



Universidade de Brasília – UnB
Campus Gama – FGA
Programa de Pós-Graduação em Engenharia Biomédica

**Segmentação de Imagens de Tomografia Computadorizada
Utilizando o Algoritmo U-Net. Estudo de Caso:
Imagem da Bexiga**

José Hevenicio do Nascimento Silva

Orientador: DR. GERARDO ANTONIO IDROBO PIZO



UNB – UNIVERSIDADE DE BRASÍLIA

FGA – FACULDADE GAMA



**Segmentação de Imagens de Tomografia Computadorizada
Utilizando o Algoritmo U-Net. Estudo de Caso:
Imagem da Bexiga**

José Hevenicio do Nascimento Silva

ORIENTADOR: DR. GERARDO ANTONIO IDROBO PIZO

**DISSERTAÇÃO DE MESTRADO EM
ENGENHARIA BIOMÉDICA**

PUBLICAÇÃO: 186A/2024

BRASÍLIA/DF, SETEMBRO DE 2024

UnB – Universidade de Brasília
FGA – Faculdade Gama
Programa de Pós-Graduação

Segmentação de Imagens de Tomografia Computadorizada
Utilizando o Algoritmo U-Net. Estudo de Caso:
Imagem da Bexiga

JOSÉ HEVENICIO DO NASCIMENTO SILVA

DISSERTAÇÃO DE MESTRADO SUBMETIDA AO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA BIOMÉDICA DA UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA BIOMÉDICA

APROVADA POR:

Dr. Gerardo Antonio Idrobo Pizo
(Orientador)

Dr. Renato Coral Sampaio
(Examinador interno) UnB

Dra. Thaína Aparecida Azevedo Tosta
(Examinador externo) Unifesp

Brasília/DF, Setembro de 2024

Ficha Catalográfica

SILVA, JOSÉ HEVENICIO DO NASCIMENTO

Segmentação de Imagens de Tomografia Computadorizada Utilizando o Algoritmo U-Net. Estudo de Caso: Imagem da Bexiga [Distrito Federal], 2024.

65p., 210 × 297 mm (FGA/UnB Gama, Mestrado em Engenharia Biomédica, 2024).

Dissertação de Mestrado em Engenharia Biomédica, Faculdade UnB Gama, Programa de Pós-Graduação em Engenharia Biomédica.

- | | |
|-------------------------------|---------------------------|
| 1. Processamento de Imagens | 2. Segmentação Automática |
| 3. Tomografia Computadorizada | 4. Aprendizado Profundo |
| I. FGA UnB/UnB. | II. Título (série) |

Referência

SILVA, JOSÉ HEVENICIO DO NASCIMENTO (2024). Segmentação de Imagens de Tomografia Computadorizada Utilizando o Algoritmo U-Net. Estudo de Caso: Imagem da Bexiga. Dissertação de mestrado em engenharia biomédica, Publicação 186A/2024, Programa de Pós-Graduação, Faculdade UnB Gama, Universidade de Brasília, Brasília, DF, 65p.

Cessão de Direitos

AUTOR: José Hevenicio do Nascimento Silva

TÍTULO: Segmentação de Imagens de Tomografia Computadorizada Utilizando o Algoritmo U-Net. Estudo de Caso: Imagem da Bexiga

GRAU: Mestre

ANO: 2024

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem a autorização por escrito do autor.

venn19@hotmail.com.br

Brasília, DF – Brasil

Resumo

O objetivo da radioterapia é levar uma determinada dose de radiação prescrita a um alvo do tratamento poupando, ao máximo os tecidos saudáveis. Contudo, ao irradiar o alvo, inevitavelmente células saudáveis também são irradiadas, podendo gerar algum tipo de toxicidade. O contorno manual de tumores e estruturas normais é uma parte muito trabalhosa e demorada no processo de planejamento de tratamento em radioterapia. Em tratamentos localizados na região pélvica, o preenchimento da bexiga pode influenciar diretamente na posição dos órgãos circundantes a ela. Uma solução é a realização da segmentação da bexiga via métodos automáticos, usando algoritmos de inteligência artificial para verificar antes do tratamento se a bexiga está cheia ou vazia, de acordo com o preenchimento dela no planejamento, eliminando dessa forma a demora nos contornos de órgão nos dias de tratamento. Este estudo investiga a aplicação de uma técnica avançada de aprendizado profundo na segmentação de imagens médicas, com foco em um estudo de caso específico: a segmentação da bexiga em imagens de Tomografia Computadorizada (TC) do tipo Cone-Beam. O objetivo principal é desenvolver uma abordagem precisa e prática para delimitar a região de interesse, contribuindo para os avanços no processamento de imagens médicas. A metodologia adotada utilizou um conjunto de imagens de tomografia computadorizada disponibilizadas pelo setor de radioterapia do Hospital Universitário de Brasília (HUB), que serviu para treinar, testar e avaliar o modelo. A base de dados é composta por 1.932 imagens (de dimensões de 256×256 pixels) de TC. A arquitetura escolhida, a U-Net, foi treinada utilizando estratégias de aumento de dados para aprimorar a sua generalização. Os resultados demonstraram a viabilidade da abordagem proposta para a segmentação da bexiga, com um coeficiente de Dice de 81%. Além disso, o desenvolvimento de um aplicativo integrado ao modelo pré-treinado poderá fornecer uma ferramenta prática e acessível para os especialistas em radioterapia. A análise qualitativa das segmentações, reforçada por exemplos visuais, destaca a eficácia do modelo em localizar e contornar a anatomia da bexiga em imagens de TC.

Palavras-chave: Processamento de Imagens, Segmentação Automática, Tomografia Computadorizada, Inteligência Artificial, Aprendizado Profundo.

Abstract

The aim of radiotherapy is to send a specific prescribed radiation dose towards a target spot for treatment, whilst avoiding damaging the surrounding healthy tissue as much as possible. However, by irradiating the spot, healthy cells are also inevitably affected, resulting in some toxicity. The manual boundary generation of tumors and normal anatomy is a long and arduous process during planning of radiotherapy treatments. In treatments focusing on the pelvic region, a filled bladder can influence directly on the position of the surrounding organs. One solution is to perform the bladder's image segmentation through automatic methods, by way of artificial intelligence algorithms, so that the bladder can be checked before treatment whether it is filled or not according with planning. This eliminates the wait on boundary generation during treatment. This study investigates the application of an advanced technique of deep learning medical image segmentation, focusing on a specific case study: the bladder's segmentation in computed tomography (CT) cone-beam images. Its main goal is to develop a precise and practical approach to delineate the region of interest, contributing to the advance of medical image processing. The adopted methodology utilized a set of CT images made available by Hospital Universitário de Brasília's (HUB) radiotherapy department, which allowed for the training, testing and evaluation of the model. The database is comprised of 1.932 256×256 pixels CT images. The chosen architecture, U-Net, was trained with the help of data augmentation techniques to improve the method's generalization. The results showed the proposed approach's viability for bladder segmentation, with a Dice coefficient of 81%. Additionally, the development of an application integrated to the pre-trained model could offer a practical and accessible tool for radiotherapy specialists. The quantitative analysis of the segmentations, supported by visualizations, highlights the model's efficacy in finding and delineating the bladder's anatomy in CT images.

Keywords: Image Processing, Automatic Segmentation, Computed Tomography, Artificial Intelligence, Deep Learning.

Sumário

1	Introdução	1
1.1	Contextualização	1
1.2	Aplicações da Inteligência Artificial na Radioterapia	1
1.3	Trabalhos Relacionados	3
1.3.1	Escolha da Arquitetura de Rede Neural Convolutacional	8
1.4	Justificativa	8
1.5	Objetivos	10
1.5.1	Objetivo Geral	10
1.5.2	Objetivos Específicos	10
1.6	Estrutura da Dissertação	10
2	Fundamentação Teórica	13
2.1	Visão Computacional	13
2.2	Processamento de Imagens	14
2.3	Segmentação de Imagens	16
2.3.1	Segmentação de Imagens Médicas	18
2.4	Inteligência Artificial	23
2.4.1	Aprendizado de Máquina	24
2.4.2	Aprendizado Profundo	32
2.4.3	Métricas de Avaliação	37
3	Metodologia	39
3.1	Aquisição do Banco de Imagens	39
3.1.1	Aumento de Dados	40

3.2	Seleção do Modelo e Detalhes da Implementação	42
3.2.1	Implementação da U-Net	43
3.2.2	Divisão da Base de Dados	44
3.2.3	Desenvolvimento e Avaliação de Desempenho do Aplicativo	44
4	Resultados e Discussões	46
4.1	Avaliação Quantitativa da Segmentação	46
4.2	Avaliação de Desempenho do Aplicativo	49
5	Conclusões e Trabalhos Futuros	51
	Lista de Referências	51

Lista de Tabelas

1.1	Resumo das características descritivas de todos os estudos incluídos na revisão da literatura.	6
2.1	Relações entre as redes biológica e artificial.	30
3.1	Parâmetros do <i>data augmentation</i>	42
4.1	Avaliação quantitativa do desempenho da segmentação	46
4.2	Comparação do modelo com outros trabalhos citados na revisão da literatura.	48

Lista de Figuras

1.1	Diagrama de Fluxo PRISMA	5
1.2	Amostra de cortes de TC da bexiga	10
1.3	Comparação da anatomia pélvica em dois pacientes	11
2.1	Esquema dos principais campos da Visão Computacional.	13
2.2	Matriz com diferentes tonalidades dos pixels na escala gray.	14
2.3	Uma imagem digital sendo representada como uma matriz de valores de pixel.	15
2.4	Exemplo de segmentação de imagem digital	16
2.5	Segmentação Semântica vs. Segmentação de Instância vs. Segmentação Panóptica	18
2.6	Segmentação dos núcleos das células de proteínas	19
2.7	Segmentação dos vasos da retina	19
2.8	Segmentação de tumor cerebral em IRM	19
2.9	Diagrama representando os subcampos da IA.	24
2.10	Classificação vs regressão	25
2.11	Representação de um neurônio biológico e um artificial	29
2.12	Várias funções de ativação	32
2.13	Diagrama representando uma rede neural profunda.	33
2.14	Diagrama representando uma Rede Neural Convolutacional.	34
2.15	Um exemplo de convolução entre uma entrada $\mathbf{I} 7 \times 7 \times 1$ e um <i>kernel</i> $\mathbf{K} 3 \times 3 \times 1$ com passo de 1.	35
2.16	Arquitetura básica da U-Net.	37
2.17	Arquitetura básica da U-Net.	38
3.1	Ilustração da TC de entrada	39

3.2	Exemplo de imagem e máscara da base de dados.	40
4.1	Comparação visual entre segmentação manual e via U-Net.	47
4.2	Interface inicial do aplicativo.	50
4.3	Resultados da segmentação da bexiga.	50

1 Introdução

1.1 Contextualização

A incidência de câncer vem crescendo consideravelmente no mundo inteiro. Na atualidade, o câncer é uma das principais causas de morte. Segundo o INCA (Instituto Nacional de Câncer), estima-se que o Brasil deverá registrar a ocorrência anual de 704 mil novos casos de câncer, ao longo do triênio 2023-2025, para cada ano [1], o que o coloca entre os países com maior incidência de câncer no mundo.

Atualmente, entre 30 e 50% dos cânceres podem ser evitados, por meio da redução de fatores de risco e da implementação de estratégias de prevenção baseadas em evidências existentes. A incidência do câncer também pode ser reduzida através da detecção precoce e do manejo de pacientes que desenvolvem câncer. Muitos cânceres têm uma alta chance de cura se diagnosticados precocemente e tratados adequadamente [2]. Uma das principais formas de tratamento é a radioterapia.

Estima-se que 60% dos pacientes oncológicos, em algum momento, necessitem de radioterapia com diferentes intenções terapêuticas [3]. Dependendo do tipo de doença e de seu estágio, pode-se realizar o tratamento com propósito de: cura; redução do volume do tumor; paliativo, minimizando alguns sintomas; ou preventivo, para evitar que células neoplásicas não detectadas, mas possivelmente presentes, se reproduzam [4].

1.2 Aplicações da Inteligência Artificial na Radioterapia

Com os recentes avanços científicos e tecnológicos, a utilização de Inteligência Artificial (IA) para aplicações na área médica e biomédica vem crescendo consideravelmente nos últimos anos. A IA pode auxiliar desde o diagnóstico de doenças (incluindo a Covid-19) ao seu tratamento, e até mesmo na realização de cirurgias [5], [6], [7], [8], [9].

Além da grande aplicabilidade, diferentes algoritmos de IA - como o Aprendizado Profundo (do inglês, *Deep Learning* - DL) - estão disponíveis e sendo implementados em processos de radioterapia e diagnóstico. Esses processos incluem o processamento e análise de imagens médicas, com aplicação no sistema de planejamento de tratamento de

paciente; na garantia de qualidade e na entrega de dose de radiação [10], [11], [12], [13].

O DL é uma classe de algoritmos do Aprendizado de Máquina (do inglês, *Machine Learning* - ML), caracterizada pelo uso de redes neurais com várias camadas de neurônios matemáticos capazes de processar dados, e reconhecimento de padrões visuais em imagens [14], [15]. Nos últimos anos, as abordagens de DL tiveram um grande impacto no campo do processamento e análise de imagens em geral, superando o estado da arte em muitas tarefas de reconhecimento visual [16], [17], [18].

Uma aplicação de DL está na segmentação não supervisionada de imagens médicas (i.e. sem requerer a intervenção de especialistas durante o processo de segmentação), pois a análise delas tornou-se uma ferramenta muito importante nas diferentes etapas do tratamento de doenças. Especialmente, ela desempenha um papel vital em imagens de Tomografia Computadorizada (TC), sendo a segmentação frequentemente a etapa inicial em um *pipeline* (sequência de etapas) de análise de imagens, pois é um componente crítico no diagnóstico e planejamento de tratamentos em radioterapia [17], [19].

Nessa especialidade, a qualidade do plano de tratamento afeta direta e significativamente os resultados do tratamento do paciente. O Sistema de Planejamento de Tratamento (SPT) desempenha um papel fundamental de garantir que as doses terapêuticas sejam calculadas e fracionadas adequadamente, de forma que a dose prescrita chegue a um volume tumoral (alvo de tratamento). Ao mesmo tempo, essa dosagem deve causar o menor dano possível nas células normais dos tecidos e dos órgãos sadios circundantes, assegurando assim que a dose aplicada será igual à prescrita pelo médico radioterapeuta [20], [21].

A segmentação de órgãos em risco em imagens de TC é uma tarefa crucial para o SPT. A segmentação manual é normalmente adotada na maioria das práticas clínicas. Contudo, essa abordagem é demorada, trabalhosa e desafiadora, principalmente devido às estruturas complexas dos órgãos, o que pode resultar em atrasos significativos no tratamento. Além disso, a segmentação manual muitas vezes é propensa a variações intra e interobservador. Por outro lado, a segmentação não supervisionada desses órgãos é útil na prática; os métodos automatizados buscam reduzir a carga de trabalho de delineamento, e unificar a definição do limite de cada órgão, fornecendo resultados eficientes e precisos, que aumenta a consistência do contorno [22], [23].

Um método para segmentação não supervisionada de imagens de TC é o uso de redes neurais profundas. Na radioterapia, as redes neurais podem processar grandes conjuntos de dados para fornecer soluções não supervisionadas, sem a necessidade de extração manual de recursos. Estudos recentes mostraram resultados promissores para contornos autosegmentados de órgãos de risco baseados em DL [24]. Comparado com os métodos tradicionais de aprendizado de máquina e visão computacional, o aprendizado

profundo tem certas vantagens em precisão e velocidade de segmentação. Portanto, o uso de aprendizado profundo para segmentar imagens médicas pode ser um auxílio para ajudar os médicos patologistas e especialistas [25].

1.3 Trabalhos Relacionados

Foi realizado em maio de 2022 uma revisão da literatura da aplicação de DL em imagens médicas. Nessa revisão, são abordados os estudos sobre esta temática, em tarefas que envolvem detecção e segmentação de órgãos na região pélvica em imagens de TC.

Este estudo foi realizado com base na estratégia de busca PICO (acrônimo para P: população/pacientes; I: intervenção; C: comparação/controle; O: desfecho/*outcome*) que é utilizado na prática baseada em evidências [26], na qual esta pesquisa consiste. Além disso, este estudo também foi conduzido de acordo com a recomendação PRISMA (do inglês, *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) [27]. Não houve um protocolo prévio para esta revisão.

Para este estudo, foram realizadas inicialmente pesquisas nas seguintes bases de dados bibliográficas: PubMed (Medline), Embase e Scopus. Ademais, foram utilizados descritores MeSH (*Medical Subject Headings*), DECs (Descritores em Ciências da Saúde) e Emtree (*Embase Thesaurus*), apenas para publicações internacionais de até 10 anos. As publicações foram selecionadas com as seguintes palavras-chave: "Câncer", "Radio-terapia", "Tomografia Computadorizada", "Aprendizado Profundo", "Pelve" e "Acurácia". Sendo a segmentação um assunto a ser melhor investigado após um amadurecimento maior do trabalho, ela não foi incluída como palavra-chave.

Com respeito aos descritores mencionados, o MeSH é um sistema de indexação de assuntos médicos e científicos desenvolvido pela *National Library of Medicine* (NLM) dos Estados Unidos. Ele é usado para categorizar e indexar artigos de revistas científicas, livros e outros recursos de informação biomédica [28]. Os descritores DECs são a versão em português do MeSH, desenvolvida pela Biblioteca Virtual em Saúde (BVS) da Organização Pan-Americana da Saúde (OPAS). Assim como o MeSH, os DECs são um vocabulário controlado usado para indexar e recuperar literatura científica na área da saúde, especialmente aquela publicada em português ou relacionada a países de língua portuguesa [29]. Finalmente o descritor Emtree é um vocabulário controlado desenvolvido pela Elsevier para indexação de artigos científicos na base de dados Embase. Semelhante ao MeSH, o Emtree é usado para categorizar e indexar a literatura biomédica, mas é específico para a base de dados Embase, que se concentra principalmente em literatura relacionada à prática baseada em evidências [30].

As pesquisas nas bases de dados foram realizadas usando as seguintes strings de busca:

- PubMed:
 - ("Radiotherapy Planning" OR "radiation therapy planning") AND ("computed tomography" OR "CT") AND ("Deep Learning" OR "organ Segmentation" OR "Autosegmentation" OR "auto-segmentation") AND ("Pelvis" OR "Pelvic Region") AND (Accuracy)
- Embase:
 - ('radiotherapy planning'/exp OR 'radiotherapy planning' OR 'radiation therapy planning') AND ('computed tomography'/exp OR 'computed tomography' OR 'ct'/exp OR 'ct') AND ('deep learning'/exp OR 'deep learning' OR 'organ segmentation' OR 'autosegmentation' OR 'auto-segmentation') AND ('pelvis'/exp OR 'pelvis' OR 'pelvic region') AND ('accuracy'/exp OR 'accuracy')
- Scopus:
 - (("Radiotherapy Planning" OR "radiation therapy planning") AND ("computed tomography" OR "CT") AND ("Deep Learning" OR "organ Segmentation" OR "Autosegmentation" OR "auto-segmentation") AND ("Pelvis" OR "Pelvic Region") AND ("Accuracy"))

Algo a destacar é que a segmentação não supervisionada é conhecida em inglês como *autosegmentation* ou *auto-segmentation*.

Após a obtenção dos artigos nas bases de dados, as duplicatas foram identificadas e removidas com o auxílio de uma ferramenta gerenciadora de referências. Os que sobraram foram selecionados para a elegibilidade, e em seguida para a triagem inicial com a análise dos títulos e resumos de cada estudo. Os artigos salvos nessa triagem foram então para a etapa de leitura dos textos completos. O passo seguinte foi aplicar os critérios de inclusão e exclusão ao longo da leitura de cada artigo.

- **Critérios de inclusão:** estudos publicados como artigos completos em inglês (a); emprego de métricas de avaliação relevantes (b); das aplicações dos modelos de aprendizado profundo (c); em segmentação de órgãos na região pélvica (d); em imagens de TC (e) na radioterapia (f).
- **Critérios de exclusão:** estudos que não são aplicados na radioterapia; estudos aplicados a imagens de ressonância magnética (visto que são diferentes das imagens de TC); segmentação de órgãos que não pertenciam à região pélvica, etc.

Com essas strings, as buscas encontraram um total de 37 estudos, sendo 9 na PubMed; 13 na Embase e 15 na Scopus. Uma vez que as duplicatas foram removidas, sobraram 17 estudos para a triagem de títulos e resumos. Após a triagem, verificou-se que cinco apresentaram resultados considerados satisfatórios, levando em conta sua relevância para a segmentação da bexiga por meio de imagens de TC e os critérios mencionados acima, e foram selecionados para leitura e análise dos textos completos. Por fim, foi concluído que os cinco estudos restantes eram considerados elegíveis para inclusão, enquanto que outros, que embora pudessem estar relacionados, seriam excluídos do trabalho por não seguirem o critério estabelecido.

A Figura 1.1 esquematiza o fluxograma PRISMA usado no trabalho geral para escolher estudos de pesquisa relevantes. Foi construída uma tabela (Tabela 1.1) detalhando as informações de cada um dos artigos selecionados, em especial as técnicas de DL usadas e as acurácias com os Coeficientes médios de Similaridades de Dados (DSC, do inglês *Data Similarity Coefficients*).

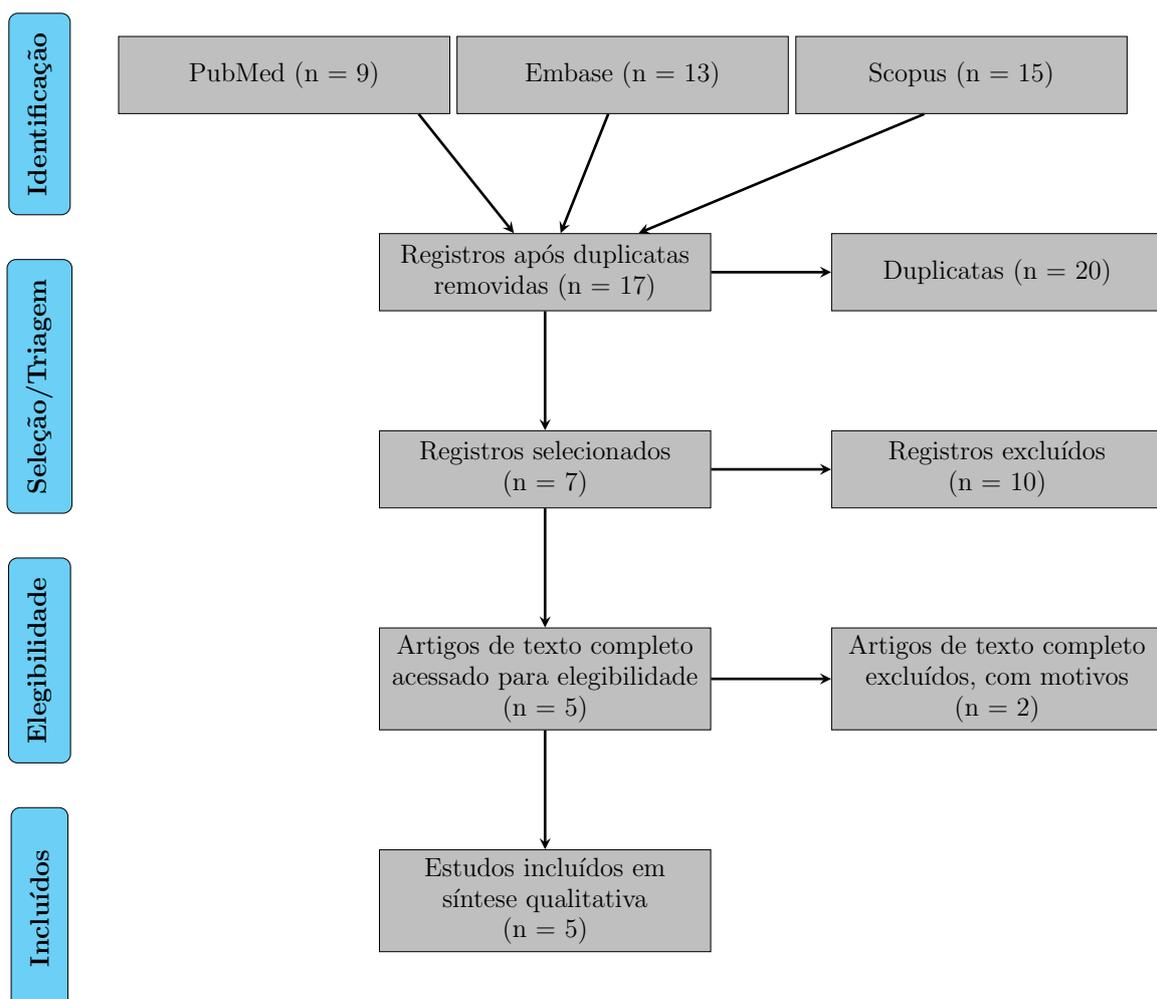


Figura 1.1. Diagrama de Fluxo PRISMA. Fonte: autor [31], adaptada de [27].

Tabela 1.1. Resumo das características descritivas de todos os estudos incluídos na revisão da literatura.

Autores	Títulos	Ano	Modelo de DL	Revista	Métricas
Rhee, Dong Joo et al [32]	<i>Automatic contouring system for cervical cancer using convolutional neural networks</i>	2020	3D V-Net	<i>Medical Physics</i>	Coef. de Dice de 0,89 para a bexiga, 0,81 para o reto, 0,94 para o fêmur esquerda, 0,93 para o fêmur direito, 0,93 para o osso pélvico.
Zhang, Zhuangzhuang et al [33]	<i>ARPM-Net: A novel CNN-based adversarial method with Markov random field enhancement for prostate and organs at risk segmentation in pelvic CT images</i>	2021	ARPM-Net	<i>Medical Physics</i>	Coef. de Dice na próstata, bexiga, reto, fêmur esquerdo e fêmur direito foram 0,88, 0,97, 0,86, 0,97 e 0,97, respectivamente.
Fu, Yabo et al [34]	<i>Pelvic multi-organ segmentation on cone-beam CT for prostate adaptive radiotherapy</i>	2020	CBCT-Net	<i>Medical Physics</i>	Coef. de Dice de 0,96, 0,91, 0,93, 0,95 e 0,95, para bexiga, próstata, reto, cabeças femorais esquerda e direita, respectivamente.
Kazemifar, Samaneh et al [35]	<i>Segmentation of the prostate and organs at risk in male pelvic CT images using deep learning</i>	2018	U-Net	<i>Biomedical Physics and Engineering Express</i>	Coef. de Dice de 0,88, 0,95 e 0,92 para a próstata, bexiga e reto, respectivamente.
Amjad A, Xu J, et al [36]	<i>General and custom deep learning autosegmentation models for organs in head and neck, abdomen, and male pelvis</i>	2022	ResUnet3D	<i>Medical Physics</i>	Coef. de Dice de 0,87, 0,95 e 0,89 para próstata, bexiga e reto, respectivamente.

Fonte: autoral [31].

Os autores Rhee, Dong Joo et al. [32] descreveram um sistema de contorno para o câncer cervical utilizando a arquitetura 3D V-Net. O objetivo do estudo foi melhorar o processo de contorno manual das estruturas anatômicas relevantes nas imagens de radioterapia do câncer cervical. O sistema proposto utilizou uma abordagem baseada em redes neurais convolucionais (do inglês *Convolutional Neural Network* - CNN), treinadas em um conjunto de dados de imagens pré-segmentadas para realizar o contorno das estruturas, como o colo do útero e os órgãos em risco. Os resultados do estudo demonstraram que o sistema de contorno alcançou uma precisão comparável ao contorno manual, reduzindo significativamente o tempo e a variabilidade interobservador.

Os autores Zhang, Zhuangzhuang et al. [33] propuseram um método inovador para a segmentação da próstata e de órgãos em risco em imagens de TC pélvica. O ARPM-Net utilizou CNNs treinadas em dados pré-segmentados e incorporou um componente adversário e o campo aleatório de Markov para melhorar a qualidade da segmentação. Os resultados demonstraram uma segmentação precisa e robusta, evidenciando o potencial dessa abordagem em aplicações clínicas relacionadas ao câncer de próstata e radioterapia pélvica.

Os autores Fu, Yabo et al. [34] descreveram um estudo para a segmentação de múltiplos órgãos pélvicos em imagens TC cone-beam utilizando a arquitetura CBCT-Net. O objetivo foi melhorar o planejamento e a precisão da radioterapia adaptativa da próstata. O estudo utilizou CNNs treinadas em um grande conjunto de dados de imagens pré-segmentadas. Essas técnicas de IA permitiram uma segmentação precisa dos órgãos relevantes, proporcionando uma abordagem eficiente e personalizada para o tratamento do câncer de próstata.

Kazemifar, Samaneh et al. [35] abordaram a segmentação da próstata e dos órgãos em risco em imagens de TC pélvica masculina utilizando a arquitetura U-Net. Os resultados obtidos demonstram que o modelo usado alcança uma precisão comparável à segmentação manual realizada por especialistas. Isso indica que a segmentação não supervisionada pode desempenhar um papel de grande importância no aprimoramento do planejamento de tratamentos de radioterapia para câncer de próstata, melhorando a eficiência e a precisão do processo de delineamento dos órgãos em risco.

Amjad A, Xu J, et al. [36] realizaram um estudo para desenvolver modelos de segmentação baseados na arquitetura ResUnet3D para órgãos nas regiões da cabeça e pescoço, abdômen e pelve masculina. Os pesquisadores utilizaram redes neurais convolucionais profundas treinadas em grandes conjuntos de dados de imagens médicas pré-segmentadas. Foram explorados modelos generalizados, treinados para segmentar órgãos nessas três regiões anatômicas, bem como modelos personalizados, adaptados para segmentar órgãos específicos dentro dessas regiões. Os resultados mostraram que tanto os modelos gene-

realizados quanto os personalizados alcançaram uma segmentação precisa dos órgãos de interesse, reduzindo a necessidade de contornos manuais e melhorando a eficiência do delineamento dos órgãos.

Os estudos realizados nos artigos selecionados demonstram consistentemente a eficácia do uso de modelos de aprendizado profundo na segmentação não supervisionada e precisa de órgãos em imagens médicas. Além disso, as abordagens propostas mostraram uma redução significativa na dependência de contornos manuais, o que pode melhorar a eficiência do planejamento de tratamentos, especialmente no contexto de cânceres cervical, de próstata e órgãos pélvicos.

1.3.1 Escolha da Arquitetura de Rede Neural Convolutacional

Como pode ser visto na revisão de literatura feita, o estado da arte em segmentação de órgão em imagens de TC utilizando modelos de IA gira em torno da arquitetura U-Net ou de uma variação dela ou uma similar a ela. Nesse contexto, foi realizada a busca pela arquitetura mais adequada para o presente projeto, seguida pela implementação da arquitetura escolhida.

Dentre tantas arquiteturas de redes neurais convolucionais disponíveis para a segmentação de imagens médicas, foi optado pela arquitetura U-Net para este trabalho [37]. Essa escolha foi fundamentada em sua eficácia comprovada, além de sua flexibilidade para adaptação a diferentes modalidades de imagem. Isso garante que o modelo possa ser facilmente ajustado para lidar com variações nas características das imagens, tornando-a uma opção robusta para a segmentação da bexiga em imagens de TC, conforme os critérios de inclusão estabelecidos na revisão conduzida.

Além disso, a U-Net tem sido amplamente adotada em diversas aplicações de segmentação médica devido à sua habilidade de lidar eficientemente com conjuntos de dados limitados, o que pode ser relevante em contextos clínicos onde a disponibilidade de grandes volumes de dados rotulados pode ser restrita. Sua capacidade de aprendizado eficaz com um número relativamente pequeno de amostras é uma vantagem significativa ao considerar a implementação em ambientes clínicos [38], [39], [40], [41], [42], [43], [44].

1.4 Justificativa

O objetivo da radioterapia é levar a dose de radiação prescrita ao alvo do tratamento e poupar ao máximo os tecidos saudáveis. Contudo, ao irradiar o alvo do tratamento, inevitavelmente células saudáveis também são irradiadas, podendo gerar algum tipo de toxicidade no paciente.

O preenchimento da bexiga em tratamentos na região pélvica pode influenciar a posição da próstata ou do útero. Manter a bexiga cheia é visto atualmente em todo o mundo como uma justificativa para melhor preservação do órgão em risco [45], pois mantê-la cheia pode potencialmente mover o intestino delgado para fora das regiões de tratamento de radiação, e resulta na diminuição da dose de radiação no intestino delgado e na toxicidade gastrointestinal. No entanto, podem ocorrer imprevisibilidades devido ao enchimento variável da bexiga durante o tratamento, em razão da má adesão dos pacientes e/ou de alterações anatômicas durante os tratamentos ou toxicidade [46], [47], [48].

Essa variação no enchimento da bexiga durante a radioterapia também leva ao deslocamento e modificação no tamanho/forma do alvo e dos órgãos em risco. Portanto, nesse tipo de tratamento, a reprodutibilidade do planejamento depende de um preenchimento da bexiga semelhante ao obtido na tomografia de planejamento [49], [50], [51], [52], [53].

Além disso, a dose de radiação absorvida pela bexiga e, portanto, a chance de toxicidade nesse órgão, também depende de seu preenchimento. Assim, a detecção e segmentação poderiam aumentar a precisão, assim como facilitar e acelerar a execução de tratamentos nos quais o preenchimento da bexiga tem relevância. A Figura 1.2 mostra as variações no formato da bexiga em cortes de TC de bexiga cheia (esquerda) e vazia (direita): (a), (b) visualização axial e (c), (d) visualização sagital [49], [50], [51], [54], [52].

A Figura 1.3 mostra a comparação da anatomia pélvica em dois pacientes separados entre o dia do planejamento de tratamento/simulação (tomografia computadorizada de quilovoltagem [TC]) e o dia do tratamento (TC de megavoltagem [MVCT]). Essas tomografias computadorizadas mostram a deformação da próstata, do reto e da bexiga, que resultaria em variação nas doses administradas [55].

O que também pode-se observar dessas imagens é que geralmente são de menor qualidade, dificultando a identificação manual de regiões para segmentação. Isso porque as imagens utilizadas neste trabalho se tratam de imagens Cone-Beam TC de megavoltagem, que apresentam um grau significativo de ruído, quando comparadas com imagens convencionais de quilovoltagem. Os resultados deste trabalho podem ser úteis para auxiliar um especialista a definir com maior segurança as regiões do órgão a ser irradiado com base em imagens de baixa qualidade.

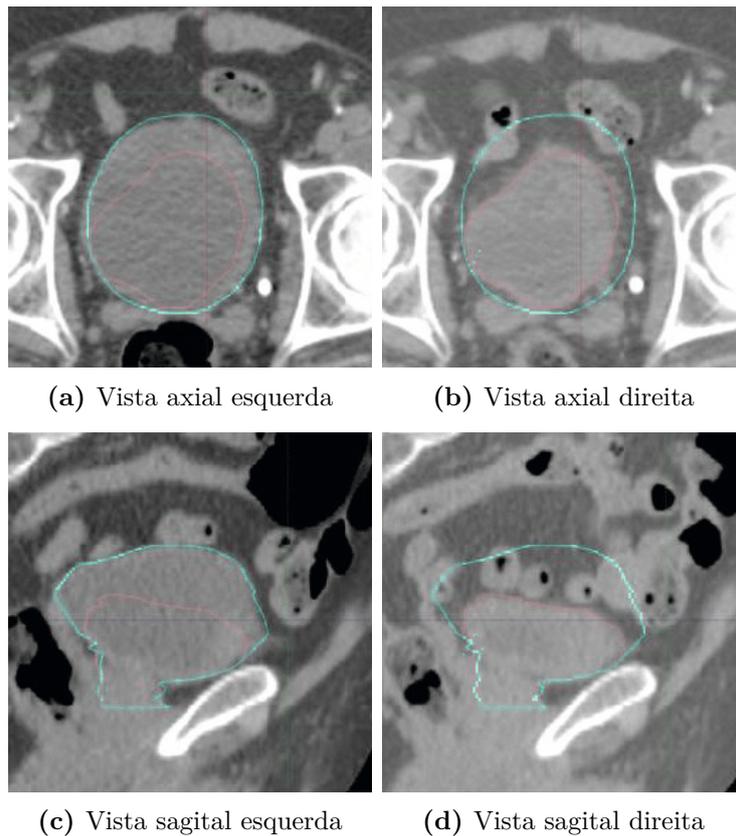


Figura 1.2. Amostra de cortes de TC da bexiga. Fonte: adaptada de [54].

1.5 Objetivos

1.5.1 Objetivo Geral

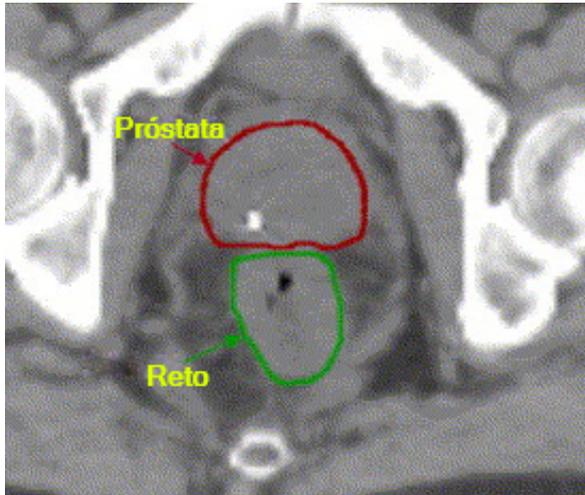
O principal objetivo deste trabalho consiste em implementar e desenvolver uma solução que auxilie e facilite os procedimentos de segmentação de órgãos localizados na região pélvica - estudo de caso: bexiga - em imagens de Tomografia Computadorizada (TC) de planejamento de tratamento na radioterapia.

1.5.2 Objetivos Específicos

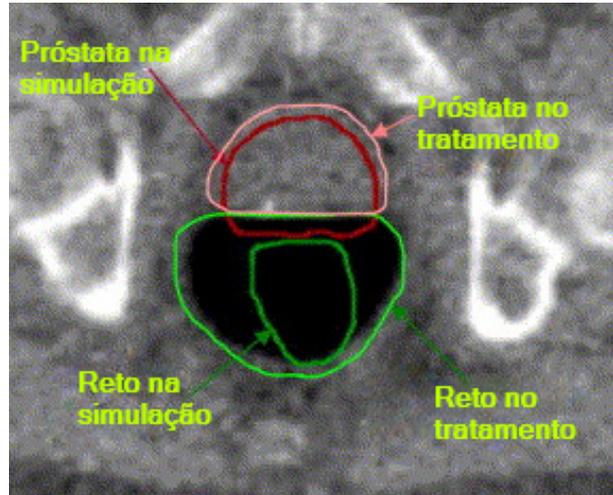
- Implementação da arquitetura U-Net;
- Desenvolvimento de um aplicativo web que realizará a segmentação automática.

1.6 Estrutura da Dissertação

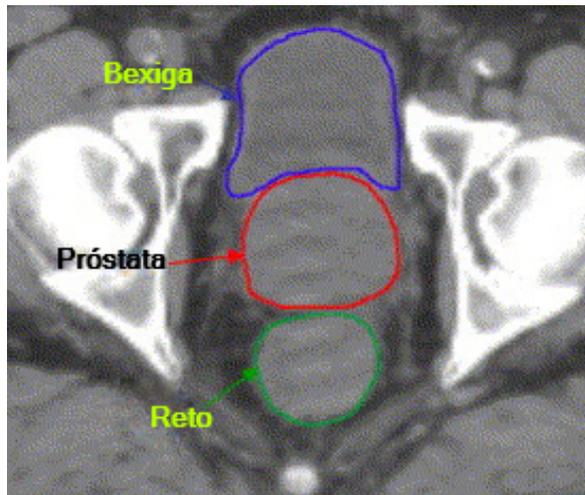
O texto desta dissertação está organizado e dividido em cinco partes. Além desta introdução, a estrutura do presente trabalho é constituída dos seguintes capítulos:



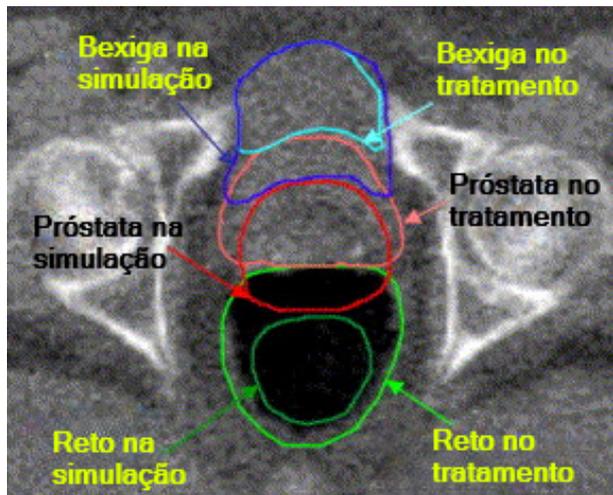
(a) Paciente #1. Simulação



(b) Paciente #1. Dia do tratamento



(c) Paciente #2. Simulação



(d) Paciente #2. Dia do tratamento

Figura 1.3. Comparação da anatomia pélvica em dois pacientes separados entre o dia do planejamento de tratamento/simulação e o dia do tratamento. Fonte: adaptada de [52].

- **2 Fundamentação Teórica:** Nesta seção, são abordados as principais teorias e conceitos relacionados ao tópico de estudo, que são: Visão Computacional, Processamento de Imagens, Segmentação de Imagens e Inteligência Artificial, estabelecendo a base teórica necessária para a compreensão das questões abordadas nesta dissertação;
- **3 Metodologia:** É apresentado em detalhes como a pesquisa foi conduzida. Isso inclui informações sobre os dados utilizados, detalhes da implementação, ferramentas utilizadas, o processo de coleta de imagens médicas, bem como o pré-processamento, treinamento, teste, validação do modelo e das métricas de avaliação. Além disso, o aplicativo desenvolvido é discutido;
- **4 Resultados e Discussões:** São apresentados os resultados alcançados pelo modelo proposto por meio de gráficos e tabelas. Em seguida, é realizada a análise e interpretação dos resultados à luz da fundamentação teórica estabelecida no Capítulo 2;
- **5 Conclusão:** É apresentada uma síntese dos principais resultados obtidos ao longo da pesquisa. São destacadas as contribuições deste estudo para o avanço do conhecimento no campo. Além disso, são discutidas as implicações práticas e teóricas dos resultados e possíveis direções para pesquisas futuras.

2 Fundamentação Teórica

2.1 Visão Computacional

A Visão Computacional (VC) é um campo científico interdisciplinar que permite computadores obterem informações de imagens digitais, vídeos e outras entradas visuais e tomem ações ou façam recomendações com base nessas informações. O objetivo da visão computacional não é apenas ver, mas também processar e fornecer resultados úteis com base na observação. Os conhecimentos obtidos pela visão computacional são então usados para automatizar tarefas ou fazer recomendações. Se a IA permite que os computadores pensem, a visão computacional permite que eles vejam, observem e compreendam [56], [57], [58].

A Figura 2.1 esquematiza vários campos de aplicações da visão computacional, que incluem reconhecimento de padrões, processamento de imagens e IA. Esses campos podem ser considerados fortemente relacionados porque, no geral, avanços em um desses campos podem levar a avanços em outros campos também [59], [60], [61].

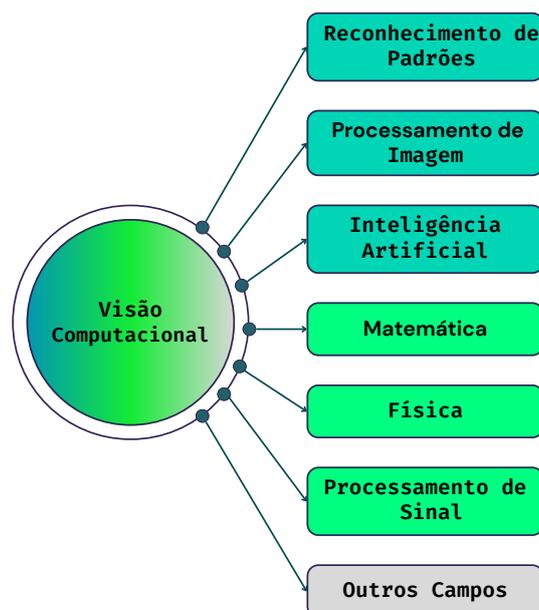


Figura 2.1. Esquema dos principais campos da VC. Fonte: autor [31].

2.2 Processamento de Imagens

Uma imagem é uma representação visual de algo (como um objeto ou entidade). Matematicamente ela pode ser definida como uma função $f(x, y)$, onde f está relacionada com a intensidade (brilho, cor ou nível de cinza) na posição (x, y) , sendo x e y respectivamente as coordenadas espaciais horizontal e vertical. O valor de f em qualquer ponto (x, y) é o valor do pixel (elemento de imagem, quadrado) naquele ponto da imagem. O pixel é a menor unidade de informação em uma imagem e possui uma quantidade finita e discreta de valor numérico, que representa sua intensidade ou nível de cinza [62].

Para imagens em tons de cinza, cada pixel tem uma intensidade que varia em números inteiros de 8 bits, fornecendo 256 tons de cinza possíveis, entre 0 e 255, sendo 0 a cor preta, a intensidade mínima, e 255 a cor branca, de intensidade máxima, e os valores intermediários são tons de cinza. A Figura 2.2 esquematiza esses conceitos [63], [64].

Já em imagens coloridas, $f(x, y)$ agora é um vetor com três valores inteiros em vez de um (3 camadas de matrizes bidimensionais) entre 0 e 255. Numa imagem na escala RGB, as cores são construídas a partir de uma combinação de vermelho (*Red*, em inglês), verde (*Green*, em inglês) e azul (*Blue*, em inglês) que a compõe [63].

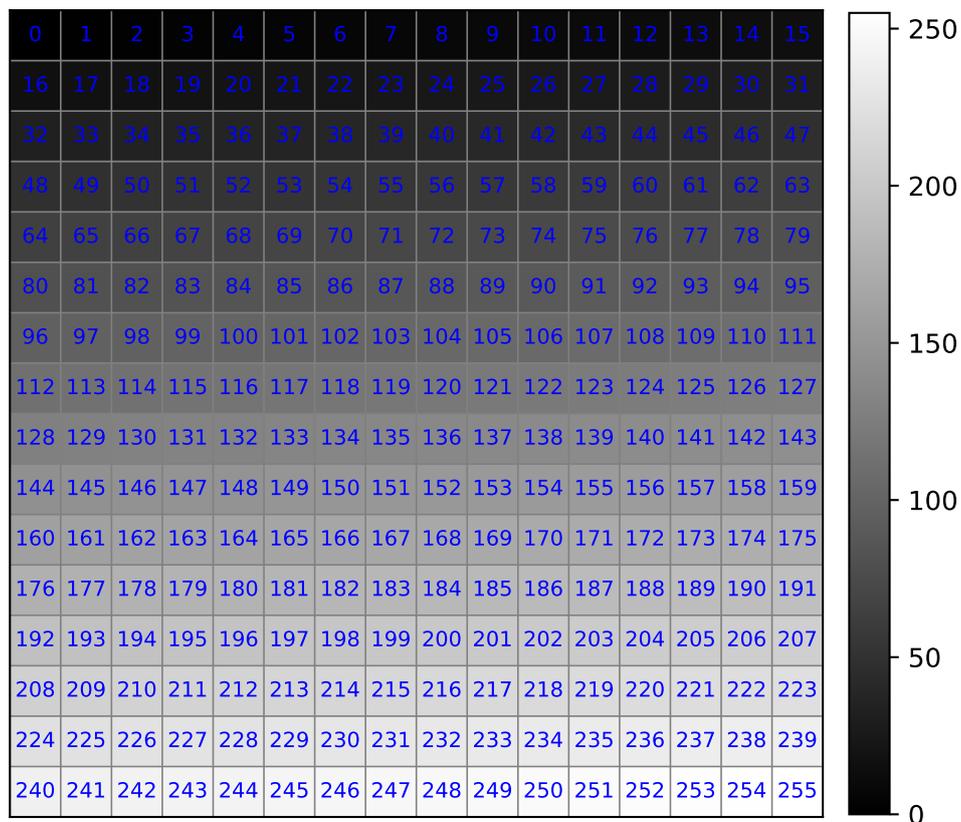


Figura 2.2. Matriz com as diferentes tonalidades dos pixels na escala de cinza. Fonte: autor [31].

Computacionalmente, a função $f(x, y)$ descreve um vetor ou uma matriz ($n \times m$) de pixels dispostos em n-linhas e m-colunas organizados bidimensionalmente (2D). Na região onde foi aplicado um *zoom* na primeira imagem da Figura 2.3, pode-se ver que a imagem é pixelada. Os números que vemos na imagem seguinte correspondem à intensidade de cada pixel da imagem, e é como o computador interpreta. Assim, uma imagem pode ser tratada como uma matriz de intensidades [62], [63], [64].

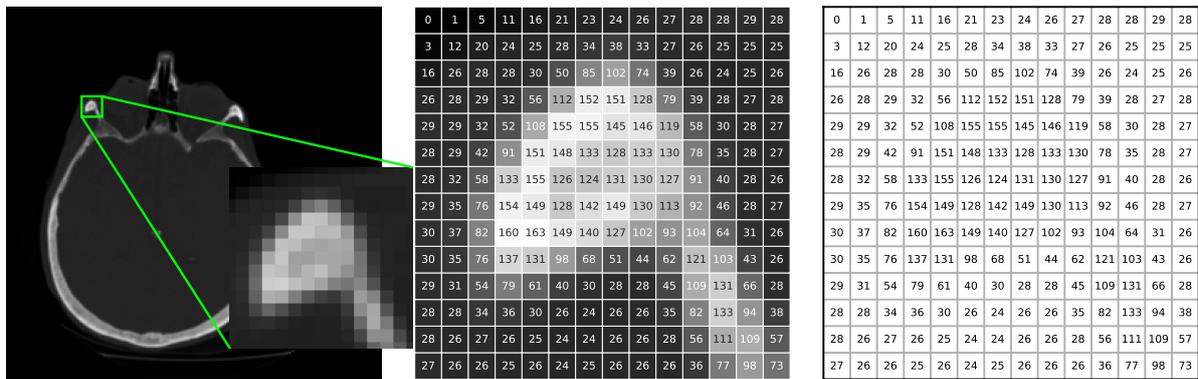


Figura 2.3. Uma imagem digital sendo representada como uma matriz de valores de pixel. Fonte: autor [31].

O processamento de imagens pode ser entendido como um conjunto de técnicas utilizado para modificar uma imagem digital, realizando certas operações nela por meio de algoritmos computacionais, seja para obter uma imagem aprimorada (em termos de qualidade), reduzir o seu tamanho (em termos de compressão de bits) ou extrair dela alguma informação útil. É um tipo de processamento de sinal cuja entrada é uma imagem, e sua saída podendo ser outra imagem ou características associadas a ela mesma [62], [63], [64].

O processamento de imagens é importante por vários motivos. Em particular, ele pode ser usado para melhorar a qualidade das imagens, torná-las mais legíveis e permitir que computadores as entendam melhor, como mencionado acima. O processamento de imagens é um subconjunto da VC, e também é uma parte fundamental da IA. Ao compreender e processar imagens, determinadas técnicas de IA podem aprender a reconhecer objetos e identificar padrões. Isso as ajuda a tomar melhores decisões e realizar tarefas com maior eficiência. Uma importante aplicação é em aprendizado profundo (DL), onde esse pré-processamento de imagem pode aumentar significativamente o desempenho de um determinado modelo. Na segmentação de imagens, esse modelo pode ser uma importante ferramenta, em especial no processamento de imagens médicas, sendo útil em diagnósticos e tratamentos, que é o foco deste trabalho [62], [63], [64].

2.3 Segmentação de Imagens

A segmentação de imagens é o processo no qual uma imagem digital é dividida em regiões ou subgrupos de pixels com atributos semelhantes (como cor, intensidade ou textura). Podendo assim separar regiões específicas (com informações importantes) do restante das regiões irrelevantes (i.e. que não contêm informação útil). Dessa forma, é possível atribuir rótulos aos pixels e classificar outros pixels de acordo com esses rótulos. Ao fazer isso, os pixels pertencentes a uma mesma categoria têm um rótulo comum atribuído a eles [62], [65], [66]. Esse processo pode reduzir a complexidade da imagem e, assim, simplificar a análise. A Figura 2.4 ilustra a segmentação de uma imagem, aplicando uma técnica de detecção de bordas para identificar as bordas ou transições abruptas de intensidade na imagem, que correspondem às fronteiras entre objetos ou regiões distintas.

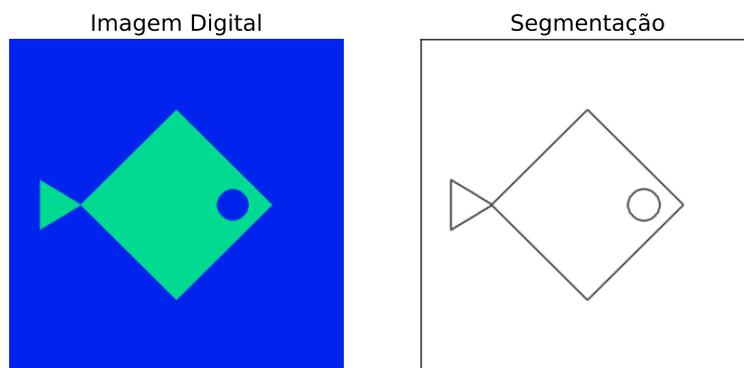


Figura 2.4. Exemplo de segmentação de imagem digital. Fonte: autor [31].

A segmentação de imagens é diferente da classificação. A classificação mostra apenas a qual classe ou classes pertence toda a imagem, enquanto a segmentação precisa identificar as informações de cada pixel da imagem [25].

Existe uma ampla variedade de técnicas de segmentação de imagens, mas não existe uma técnica padrão que possa produzir resultados satisfatórios para todas as aplicações de imagem [67]. Algoritmos de segmentação geralmente são baseados em duas categorias básicas em termos de valores na escala cinza: ou considera-se a semelhança entre os níveis de cinza (similaridade) ou entre as suas diferenças (descontinuidade) [62].

As principais abordagens usadas na primeira categoria são segmentações baseadas em regiões, que utilizam informações de intensidade de pixels para identificar regiões de imagens que possuem valores de intensidade semelhantes. Alguns exemplos de algoritmos de segmentação baseados em similaridade incluem os métodos baseados em limiarização, como o método de *Otsu*; técnicas de segmentação baseadas em região, como o método de crescimento de regiões; e técnicas de agrupamento (*clustering*), como o *k-means*, onde os pixels são agrupados em *clusters* com base em suas intensidades [62].

Já a segmentação baseada na segunda categoria utiliza informações de bordas e mu-

danças abruptas nos valores de intensidade dos pixels para identificar regiões diferentes na imagem. Alguns exemplos de algoritmos de segmentação baseados em descontinuidade incluem a detecção de linhas, como a transformada de *Hough*; técnicas de segmentação baseadas em contornos, como a técnica de *watershed*; e a detecção de bordas como o *canny*, que identifica pontos de transição entre valores de intensidade [62].

Ambas as categorias têm seus pontos fortes e fracos e podem ser usadas em diferentes situações, dependendo da imagem que está sendo segmentada e do resultado desejado. Por exemplo, a segmentação baseada em similaridade pode ser útil para segmentar objetos em uma imagem com fundo uniforme, enquanto a segmentação baseada em descontinuidade pode ser útil para segmentar objetos com bordas distintas em um fundo complexo. Além disso, é comum que os algoritmos de segmentação combinem as duas abordagens para obter resultados melhores, como em algoritmos de segmentação híbrida [62].

O processo de segmentação de imagem é dividido em segmentação semântica, de instância e panóptica, de acordo com as diferentes granularidades grossas e finas da segmentação. A segmentação de imagens médicas é considerada uma tarefa de segmentação semântica [66].

A segmentação semântica desempenha o processo de atribuir um rótulo categórico a cada pixel da imagem, e ela se refere à classificação de pixels em classes “semânticas”. Ou seja, os pixels pertencentes a uma determinada classe são simplesmente atribuídos para essa classe, e os pixels pertencentes a classes diferentes são separados (segregados), pois nenhuma outra informação ou contexto é levado em consideração. Por outro lado, a segmentação de instância pode segregar regiões contendo pixels da mesma classe enquanto ignora outras classes. Ou seja, esse processo localiza e delinea distintamente cada objeto de interesse em uma imagem, classificando os pixels em categorias com base em “instâncias” em vez de classes [68]. Já a segmentação panóptica - a tarefa de segmentação desenvolvida mais recentemente - pode ser expressa como a combinação de segmentação semântica e segmentação de instância. Ela dá um rótulo categórico exclusivo para as classes de pixels, assim como suas instâncias. Dessa forma, cada pixel é codificado com dois valores, a saber, o rótulo da classe a que pertence e o número da instância [68].

A Figura 2.5 ilustra a aplicação dos três processos de segmentação para segmentar a imagem fornecida. É possível observar que enquanto a segmentação semântica segmenta um amplo limite de objetos pertencentes a uma classe específica, a segmentação de instância fornece um mapa de segmento para cada objeto visualizado na imagem, sem qualquer ideia da classe à qual o objeto pertence. Por fim, a segmentação panóptica pode ser vista como a conjugação de tarefas de segmentação de instância e semântica. A segmentação panóptica fornece os mapas segmentados de todos os objetos de qualquer classe específica presentes na imagem [69].

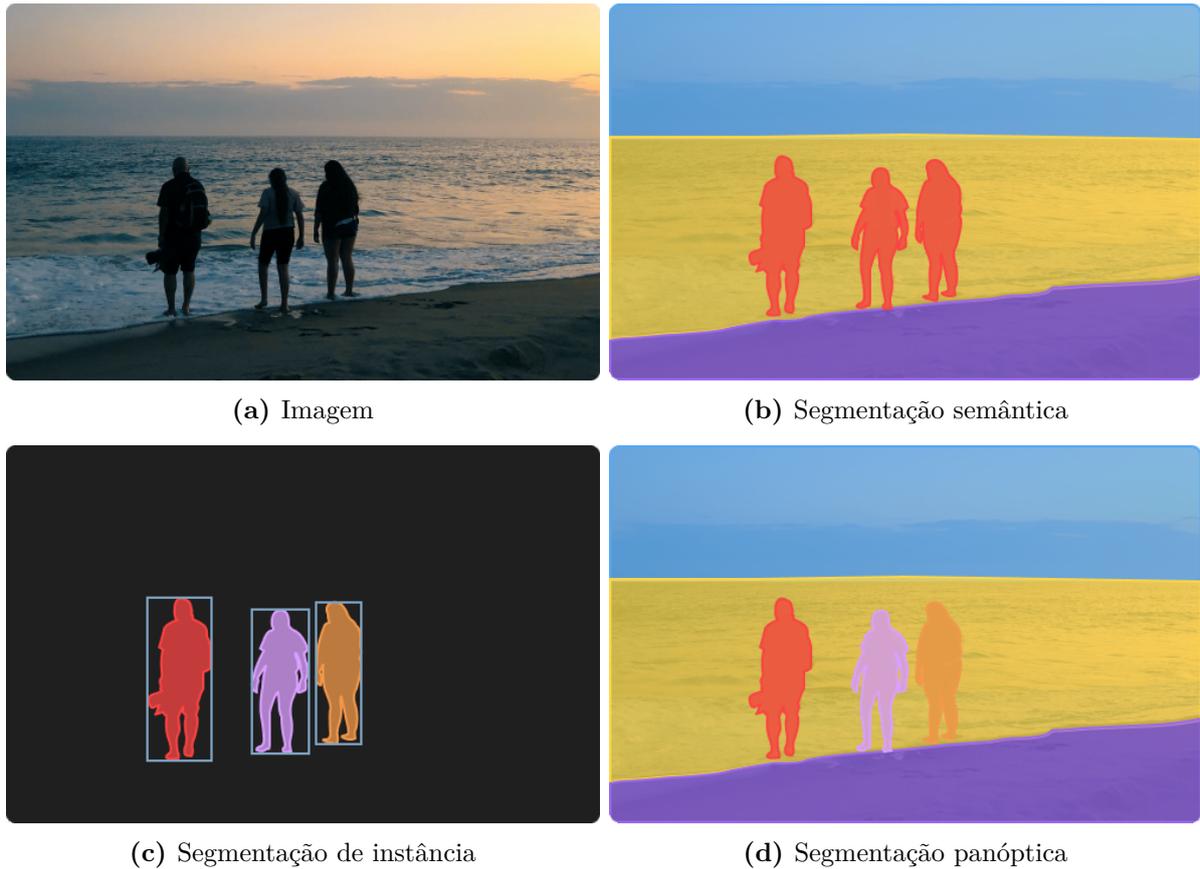


Figura 2.5. Segmentação Semântica vs. Segmentação de Instância vs. Segmentação Panóptica. Fonte: adaptada de [69].

2.3.1 Segmentação de Imagens Médicas

A imagiologia médica oferece informações fundamentais sobre a anatomia e função dos órgãos, bem como visa detectar estados de doenças por meio de diagnósticos, e auxiliar nos tratamentos [70].

A segmentação de imagens desempenha um papel fundamental em muitas aplicações de imagens médicas, facilitando o delineamento de estruturas anatômicas (o que permite uma análise mais precisa dos dados anatômicos, isolando apenas as áreas necessárias) e outras regiões de interesse.

As Figuras 2.6, 2.7 e 2.8 ilustram alguns exemplos de segmentação de imagens médicas em diferentes áreas e aplicações. A segmentação dos núcleos de células de proteínas humanas em imagens microscópicas é usada em várias aplicações em biologia celular e medicina, como na análise e interpretação dessas imagens, permitindo a quantificação da expressão gênica, além da análise da distribuição de proteínas nos tecidos e a identificação de padrões de expressão relacionados a diferentes condições biológicas [71]. Da mesma forma, a segmentação dos vasos da retina é usada no diagnóstico e tratamento de várias doenças oculares, como retinopatia diabética, a degeneração macular relacionada à idade

e glaucoma [72], [73]. Já a segmentação de tumores cerebrais a partir de imagens de ressonância magnética (IRM) é essencial para o diagnóstico, tratamento e monitoramento de pacientes com doenças neurológicas. Ela permite a identificação precisa dos tumores, o planejamento do tratamento e a avaliação da eficácia das terapias [74].

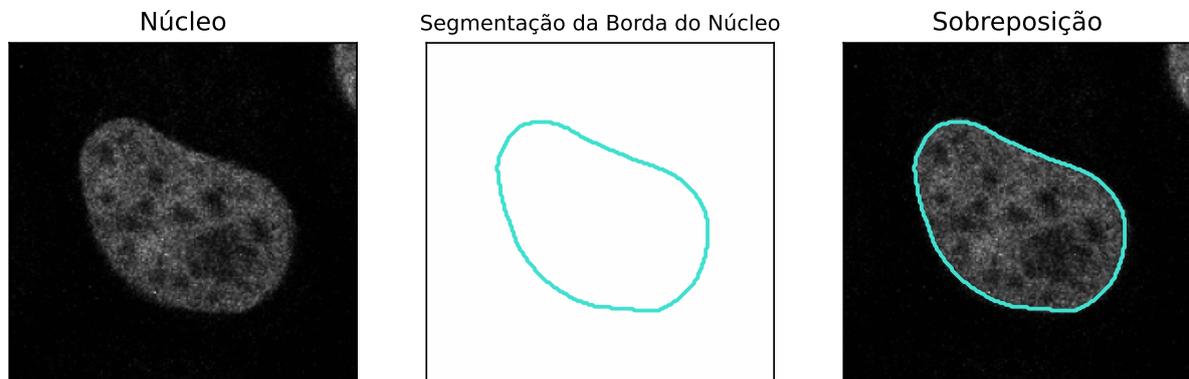


Figura 2.6. Segmentação dos núcleos das células. Fonte: [31], adaptada de [71].



Figura 2.7. Segmentação dos vasos da retina. Fonte: [31], adaptada de [72].

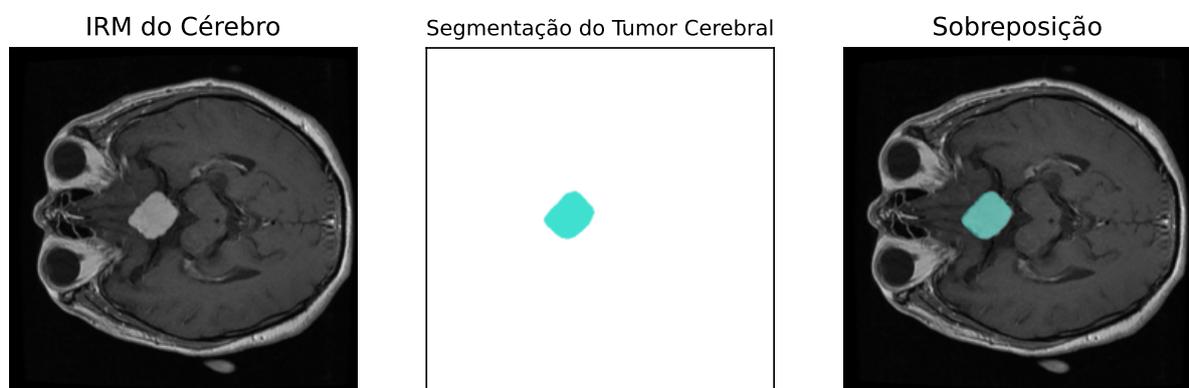


Figura 2.8. Segmentação de tumor cerebral em IRM. Fonte: [31], adaptada de [74].

Ao trabalharmos com imagens médicas, o objetivo da segmentação se baseia principalmente em [75]:

- Quantificar volumes de tecido;
- Auxiliar no diagnóstico;
- Compreender e estudar a estrutura anatômica;
- Identificar a região de interesse (ROI, do inglês, *Region Of Interest*), ou seja, localizar anomalias, como uma lesão ou um tumor;
- Medir o volume do tecido para quantificar o crescimento do tumor ou sua diminuição pós-tratamento;
- Auxiliar no planejamento de tratamento por radioterapia, como por exemplo no cálculo da dose de radiação;
- Auxiliar em cirurgia assistida por computador.

As diferentes técnicas ou abordagens para realizar a segmentação de imagens médicas podem ser classificadas em três categorias, sendo elas a segmentação manual, semiautomática e automática. Cada abordagem tem suas vantagens e desvantagens, e a sua escolha depende do tipo de imagem e do objetivo do diagnóstico.

2.3.1.1 Método de Segmentação Manual

A segmentação manual de imagens é uma parte rotineira dos fluxos de trabalho de radioterapia. Esse método de segmentação requer que os especialistas determinem primeiro a ROI e, em seguida, desenhem manualmente limites precisos ao redor da ROI (delineando regiões como volumes tumorais e órgãos em risco) anotando corretamente cada um dos pixels da imagem de TC. A precisão e reprodutibilidade das ROIs é um aspecto importante dos protocolos de controle de qualidade [68], [76]. No caso da radioterapia, os especialistas são radioterapeutas que têm conhecimento anatômico e fisiológico adquirido por meio de treinamento e experiência.

O método de segmentação manual é necessário pois fornece as imagens rotuladas para o treinamento adicional dos métodos de segmentação semiautomática e automática, como também para verificar a qualidade dos resultados desses métodos. Contudo, este método é demorado e propenso a variabilidade inter e intra-observador, onde diferentes especialistas podem segmentar a mesma imagem de forma diferente [68], [76], [77], [78].

2.3.1.2 Método de Segmentação Semiautomática

A segmentação semiautomática tenta resolver alguns dos problemas relacionados à segmentação manual. Os métodos de segmentação semiautomática envolvem a interação do usuário com algoritmos automatizados para produzir resultados de segmentação com maior precisão. Nessa abordagem, o radioterapeuta define um conjunto inicial de regiões ou objetos de interesse na imagem e, em seguida, o software utiliza algoritmos para segmentar o restante da imagem. Essa interação do usuário requer três propósitos principais: inicialização, intervenção ou *feedback*, resposta e avaliação [68], [77], [79].

A etapa de inicialização pode envolver a verificação, seleção ou edição manual dos limites da ROI inicial aproximada (para reduzir o erro de segmentação) que é usada posteriormente para segmentar toda a região em questão. Além da inicialização, os algoritmos automatizados podem ser direcionados para um resultado desejado durante o processo, recebendo *feedbacks* e fornecendo ajustes em resposta. Além disso, o usuário pode avaliar os resultados e modificar ou repetir o processo se não estiver satisfeito. A segmentação semiautomática pode reduzir o tempo de delineamento ou contorno, assim como a variabilidade intra-observador [78], [79], [80].

2.3.1.3 Método de Segmentação Automática

A segmentação automática é a abordagem mais avançada e usa algoritmos de IA para realizar a segmentação de imagens médicas sem intervenção ou interação do usuário. Esses algoritmos podem ser treinados usando uma grande quantidade de dados e podem identificar com precisão as regiões ou objetos de interesse na imagem. A segmentação automática é rápida, eficiente e pode ser realizada em tempo real, permitindo delineamentos mais precisos e rápidos e pode aumentar a consistência do contorno entre os radioterapeutas, reduzindo ainda mais a variabilidade intra-observador e inter-observador [68], [78], [79], [80].

Esses algoritmos de IA podem ser divididos em duas categorias principais: algoritmos baseados em aprendizado supervisionado e baseados em aprendizado não supervisionado. As técnicas baseadas em aprendizado supervisionado envolvem o treinamento de um modelo de IA usando um conjunto de dados rotulados. Nesse processo, as imagens são pré-processadas e transformadas em recursos que são usados pelo modelo para aprender a segmentar as áreas de interesse. O modelo é treinado para mapear as características das imagens para os rótulos correspondentes, o que permite a segmentação automática de imagens novas e não rotuladas [78], [79], [80].

As técnicas baseadas em aprendizado não supervisionado não usam um conjunto de dados rotulados. Em vez disso, essas técnicas exploram as características das ima-

gens para segmentar as áreas de interesse. Essas técnicas podem incluir métodos de agrupamento, como o *k-means*, ou técnicas baseadas em redes neurais, como as redes autocodificadoras [78], [79], [80].

Os passos básicos para realizar a segmentação automática de imagens médicas na radioterapia incluem a aquisição de dados, pré-processamento, seleção de método de segmentação, treinamento do modelo, validação do modelo, aplicação do modelo (segmentação), pós-processamento e verificação. Cada etapa requer cuidadosa consideração e seleção das técnicas apropriadas para obter uma segmentação precisa e confiável. Abaixo são detalhados cada um destes passos [81], [82], [25], [83], [84].

- **Aquisição de imagens médicas:** o primeiro passo na segmentação automática de imagens médicas na radioterapia é a aquisição das imagens médicas em questão. Isso pode ser feito por meio de técnicas como TC, IRM ou radiografia;
- **Anotação dos dados:** em seguida, as imagens médicas devem ser anotadas por especialistas para identificar as regiões de interesse que devem ser segmentadas. Essas anotações são feitas manualmente usando ferramentas de anotação;
- **Pré-processamento de imagens:** este é o passo seguinte, onde os dados anotados devem ser preparados para treinamento do modelo de IA, i.e., a imagem será pré-processada antes de ser segmentada. O pré-processamento é realizado para melhorar a qualidade das imagens e remover artefatos que possam afetar a precisão da segmentação, por meio de processos como a remoção de ruído ou correção de distorção;
- **Seleção do método de segmentação:** nesta etapa, o método de segmentação mais adequado é escolhido com base na imagem médica e nas características da região a ser segmentada;
- **Divisão dos dados em conjuntos de treinamento, validação e teste:** o conjunto de treinamento é usado para treinar o modelo, enquanto que o conjunto de validação é usado para ajustar os hiperparâmetros do modelo e evitar o sobreajuste (*overfitting*), e o conjunto de teste é usado para avaliar o desempenho do modelo após o treinamento;
- **Treinamento do modelo de IA:** o modelo de IA é treinado em um grande conjunto de imagens médicas rotuladas manualmente, para que possa aprender a identificar e separar diferentes estruturas anatômicas. O treinamento é um processo iterativo em que o modelo é ajustado até que seja capaz de segmentar as imagens com alta precisão. Durante o treinamento, os pesos dos neurônios do modelo são atualizados por meio da minimização de uma função de perda que mede a diferença

entre as máscaras segmentadas geradas pelo modelo e as máscaras reais rotuladas pelos especialistas;

- **Validação do modelo:** uma vez que o modelo é treinado, ele é validado em um conjunto de dados de teste para avaliar sua precisão e desempenho na segmentação de novas imagens. Para isso, é usada uma base de imagens no conjunto de dados de teste que são diferentes daquelas usadas para o treinamento. Isso ajuda a avaliar a capacidade do modelo de generalizar para novos dados e detectar problemas de *overfitting*. A precisão do modelo é avaliada em comparação com as segmentações manuais de referência;
- **Aplicação do modelo:** esta é a etapa da segmentação automática, onde o modelo treinado e validado é aplicado em novas imagens médicas para segmentar as estruturas anatômicas de interesse. O resultado da segmentação é então revisado e refinado, se necessário, por um especialista em radioterapia;
- **Pós-processamento:** após a segmentação automática, é realizado o pós-processamento nas regiões segmentadas para remover quaisquer artefatos ou ruídos remanescentes e melhorar a precisão da segmentação;
- **Avaliação da segmentação:** finalmente, é importante avaliar a precisão da segmentação automática e comparar com a segmentação manual. A avaliação pode ser realizada usando várias métricas de desempenho, como a sensibilidade, especificidade, precisão, coeficiente de correlação, índice de Jaccard e coeficiente de Dice.

É importante notar que a qualidade da segmentação automática depende da qualidade dos dados de treinamento e da escolha adequada das características da imagem. Além disso, a validação e o refinamento dos resultados da segmentação são críticos para garantir a precisão e a confiabilidade da técnica.

2.4 Inteligência Artificial

A IA pode ser definida como a ciência e engenharia de criar agentes inteligentes, ou seja, sistemas que são capazes de perceber seu ambiente e tomar decisões para maximizar suas chances de atingir tarefas ou objetivos pré-determinados [85]. Esses sistemas podem ser tanto máquinas quanto softwares, e envolve o uso de técnicas computacionais para automatizar tarefas que, de outra forma, iriam requerer inteligência humana para serem executadas, como reconhecimento de fala, visão computacional, tomada de decisão e aprendizado automático [86].

A IA é o campo que abrange tanto o aprendizado de máquina quanto o aprendizado profundo, sendo o primeiro um subcampo da IA e o segundo um subcampo do primeiro. A Figura 2.9 representa esse esquema.

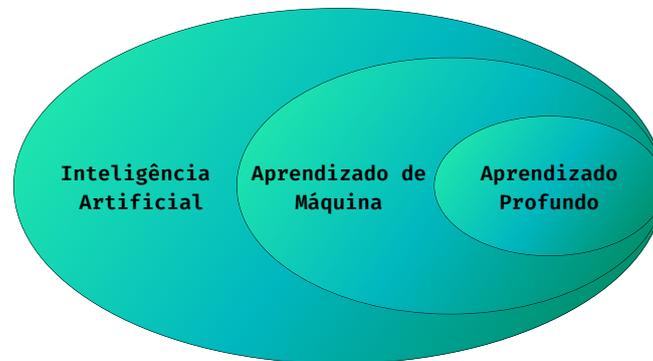


Figura 2.9. Diagrama representando os subcampos da IA. Fonte: autor [31], adaptada de [87].

2.4.1 Aprendizado de Máquina

O aprendizado de máquina (ML) é um ramo da IA que permite que as máquinas aprendam e melhorem suas performances em tarefas específicas a partir da experiência e dos dados [88]. Em outras palavras, é a capacidade de construir algoritmos que podem analisar e aprender com os dados, sem terem códigos especializados escritos manualmente para esse propósito. O objetivo principal do ML é criar modelos que possam prever resultados ou comportamentos futuros com base em dados históricos [89].

O processo de aprendizado envolve a identificação de padrões e relações entre variáveis nos dados, e a utilização dessas informações para fazer previsões sobre novos dados [90]. Uma das principais características desse processo é a capacidade de melhorar o desempenho em uma tarefa ao longo do tempo, utilizando modelos estatísticos que são ajustados automaticamente à medida que o sistema é exposto a mais dados [90], [91], [87].

No entanto, é importante destacar que o sucesso do aprendizado de máquina depende da qualidade dos dados utilizados para treinar os modelos [88], [92]. Dessa forma, é fundamental garantir a transparência e justiça dos sistemas, a fim de evitar a reprodução de preconceitos e discriminações presentes nos dados utilizados para o treinamento [88], [92].

Existem três categorias principais nas quais os modelos de ML se enquadram: Aprendizado Supervisionado, Aprendizado Não Supervisionado e Aprendizado por Reforço. O Aprendizado Supervisionado é usado quando os dados de treinamento incluem exemplos rotulados. O Aprendizado Não Supervisionado é usado quando os dados de treinamento não possuem rótulos. Já o Aprendizado por Reforço é usado para ensinar uma máquina a tomar decisões em um ambiente dinâmico, com base nas recompensas e penalidades que recebe [87].

2.4.1.1 Aprendizado Supervisionado

O aprendizado supervisionado é uma das técnicas mais utilizadas em ML. Nesse método, o modelo é treinado com exemplos rotulados com o objetivo de aprender a função que mapeia os dados de entrada para uma saída desejada, baseada nos rótulos fornecidos durante o treinamento [90], [93]. Esse tipo de aprendizado é muito utilizado em problemas de classificação e regressão, onde o objetivo é prever a classe ou valor de uma variável de saída com base em um conjunto de variáveis de entrada:

- **Classificação:** referem-se a algoritmos que abordam problemas de classificação em que a variável de saída é categórica. Por exemplo, sim ou não, verdadeiro ou falso, masculino ou feminino, etc [94];
- **Regressão:** a distinção mais significativa entre classificação e regressão é que a classificação prevê rótulos de classes distintas (discretas), enquanto a regressão facilita a previsão de uma quantidade contínua. Exemplos incluem previsão do tempo, análise de tendências de mercado, etc [94].

A Figura 2.10 mostra um exemplo de como a classificação é diferente com modelos de regressão.

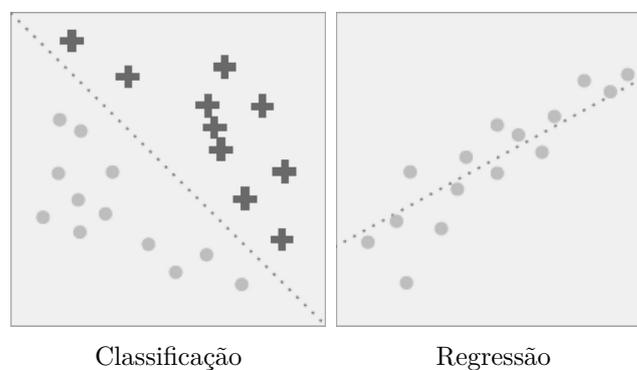


Figura 2.10. Classificação vs Regressão. Na classificação, a linha pontilhada representa um limite linear que separa as duas classes; na regressão, a linha pontilhada modela a relação linear entre as duas variáveis. Fonte: adaptada de [94].

O processo de treinamento envolve a apresentação do modelo a um conjunto de dados de treinamento, composto por pares de entrada e saída. O modelo ajusta seus parâmetros para minimizar o erro entre as previsões e os rótulos esperados, utilizando algoritmos de otimização como o gradiente descendente [87].

Um exemplo clássico de aplicação do aprendizado supervisionado é o reconhecimento de imagens, no qual o modelo é treinado com imagens rotuladas e, em seguida, é capaz de classificar novas imagens [95]. O aprendizado supervisionado é aplicado em diversas áreas, como detecção de fraude em transações financeiras [96] e diagnóstico médico [97].

Os algoritmos de aprendizado supervisionado mais populares incluem árvores de decisão, regressão linear, regressão logística, máquinas de vetores de suporte (*Support Vector Machines* - SVM, em inglês) e redes neurais artificiais (RNAs) [90]. Cada um desses algoritmos tem suas próprias vantagens e desvantagens, e a escolha do algoritmo depende do problema em questão.

Um dos maiores desafios do aprendizado supervisionado é lidar com o *overfitting*, que ocorre quando o modelo se ajusta aos dados de treinamento, mas não é generalizável para novos dados [93]. Várias técnicas, como validação cruzada e regularização, são usadas para evitar o *overfitting*.

2.4.1.2 Aprendizado não Supervisionado

O aprendizado não supervisionado é outra técnica de aprendizado de máquina, no qual o modelo é treinado com dados não rotulados. Ao contrário do aprendizado supervisionado, não há saídas desejadas fornecidas durante o treinamento. O objetivo é encontrar estruturas ou padrões nos dados de entrada sem a necessidade de rótulos prévios [91].

De acordo com [90] e [98], o aprendizado não supervisionado é utilizado em problemas de agrupamento (*clustering*, em inglês), redução de dimensionalidade e detecção de anomalias:

- **Agrupamento:** é usado para agrupar instâncias de dados similares em *clusters* [93];
- **Redução de dimensionalidade:** é usada para encontrar uma representação de baixa dimensionalidade dos dados que mantenha as informações relevantes, pois a redução de dimensionalidade permite que os dados sejam representados em um espaço de menor dimensão sem perda significativa de informação [93];
- **Detecção de anomalias:** é uma técnica de análise de dados que tem como objetivo identificar padrões incomuns ou desvios significativos em um conjunto de dados. A detecção de anomalias pode ser útil em várias aplicações, como na detecção de

fraudes, na análise de dados de segurança e na análise de sentimentos em textos [99].

Sobre os algoritmos mais comuns utilizados no aprendizado não supervisionado, destacam-se o K-Means, a análise de componentes principais (*Principal Component Analysis* - PCA, do inglês), o agrupamento hierárquico e os algoritmos de associação. Os algoritmos de agrupamento hierárquico e K-Means são técnicas amplamente utilizadas para identificar grupos em conjuntos de dados. Enquanto o K-Means é uma técnica de agrupamento baseada em centroides, o agrupamento hierárquico constrói uma hierarquia de grupos de forma iterativa. Já a PCA é uma técnica de redução de dimensionalidade que transforma um conjunto de variáveis correlacionadas em um conjunto menor de variáveis não correlacionadas. Por fim, os algoritmos de associação são usados para identificar padrões frequentes em dados transacionais [90], [93].

Um exemplo de aplicação do aprendizado não supervisionado é a análise de mercado para identificar grupos de clientes com interesses semelhantes. Outra aplicação é a análise de redes sociais para identificar comunidades de usuários com interesses em comum [100].

2.4.1.3 Aprendizado por Reforço

O aprendizado por reforço é outra técnica de ML, na qual o modelo é treinado a partir de interações de um agente com um ambiente dinâmico. Nessa abordagem, o agente aprende a tomar decisões ao receber *feedback* positivo ou negativo na forma de recompensas ou penalidades, o que permite ao agente atualizar seu modelo interno do ambiente e sua estratégia de tomada de decisão. O objetivo do agente é maximizar uma medida cumulativa de recompensa ao longo do tempo, o que exige que ele aprenda a identificar as ações que levam a recompensas mais altas e evite as ações que levam a penalidades [101].

No *feedback* positivo, o agente recebe uma recompensa por suas ações corretas, o que aumenta a probabilidade de que ele repita essas ações no futuro. Por exemplo, se um agente está jogando um jogo e recebe uma pontuação alta por completar uma tarefa, ele é recompensado e incentivado a repetir essa tarefa no futuro. Já no *feedback* negativo, o agente recebe uma punição por suas ações incorretas, o que diminui a probabilidade de que ele repita essas ações no futuro. Por exemplo, se um agente está jogando um jogo e perde pontos por cometer um erro, ele é punido e incentivado a evitar esse erro no futuro.

O aprendizado por reforço é particularmente útil em situações em que a tomada de decisão requer ações sequenciais e que o sucesso depende não apenas da ação atual, mas também das ações futuras [91].

Existem várias técnicas e algoritmos de aprendizado por reforço que podem ser usados

para treinar um agente a tomar decisões em diferentes tipos de ambientes. Os algoritmos mais populares incluem *Q-learning* e *Policy Gradient* [102]. O algoritmo *Q-learning* utiliza uma tabela de valores de ação para aprender a escolher a melhor ação em cada estado do ambiente [103]. Já o *Policy Gradient* é um método de aprendizado por reforço que busca otimizar diretamente a política de ação do agente, sem a necessidade de estimar o valor de cada ação. Ele faz isso por meio do cálculo dos gradientes da função objetivo em relação aos parâmetros da política, que são usados para atualizar esta por meio do método de gradiente ascendente [101]. Outro algoritmo comum é o SARSA, uma variação do *Q-learning* que leva em consideração a política de ação atual do agente, permitindo o aprendizado de políticas comportamentais mais suaves [104].

O aprendizado por reforço tem sido amplamente aplicado em jogos, robótica e controle de processos. Por exemplo, o AlphaGo, um programa de computador desenvolvido pela Google DeepMind, utilizou aprendizado por reforço para derrotar campeões mundiais no jogo de tabuleiro Go [105]. Outra aplicação é o controle de robôs para realizar tarefas complexas, como caminhar e manipular objetos [106].

2.4.1.4 Redes Neurais Artificiais

As RNAs podem ser definidas como um conjunto de algoritmos de aprendizado de máquina que se inspiram na estrutura e funcionamento do cérebro humano, com o objetivo de resolver problemas complexos utilizando modelos matemáticos baseados na estrutura e função dos neurônios biológicos [107]. A aprendizagem em sistemas biológicos envolve ajustes nas conexões sinápticas que existem entre os neurônios [108]. Isso também é verdade para as RNAs. Elas são compostas por camadas de neurônios artificiais altamente interconectados que têm uma função de ativação e podem ser conectados a outros neurônios, formando assim uma rede complexa de interações [109]. Essas interações ocorrem através de pesos sinápticos, que são ajustados durante o processo de aprendizado, permitindo que a rede neural possa ser treinada para realizar tarefas específicas, como classificação, previsão e reconhecimento de padrões em dados complexos [87], [110].

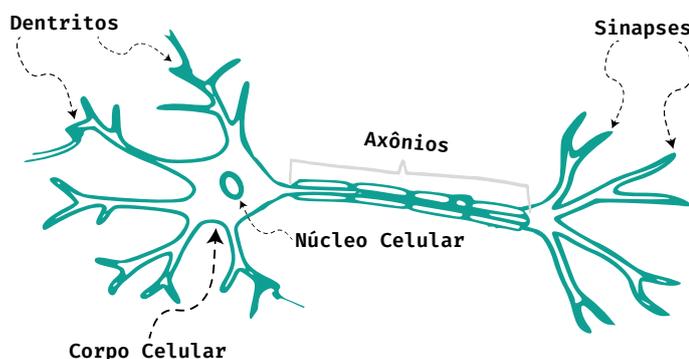
O aprendizado em RNAs é realizado por meio da otimização de uma função de perda, que mede o erro da rede em relação ao resultado esperado [110]. A escolha da função de perda e do algoritmo de otimização depende do problema em questão e eles podem ser selecionados de acordo com o conhecimento prévio sobre o domínio do problema [111].

2.4.1.5 Neurônios Biológicos e Artificiais

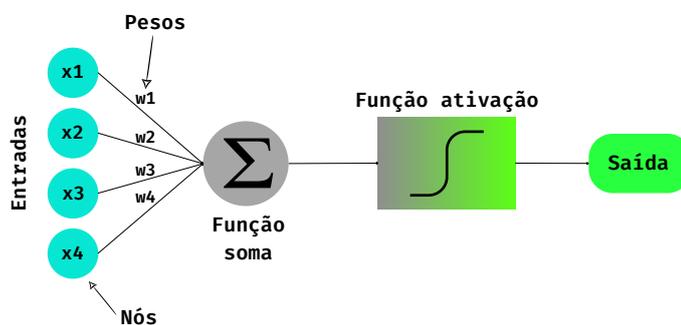
Os neurônios biológicos são as unidades básicas do sistema nervoso biológico, responsáveis por processar e transmitir informações por meio de sinais elétricos e químicos. Eles

possuem três principais componentes: dendritos, corpo celular e axônio. A Figura 2.11a esquematiza esses conceitos. Os dendritos recebem os sinais de outros neurônios; o corpo celular por sua vez integra esses sinais e o axônio transmite o sinal para outros neurônios [108], [112], [113].

Já os neurônios artificiais são as unidades básicas das redes neurais artificiais. Eles são modelados matematicamente para simular o comportamento dos neurônios biológicos. Cada neurônio artificial recebe entradas ponderadas ($x_i \times w_i$), as soma e aplica uma função de ativação para produzir uma saída, que é transmitida para outros neurônios na rede. A Figura 2.11b esquematiza uma RNA simples (chamada *perceptron*), com 4 neurônios na camada de entrada e um único neurônio na camada de saída, onde x_i é a i -ésima entrada binária e w_i é o peso sináptico (conexão) associado à entrada x_i . A função de ativação pode ser linear ou não-linear, dependendo do objetivo da rede neural [107], [114], [115].



(a) Estrutura de um neurônio biológico.



(b) Modelo de um neurônio artificial.

Figura 2.11. Representação de um neurônio biológico e um artificial. Fonte: autor [31], adaptada de [108].

Segue uma breve descrição de cada um dos componentes fundamentais do neurônio artificial (Figura 2.11b):

- **Entradas:** são os dados ou sinais que são fornecidos à rede neural. Eles representam as informações de entrada para a rede, que podem ser valores numéricos, vetores ou matrizes, dependendo da natureza do problema;
- **Pesos:** são parâmetros associados às conexões entre os neurônios em uma rede neural. Cada peso é atribuído a uma conexão, que determina a importância relativa da entrada correspondente para a saída do neurônio. Os pesos são ajustados durante o processo de treinamento da rede para otimizar o desempenho do modelo;
- **Função soma:** é o resultado da multiplicação das entradas pelos seus respectivos pesos e a subsequente soma dos produtos ponderados. É uma etapa intermediária na computação de um neurônio e representa a combinação linear das entradas ponderadas;
- **Função de ativação:** é uma função matemática aplicada à soma ponderada das entradas de um neurônio.

A Tabela 2.1 mostra a relação entre rede neural biológica e rede neural artificial.

Tabela 2.1. Relações entre as redes biológica e artificial.

Rede Neural Biológica	Rede Neural Artificial
Dendritos	Entradas
Núcleo celular	Nós
Sinapse	Pesos
Axônio	Saída

2.4.1.6 Função de Ativação

A função de ativação em redes neurais é uma função matemática que é aplicada a cada neurônio em uma rede neural. Ela determina a saída do neurônio com base em sua entrada ponderada. A função de ativação introduz a não-linearidade necessária para que a rede neural possa aprender a representar relações complexas entre os dados [116]. As funções de ativação mais comumente utilizadas na literatura são a função *Identity* (Identidade); *Sign* (Sinal); *Sigmoid* (Sigmóide); *Tanh* (Tangente Hiperbólica); *ReLU* (Unidade Linear Retificada); *HardTanh*. Cada função de ativação é descrita em termos de sua fórmula matemática, faixa de valores de entrada e saída, e propriedades associadas.

A função de ativação **Identity** retorna a entrada sem nenhuma alteração. Ela é uma função linear que preserva os valores originais, e é comumente usada em camadas de saída quando a rede neural é usada para tarefas de regressão [91], [87], [107]. Essa função é representada matematicamente por:

$$f(x) = x. \quad (2.1)$$

A função de ativação **Sign** atribui o valor +1, 0 e -1, se a entrada for maior, igual ou menor que zero, respectivamente. Ela mapeia a entrada para valores discretos, sendo útil em certos contextos, como classificação binária [91], [87], [107]. A função Sign(x) é definida matematicamente como:

$$\text{Sign}(x) = \begin{cases} 1, & \text{se } x > 0, \\ 0, & \text{se } x = 0, \\ -1, & \text{se } x < 0. \end{cases} \quad (2.2)$$

A função de ativação **Sigmoid** é uma função sigmoide que mapeia a entrada para um valor no intervalo entre 0 e 1. Ela é amplamente utilizada em redes neurais por sua propriedade de mapear valores para probabilidades. A função *Sigmoid* é representada pela equação:

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \quad (2.3)$$

A função de ativação **Tanh** é uma função sigmoide que mapeia a entrada para um valor no intervalo entre -1 e 1. Ela é simétrica em relação ao ponto zero e permite uma maior variação nos valores de saída em comparação com a função *Sigmoid*. A função Tanh(x) é definida matematicamente pela Equação 2.4:

$$\text{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = 2\sigma(2x) - 1. \quad (2.4)$$

A função de ativação **ReLU** (Unidade Linear Retificada, em inglês *Rectified Linear Unit*) retorna zero para valores negativos e a própria entrada para valores positivos. Ela é uma função não linear amplamente usada em redes neurais devido à sua eficiência computacional e capacidade de lidar com o problema do desaparecimento do gradiente. A função ReLU(x) é definida matematicamente como:

$$\text{ReLU}(x) = \max(0, x). \quad (2.5)$$

A função de ativação **HardTanh** é uma versão "hard"(limitada) da função Tanh. Ela retorna -1 para valores menores que -1, 1 para valores maiores que 1 e a própria entrada para valores dentro desse intervalo. A função $\text{HardTanh}(x)$ é definida matematicamente como:

$$\text{HardTanh}(x) = \max(-1, \min(1, x)). \quad (2.6)$$

As representações gráficas de todas as funções de ativação mencionadas acima são ilustradas na Figura 2.12.

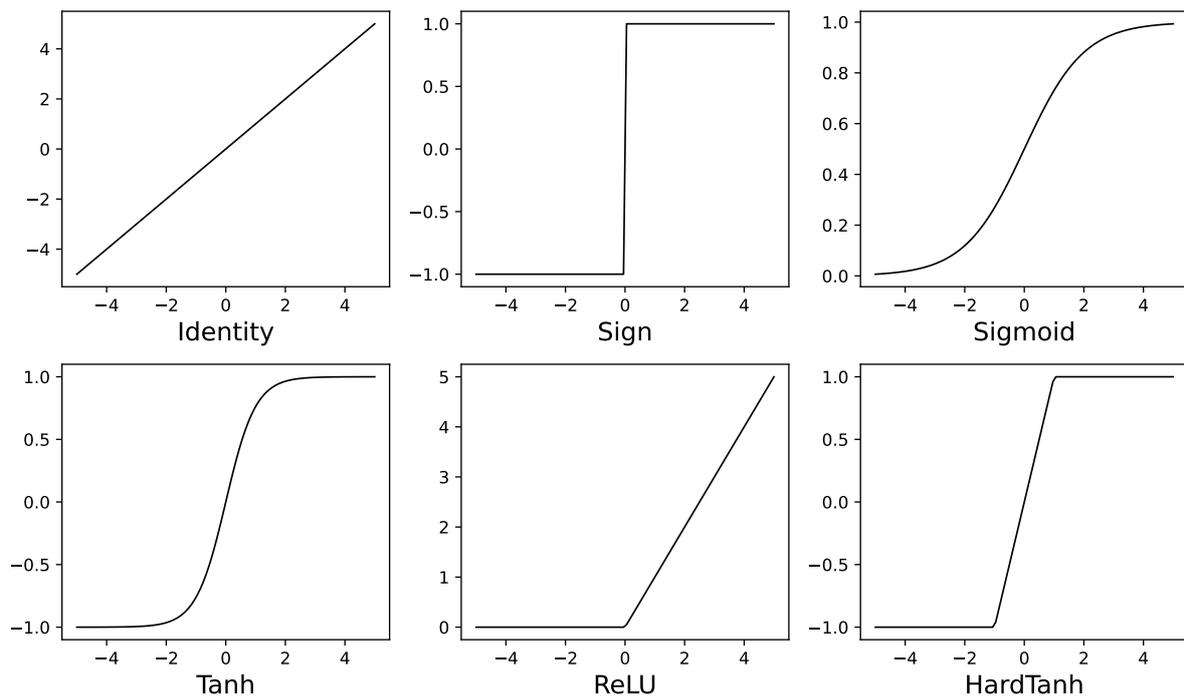


Figura 2.12. Várias funções de ativação. Fonte: autor [31], adaptada de [116].

2.4.2 Aprendizado Profundo

O DL, como mostra a Figura 2.9, é um subcampo do ML que busca construir e treinar modelos de redes neurais profundas [87], [115]. O objetivo central do DL é capacitar máquinas para realizar tarefas de alto nível, como reconhecimento de padrões, processamento de linguagem natural, processamento de imagens e tomada de decisões [115].

Ao contrário das abordagens tradicionais de ML, que exigem a engenharia manual de recursos (características dos dados), o DL permite que os modelos aprendam de forma automática as melhores representações diretamente dos dados brutos [115].

Isso é possível devido à presença de múltiplas camadas ocultas nas arquiteturas das

redes neurais profundas, que são capazes de aprender e extrair representações complexas e hierárquicas de dados. Cada camada oculta extrai características relevantes com base nas informações fornecidas pela camada anterior, permitindo uma progressiva capacidade de abstração e generalização dos dados [87].

A Figura 2.13 esquematiza a arquitetura de uma rede neural profunda típica que contém duas camadas ocultas, além de uma camada de entrada e uma de saída [116].

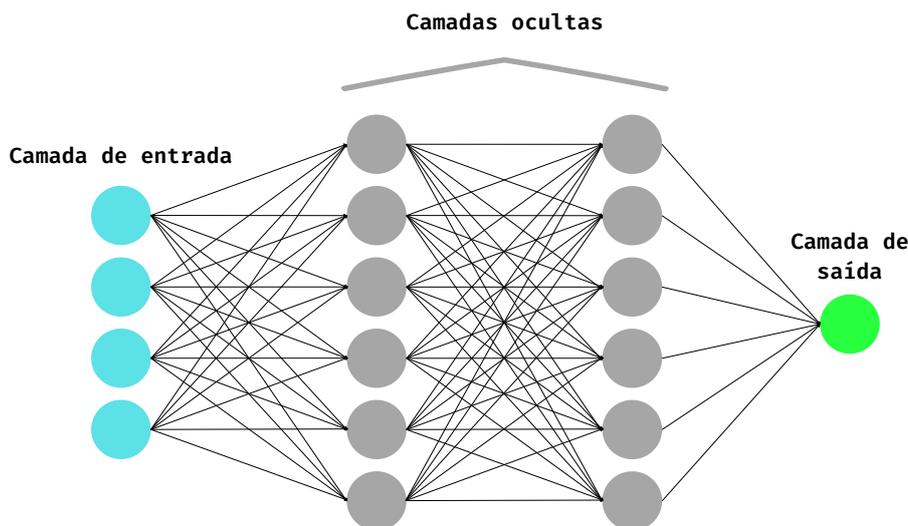


Figura 2.13. Diagrama representando uma rede neural profunda. Fonte: autor [31], adaptada de [116].

O treinamento de uma rede neural profunda envolve a apresentação dos dados de treinamento à rede e ao ajuste iterativo dos pesos das conexões sinápticas entre os neurônios, com o objetivo de minimizar uma função de custo que mede a discrepância entre as saídas da rede e os rótulos esperados [87].

2.4.2.1 Rede Neural Convolutacional

A rede neural convolutacional (*Convolutional Neural Network* - CNN, em inglês) é uma arquitetura de aprendizado profundo que foi projetada para processar dados de natureza espacial, como imagens. O principal benefício da CNN em comparação com outras arquiteturas é que ela identifica automaticamente as características relevantes sem qualquer supervisão humana [95], [117], [118]. A CNN é composta por camadas específicas, cada uma desempenhando uma função essencial no aprendizado e extração de características [95].

A ordem das camadas em uma CNN pode variar dependendo da arquitetura específica, mas geralmente segue a especificada na Figura 2.14.

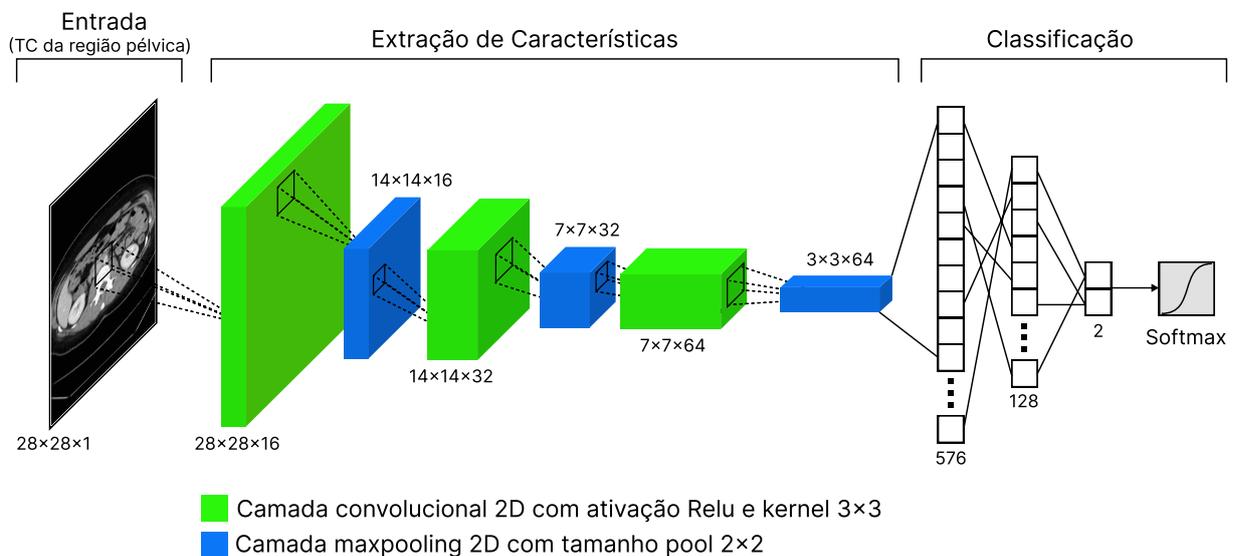


Figura 2.14. Diagrama representando uma Rede Neural Convolucional. Fonte: autor [31], adaptada de [116].

- **Camada de Entrada:** é responsável por receber os dados de entrada, que geralmente são imagens. As imagens são representadas como tensores de pixels, onde cada valor representa a intensidade do pixel em uma determinada posição. A camada de entrada prepara e alimenta esses dados para as próximas camadas da rede [115], [116], [95];
- **Camadas de Convolução:** as camadas de convolução são a essência das CNNs. Cada camada de convolução é composta por um conjunto de filtros (ou *kernels*) que deslizam pela imagem, calculando a convolução entre o filtro e uma região localizada na imagem. Essa operação de convolução resulta em mapas de características que destacam a presença de características específicas, como bordas, texturas ou padrões em diferentes partes da imagem. Cada filtro aprende a detectar um tipo particular de característica, e a rede combina múltiplos filtros para capturar uma variedade de características em várias escalas e níveis de abstração [115], [116], [95]. A Figura 2.15 esquematiza esse conceito;
- **Camadas de Ativação:** após cada camada de convolução, é aplicada uma função de ativação não linear para introduzir não linearidade e melhorar a capacidade de aprendizado na rede. A função de ativação é aplicada elemento a elemento aos mapas de características gerados pela convolução. A função ReLU é amplamente usada nas CNNs [115], [116], [95];

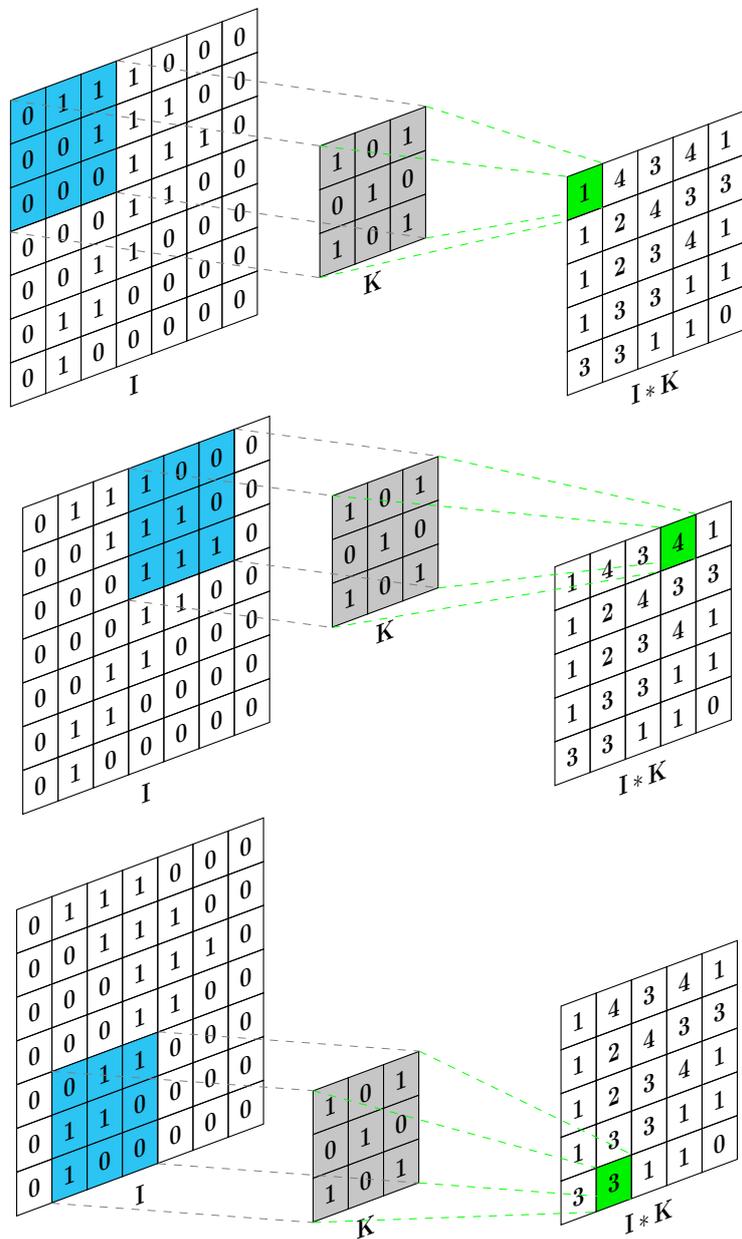


Figura 2.15. Um exemplo de convolução entre entrada I ($7 \times 7 \times 1$) e um *kernel* K ($3 \times 3 \times 1$) com passo de 1. Fonte: autor [31], adaptada de [116].

- **Camadas de *Pooling*:** estas camadas são responsáveis por reduzir a dimensionalidade espacial dos mapas de características gerados pelas camadas de convolução. O *pooling* é uma técnica que divide o mapa de características em regiões e realiza uma operação, como selecionar o valor máximo (*max pooling*) ou calcular a média (*average pooling*), para obter um valor representativo para cada região. Essa redução da resolução espacial permite que a rede se torne mais robusta a pequenas variações na posição das características e também reduz o número de parâmetros na rede, tornando-a mais eficiente computacionalmente [115], [116], [95];
- **Camadas Totalmente Conectadas:** estas camadas finais são responsáveis por mapear as características extraídas para as classes de destino. Elas incorporam a capacidade de generalização da rede neural, permitindo que a rede tome decisões finais com base nas características aprendidas durante o treinamento [115], [116], [95];
- **Camada de Saída:** esta é a última camada da CNN e produz as previsões finais do modelo. Dependendo da tarefa em questão, a camada de saída pode ter diferentes configurações [115], [116], [95].

2.4.2.2 U-Net

A U-Net é uma arquitetura de rede neural projetada principalmente para a segmentação de imagens biomédicas [37]. A estrutura fundamental da U-Net é composta por dois caminhos distintos, como demonstra a Figura 2.16. O primeiro caminho é conhecido como caminho de contração, também chamado de codificador ou caminho de análise. Esse caminho se assemelha a uma rede de convolução tradicional e tem a função de extrair informações de classificação das imagens. O segundo caminho é chamado de caminho de expansão, também referido como decodificador ou caminho de síntese. Esse caminho consiste em operações de convolução ascendente e concatenações com características do caminho de contração. Essa expansão possibilita que a rede aprenda informações de classificação de nível localizado. Além disso, o caminho de expansão também aumenta a resolução da saída [37]. Finalmente, após o caminho de expansão, a camada convolucional final produz uma imagem totalmente segmentada. Isso permite que a U-Net crie segmentações detalhadas e precisas das imagens de entrada [119].

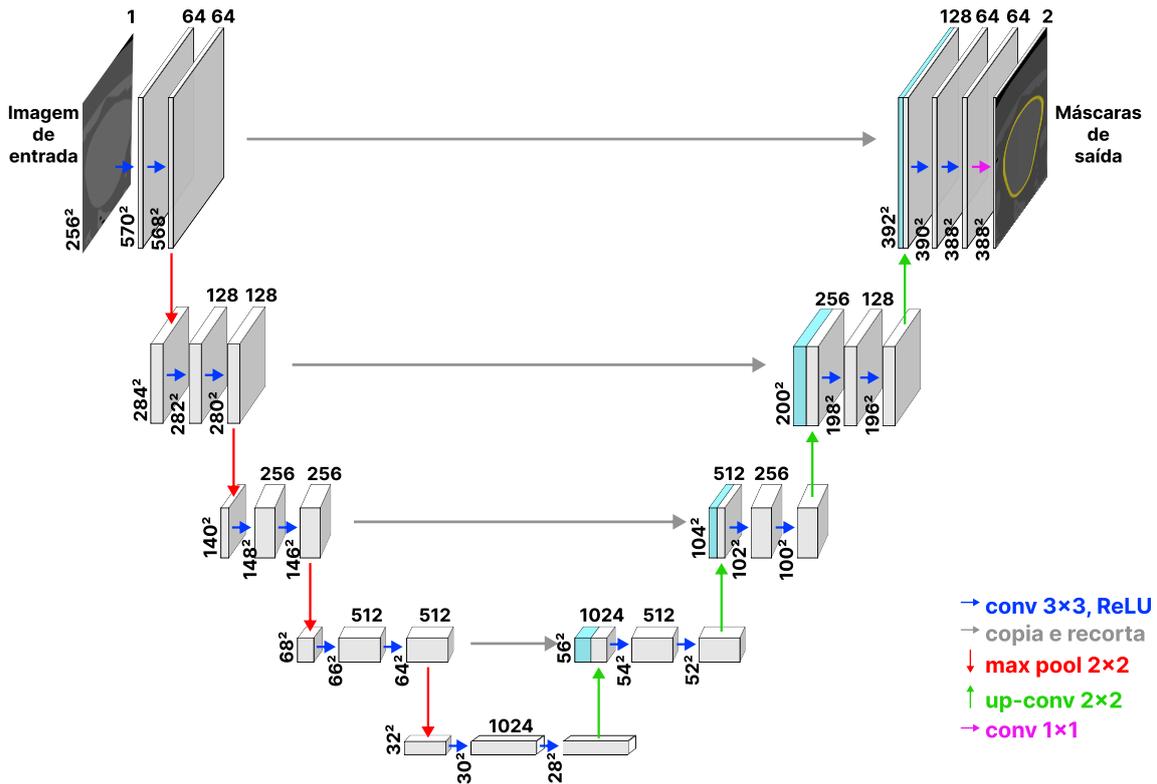


Figura 2.16. Arquitetura básica da U-Net. Fonte: autor [31], adaptada de [37].

2.4.3 Métricas de Avaliação

Nesta seção, é explorada a métrica de avaliação amplamente utilizada para medir a qualidade das segmentações em imagens médicas. A métrica desempenha um papel fundamental na análise crítica das segmentações geradas e na avaliação de desempenho do modelo. A métrica a ser discutida é o coeficiente de Dice (do inglês, *Dice Coefficient*).

Esse coeficiente é usado para avaliar a sobreposição espacial entre a segmentação prevista pelo modelo e a segmentação de referência (*ground truth*) [120]. Quanto mais próximo o valor do coeficiente de Dice de 1, melhor a sobreposição entre as segmentações [121]. Sua definição matemática é apresentada pela Equação 2.7:

$$\text{Dice} = \frac{2 \times |S \cap G|}{|S| + |G|}. \quad (2.7)$$

onde S representa a segmentação prevista pelo modelo; G representa a segmentação de referência (*ground truth*); $S \cap G$ é o tamanho da interseção entre S e G ; $|S|$ é o tamanho de S ; $|G|$ é o tamanho de G .

A Figura 2.17 ilustra a representação da sobreposição espacial e reprodutibilidade do coeficiente de Dice.

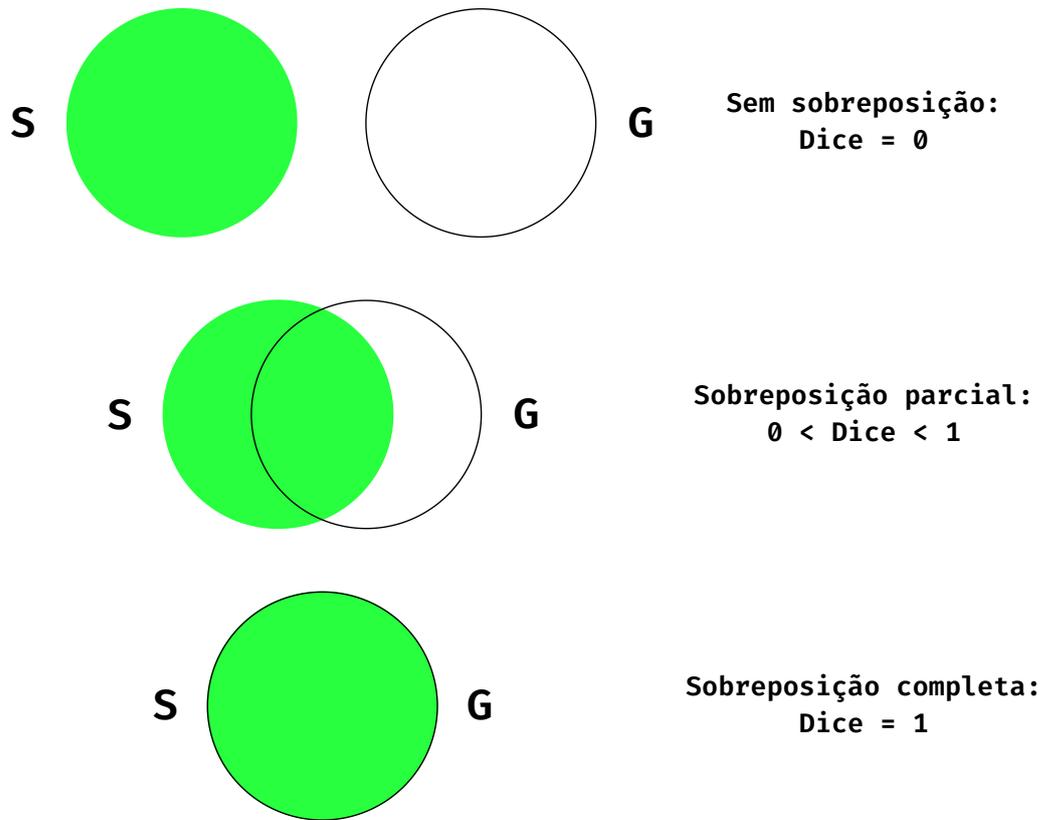


Figura 2.17. Sobreposição espacial das segmentações alvo G e S. Fonte: autor [31], adaptada de [120].

3 Metodologia

A metodologia deste trabalho consistiu das seguintes etapas: Aquisição do banco de dados de imagens médicas contendo as imagens de TC da bexiga e as máscaras correspondentes; Seleção, implementação, treino, teste e avaliação do modelo e Desenvolvimento e avaliação de desempenho do aplicativo.

A Figura 3.1 é um esquema que representa a ideia principal do objetivo do trabalho.

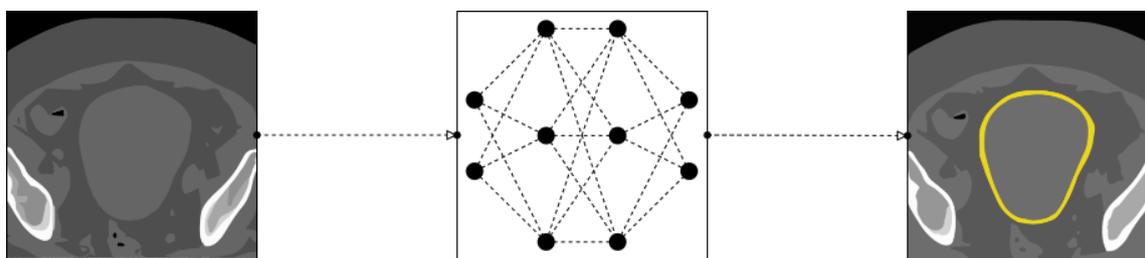


Figura 3.1. Ilustração da TC de entrada. Fonte: [35].

3.1 Aquisição do Banco de Imagens

As imagens foram adquiridas a partir de um conjunto de dados fornecido pelo setor de radioterapia do Hospital Universitário de Brasília (HUB), composto de 1.932 imagens de TC e 1.932 máscaras, sendo estas imagens dos contornos da estrutura da bexiga correspondentes. Todos os contornos foram desenhados manualmente por três especialistas.

São imagens tomográficas do tipo *cone beam*, obtidas com feixes de 6MV produzidos no acelerador linear *Halcyon* (*Varian Medical System* - Palo Alto, CA). Esse tipo de imagem possui mais ruído do que as convencionais por dois motivos: os feixes são do tipo *cone beam* e de megavoltagem. A energia é da ordem de MeV, diferente da energia usada em radiodiagnóstico, que é da ordem de KeV. Isso gera mais espalhamento (e portanto menor resolução na imagem resultante), pois feixes de MeV tendem a gerar mais espalhamento Compton. Todas as imagens foram adquiridas com matriz de 256×256 pixels e espessura de corte de 2mm (resolução $0.11\text{cm} \times 0.11\text{cm} \times 0.2\text{cm}$). A Figura 3.2 mostra alguns exemplos de imagem, a máscara correspondente e a sobreposição das duas.

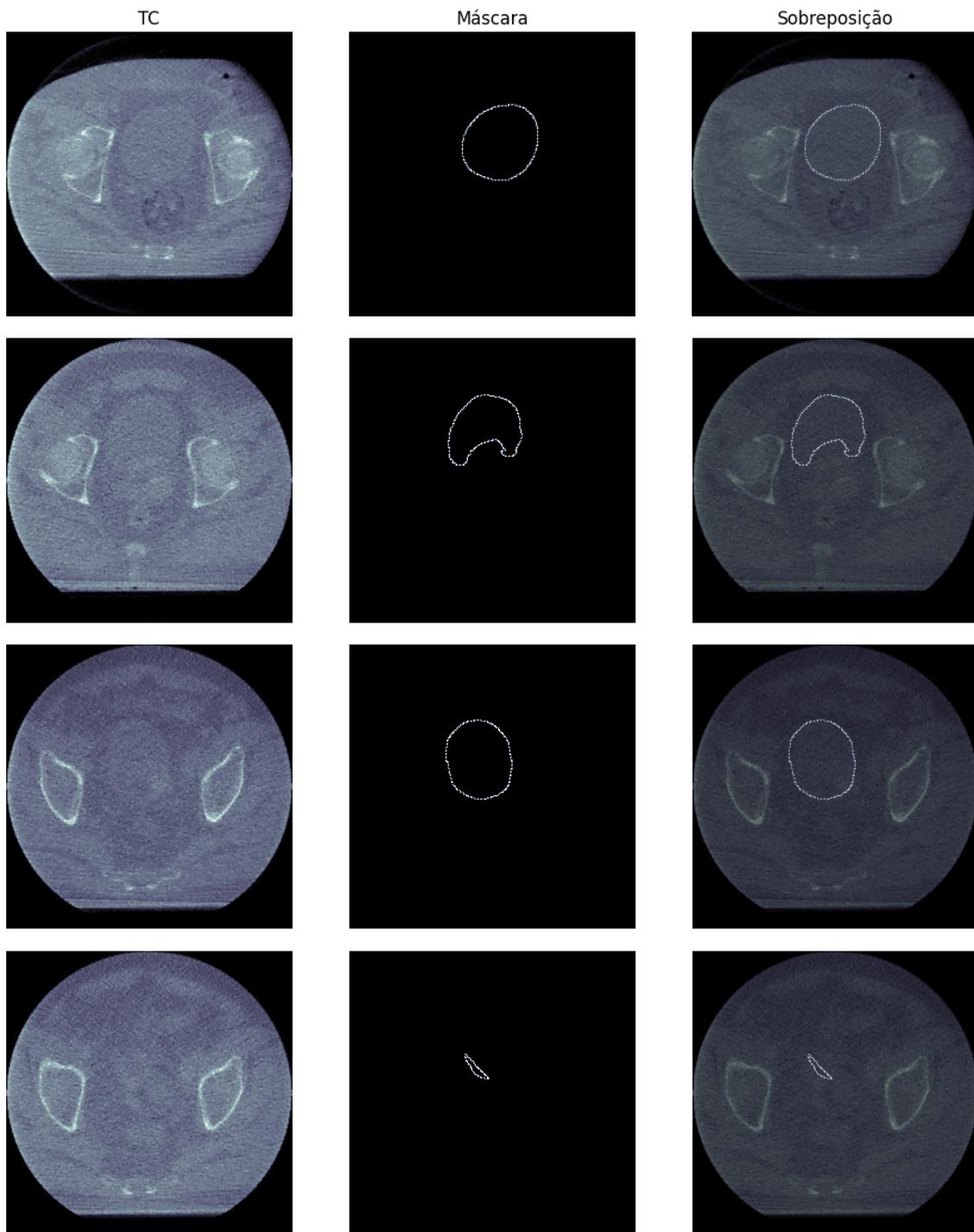


Figura 3.2. Exemplo de imagem e máscara da base de dados.

3.1.1 Aumento de Dados

Por se tratar de uma base de imagens limitada e pequena, foi adotado um recurso de aumento de dados (*data augmentation*) para aumentar a quantidade de imagens disponíveis [38], [122], [123], [124], [125], [126].

O treinamento eficaz de modelos de aprendizado profundo, como a U-Net, geralmente

requer um volume significativo de dados rotulados. Métodos de aumento de dados, como rotações e inversões, produzem as propriedades desejadas de invariância e robustez da rede resultante. Além de inversões e rotações, o aumento de dados incluiu variações de valores de cinza e deformações elásticas aleatórias às amostras de treinamento, sendo recomendadas como melhores práticas gerais para redes neurais convolucionais [37], [127].

A ampliação do conjunto de dados permite expandir o conjunto de dados original, criando variações artificiais nas imagens existentes. O aumento de dados é uma técnica de extrema importância que ajuda a melhorar a capacidade do modelo de generalizar padrões aprendidos a partir de um conjunto de dados limitado, reduzindo o risco de *overfitting* e melhorando o desempenho em dados de teste [38], [128], [129], [130], [131].

Em cada uma das imagens foram realizadas 3 transformações escolhidas de forma aleatória dentre as 6 transformações listadas abaixo, resultando em um aumento de dados para 7.728 imagens, com 4 imagens de uma mesma fonte.

3.1.1.1 Detalhes das Transformações Utilizadas:

- **CLAHE** (*Contrast Limited Adaptive Histogram Equalization*): Desenvolvida para melhorar o contraste em regiões específicas de uma imagem, a transformação CLAHE foi escolhida para realçar detalhes sutis nas estruturas de interesse. A aplicação da CLAHE visa melhorar a visibilidade de detalhes finos e sutis nas imagens, garantindo que o modelo possa aprender padrões em regiões com variação de contraste [132];
- **CenterCrop**: Esta transformação envolve o corte centrado da imagem para focar em regiões de interesse, mantendo a relevância das estruturas mesmo em diferentes tamanhos de entrada. A transformação *CenterCrop* mantém o foco nas estruturas de interesse, mesmo quando as imagens têm tamanhos variados [133];
- **HorizontalFlip e VerticalFlip**: Estas transformações são aplicadas para criar variações horizontais e verticais nas imagens, simulando diferentes perspectivas e orientações das estruturas. Essas transformações simulam variações naturais na orientação das estruturas, garantindo que o modelo seja robusto a diferentes perspectivas e orientações [134], [135];

- ***Transpose***: Esta operação de transposição troca as posições de linhas e colunas, adicionando variações geométricas às imagens. Assim como as transformações horizontais e verticais simulam variações naturais na orientação das estruturas, a transposição garante que o modelo seja robusto a diferentes perspectivas e orientações [136];
- ***RandomRotate90***: A rotação aleatória de 90 graus amplia ainda mais as perspectivas, treinando o modelo para reconhecer e segmentar estruturas em diferentes orientações. Rotações aleatórias ampliam ainda mais a diversidade das imagens, preparando o modelo para segmentar estruturas em qualquer orientação [137].

Os parâmetros das transformações utilizadas no *data augmentation* podem ser visualizadas na Tabela 3.1:

Tabela 3.1. Parâmetros do *data augmentation*.

Transformação	Probabilidade
CLAHE	0,5
<i>CenterCrop</i>	0,5
<i>HorizontalFlip</i>	0,5
<i>VerticalFlip</i>	0,5
<i>Transpose</i>	0,5
<i>RandomRotate90</i>	0,5

3.2 Seleção do Modelo e Detalhes da Implementação

A escolha do modelo foi feita com base nos resultados observados na revisão bibliográfica feita, em que resultados com um coeficiente de Dice significativo na região estudada eram produzidos por uma segmentação feita na arquitetura U-Net, e portanto ela serviu como um bom primeiro passo para se trabalhar com as imagens da base de dados. A rede neural convolucional e o pré-processamento dos dados foram implementados na linguagem de programação Python 3.10, devido à sua ampla disponibilidade, versatilidade e suporte para a execução de ferramentas e bibliotecas necessárias para o desenvolvimento dos modelos e algoritmos propostos. Além disso, foi utilizado o framework TensorFlow 2.0 e bibliotecas como Keras 2.12.0, NumPy 1.25.0, OpenCV 4.7.0.72 e Matplotlib 3.7.1. Para o desenvolvimento e execução dos experimentos, foi utilizado o ambiente de programação online Google Colab. O aplicativo foi desenvolvido usando Streamlit 1.26.0, que também é uma biblioteca Python.

3.2.1 Implementação da U-Net

A implementação da U-Net foi feita do zero em Python, Tensorflow e Keras no ambiente Colab, usando os seguintes hiperparâmetros:

- **Função de Ativação:** Função de ativação ReLU para promover a não-linearidade na rede [138], [139];
- **Otimizador:** Otimizador Adam para ajustar os pesos da rede durante o treinamento, permitindo uma convergência mais rápida [140];
- **Inicialização de Pesos:** Padrão da biblioteca, buscando um equilíbrio entre convergência e prevenção de gradientes nulos ou explosivos [139].

Já a configuração dos hiperparâmetros de treinamento foram:

- **Taxa de Aprendizado (*learning rate*):** Foi definida como $1e^{-4}$, um valor comumente utilizado para tarefas de segmentação [139];
- **Número de Épocas (*epochs*):** 100 épocas para garantir uma exploração abrangente do espaço de treinamento. A escolha desse número foi devido ao tamanho do conjunto de dados, a capacidade de computação disponível e considerações sobre estabilidade e eficiência no treinamento [139], [141];
- **Tamanho do Lote (*batch size*):** 32 para equilibrar eficiência computacional e estabilidade no treinamento, e é comumente encontrada na literatura e é amplamente utilizada na prática [87], [139];
- **Callbacks:** Foi implementada uma variedade de *callbacks*, como *ModelCheckpoint* para salvar o melhor modelo; *ReduceLROnPlateau* para ajustar dinamicamente a taxa de aprendizado; *CSVLogger* para registrar métricas em um arquivo CSV; *TensorBoard* para análise visual e *EarlyStopping* para evitar *overfitting*.

Já os hiperparâmetros específicos para segmentação da bexiga são:

- **Tamanho da Imagem de Entrada:** O tamanho da imagem de entrada foi definido como 256×256 pixels, uma escolha comum para imagens médicas que visa equilibrar detalhes e eficiência computacional [83], [122]. As imagens de entrada foram obtidas com estas dimensões, porém o modelo foi também tentativamente treinado com dimensões maiores e menores. Contudo, as performances não foram boas para estes casos. Ao reduzir a dimensão, o custo computacional era menor, porém o modelo treinado não conseguia identificar o local a ser segmentado. Ao

aumentar a dimensão, o custo computacional foi muito grande, não permitindo que o treinamento fosse concluído;

- **Número de Classes:** Considerando a tarefa de segmentação binária (bexiga e não-bexiga), foi configurado um único canal na camada de saída [87], [139].

Já a métrica e a função de perda específicas escolhidas para avaliar o desempenho da segmentação foram:

- **Coefficiente de Dice:** Por ser amplamente usada na avaliação da segmentação de imagens médicas [32], [33], [34], [35], [36];
- **Função de perda (*Loss Function*):** A função de perda escolhida, *dice_coef_loss*, alinha-se às características específicas da tarefa de segmentação, incentivando a convergência para máscaras mais precisas [142].

3.2.2 Divisão da Base de Dados

A estratégia de divisão da base de dados, que consiste em 7.728 imagens distribuídas entre os conjuntos de treino, validação e teste na proporção de 70%, 20% e 10% respectivamente (equivalente a 5409 para treino, 2.087 para validação e 232 para teste), foi escolhida com base em práticas comuns na comunidade científica [143],[144], [145], [126]. A alocação de 70% para o conjunto de treinamento é para assegurar que o modelo seja exposto a uma variabilidade significativa durante a fase de aprendizado. A reserva de 10% para o conjunto de teste permite uma avaliação do desempenho do modelo em dados não vistos, enquanto os 20% destinados ao conjunto de validação apoiam a atualização de parâmetros durante o treinamento. Essa abordagem balanceada visa otimizar a generalização e a eficácia do modelo, seguindo diretrizes estabelecidas na literatura científica [144], [145], [146], [126].

3.2.3 Desenvolvimento e Avaliação de Desempenho do Aplicativo

Nesta seção, será descrito o processo de desenvolvimento do aplicativo web, destacando a integração da arquitetura U-Net pré-treinada para a segmentação da bexiga em imagens de TC. Será abordada também a estratégia de avaliação de desempenho do aplicativo.

3.2.3.1 Desenvolvimento do Aplicativo

O aplicativo foi desenvolvido utilizando a biblioteca Streamlit, uma ferramenta eficaz para a criação de interfaces web interativas em Python. A interface do aplicativo foi projetada para ser acessível, permitindo a fácil interação com usuários, independentemente de sua experiência em processamento de imagens médicas.

Foi realizada a integração da U-Net pré-treinada (treinada com a metodologia descrita na seção anterior) ao aplicativo, permitindo que os usuários façam o *upload* de suas próprias imagens de TC diretamente no aplicativo. A U-Net então realiza a segmentação da bexiga, destacando a região de interesse nas imagens.

O aplicativo apresenta as seguintes funcionalidades principais:

- **Upload de Imagens:** Os usuários podem fazer o *upload* de imagens diretamente no aplicativo, onde os usuários foram os físicos radioterapeutas especialistas na área de segmentação de imagens médicas;
- **Segmentação Automática:** Utilizando a arquitetura U-Net pré-treinada, o aplicativo realiza a segmentação automática da bexiga nas imagens fornecidas;
- **Visualização Instantânea:** As imagens segmentadas são apresentadas instantaneamente aos usuários, proporcionando uma visualização imediata dos resultados, com o tempo de segmentação variando entre 5 e 10 segundos por imagem.

Já a avaliação de desempenho do aplicativo se baseou em dois principais critérios:

- **Eficácia na Segmentação:** A precisão na segmentação da bexiga foi avaliada utilizando um conjunto diversificado de imagens;
- **Tempo de Resposta:** O tempo necessário para realizar a segmentação foi medido para garantir que o aplicativo ofereça uma experiência eficiente de usuário;

4 Resultados e Discussões

4.1 Avaliação Quantitativa da Segmentação

A avaliação quantitativa do desempenho do modelo na tarefa de segmentação é apresentada por meio de uma métrica específica, fornecendo informações valiosas sobre a eficácia da abordagem proposta. A métrica deste trabalho inclui:

Tabela 4.1. Avaliação quantitativa do desempenho da segmentação

Métrica	Valor
Coefficiente de Dice	0,81

- **Coefficiente de DICE:** O coeficiente de Dice apresentou um valor de 81%. Essa métrica varia de 0 a 1, onde 1 indica uma concordância significativa, sugerindo que o modelo foi capaz de realizar a segmentação com alta precisão. Portanto, o valor obtido demonstra uma boa concordância entre as máscaras segmentadas pelo modelo e as máscaras de referência;

A Figura 4.1 apresenta um exemplo visual de segmentação para ilustrar a precisão do modelo. Ficam em destaque a seguir casos representativos de imagens de TC antes e depois da aplicação do modelo, comparando as máscaras de segmentação com as máscaras de referência.

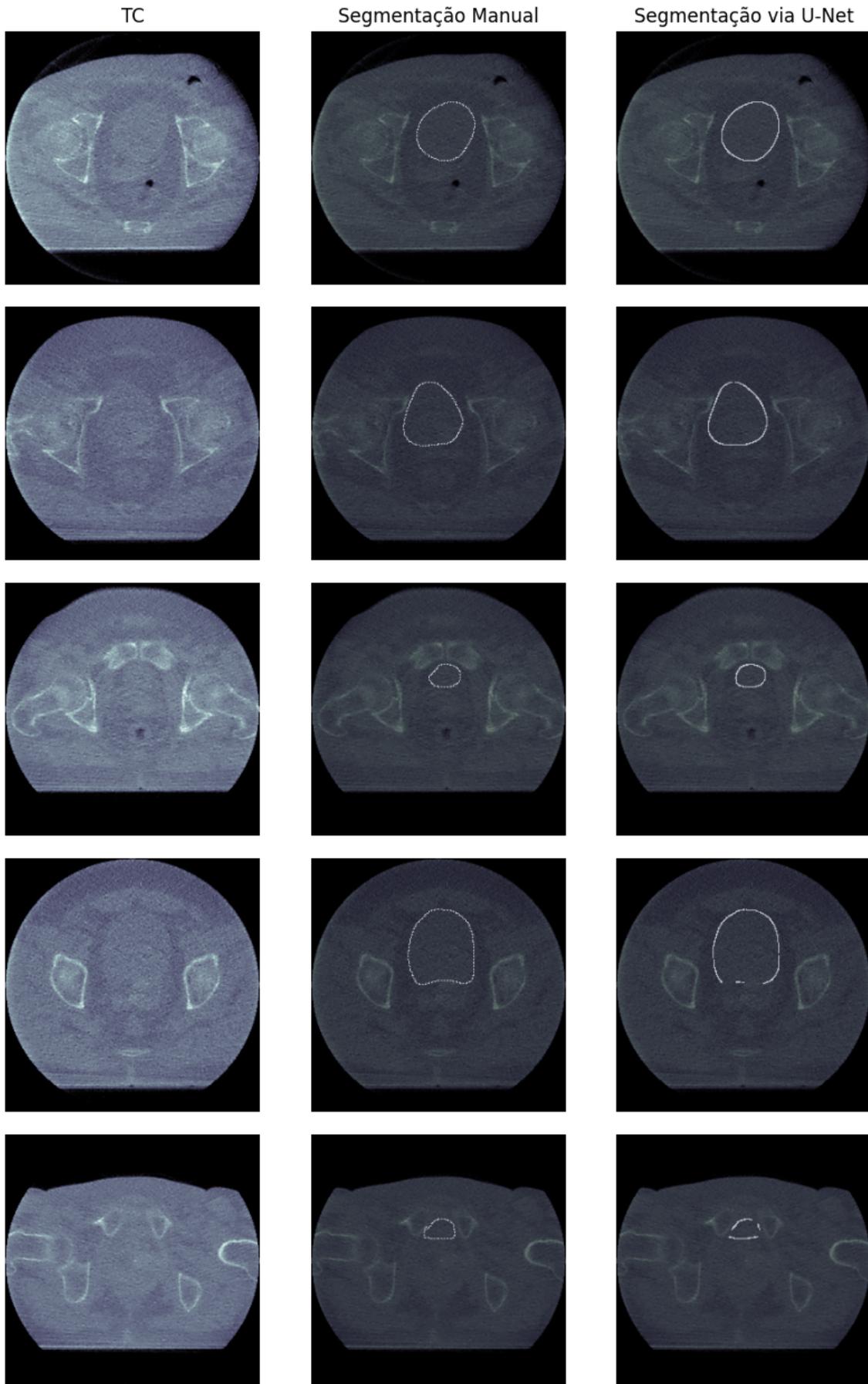


Figura 4.1. Comparação entre segmentação manual e via U-Net.

É possível observar na Tabela 4.2 que, quando comparada com os modelos propostos pela revisão de literatura realizada, a diferença de performance é pequena. No entanto, a U-Net implementada ainda obteve um resultado menor, visto que os valores podem variar em pequena escala de acordo com diversas variáveis, como: o tamanho do banco de imagens; a qualidade do conjunto de imagens (foi mencionado anteriormente que as imagens usadas são mais ruidosas que as convencionais); as máscaras usadas não estarem com um delineamento contínuo e sim em uma nuvem de pontos coordenados em torno da ROI; a divisão dos dados, o modelo computacional e a própria natureza estocástica dos modelos. O valor obtido do coeficiente reflete diretamente na precisão da segmentação da bexiga. Quando as duas últimas segmentações da Figura 4.1 são analisadas visualmente, pode-se observar que as bexigas das respectivas imagens não foram totalmente segmentadas.

Tabela 4.2. Comparação do modelo com outros trabalhos citados na revisão da literatura.

Autores	Modelo	Métricas
Implementado pelo autor	U-Net	Coef. de Dice de 0,81 para a bexiga
Rhee, Dong Joo et al [32]	3D V-Net	Coef. de Dice de 0,89 para a bexiga
Zhang, Zhuangzhuang et al [33]	ARPM-Net	Coef. de Dice de 0,97 para a bexiga
Fu, Yabo et al [34]	CBCT-Net	Coef. de Dice de 0,96 para a bexiga
Kazemifar, Samaneh et al [35]	U-Net	Coef. de Dice de 0,95 para a bexiga
Amjad A, Xu J, et al [36]	ResUnet3D	Coef. de Dice de 0,95 para a bexiga

Apesar do desempenho geral não ser melhor, o modelo proposto é mais simples. Foi possível alcançar uma performance similar devido ao maior pré-processamento das imagens, além da aplicação da técnica de aumento de dados e da redução da quantidade de ruídos e de informações desnecessárias. Conseqüentemente, a qualidade dos dados de treinamento pôde ser melhorada.

Além dos resultados obtidos na segmentação da bexiga, a U-Net tem mostrado bom desempenho em diversos contextos, incluindo a segmentação de vasos sanguíneos na retina, bordas de núcleos celulares, e lesões de melanoma. Nesses casos, a alta acurácia está associada à maior resolução e qualidade das imagens, que proporcionam bordas mais definidas e menos ruído, facilitando o aprendizado do modelo. A performance mais baixa observada neste estudo pode ser atribuída à qualidade inferior e ao ruído presente nas imagens de TC, que dificultam o delineamento da região de interesse, impactando o coeficiente de Dice. Isso destaca a importância de imagens de alta qualidade para uma segmentação precisa, especialmente em contextos médicos.

4.2 Avaliação de Desempenho do Aplicativo

Para alcançar os objetivos propostos, foi desenvolvido um aplicativo web, projetado para a utilização prática da arquitetura U-Net pré-treinada para realizar a segmentação da bexiga. Para a construção do aplicativo foram utilizadas algumas ferramentas descritas no capítulo anterior.

O objetivo do aplicativo é permitir que o usuário carregue uma ou mais imagens de dimensões 256×256 pixels do diretório específico no computador. Após o *upload*, a imagem o aplicativo processa e realiza a segmentação e exibe o resultado, sendo possível realizar o *download* da imagem segmentada.

A Figura [4.2] destaca a interface inicial do aplicativo (*Home*). Há também há um breve texto explicando do que se trata a ferramenta, o seu propósito e como usá-la, além de uma animação ilustrativa da realização da segmentação. No menu principal tem dois botões: "*Home*" e "*Upload Image*".

Ao clicar no botão "*Upload Image*" o usuário é direcionado para a interface ilustrada na Figura [4.3], onde pode carregar uma ou mais imagens (limitada a 200MB) para que possa realizar a segmentação ou segmentações desejadas, clicando no botão "*Browse files*".

Após selecionar a imagem desejada, o aplicativo exibe instantaneamente os resultados da segmentação, que incluem: a imagem original, predição e a sobreposição. A Figura [4.3] também apresenta um exemplo de segmentação realizada pelo aplicativo.

Os resultados visuais do aplicativo foram comparados com estudos anteriores revisados na literatura, destacando a eficácia do modelo proposto na segmentação da bexiga, considerando a qualidade variável das imagens de treinamento por serem imagens *Cone-Beam*, já mencionadas como sendo naturalmente de menor qualidade e pouco utilizadas na literatura consultada.

O aplicativo foi testado de forma não formal por três Físicos Médicos especialistas em Radioterapia. Estes usuários exploraram a ferramenta em um contexto prático, fornecendo *feedback* qualitativo sobre a interface, usabilidade e precisão dos resultados de segmentação. Embora o teste não tenha seguido um protocolo formal estruturado, os especialistas indicaram que o aplicativo apresenta potencial para ser uma ferramenta útil no ambiente clínico, especialmente por sua simplicidade de uso e pela clareza dos resultados fornecidos.

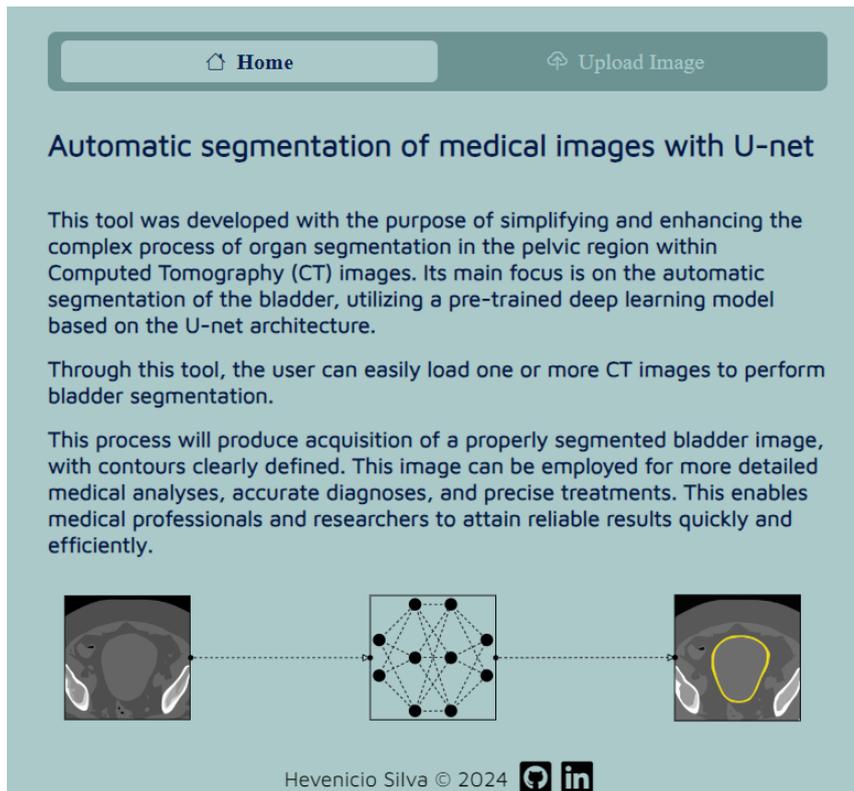


Figura 4.2. Interface inicial do aplicativo, pela aba *Home*.

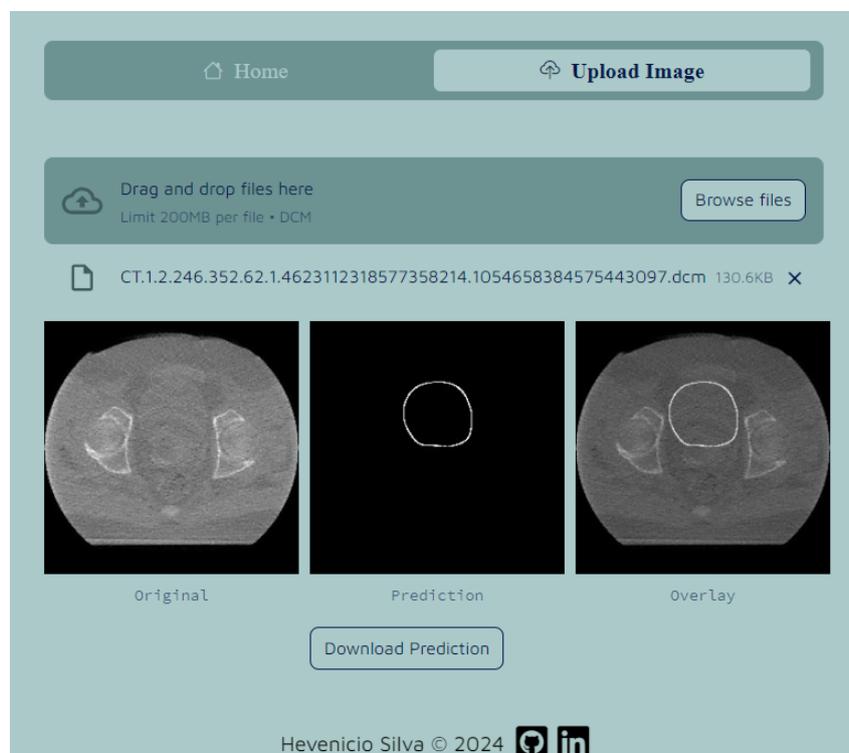


Figura 4.3. Resultados da segmentação da bexiga, pela aba *Upload Image*.

5 Conclusões e Trabalhos Futuros

Esta dissertação explorou de forma abrangente a aplicação do modelo de aprendizado profundo por meio da arquitetura U-Net na segmentação de imagens de tomografia computadorizada para a identificação da região da bexiga. Ao longo desta pesquisa, foram abordados aspectos fundamentais, desde a revisão bibliográfica das técnicas mais recentes em segmentação de imagens médicas até o desenvolvimento prático de um aplicativo web funcional.

Os resultados obtidos revelam um bom desempenho da arquitetura U-Net na tarefa de segmentação, evidenciado pela métrica quantitativa do Coeficiente de Dice e também ao comparar com os resultados de outras arquiteturas. As concordâncias entre as máscaras preditas e as máscaras de referência validam a performance do modelo em destacar com precisão a região da bexiga nas imagens de TC.

É esperado que o aplicativo web desenvolvido possa ser aprimorado e se torne uma ferramenta ainda mais prática e acessível para os especialistas em Radioterapia. A interface considerada acessível pelos usuários, juntamente com o desempenho da arquitetura U-Net, representam uma contribuição na aplicação de técnicas de aprendizado profundo na prática clínica.

Portanto, os resultados obtidos destacam que os objetivos propostos foram alcançados.

Contudo, é importante reconhecer as limitações do estudo, incluindo a dependência da qualidade do conjunto de dados e possíveis desafios em cenários clínicos específicos. Essas considerações ressaltam a importância de pesquisas futuras para aprimorar as capacidades do modelo em condições mais diversas, como treinar o modelo com imagens de dimensão maior; treinar o modelo para segmentar outra região; aplicar outras técnicas de pré-processamento e de aumento de dados, assim como usar outra arquitetura.

Em suma, esta dissertação contribui para o campo emergente da segmentação médica usando o aprendizado profundo, oferecendo uma solução prática para a identificação não-supervisionada da bexiga em imagens de TC *Cone-Beam*. O potencial clínico deste trabalho representa um passo significativo em direção à integração de tecnologias de IA na prática médica.

Lista de Referências

- [1] Instituto Nacional de Câncer - INCA. *ESTIMATIVA 2023 - Incidência de Câncer no Brasil*. 23 nov. 2022. Disponível em: <<https://www.gov.br/inca/pt-br/assuntos/noticias/2022/inca-estima-704-mil-casos-de-cancer-por-ano-no-brasil-ate-2025>>. Acesso em: 2 janeiro 2023. (Citado na página: [1])
- [2] World Health Organization - WHD. *REDUCING the cancer burden*. 3 fev. 2022. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/cancer>>. Acesso em: 2 janeiro 2023. (Citado na página: [1])
- [3] BENTZEN, S. M. et al. *Radiation Oncology Advances*. 1st. ed. New York: Springer Science, 2008. (Citado na página: [1])
- [4] GUNDERSON, L. L.; TEPPER., J. E. *Clinical Radiation Oncology*. 3rd. ed. Carolina: Saunders North, 2012. (Citado na página: [1])
- [5] KALET, A. M.; LUK, S. M. H.; PHILLIPS, M. H. Radiation therapy quality assurance tasks and tools: The many roles of machine learning. *Medical Physics*, v. 47, n. 5, p. e168–e177, 2020. Disponível em: <<https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.13445>>. (Citado na página: [1])
- [6] DE FAUW, J. et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature medicine*, v. 24, 2018. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/30104768/>>. (Citado na página: [1])
- [7] ZHANG, K. et al. Clinically applicable ai system for accurate diagnosis, quantitative measurements, and prognosis of covid-19 pneumonia using computed tomography. *Cell*, v. 181, 2020. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/32416069/>>. (Citado na página: [1])
- [8] WANG, C. et al. Artificial intelligence in radiotherapy treatment planning: Present and future. *Technology in Cancer Research & Treatment*, v. 18, 2019. Disponível em: <<https://doi.org/10.1177/1533033819873922>>. (Citado na página: [1])
- [9] GREENFIELD, D. *ARTIFICIAL Intelligence in Medicine: Applications, implications, and limitations*. Science in the News - Harvard, jun. 2019. Disponível em: <<http://sitn.hms.harvard.edu/flash/2019/artificial-intelligence-in-medicine-applications-implications-and-limitations/>>. (Citado na página: [1])
- [10] WALL, P. D.; FONTENOT, J. D. Application and comparison of machine learning models for predicting quality assurance outcomes in radiation therapy treatment planning. *Informatics in Medicine Unlocked*, v. 18, p. 100292, 2020.

ISSN 2352-9148. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S2352914819303661>>. (Citado na página: [2])

[11] ERICKSON, B. J. et al. Machine learning for medical imaging. *RadioGraphics*, v. 37, n. 2, p. 505–515, 2017. Disponível em: <<https://doi.org/10.1148/rg.2017160130>>. (Citado na página: [2])

[12] LEE, H.; CHEN, Y.-P. P. Image based computer aided diagnosis system for cancer detection. *Expert Systems with Applications*, v. 42, n. 12, p. 5356 – 5365, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417415000986>>. (Citado na página: [2])

[13] SIDDIQUE, S.; CHOW, J. C. Artificial intelligence in radiotherapy. *Reports of Practical Oncology & Radiotherapy*, v. 25, n. 4, p. 656 – 666, 2020. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1507136720300444>>. (Citado na página: [2])

[14] DU, G. et al. Medical image segmentation based on U-Net: A review. *Journal of Imaging Science and Technology*, mar. 2020. Disponível em: <https://www.researchgate.net/publication/339929534_Medical_Image_Segmentation_based_on_U-Net_A_Review#:~:text=Conclusions%20The%20nnU%2DNet%20is,a%20modestly%20sized%20training%20set.> (Citado na página: [2])

[15] NICHOLS, J.; HERBERT, C.; BAKER, M. Machine learning: applications of artificial intelligence to imaging and diagnosis. *Biophysical reviews*, v. 11(1), p. 111–118, 2019. (Citado na página: [2])

[16] HE, K. et al. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. p. 770–778. (Citado na página: [2])

[17] GSAXNER, C. et al. Exploit fully automatic low-level segmented pet data for training high-level deep learning algorithms for the corresponding ct data. *PLoS ONE*, Public Library of Science, v. 14, n. 3, mar 2019. (Citado na página: [2])

[18] CARDENAS, C. E. et al. Advances in auto-segmentation. *Seminars in Radiation Oncology*, v. 29, n. 3, p. 185–197, 2019. ISSN 1053-4296. Adaptive Radiotherapy and Automation. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1053429619300104>>. (Citado na página: [2])

[19] BRODY, H. Medical imaging. *Nature*, Nature Publishing Group UK London, v. 502, n. 7473, p. S81–S81, 2013. (Citado na página: [2])

[20] YORIYAZ, H.; FONSECA, G. P.; BELLEZZO, M. Sistemas de planejamento em radioterapia. *Revista Brasileira de Física Médica*, v. 13, n. 1, p. 92–98, set. 2019. Disponível em: <<https://www.rbfm.org.br/rbfm/article/view/512>>. (Citado na página: [2])

[21] SOUZA, C. N. d.; MONTI, C. R.; SIBATA, C. H. Recomendações para se evitar grandes erros de dose em tratamentos radioterapêuticos. *Radiologia Brasileira*, Publicação do Colégio Brasileiro de Radiologia e Diagnóstico por Imagem, v. 34, n. 1, p. 29–37, jan. 2001. ISSN 0100-3984. Disponível em:

<<https://www.scielo.br/j/rb/a/9HhgvXhRBrJP6pGV4CRhhBz/>>. (Citado na página: [2])

[22] SHARP, G. et al. Vision 20/20: Perspectives on automated image segmentation for radiotherapy. *Medical physics*, v. 41, n. 5, p. 050902, maio 2014. (Citado na página: [2])

[23] LIU, X. et al. Review of deep learning based automatic segmentation for lung cancer radiotherapy. *Frontiers in Oncology*, v. 11, jul. 2021. ISSN 2234-943X. Disponível em: <<https://www.frontiersin.org/article/10.3389/fonc.2021.717039>>. (Citado na página: [2])

[24] WONG, J. et al. Comparing deep learning-based auto-segmentation of organs at risk and clinical target volumes to expert inter-observer variability in radiotherapy planning. *Radiotherapy and Oncology*, v. 144, p. 152–158, 2020. ISSN 0167-8140. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167814019334668>>. (Citado na página: [2])

[25] LIU, X. et al. A review of deep-learning-based medical image segmentation methods. *Sustainability*, MDPI AG, v. 13, n. 3, p. 1224, Jan 2021. ISSN 2071-1050. Disponível em: <<http://dx.doi.org/10.3390/su13031224>>. (Citado 3 vezes nas páginas: [3, 16 e 22])

[26] SANTOS, C. M. da C.; PIMENTA, C. A. de M.; NOBRE, M. R. C. The pico strategy for the research question construction and evidence search. *Revista Latino-Americana de Enfermagem*, v. 15, n. 3, p. 508–511, maio-jun. 2007. ISSN 1518-8345. (Citado na página: [3])

[27] MOHER, D. et al. Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *PLOS Medicine*, Public Library of Science, v. 6, n. 7, p. 1–6, set. 2009. Disponível em: <<https://doi.org/10.1371/journal.pmed.1000097>>. (Citado 2 vezes nas páginas: [3, 5])

[28] U.S. National Library of Medicine. *Medical Subject Headings (MeSH)*. 1960. <<https://www.nlm.nih.gov/mesh/meshhome.html>>. (Citado na página: [3])

[29] Biblioteca Virtual em Saúde. *Descritores em Ciências da Saúde (DeCS)*. 2023. <<https://decs.bvsalud.org/>>. (Citado na página: [3])

[30] Elsevier. *Emtree*. 1988. <<https://www.elsevier.com/solutions/embase-biomedical-research/embase-coverage-and-content>>. (Citado na página: [3])

[31] SILVA, H. *Ilustrações da Dissertação*. 2023. Disponível em: <<https://hevenicio.github.io/index.html>>. (Citado 15 vezes nas páginas: [5, 6, 13, 14, 15, 16, 19, 24, 29, 32, 33, 34, 35, 37 e 38])

[32] RHEE, D. J. et al. Automatic contouring system for cervical cancer using convolutional neural networks. *Medical Physics*, John Wiley and Sons Ltd, v. 47, n. 11, p. 5648–5658, nov. 2020. ISSN 00942405. (Citado 4 vezes nas páginas: [6, 7, 44 e 48])

- [33] ZHANG, Z. et al. Arpm-net: A novel cnn-based adversarial method with markov random field enhancement for prostate and organs at risk segmentation in pelvic ct images. *Medical Physics*, John Wiley and Sons Ltd, v. 48, n. 1, p. 227–237, jan. 2021. ISSN 00942405. (Citado 4 vezes nas páginas: [6, 7, 44 e 48])
- [34] FU, Y. et al. Pelvic multi-organ segmentation on cone-beam CT for prostate adaptive radiotherapy. *Medical Physics*, John Wiley and Sons Ltd., v. 47, n. 8, p. 3415–3422, ago. 2020. ISSN 00942405. (Citado 4 vezes nas páginas: [6, 7, 44 e 48])
- [35] KAZEMIFAR, S. et al. Segmentation of the prostate and organs at risk in male pelvic CT images using deep learning. *Biomedical Physics and Engineering Express*, Institute of Physics Publishing, v. 4, n. 5, jul. 2018. ISSN 20571976. (Citado 5 vezes nas páginas: [6, 7, 39, 44 e 48])
- [36] AMJAD, A. et al. General and custom deep learning autosegmentation models for organs in head and neck, abdomen, and male pelvis. *Medical Physics*, John Wiley and Sons Ltd, v. 49, n. 3, p. 1686–1700, mar. 2022. ISSN 00942405. (Citado 4 vezes nas páginas: [6, 7, 44 e 48])
- [37] RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015. Disponível em: <<http://arxiv.org/abs/1505.04597>>. (Citado 4 vezes nas páginas: [8, 36, 37 e 41])
- [38] BACCOUCH, W. et al. A comparative study of CNN and U-Net performance for automatic segmentation of medical images: application to cardiac mri. *Procedia Computer Science*, v. 219, p. 1089–1096, 2023. ISSN 1877-0509. CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN – International Conference on Project MANAgement / HCist – International Conference on Health and Social Care Information Systems and Technologies 2022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050923003976>>. (Citado 3 vezes nas páginas: [8, 40 e 41])
- [39] ILHAN, A. et al. Covid-19 lung ct image segmentation using localization and enhancement methods with U-Net. *Procedia Computer Science*, v. 218, p. 1660–1667, 2023. ISSN 1877-0509. International Conference on Machine Learning and Data Engineering. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050923001448>>. (Citado na página: [8])
- [40] CHEN, G. et al. Rethinking the unpretentious u-net for medical ultrasound image segmentation. *Pattern Recognition*, v. 142, p. 109728, 2023. ISSN 0031-3203. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0031320323004260>>. (Citado na página: [8])
- [41] ZHANG, J. et al. St-unet: Swin transformer boosted U-Net with cross-layer feature enhancement for medical image segmentation. *Computers in Biology and Medicine*, v. 153, p. 106516, 2023. ISSN 0010-4825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010482522012240>>. (Citado na página: [8])
- [42] WANG, S. et al. Stacked dilated convolutions and asymmetric architecture for U-Net-based medical image segmentation. *Computers in Biology and Medicine*, v. 148,

p. 105891, 2022. ISSN 0010-4825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010482522006308>>. (Citado na página: [8])

[43] WANG, Z.; ZOU, Y.; LIU, P. X. Hybrid dilation and attention residual U-Net for medical image segmentation. *Computers in Biology and Medicine*, v. 134, p. 104449, 2021. ISSN 0010-4825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010482521002432>>. (Citado na página: [8])

[44] LI, C.; LI, X.; ZHOU, R. Cochlear ct image segmentation based on u-net neural network. *Journal of Radiation Research and Applied Sciences*, v. 16, n. 2, p. 100560, 2023. ISSN 1687-8507. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1687850723000389>>. (Citado na página: [8])

[45] GALLITTO, M. et al. Factors contributing to bladder filling variability in patients receiving radiotherapy for prostate cancer. *International Journal of Radiation Oncology, Biology, Physics*, Elsevier, v. 111, n. 3, p. e523, nov. 2021. ISSN 0360-3016. Disponível em: <<https://doi.org/10.1016/j.ijrobp.2021.07.1429>>. (Citado na página: [9])

[46] ACKLIN-WEHNERT, S.; AL., C. et. Toxicity and outcomes of moderately hypofractionated radiation for prostate cancer with seminal vesicle involvement. *Advances in Radiation Oncology*, Elsevier, v. 8, n. 5, p. 101252, set. - out. 2023. ISSN 2452-1094. Disponível em: <<https://doi.org/10.1016/j.adro.2023.101252>>. (Citado na página: [9])

[47] CHAUHAN, K. et al. Assessment of bladder filling during prostate cancer radiation therapy with ultrasound and cone-beam ct. *Frontiers in Oncology*, v. 13, p. 1200270, jul. 2023. (Citado na página: [9])

[48] NASSER, N. J. et al. Maintaining consistent bladder filling during external beam radiotherapy for prostate cancer. *Technical Innovations & Patient Support in Radiation Oncology*, v. 17, p. 1–4, jan. 2021. (Citado na página: [9])

[49] MAYYAS, E. et al. Evaluation of multiple image-based modalities for image-guided radiation therapy (igrt) of prostate carcinoma: a prospective study. *Med Phys*, v. 40, n. 4, p. 041707, abr. 2013. (Citado na página: [9])

[50] HAWORTH, A. et al. Prostate bed radiation therapy: the utility of ultrasound volumetric imaging of the bladder. *Clin Oncol (R Coll Radiol)*, v. 26, n. 12, p. 789–796, dez. 2014. Epub 2014 set. 18. (Citado na página: [9])

[51] CHEN, Z. et al. Dosimetric impact of different bladder and rectum filling during prostate cancer radiotherapy. *Radiation Oncology*, BioMed Central, v. 11, n. 1, p. 103, ago. 2016. ISSN 1748-717X. Disponível em: <<https://doi.org/10.1186/s13014-016-0681-z>>. (Citado na página: [9])

[52] KUPELIAN, P. A. et al. Daily variations in delivered doses in patients treated with radiotherapy for localized prostate cancer. *International Journal of Radiation Oncology*Biological*Physics*, v. 66, n. 3, p. 876–882, 2006. ISSN 0360-3016. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0360301606009795>>. (Citado 2 vezes nas páginas: [9, 11])

- [53] ROESKE, J. C. et al. Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy. *International Journal of Radiation Oncology, Biology, Physics*, Elsevier, v. 33, n. 5, p. 1321–1329, dez. 1995. ISSN 0360-3016. Disponível em: <[https://doi.org/10.1016/0360-3016\(95\)00225-1](https://doi.org/10.1016/0360-3016(95)00225-1)>. (Citado na página: [9])
- [54] JUNEJA, P. et al. Adaptive radiotherapy for bladder cancer using deformable image registration of empty and full bladder. In: MAGJAREVIĆ, R. (Ed.). *IFMBE Proceedings*. Jaffray, D. (eds) World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada: Springer, 2015. v. 51. (Citado 2 vezes nas páginas: [9, 10])
- [55] DUTTA, S. et al. Dosimetric impact of variable bladder filling on imrt planning for locally advanced carcinoma cervix. *Journal of the Egyptian National Cancer Institute*, v. 32, n. 1, p. 29–37, jul. 2020. Disponível em: <<https://doi.org/10.1186/s43046-020-00033-5>>. (Citado na página: [9])
- [56] MA, Z. et al. IEEE Access Special Section Editorial: Recent advantages of computer vision. *IEEE Access*, v. 6, p. 31481–31485, 2018. (Citado na página: [13])
- [57] International Business Machines Corporation (IBM). *WHAT is computer vision?* 2019. Disponível em: <<https://www.ibm.com/topics/computer-vision#anchor--1327322358>>. Acesso em: 3 janeiro 2023. (Citado na página: [13])
- [58] NVIDIA. *COMPUTER Vision*. 2020. Disponível em: <<https://developer.nvidia.com/computer-vision>>. Acesso em: 3 janeiro 2023. (Citado na página: [13])
- [59] VIJ, R.; ARORA, S. Computer vision with deep learning techniques for neurodegenerative diseases analysis using neuroimaging: A survey. In: *International Conference on Innovative Computing and Communications*. Shaheed Sukhdev College of Business Studies, University of Delhi, Delhi, India, February 19–20: Springer, 2022. p. 179–189. ISBN 978-981-16-2596-1. (Citado na página: [13])
- [60] IBRAHIM, M. R.; HAWORTH, J.; CHENG, T. Understanding cities with machine eyes: A review of deep computer vision in urban analytics. *Cities*, v. 96, p. 102481, 2020. ISSN 0264-2751. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0264275119308443>>. (Citado na página: [13])
- [61] NÓBREGA, R. *Building interactive spatial and temporal models using multimodal data*. Tese (Doutorado em Ciência da Computação) — Departamento de Informática, Faculdade de Ciências e Tecnologia - FCT, Universidade Nova de Lisboa, Portugal, jan. 2009. (Citado na página: [13])
- [62] GONZALEZ, R. C.; WOODS, R. E. *Digital image processing*. 3rd. ed. [S.l.]: Pearson, 2007. (Citado 4 vezes nas páginas: [14, 15, 16 e 17])
- [63] SIDDIQI, A.; AL-LAWATI, M.; BOULBRACHENE, M. *Modern Engineering Mathematics*. Taylor & Francis, 2017. ISBN 9781498712057. Disponível em: <<https://books.google.com.br/books?id=vFQdtAEACAAJ>>. (Citado 2 vezes nas páginas: [14, 15])

- [64] Stanford Artificial Intelligence Laboratory's Outreach Summer camp (SAILORS). *Tutorial 1: Image Filtering*. 2015. Disponível em: <<https://ai.stanford.edu/~syyeung/cvweb/tutorial1.html>>. Acesso em: 20 dezembro 2022. (Citado 2 vezes nas páginas: [14, 15])
- [65] SONKA, M.; FITZPATRICK, J. *Handbook of Medical Imaging, Volume 2: Medical Image Processing and Analysis*. SPIE Press, 2000. (Handbook of medical imaging). ISBN 9780819477606. Disponível em: <<https://books.google.com.br/books?id=qVDFSAAACAAJ>>. (Citado na página: [16])
- [66] BANKMAN, I. N. *Handbook of Medical Image Processing and Analysis*. 2nd. ed. Burlington: Academic Press, 2009. ISBN 978-0-12-373904-9. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B978012373904950012X>>. (Citado 2 vezes nas páginas: [16, 17])
- [67] FARAG, A. A. et al. *Advanced Segmentation Techniques*. Boston, MA: Springer US, 2005. 479-533 p. ISBN 978-0-306-48551-0. Disponível em: <https://doi.org/10.1007/0-306-48551-6_9>. (Citado na página: [16])
- [68] HAQUE, I. R. I.; NEUBERT, J. Deep learning approaches to biomedical image segmentation. *Informatics in Medicine Unlocked*, v. 18, p. 100297, 2020. ISSN 2352-9148. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S235291481930214X>>. (Citado 3 vezes nas páginas: [17, 20 e 21])
- [69] BANDYOPADHYAY, H. *AN INTRODUCTION to Image Segmentation: Deep Learning vs. Traditional*. 2021. V7Labs. Disponível em: <<https://www.v7labs.com/blog/image-segmentation-guide#applications-of-image-segmentation>>. (Citado 2 vezes nas páginas: [17, 18])
- [70] GHARAJEH, M. S. Chapter eight - biological big data analytics. In: RAJ, P.; DEKA, G. C. (Ed.). *A Deep Dive into NoSQL Databases: The Use Cases and Applications*. Elsevier, 2018, (Advances in Computers, v. 109). p. 321–355. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0065245817300335>>. (Citado na página: [18])
- [71] MIURA, K. Measurements of intensity dynamics at the periphery of the nucleus. In: *Bioimage Data Analysis Workflows*. Springer, Cham, 2020. p. 9–32. ISBN 978-3-030-22386-1. Disponível em: <https://doi.org/10.1007/978-3-030-22386-1_2>. (Citado 2 vezes nas páginas: [18, 19])
- [72] WANG, Y. et al. Robust vessel segmentation in fundus images. *International Journal of Biomedical Imaging*, Hindawi Publishing Corporation, v. 2013, p. 154860, dez. 2013. ISSN 1687-4188. Disponível em: <<https://doi.org/10.1155/2013/154860>>. (Citado na página: [19])
- [73] GALDRAN, A. et al. State-of-the-art retinal vessel segmentation with minimalistic models. *Scientific Reports*, Nature Publishing Group, v. 12, n. 1, p. 6174, abr. 2022. ISSN 2045-2322. Disponível em: <<https://doi.org/10.1038/s41598-022-09675-y>>. (Citado na página: [19])

- [74] GUPTA, A. et al. Brain tumor segmentation from mri images using deep learning techniques. ArXiv, abs/2305.00257, 2023. Disponível em: <<https://doi.org/10.48550/arXiv.2305.00257>>. (Citado na página: [19])
- [75] SHARMA, N.; AGGARWAL, L. M. Automated medical image segmentation techniques. *Journal of Medical Physics / Association of Medical Physicists of India*, v. 35, p. 3–14, 2010. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2825001/>>. (Citado na página: [20])
- [76] HARRISON, K. et al. Machine learning for auto-segmentation in radiotherapy planning. *Clinical Oncology*, v. 34, n. 2, p. 74–88, 2022. ISSN 0936-6555. Artificial Intelligence in Radiation Therapy. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0936655521004635>>. (Citado na página: [20])
- [77] IŞIN, A.; DIREKOĞLU, C.; ŞAH, M. Review of mri-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science*, v. 102, p. 317–324, 2016. ISSN 1877-0509. 12th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2016, 29-30 August 2016, Vienna, Austria. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S187705091632587X>>. (Citado 2 vezes nas páginas: [20, 21])
- [78] STARMANS, M. P. et al. Chapter 18 - Radiomics: Data mining using quantitative medical image features. In: ZHOU, S. K.; RUECKERT, D.; FICHTINGER, G. (Ed.). *Handbook of Medical Image Computing and Computer Assisted Intervention*. Academic Press, 2020, (The Elsevier and MICCAI Society Book Series). p. 429–456. ISBN 978-0-12-816176-0. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780128161760000235>>. (Citado 3 vezes nas páginas: [20, 21 e 22])
- [79] CASTIGLIONI, I. et al. AI applications to medical images: From machine learning to deep learning. *Physica Medica*, v. 83, p. 9–24, 2021. ISSN 1120-1797. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1120179721000946>>. (Citado 2 vezes nas páginas: [21, 22])
- [80] PARMAR, C. et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLOS ONE*, Public Library of Science, v. 9, n. 7, p. 1–8, jul. 2014. Disponível em: <<https://doi.org/10.1371/journal.pone.0102107>>. (Citado 2 vezes nas páginas: [21, 22])
- [81] MOSTAFA, A. M.; ZAKARIAH, M.; ALDAKHEEL, E. A. Brain tumor segmentation using deep learning on mri images. *Diagnostics*, v. 13, n. 9, 2023. ISSN 2075-4418. Disponível em: <<https://www.mdpi.com/2075-4418/13/9/1562>>. (Citado na página: [22])
- [82] WANG, R. et al. Medical image segmentation using deep learning: A survey. *IET Image Processing*, v. 16, n. 5, p. 1243–1267, 2022. Disponível em: <<https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/ipr2.12419>>. (Citado na página: [22])
- [83] MINH, T. N. et al. Automated image data preprocessing with deep reinforcement learning. *CoRR*, abs/1806.05886, 2018. Disponível em: <<http://arxiv.org/abs/1806.05886>>. (Citado 2 vezes nas páginas: [22, 43])

- [84] SHI, F. et al. Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for covid-19. *IEEE Reviews in Biomedical Engineering*, v. 14, p. 4–15, 2021. (Citado na página: [22])
- [85] NORVIG, P.; RUSSELL, S. J. *Inteligência Artificial: Tradução da 3ª Edição*. [S.l.]: Elsevier, 2013. 1016 p. ISBN 978-85-352-3701-6. (Citado na página: [23])
- [86] JOINER, I. A. Chapter 1 - Artificial Intelligence: AI is nearby. In: *Emerging Library Technologies*. Chandos Publishing, 2018, (Chandos Information Professional Series). p. 1–22. ISBN 978-0-08-102253-5. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B9780081022535000022>>. (Citado na página: [23])
- [87] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. MIT Press, 2016. (Adaptive computation and machine learning). ISBN 9780262035613. Disponível em: <<https://books.google.co.in/books?id=Np9SDQAAQBAJ>>. (Citado 9 vezes nas páginas: [24, 25, 26, 28, 31, 32, 33, 43 e 44])
- [88] MITCHELL, T. M. *Machine learning*. New York: McGraw-hill, 1997. v. 1. (Citado na página: [24])
- [89] GÉRON, A. *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems*. Sebastopol, California: O'Reilly Media, 2017. ISBN 978-1491962299. (Citado na página: [24])
- [90] ALPAYDIN, E. *Introduction to Machine Learning*. Cambridge, MA: MIT press, 2010. (Citado 4 vezes nas páginas: [24, 25, 26 e 27])
- [91] BISHOP, C. M. *Pattern recognition and machine learning*. New York: Springer, 2006. (Citado 4 vezes nas páginas: [24, 26, 27 e 31])
- [92] MURPHY, K. P. *Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series)*. Cambridge, MA: The MIT Press, 2012. (Citado na página: [24])
- [93] HASTIE, T.; TIBSHIRAN, R.; FRIEDMAN, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2016. Disponível em: <<https://web.stanford.edu/~hastie/ElemStatLearn/>>. (Citado 3 vezes nas páginas: [25, 26 e 27])
- [94] SARKER, I. H. Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, Springer, v. 2, n. 3, p. 160, 2021. ISSN 2661-8907. Publicado online: 22 março 2021. Disponível em: <<https://doi.org/10.1007/s42979-021-00592-x>>. (Citado na página: [25])
- [95] KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F. et al. (Ed.). *Advances in neural information processing systems*. Lake Tahoe, NV: Curran Associates, Inc., 2012. v. 25, p. 1097–1105. ISBN 9781627480031. Disponível em: <<http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks>>. (Citado 4 vezes nas páginas: [26, 33, 34 e 36])

- [96] POZZOLO, A. D. et al. Credit card fraud detection: A realistic modeling and a novel learning strategy. *IEEE Trans. Neural Networks Learn. Syst.*, v. 29, n. 8, p. 3784–3797, 2018. Disponível em: <http://dblp.uni-trier.de/db/journals/tnn/tnn29.html#PozzoloBCAB18>. (Citado na página: [26])
- [97] ESTEVA, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, Macmillan Publishers Limited, part of Springer Nature. All rights reserved., v. 542, p. 115–118, jan. 2017. Disponível em: <http://dx.doi.org/10.1038/nature21056>. (Citado na página: [26])
- [98] BENGIO, Y.; COURVILLE, A.; VINCENT, P. Representation learning: A review and new perspectives. ArXiv, abs/1206.5538, 2012. Disponível em: <http://arxiv.org/abs/1206.5538>. (Citado na página: [26])
- [99] HINTON, G. E.; SALAKHUTDINOV, R. R. Reducing the dimensionality of data with neural networks. *Science*, v. 313, n. 5786, p. 504–507, jul. 2006. Disponível em: <http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed&uid=16873662&cmd=showdetailview&indexed=google>. (Citado na página: [27])
- [100] WU, X. et al. Data mining with big data. *IEEE Trans. Knowl. Data Eng.*, v. 26, n. 1, p. 97–107, 2014. Disponível em: <http://dblp.uni-trier.de/db/journals/tkde/tkde26.html#WuZW014>. (Citado na página: [27])
- [101] SUTTON, R. S.; BARTO, A. G. *Reinforcement Learning: An Introduction*. 2nd. ed. The MIT Press, 2018. Disponível em: <http://incompleteideas.net/book/the-book-2nd.html>. (Citado 2 vezes nas páginas: [27, 28])
- [102] MNIH, V. et al. Human-level control through deep reinforcement learning. *Nature*, v. 518, n. 7540, p. 529–533, 2015. ISSN 14764687. Disponível em: <https://doi.org/10.1038/nature14236>. (Citado na página: [28])
- [103] WATKINS, C. J. C. H.; DAYAN, P. Q-learning. *Machine Learning*, v. 8, n. 3, p. 279–292, maio 1992. ISSN 1573-0565. Disponível em: <https://doi.org/10.1007/BF00992698>. (Citado na página: [28])
- [104] RUMMERY, G. A.; NIRANJAN, M. *On-Line Q-Learning Using Connectionist Systems*. Cambridge, England, 1994. (Citado na página: [28])
- [105] SILVER, D. et al. Mastering the game of Go with deep neural networks and tree search. *Nature*, v. 529, n. 7587, p. 484–489, 2016. ISSN 14764687. Disponível em: <https://doi.org/10.1038/nature16961>. (Citado na página: [28])
- [106] KOBER, J.; BAGNELL, J. A.; PETERS, J. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, SAGE Publications, v. 32, n. 11, p. 1238–1274, ago. 2013. Disponível em: <https://doi.org/10.1177/2F0278364913495721>. (Citado na página: [28])
- [107] HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. New York: Macmillan, 1994. (Citado 3 vezes nas páginas: [28, 29 e 31])
- [108] STERGIOU, C.; SIGANOS, D. Neural networks. 2014. Disponível em: <http://srii.sou.edu.ge/neural-networks.pdf>. (Citado 2 vezes nas páginas: [28, 29])

- [109] ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, v. 65, n. 6, p. 386–408, 1958. (Citado na página: [28])
- [110] BISHOP, C. M. *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press, 1995. (Citado na página: [28])
- [111] JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. *Science*, v. 349, n. 6245, p. 255–260, 2015. Disponível em: <<https://www.science.org/doi/abs/10.1126/science.aaa8415>>. (Citado na página: [28])
- [112] KANDEL, E. R. et al. The brain and behavior. In: SYDOR, A.; LEBOWITZ, H. (Ed.). *Principles of Neural Science*. New York, NY: McGraw-Hill Education, 2014. Disponível em: <accessbiomedicalscience.mhmedical.com/content.aspx?aid=1141370944>. (Citado na página: [29])
- [113] PURVES, D. et al. *Neuroscience*. Sunderland, MA: Sinauer Associates, Inc, 2004. (Citado na página: [29])
- [114] MCCULLOCH, W. S.; PITTS, W. H. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, v. 5, p. 115–133, 1943. (Citado na página: [29])
- [115] LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015. Disponível em: <<https://doi.org/10.1038/nature14539>>. (Citado 4 vezes nas páginas: [29, 32, 34 e 36])
- [116] AGGARWAL, C. C. An introduction to neural networks. In: *Neural Networks and Deep Learning: A Textbook*. Springer Cham, 2018. p. 1–52. ISBN 978-3-319-94463-0. Disponível em: <https://doi.org/10.1007/978-3-319-94463-0_1>. (Citado 6 vezes nas páginas: [30, 32, 33, 34, 35 e 36])
- [117] SARKER, I. H. Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, v. 2, n. 6, p. 420, ago. 2021. ISSN 2661-8907. Disponível em: <<https://doi.org/10.1007/s42979-021-00815-1>>. (Citado na página: [33])
- [118] ALZUBAIDI, L. et al. Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. *Journal of Big Data*, v. 8, n. 1, p. 53, mar. 2021. ISSN 2196-1115. Disponível em: <<https://doi.org/10.1186/s40537-021-00444-8>>. (Citado na página: [33])
- [119] LITJENS, G. et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, v. 42, p. 60–88, 2017. ISSN 1361-8415. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1361841517301135>>. (Citado na página: [36])
- [120] ZOU, K. H. et al. Statistical validation of image segmentation quality based on a spatial overlap index. *Academic Radiology*, v. 11, n. 2, p. 178–189, 2004. (Citado 2 vezes nas páginas: [37, 38])

- [121] POPOVIC, A. et al. Statistical validation metric for accuracy assessment in medical image segmentation. *International Journal of Computer Assisted Radiology and Surgery*, v. 2, n. 3, p. 169–181, 2007. ISSN 1861-6429. Disponível em: <<https://doi.org/10.1007/s11548-007-0125-1>>. (Citado na página: [37])
- [122] SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of Big Data*, v. 6, n. 1, p. 60, 2019. ISSN 2196-1115. Disponível em: <<https://doi.org/10.1186/s40537-019-0197-0>>. (Citado 2 vezes nas páginas: [40, 43])
- [123] GARCEA, F. et al. Data augmentation for medical imaging: A systematic literature review. *Computers in Biology and Medicine*, v. 152, p. 106391, 2023. ISSN 0010-4825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S001048252201099X>>. (Citado na página: [40])
- [124] NEMOTO, T. et al. Effects of sample size and data augmentation on U-Net-based automatic segmentation of various organs. *Radiological Physics and Technology*, Springer, v. 14, n. 3, p. 318–327, set. 2021. ISSN 1865-0341. Disponível em: <<https://doi.org/10.1007/s12194-021-00630-6>>. (Citado na página: [40])
- [125] CHEN, C.; FAN, L. Scene segmentation of remotely sensed images with data augmentation using u-net++. In: IEEE COMPUTER SOCIETY. *2021 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI)*. Shanghai, China, 2021. p. 201–205. (Citado na página: [40])
- [126] MONTAHA, S. et al. Mnet-10: A robust shallow convolutional neural network model performing ablation study on medical images assessing the effectiveness of applying optimal data augmentation technique. *Frontiers in Medicine*, v. 9, 2022. ISSN 2296-858X. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fmed.2022.924979>>. (Citado 2 vezes nas páginas: [40, 44])
- [127] LIVNE, M. et al. A U-Net deep learning framework for high performance vessel segmentation in patients with cerebrovascular disease. *Frontiers in Neuroscience*, v. 13, 2019. ISSN 1662-453X. Disponível em: <<https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2019.00097>>. (Citado na página: [41])
- [128] TAYLOR, L.; NITSCHKE, G. Improving deep learning using generic data augmentation. ArXiv, abs/1708.06020, ago. 2017. (Citado na página: [41])
- [129] EATON-ROSEN, Z. et al. Improving data augmentation for medical image segmentation. Semantic Scholar, ID: 14007201, abr. 2018. Disponível em: <<https://api.semanticscholar.org/CorpusID:14007201>>. (Citado na página: [41])
- [130] ALEXAKIS, E.; ARMENAKIS, C. Evaluation of UNet and UNet++ architectures in high resolution image change detection applications. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2020. Disponível em: <<https://api.semanticscholar.org/CorpusID:221694307>>. (Citado na página: [41])
- [131] MÜLLER, D.; KRAMER, F. MIScnn: a framework for medical image segmentation with convolutional neural networks and deep learning. *BMC Medical Imaging*, BioMed Central, v. 21, n. 1, p. 12, jan. 2021. ISSN 1471-2342. Disponível em: <<https://doi.org/10.1186/s12880-020-00543-7>>. (Citado na página: [41])

- [132] Alumentations Development Team. *ALBUMENTATIONS Documentation: API Reference - CLAHE*. 2022. Acessado em: 27 fevereiro 2024. Disponível em: https://alumentations.ai/docs/api_reference/augmentations/transforms/#alumentations.augmentations.transforms.CLAHE. (Citado na página: [41])
- [133] Alumentations Development Team. *ALBUMENTATIONS Documentation: API Reference - CenterCrop*. 2022. Acessado em: 27 fevereiro 2024. Disponível em: https://alumentations.ai/docs/api_reference/augmentations/crops/transforms/#alumentations.augmentations.crops.transforms.CenterCrop. (Citado na página: [41])
- [134] Alumentations Development Team. *ALBUMENTATIONS Documentation: API Reference - HorizontalFlip*. 2022. Acessado em: 27 fevereiro 2024. Disponível em: https://alumentations.ai/docs/api_reference/augmentations/transforms/#alumentations.augmentations.transforms.HorizontalFlip. (Citado na página: [41])
- [135] Alumentations Development Team. *ALBUMENTATIONS Documentation: API Reference - VerticalFlip*. 2022. Acessado em: 27 fevereiro 2024. Disponível em: https://alumentations.ai/docs/api_reference/augmentations/transforms/#alumentations.augmentations.transforms.VerticalFlip. (Citado na página: [41])
- [136] Alumentations Development Team. *ALBUMENTATIONS Documentation: Experimental API Reference - TransposeSymmetricKeypoints*. 2022. Acessado em: 27 fevereiro 2024. Disponível em: https://alumentations.ai/docs/experimental/api_reference/augmentations/transforms/#alumentations_experimental.augmentations.transforms.TransposeSymmetricKeypoints. (Citado na página: [42])
- [137] Alumentations Development Team. *ALBUMENTATIONS Documentation: API Reference - RandomRotate90*. 2022. Acessado em: 27 fevereiro 2024. Disponível em: https://alumentations.ai/docs/api_reference/augmentations/geometric/rotate/#alumentations.augmentations.geometric.rotate.RandomRotate90. (Citado na página: [42])
- [138] HANIN, B. Universal function approximation by deep neural nets with bounded width and relu activations. *Mathematics*, v. 7, n. 10, 2019. ISSN 2227-7390. Disponível em: <https://www.mdpi.com/2227-7390/7/10/992>. (Citado na página: [43])
- [139] GAILLOCHET, M.; DESROSIERS, C.; LOMBAERT, H. Active learning for medical image segmentation with stochastic batches. *Medical Image Analysis*, v. 90, p. 102958, 2023. ISSN 1361-8415. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1361841523002189>. (Citado 2 vezes nas páginas: [43, 44])
- [140] KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. ArXiv, abs/1412.6980, jan. 2017. Disponível em: <https://arxiv.org/abs/1412.6980>. (Citado na página: [43])
- [141] BENGIO, Y. Practical recommendations for gradient-based training of deep architectures. ArXiv, abs/1206.5533, 2012. Disponível em: <https://arxiv.org/abs/1206.5533>. (Citado na página: [43])

- [142] YE, Y. et al. Image segmentation using improved U-Net model and convolutional block attention module based on cardiac magnetic resonance imaging. *Journal of Radiation Research and Applied Sciences*, v. 17, n. 1, p. 100816, 2024. ISSN 1687-8507. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1687850723002947>>. (Citado na página: [44])
- [143] TAJBAKHSI, N. et al. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. *Medical Image Analysis*, v. 63, p. 101693, 2020. ISSN 1361-8415. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S136184152030058X>>. (Citado na página: [44])
- [144] MURAINA, I. Ideal dataset splitting ratios in machine learning algorithms: General concerns for data scientists and data analysts. In: *7th International Mardin Artuklu Scientific Researches Conference*. Mardin, Turkey: EPALE - Electronic Platform for Adult Learning in Europe, 2022. (Citado na página: [44])
- [145] CHU, T. et al. Improving weakly supervised lesion segmentation using multi-task learning. In: HEINRICH, M. et al. (Ed.). *Proceedings of the Fourth Conference on Medical Imaging with Deep Learning*. PMLR, 2021. (Proceedings of Machine Learning Research, v. 143), p. 60–73. Disponível em: <<https://proceedings.mlr.press/v143/chu21a.html>>. (Citado na página: [44])
- [146] BOLOCAN, V.-O. et al. Convolutional neural network model for segmentation and classification of clear cell renal cell carcinoma based on multiphase ct images. *Journal of Imaging*, v. 9, n. 12, 2023. ISSN 2313-433X. Disponível em: <<https://www.mdpi.com/2313-433X/9/12/280>>. (Citado na página: [44])