA USA, v. 17, n. 5, p. e0266799, 2022. [32](#)

MURTHY, S.; RICE, T. The solitary pulmonary nodule: a primer on differential diagnosis. v. 14, n. 3, p. 239–249, 2002. [35](#)

NAIDICH, D. P.; BANKIER, A. A.; MACMAHON, H.; SCHAEFER-PROKOP, C. M.; PISTOLESI, M.; GOO, J. M.; MACCHIARINI, P.; CRAPO, J. D.; HEROLD, C. J.; AUSTIN, J. H.; TRAVIS, W. D. Recommendations for the management of subsolid pulmonary nodules detected at ct: a statement from the fleischner society. **Radiology**, Radiological Society of North America, Inc., v. 266, n. 1, p. 304–317, 2013. [29](#), [36](#)

NATHAN, M.; COLLINS, V.; ADAMS, R. Differentiation of benign and malignant pulmonary nodules by growth rate. **Radiology**, The Radiological Society of North America, v. 79, n. 2, p. 221–232, 1962. [35](#)

NAYYAR, A.; GADHAVI, L.; ZAMAN, N. Machine learning in healthcare: review, opportunities and challenges. **Machine learning and the internet of medical things in healthcare**, Elsevier, p. 23–45, 2021. [76](#)

NESTEROV, Y. A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$. In: **Dokl. Akad. Nauk. SSSR**. URSS: SSSR, 1983. v. 269, n. 3, p. 543. [60](#)

OKTAY, O.; SCHLEMPER, J.; FOLGOC, L. L.; LEE, M.; HEINRICH, M.; MISAWA, K.; MORI, K.; MCDONAGH, S.; HAMMERLA, N. Y.; KAINZ, B.; GLOCKER, B.; RUECKERT, D. Attention u-net: Learning where to look for the pancreas. **arXiv preprint arXiv:1804.03999**, 2018. [31](#), [32](#)

OLIVA, A.; TORRALBA, A. Modeling the shape of the scene: A holistic representation of the spatial envelope. **International journal of computer vision**, Springer, v. 42, p. 145–175, 2001. [31](#)

ONOFREY, J. A.; CASETTI-DINESCU, D. I.; LAURITZEN, A. D.; SARKAR, S.; VENKATARAMAN, R.; FAN, R. E.; SONN, G. A.; SPREngle, P. C.; STAIB, L. H.; PAPADEMETRIS, X. Generalizable Multi-Site Training and Testing Of Deep Neural Networks Using Image Normalization. In: **IEEE. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)**. Venice, Italy, 2019. [47](#)

OULEFKI, A.; AGAIAN, S.; TRONGTIRAKUL, T.; LAOUAR, A. K. Automatic COVID-19 lung infected region segmentation and measurement using CT-scans images. **Pattern Recognition**, Elsevier BV, v. 114, p. 107747, 6 2021. ISSN 0031-3203. [45](#)

PARKHI, O. M.; VEDALDI, A.; ZISSERMAN, A.; JAWAHAR, C. Cats and dogs. p. 3498–3505, 2012. [31](#)

PASZKE, A.; GROSS, S.; MASSA, F.; LERER, A.; BRADBURY, J.; CHANAN, G.; KILLEEN, T.; LIN, Z.; GIMELSHEIN, N.; ANTIGA, L.; DESMAISON, A.; KOPF, A.; YANG, E.; DEVITO, Z.; RAISON, M.; TEJANI, A.; CHILAMKURTHY, S.; STEINER, B.; FANG, L.; BAI, J.; CHINTALA, S. Pytorch: An imperative style, high-performance deep learning library. In: WALLACH, H.; LAROCHELLE, H.; BEYGELZIMER, A.; ALCHÉ-BUC, F. d'; FOX, E.; GARNETT, R. (Ed.). **Advances in Neural Information Processing**

Systems. Vancouver, British Columbia, Canada: Curran Associates, Inc., 2019. v. 32. [59](#)

PHAM, D. L.; XU, C.; PRINCE, J. L. Current methods in medical image segmentation. **Annual review of biomedical engineering**, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 2, n. 1, p. 315–337, 2000. [36](#), [37](#)

_____. A survey of current methods in medical image segmentation. **Annual review of biomedical engineering**, v. 2, n. 3, p. 315–337, 2000. [30](#)

PIANYKH, O. S.; PIANYKH, O. S. Dicom security. Springer, 2012. [39](#)

POWERS, D. M. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. **arXiv preprint arXiv:2010.16061**, 2020. [42](#)

PUNN, N. S.; AGARWAL, S. Modality specific u-net variants for biomedical image segmentation: a survey. **Artificial Intelligence Review**, Springer, v. 55, n. 7, p. 5845–5889, 2022. [46](#)

QINHUA, H.; L., F. D. F. S.; G., H.; S., S. A.; FRANCISCO, H. S. S.; TAO, H.; P., P. R. F. An effective approach for CT lung segmentation using mask region-based convolutional neural networks. **Artif. Intell. Medicine**, 2020. [44](#), [45](#)

REINHOLD, J. C.; DEWEY, B. E.; CARASS, A.; PRINCE, J. L. Evaluating the impact of intensity normalization on MR image synthesis. In: ANGELINI, E. D.; LANDMAN, B. A. (Ed.). **Medical Imaging 2019: Image Processing**. San Diego, California, USA, 2019. p. 126. [47](#)

REY, D.; NEUHÄUSER, M. Wilcoxon-signed-rank test. In: _____. **International Encyclopedia of Statistical Science**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 1658–1659. ISBN 978-3-642-04898-2. [74](#)

RICKY, R. S. nnu-Net: Further Automating Biomedical Image Autosegmentation. **Radiology: Imaging Cancer**, 2021. [47](#), [49](#), [50](#)

ROBBINS, H.; MONRO, S. A stochastic approximation method. **The annals of mathematical statistics**, JSTOR, p. 400–407, 1951. [60](#)

- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. p. 234–241, 2015. [30](#), [31](#), [37](#), [42](#), [49](#)
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **nature**, Nature Publishing Group UK London, v. 323, n. 6088, p. 533–536, 1986. [41](#), [43](#)
- RUSSAKOVSKY, O.; DENG, J.; SU, H.; KRAUSE, J.; SATHEESH, S.; MA, S.; HUANG, Z.; KARPATY, A.; KHOSLA, A.; BERNSTEIN, M.; BERG, A. C.; FEI-FEI, L. Imagenet large scale visual recognition challenge. **International journal of computer vision**, Springer, v. 115, p. 211–252, 2015. [31](#)
- RUSSELL, S. J.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 4. ed. Upper Saddle River, NJ: Pearson Education, 2020. 1136 p. (Pearson Series in Artificial Intelligence). ISBN 978-0134610993. [41](#)
- SCHOLTEN, E. T.; HOOP, B. de; JACOBS, C.; VORST, S. van Amelsvoort-van de; KLAVEREN, R. J. van; OUDKERK, M.; VLIEGENTHART, R.; KONING, H. J. de; AALST, C. M. van der; MALI, W. T. M.; GIETEMA, H. A.; PROKOP, M.; GINNEKEN, B. van; JONG, P. A. de. Semi-automatic quantification of subsolid pulmonary nodules: comparison with manual measurements. **PLoS One**, Public Library of Science, v. 8, n. 11, p. e80249, 2013. [32](#)
- SETAREHDAN, S. K.; SINGH, S. Advanced algorithmic approaches to medical image segmentation: state-of-the-art applications in cardiology, neurology, mammography and pathology. Springer Science & Business Media, 2001. [36](#)
- SHALEV-SHWARTZ, S.; BEN-DAVID, S. **Understanding machine learning: From theory to algorithms**. Cambridge, UK: Cambridge university press, 2014. [41](#)
- SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). **Biometrika**, Oxford University Press, v. 52, n. 3-4, p. 591–611, 1965. [74](#)
- SHARMA, N.; AGGARWAL, L. M. Automated medical image segmentation techniques. **Journal of medical physics/Association of Medical Physicists of India**, Wolters Kluwer–Medknow Publications, v. 35, n. 1, p. 3, 2010. [30](#), [37](#)

SHINOHARA, R. T.; SWEENEY, E. M.; GOLDSMITH, J.; SHIEE, N.; MATEEN, F. J.; CALABRESI, P. A.; JARSO, S.; PHAM, D. L.; REICH, D. S.; CRAINICEANU, C. M. Statistical normalization techniques for magnetic resonance imaging. **NeuroImage: Clinical**, Elsevier BV, v. 6, p. 9–19, 2014. ISSN 2213-1582. [47](#)

SOUZA, L. F.; HOLANDA, G.; SILVA, F. H.; ALVES, S. S.; FILHO, P. P. Automatic lung segmentation in CT images using mask R-CNN for mapping the feature extraction in supervised methods of machine learning using transfer learning. **International Journal of Hybrid Intelligent Systems**, SAGE Publications, v. 16, n. 4, p. 189–205, feb 5 2021. ISSN 1448-5869. [45](#)

STEENBAKKERS, R. J. H. M.; DUPPEN, J. C.; FITTON, I.; DEURLOO, K. E. I.; ZIJP, L.; UITTERHOEVE, A. L. J.; RODRIGUS, P. T. R.; KRAMER, G. W. P.; BUSSINK, J.; JAEGER, K. D.; BELDERBOS, J. S. A.; HART, A. A. M.; NOWAK, P. J. C. M.; HERK, M. van; RASCH, C. R. N. Observer variation in target volume delineation of lung cancer related to radiation oncologist–computer interaction: a 'big brother' evaluation. **Radiotherapy and Oncology**, Elsevier, v. 77, n. 2, p. 182–190, 2005. [29](#)

STEHMAN, S. V. Selecting and interpreting measures of thematic classification accuracy. **Remote sensing of Environment**, Elsevier, v. 62, n. 1, p. 77–89, 1997. [42](#)

STONE, M. Cross-validatory choice and assessment of statistical predictions. **Journal of the royal statistical society: Series B (Methodological)**, Wiley Online Library, v. 36, n. 2, p. 111–133, 1974. [44](#)

SUDRE, C. H.; LI, W.; VERCAUTEREN, T.; OURSELIN, S.; CARDOSO, M. J. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: SPRINGER. **Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3**. Québec City, QC, Canada, 2017. p. 240–248. [42](#), [43](#)

SUNG, H.; FERLAY, J.; SIEGEL, R. L.; LAVERSANNE, M.; SOERJOMATARAM, I.; JEMAL, A.; BRAY, F. Global cancer statistics 2020:

Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: a cancer journal for clinicians**, Wiley Online Library, v. 71, n. 3, p. 209–249, 2021. [29](#)

T., N. A.; T., D.; S., M.; P., L.; YURIE, S.; A., B.; C., B.; S., K. Accuracy and precision of manual segmentation of the maxillary sinus in MR images-a method study. **British Journal of Radiology**, 2018. [43](#)

TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: PMLR. **International conference on machine learning**. Long Beach, CA, USA, 2019. p. 6105–6114. [31](#)

TANG, H.; ZHANG, C.; XIE, X. Nodulenet: Decoupled false positive reduction for pulmonary nodule detection and segmentation. In: SPRINGER. **Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22**. Shenzhen, China, 2019. p. 266–274. [76](#)

TANIMOTO, T. T. Elementary mathematical theory of classification and prediction. **International Business Machines Corporation (IBM Report)**, McGraw-HillNew York, 1958. [42](#), [43](#)

TEAM, N. L. S. T. R. Reduced lung-cancer mortality with low-dose computed tomographic screening. **New England Journal of Medicine**, Mass Medical Soc, v. 365, n. 5, p. 395–409, 2011. [29](#)

V., Y.; I., V. Boundary overlap for medical image segmentation evaluation. **Medical Imaging**, 2017. [43](#)

VALIANT, L. G. A theory of the learnable. **Communications of the ACM**, ACM New York, NY, USA, v. 27, n. 11, p. 1134–1142, 1984. [41](#)

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, Ł.; POLOSUKHIN, I. Attention is all you need. **Advances in neural information processing systems**, v. 30, 2017. [31](#)

WORTH, A. J.; MAKRIS, N.; JR, V. S. C.; KENNEDY, D. N. Neuroanatomical segmentation in mri: technological objectives. **International Journal of Pattern Recognition and Artificial Intelligence**, World Scientific, v. 11, n. 08, p. 1161–1187, 1997. [30](#)

YANIV, Z.; LOWEKAMP, B. C.; JOHNSON, H. J.; BEARE, R. Simpleitk image-analysis notebooks: a collaborative environment for education and reproducible research. **Journal of digital imaging**, Springer, v. 31, n. 3, p. 290–303, 2018. [56](#)

YEPES-CALDERON, F.; MCCOMB, J. G. Manual segmentation errors in medical imaging. proposing a reliable gold standard. In: SPRINGER. **International Conference on Applied Informatics**. Honolulu, Hawaii, USA, 2019. p. 230–241. [30](#)

YOON, S. H.; KIM, J.; LEE, K. J.; NAM, C.-M.; KIM, J.; LEE, K. H.; LEE, K. W. Volumetric analysis of pulmonary nodules: reducing the discrepancy between the diameter-based volume calculation and voxel-counting method. **Quantitative Imaging in Medicine and Surgery**, AME Publications, v. 12, n. 3, p. 1674, 2022. [30](#), [32](#)

YU, A. C.; MOHAJER, B.; ENG, J. External validation of deep learning algorithms for radiologic diagnosis: a systematic review. **Radiology: Artificial Intelligence**, Radiological Society of North America, v. 4, n. 3, p. e210064, 2022. [33](#)

ZHAO, B.; JAMES, L. P.; MOSKOWITZ, C. S.; GUO, P.; GINSBERG, M. S.; LEFKOWITZ, R. A.; QIN, Y.; RIELY, G. J.; KRIS, M. G.; SCHWARTZ, L. H. Evaluating variability in tumor measurements from same-day repeat ct scans of patients with non-small cell lung cancer. **Radiology**, Radiological Society of North America, Inc., v. 252, n. 1, p. 263–272, 2009. [57](#)

ZHOU, Z.; SIDDIQUEE, M. M. R.; TAJBAKHSI, N.; LIANG, J. Unet++: A nested u-net architecture for medical image segmentation. p. 3–11, 2018. [31](#), [32](#)