



UNIVERSIDADE FEDERAL DO MARANHÃO - UFMA
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
DOUTORADO EM CIÊNCIA DA COMPUTAÇÃO - ASSOCIAÇÃO UFMA/UFPI

LUANA BATISTA DA CRUZ

SEGMENTAÇÃO AUTOMÁTICA DE RINS E TUMORES RENAIOS EM
IMAGENS DE TOMOGRAFIA COMPUTADORIZADA BASEADA EM
APRENDIZADO PROFUNDO 2.5D

SÃO LUÍS - MA

2022

LUANA BATISTA DA CRUZ

**SEGMENTAÇÃO AUTOMÁTICA DE RINS E TUMORES RENAIOS EM
IMAGENS DE TOMOGRAFIA COMPUTADORIZADA BASEADA EM
APRENDIZADO PROFUNDO 2.5D**

Tese apresentada ao Curso de Doutorado em Ciência da Computação - Associação UFMA/UFPI como parte dos requisitos necessários para obtenção do grau de Doutora em Ciência da Computação.

Orientador: Prof. Dr. Aristófanes Corrêa Silva

Coorientador: Prof. Dr. João Dallyson Sousa de Almeida

SÃO LUÍS - MA

2022

Tese de autoria de Luana Batista da Cruz, sob o título “**Segmentação automática de rins e tumores renais em imagens de tomografia computadorizada baseada em aprendizado profundo 2.5D**”, apresentada ao Curso de Doutorado em Ciência da Computação - Associação UFMA/UFPI, como parte dos requisitos para obtenção do título de Doutora em Ciência da Computação, aprovada em _____ de _____ de _____ pela comissão julgadora constituída pelos doutores:

Prof. Dr. Aristófanes Corrêa Silva
Orientador
Universidade Federal do Maranhão

Prof. Dr. João Dallyson Sousa de Almeida
Coorientador
Universidade Federal do Maranhão

Profa. Dra. Aura Conci
Examinadora Externa
Universidade Federal Fluminense

Prof. Dr. Leandro Augusto Frata Fernandes
Examinador Externo
Universidade Federal Fluminense

Prof. Dr. Anselmo Cardoso de Paiva
Examinador Interno
Universidade Federal do Maranhão

Prof. Dr. Kelson Rômulo Teixeira Aires
Examinador Interno
Universidade Federal do Piauí

Ao meu pai (in memoriam) e mãe, irmão, família e amigos.

AGRADECIMENTOS

Em primeiro lugar, a Deus, por me iluminar e guiar. Obrigada Paizin, por mais uma benção concedida mesmo não sendo merecedora.

Aos meus pais, José Batista da Cruz (*in memoriam*) e Isabel Maria da Conceição, meus doutores da vida. Obrigada por todo amor e apoio imprescindível na minha vida desde sempre. Aos demais familiares, pelos incentivos e carinhos constantes. Em especial meu irmão José Lucas Batista da Cruz.

Ao meu orientador Aristófanes Corrêa Silva, por todos os conselhos, ensinamentos, confiança e credibilidade depositada em mim desde o início desta jornada. Ao meu coorientador João Dallyson Sousa de Almeida, pela ajuda, disponibilidade e conhecimento. A estes sempre terei respeito e admiração. Muito obrigada, professores!

Aos meus professores, pelo conhecimento que me transmitiram. Em especial, ao professor Anselmo Cardoso de Paiva, por sua contínua assistência durante minha pesquisa de mestrado e doutorado. Eu sou grata!

Aos meus amigos e colegas do NCA, pelos conselhos, ajudas, companheirismo, pela contribuição na minha formação pessoal e profissional. Em especial, a Domingos Alves Dias Junior, meu companheiro de estudo e vida. Obrigada por todos os momentos disponibilizados a me ajudar e suportar minhas tensões. Você faz parte dessa conquista!

À UFMA pela oportunidade concedida. E a CAPES pelo apoio financeiro durante o doutorado.

Agradeço a todos que de alguma forma contribuíram para minha formação acadêmica.

Muito obrigada, que Deus os abençoe sempre!

“Confia ao SENHOR as tuas obras, e os teus desígnios serão estabelecidos.”

Provérbios 16:3.

RESUMO

O câncer renal é um problema de saúde pública que afeta milhares de pessoas em todo o mundo. A segmentação precisa dos rins e dos tumores renais pode ajudar os especialistas a diagnosticar doenças e melhorar o planejamento do tratamento, o que é importante na prática clínica. No entanto, devido à heterogeneidade dos rins e dos tumores renais, a segmentação manual é um processo demorado e sujeito a variabilidade entre os especialistas. Devido a esse trabalho árduo, técnicas computacionais, como redes neurais convolucionais (*Convolutional Neural Networks - CNNs*), tornaram-se populares em tarefas de segmentação automática de rins e tumores renais. As redes tridimensionais (3D) apresentam bom desempenho em tarefas de segmentação, mas são complexas e apresentam altos custos computacionais. Assim, as redes bidimensionais são as mais usadas devido ao consumo de memória relativamente baixo, mas não exploram os recursos 3D. Portanto, nesta tese, redes 2.5D, que equilibram o consumo de memória e a complexidade do modelo, são propostas para auxiliar médicos especializados na detecção de rins e tumores renais em tomografia computadorizada (TC). Estas redes estão inseridas em um método proposto organizado em quatro etapas: (1) pré-processamento da base de imagens; (2) segmentação inicial dos rins e tumores renais usando os modelos ResUNet 2.5D e DeepLabv3+ 2.5D, respectivamente; (3) reconstrução dos tumores renais usando operação binária; e (4) redução de falsos positivos usando técnicas de processamento de imagens. O método proposto foi avaliado em 210 TCs da base de imagens KiTS19. Na segmentação dos rins, apresentou 97,45% de Dice, 95,05% de Jaccard, 99,95% de acurácia, 98,44% de sensibilidade e 99,96% de especificidade. Na segmentação dos tumores renais, foi obtido 84,06% de Dice, 75,04% de Jaccard, 99,94% de acurácia, 88,33% de sensibilidade e 99,95% de especificidade. De maneira geral, os resultados fornecem fortes evidências de que o método proposto é uma ferramenta com potencial para auxiliar no diagnóstico da doença.

Palavras-chave: Câncer de rim, Segmentação de rim, Segmentação de tumores renais, Redes neurais convolucionais, Aprendizado profundo, Tomografia computadorizada, Imagens médicas.

ABSTRACT

Kidney cancer is a public health problem that affects thousands of people worldwide. The precise segmentation of kidneys and kidney tumors can help medical specialists to diagnose diseases and improve treatment planning, which is highly required in clinical practice. However, because of the heterogeneity of kidneys and kidney tumors, manually segmenting is a time-consuming process and subject to variability among specialists. Because of this hard work, computational techniques, such as convolutional neural networks (CNNs), have become popular in automatic medical image segmentation tasks. Three-dimensional (3D) networks have a high segmentation capacity, but they are complex and have high computational costs. Thus, two-dimensional networks are the most used owing to the relatively low memory consumption, but they do not exploit 3D features. Therefore, in this thesis, 2.5D networks, that balances memory consumption and model complexity, are proposed to doctors specialized in the detection of kidneys and kidney tumors in computed tomography (CT). These networks are inserted in a proposed method organized in four steps: (1) image base pre-processing; (2) initial segmentation of kidneys and kidney tumors using ResUNet 2.5D and DeepLabv3+ 2.5D models, respectively; (3) reconstruction of kidney tumors using binary operation; and (4) reduction of false positives using image processing techniques. The proposed method was evaluated in 210 CTs from the KiTS19 image base. In the segmentation of the kidneys, it presented 97.45% of Dice, 95.05% of Jaccard, 99.95% of accuracy, 98.44% of sensitivity and 99.96% of specificity. In the segmentation of renal tumors, 84.06% of Dice, 75.04% of Jaccard, 99.94% accuracy, 88.33% sensitivity and 99.95% specificity were obtained. Overall, the results provide strong evidence that the proposed method is a powerful tool to help diagnose the disease.

Keywords: Kidney cancer, Kidney segmentation, Kidney tumor segmentation, Convolutional neural networks, Deep learning, Computed tomography, Medical images.

LISTA DE ILUSTRAÇÕES

Figura 1 – Estrutura interna do rim.	32
Figura 2 – Equipamento de TC.	34
Figura 3 – Rede <i>Multilayer Perceptron</i>	38
Figura 4 – Exemplo de uma arquitetura CNN.	42
Figura 5 – Ilustração de uma operação de convolução.	43
Figura 6 – Ilustração da camada de subamostragem.	43
Figura 7 – Arquitetura U-Net.	45
Figura 8 – Arquitetura DeepLabv3+.	47
Figura 9 – Matriz de confusão.	49
Figura 10 – Etapas do método proposto.	52
Figura 11 – Método proposto para agrupar os exames semelhantes.	53
Figura 12 – Exemplo do método do cotovelo para seleção o número ideal de K-grupos.	55
Figura 13 – Pré-processamento: (a) especificação do histograma; (b) janelamento.	56
Figura 14 – Pré-processamento: (a) imagem original; (b) janelamento.	57
Figura 15 – Arquitetura ResUNet-101.	58
Figura 16 – Abordagem 2.5D: exemplo de imagens de entrada.	60
Figura 17 – Arquitetura DPN-131.	62
Figura 18 – Arquitetura DeepLabv3+.	63
Figura 19 – Delimitação da região de interesse.	63
Figura 20 – Candidatos a tumores renais: (a) marcação do especialista; (b) candidatos de tumores renais na região renal; (c) candidatos de tumores renais na região abdominal.	65
Figura 21 – Exemplos de casos difíceis. Rins marcados em azul e tumores renais em vermelho.	65
Figura 22 – Etapas da reconstrução de tumores renais.	66
Figura 23 – União da segmentação inicial dos rins e reconstrução dos tumores renais.	67
Figura 24 – Resultado final da segmentação de rins após pós-processamento.	68
Figura 25 – Intersecção da segmentação final dos rins e reconstrução dos tumores renais.	69
Figura 26 – Resultado final da segmentação de tumores renais após a aplicação do pós-processamento.	69

Figura 27 – Informações da base de imagens: fatia com marcações de rins (azul) e tumores (vermelho).	72
Figura 28 – Grupos de tumores encontrados com a aplicação do método a partir dos conjuntos de treinamento e validação.	74
Figura 29 – Etapas da segmentação dos rins. Marcação do especialista (verde) e segmentações do método (azul - rins, vermelho - tumores renais).	84
Figura 30 – Etapas da segmentação dos tumores renais. Marcação do especialista (verde) e segmentação do método (vermelho).	86
Figura 31 – Estudo de caso 1. Segmentações dos rins (cinza) e reconstrução dos tumores renais (verde).	94
Figura 32 – Estudo de caso 2. Segmentações dos rins (cinza) e reconstrução dos tumores renais (verde).	94
Figura 33 – Estudo de caso 3. Segmentação dos rins (cinza) e reconstrução dos tumores renais (verde).	95
Figura 34 – Estudo de caso 1. Segmentação dos rins (cinza), segmentação dos tumores renais (vermelho) e reconstrução dos tumores renais (verde).	96
Figura 35 – Estudo de caso 2. Segmentação dos rins (cinza), segmentação dos tumores renais (vermelho) e reconstrução dos tumores renais (verde).	97

LISTA DE TABELAS

Tabela 1 – Resumo dos trabalhos relacionados.	29
Tabela 2 – Aplicação dos pré-processamentos para segmentação dos rins e candidatos a tumores renais.	51
Tabela 3 – Total de características extraídas em cada modelo de CNN.	73
Tabela 4 – Melhores grupos e suas inércias encontrados para os modelos usando o método do cotovelo.	73
Tabela 5 – Resultados iniciais das etapas do método proposto.	76
Tabela 6 – Resultados finais do método proposto.	77
Tabela 7 – Resultados da etapa de segmentação inicial dos rins.	78
Tabela 8 – Resultados da etapa de segmentação de candidatos a tumores renais na região renal.	80
Tabela 9 – Resultados da etapa de segmentação de candidatos a tumores renais na região abdominal.	81
Tabela 10 – Resultados da etapa de reconstrução dos tumores renais.	82
Tabela 11 – Resultados da etapa de segmentação final dos rins.	83
Tabela 12 – Resultados da etapa de segmentação final dos tumores renais.	85
Tabela 13 – Resultado do método proposto com/sem balanceamento de fatias. . .	87
Tabela 14 – Resultado do método proposto com distribuição aleatória e proporcional.	89
Tabela 15 – Resultados da segmentação de rins com e sem a etapa de reconstrução tumoral.	89
Tabela 16 – Resultados da segmentação de tumores renais aplicando a máscara do método e do especialista.	90
Tabela 17 – Resumo dos diversos experimentos testados.	91
Tabela 18 – Comparação dos resultados do método com trabalhos relacionados à segmentação de rins.	98
Tabela 19 – Comparação dos resultados do método com trabalhos relacionados à segmentação de tumores renais.	100
Tabela 20 – Comparação dos resultados do método proposto com os trabalhos relacionados à segmentação de rins e tumores renais.	102
Tabela 21 – Produções científicas em relação ao método proposto para segmentação de rins e tumores renais.	111

Tabela 22 – Produções científicas em outras aplicações de processamento de imagens e visão computacional.	112
--	-----

LISTA DE ABREVIATURAS E SIGLAS

3D-MS-RFCNN	<i>Multi-Scale Residual Fully Convolutional Neural Network</i>
Acc	Acurácia
ASPP	<i>Atrous Spatial Pyramid Pooling</i>
CAD	<i>Computer Aided Diagnosis</i>
CADx	<i>Computer Aided Detection</i>
CNN	<i>Convolutional Neural Network</i>
DenseNet	<i>Densely Convolutional Network</i>
DNA	<i>Deoxyribonucleic Acid</i>
DPN	<i>Dual Path Network</i>
DRLSE	<i>Distance Regularized Level-Sets Evolution</i>
Esp	Especificidade
FCN	<i>Fully Convolutional Network</i>
FN	Falso Negativo
FP	Falso Positivo
FSGAN	<i>Feature Sharing Generative Adversarial Network</i>
GPU	<i>Graphics Processing Unit</i>
HU	<i>Hounsfield Units</i>
IA	Inteligência Artificial
IoU	<i>Intersection over Union</i>
Jacc	Jaccard
MB-FSGAN	<i>Multi-Branch Feature Sharing Generative Adversarial Network</i>
MLP	<i>Multilayer Perceptron</i>

MSFE	<i>Multi-Scale Feature Extractor</i>
MSS	<i>Multi-Scale Supervised</i>
NMF	<i>Nonnegative Matrix Factorization</i>
PPM	<i>Pyramid Pooling Module</i>
ReLU	<i>Rectified Linear Units</i>
ResNet	<i>Residual Network</i>
RM	Ressonância Magnética
RNAs	Redes Neurais Artificiais
SE	<i>Squeeze-and-Excitation</i>
Sen	Sensibilidade
SERU	SE-ResNeXT U-Net
SIFCM	<i>Spatial Intuitionistic Fuzzy C-Means Clustering</i>
SSCNs	<i>Submanifold Sparse Convolutional Networks</i>
SSN	<i>Sparse Segmentation Network</i>
TC	Tomografia Computadorizada
TPE	<i>Tree of Parzen Estimators</i>
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
VTP	<i>Voxel-to-Point</i>

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Objetivo Geral	19
1.2	Objetivos Específicos	19
1.3	Organização do Trabalho	20
2	TRABALHOS RELACIONADOS	21
2.1	Segmentação de Rins	21
2.2	Segmentação de Tumores Renais	24
2.3	Segmentação de Rins e Tumores Renais	25
2.4	Considerações Finais	29
3	FUNDAMENTAÇÃO TEÓRICA	31
3.1	Rins e Tumores Renais	31
3.2	Tomografia Computadorizada	33
3.3	Processamento Digital de Imagens	34
3.4	Especificação do Histograma	35
3.5	Redes Neurais Artificiais	36
3.5.1	Neurônio artificial	37
3.5.2	Redes <i>Multilayer Perceptron</i>	37
3.5.3	<i>Backpropagation</i>	39
3.6	Aprendizado Profundo	40
3.6.1	Redes Neurais Convolucionais	41
3.6.1.1	Camada de Convolução	42
3.6.1.2	Camada de Subamostragem	43
3.6.1.3	Camada Completamente Conectada	44
3.7	Arquitetura U-Net	44
3.8	Arquitetura DeepLabv3+	46
3.9	<i>Data Augmentation</i>	47
3.10	Métricas de Desempenho	48
3.11	Considerações Finais	50
4	MÉTODO PROPOSTO	51

4.1	Pré-processamento	51
4.1.1	Distribuição Proporcional da Base de Imagens	52
4.1.2	Pré-processamentos Aplicados a Segmentação dos Rins	56
4.1.3	Pré-processamentos Aplicados à Segmentação Inicial de Candidatos a Tumores Renais	57
4.2	Segmentação Inicial	57
4.2.1	Segmentação dos rins usando a ResUNet	58
4.2.1.1	Treinamento da ResUNet 2.5D	59
4.2.2	Segmentação Inicial de Candidatos a Tumores Renais	60
4.2.2.1	Candidatos de Tumores Renais na Região Renal	61
4.2.2.1.1	Treinamento da DeepLabv3+ 2.5D	62
4.2.2.2	Candidatos de Tumores Renais na Região Abdominal	64
4.3	Reconstrução dos Tumores Renais	65
4.4	Redução de Falsos Positivos	66
4.4.1	Redução de Falsos Positivos para os Rins	67
4.4.2	Redução de Falsos Positivos para os Tumores Renais	68
4.5	Considerações Finais	69
5	RESULTADOS	71
5.1	Base de Imagens	71
5.1.1	Distribuição Proporcional da Base de Imagens	72
5.1.2	Preparação da Base de Imagens	74
5.2	Configuração Experimental	74
5.3	Resultados Gerais do Método Proposto	75
5.4	Segmentação Inicial dos Rins	78
5.5	Segmentação Inicial de Candidatos a Tumores Renais	79
5.5.1	Candidatos a Tumores Renais na Região Renal	79
5.5.2	Candidatos a Tumores Renais na Região Abdominal	80
5.6	Reconstrução dos Tumores Renais	81
5.6.1	Redução de Falsos Positivos para os Rins	82
5.6.2	Redução de Falsos Positivos para os Tumores Renais	84
5.6.3	Método Proposto com/sem Balanceamento de Fatias	87

5.6.4	Método Proposto com Distribuição Aleatória e Proporcional da Base de Imagens	88
5.6.5	Segmentação de Rins sem/com a Etapa de Reconstrução de Tumores Renais	89
5.6.6	Segmentação de Tumores Renais aplicando a Máscara do Método/Especialista	90
5.7	Diversos Experimentos Testados	91
5.8	Considerações Finais	92
6	DISCUSSÃO	93
6.1	Estudos de Caso – Segmentação de Rins	93
6.1.1	Estudo de Caso 1 – Paciente case_00018	93
6.1.2	Estudo de Caso 2 – Paciente case_00149	93
6.1.3	Estudo de Caso 3 – Paciente case_00045	94
6.2	Estudos de Caso – Segmentação de Tumores Renais	95
6.2.1	Estudo de Caso 1 – Paciente case_00012	95
6.2.2	Estudo de Caso 2 – Paciente case_00028	96
6.3	Comparação do Método Proposto com os Trabalhos Relacionados	97
6.3.1	Segmentação de Rins	97
6.3.2	Segmentação de Tumores Renais	100
6.3.3	Segmentação de Rins e Tumores Renais	102
6.4	Aspectos Importantes do Método Proposto	105
7	CONCLUSÃO	109
7.1	Contribuições	109
7.2	Trabalhos Futuros	110
7.3	Produções Científicas	111
	REFERÊNCIAS	113

1 INTRODUÇÃO

Os rins são um par de órgãos localizados na parte posterior do abdômen e são protegidos pela caixa torácica. Esses órgãos são cercados por uma fina camada de tecido conjuntivo e por uma camada de gordura. Além disso, são responsáveis por filtrar o sangue das artérias renais, retirando o excesso de água, sal e resíduos excretando-os pela urina. São também órgãos que desempenham um papel fundamental no metabolismo da vitamina D, na fabricação de hormônios que estimulam a produção de hemárias e na regulação da pressão arterial (SOCIETY, 2017; FUND, 2019).

Portanto, os rins são extremamente importantes para o funcionamento do organismo. No entanto, alguns fatores de risco, como tabagismo, obesidade e hipertensão, podem afetar a função renal e levar ao câncer renal. Nesse sentido, mudanças no ácido desoxirribonucleico (*Deoxyribonucleic Acid - DNA*) que ativam oncógenes ou genes supressores de tumor podem fazer com que as células renais percam sua função primária e começem a se multiplicar rapidamente, causando tumores renais (SOCIETY, 2017; UK, 2019; HEALTH, 2021).

O câncer de rim (carcinoma de células renais) é um tipo de câncer que surge das células renais e pode crescer lentamente ou de forma mais agressiva. O câncer de rim geralmente cresce como uma única massa, mas outros tumores podem surgir em um ou nos dois rins. Em um estágio mais avançado, as células malignas podem cair na corrente sanguínea e afetar outros órgãos (SEYFRIED; HUYSENTRUYT, 2013; SHUCH et al., 2015). Este processo é chamado de metástase e afeta mais de 90% dos casos de câncer renal (SOCIETY, 2017; GHOSN et al., 2019). O tipo mais prevalente é chamado de carcinoma de células claras ou renais e representa cerca de 60-80% de todos os cânceres renais (BRAY et al., 2018; XU; ATKINS; MCDERMOTT, 2020).

Entre os tipos de câncer, o câncer renal é considerado o câncer mais letal do trato urinário (FERLAY et al., 2015), e a 16^a causa mais comum de morte por câncer (FUND, 2019). Embora 59% dos casos ocorram em países mais desenvolvidos, as taxas de sobrevivência são relativamente altas, enquanto em países de baixa renda as taxas de sobrevivência são baixas, pois a detecção de tumores ocorre em estágios mais avançados (FUND, 2019). Além disso, o câncer renal é o 9º câncer mais comum em homens e o 14º mais comum em mulheres (ZNAOR et al., 2015; ROSSI et al., 2018). De acordo

com as últimas pesquisas estatísticas, mais de 400.000 novos casos de câncer renal foram diagnosticados e quase 180.000 mortes relacionadas foram relatadas em 2018 (BRAY et al., 2018; FUND, 2019).

O diagnóstico precoce é fundamental para o tratamento adequado do câncer renal e aumenta as chances de cura. Além disso, o sucesso da radioterapia depende da segmentação precisa do tecido-alvo, pois são fornecidas informações estruturais sobre irregularidades na forma e medidas do tamanho dos tumores renais que podem ser usadas por especialistas para analisar condições clínicas graves (DALLAL et al., 2017). Para isso, a tomografia computadorizada (TC) é considerada o padrão ouro entre todas as modalidades de imagem disponíveis para investigação de tumores renais, especialmente em seu estágio inicial (KAUR; JUNEJA, 2016). A segmentação de tumores renais a partir de imagens de TC é uma tarefa essencial para o diagnóstico, tratamento e planejamento cirúrgico na urologia (YANG et al., 2014).

Devido à importância clínica em geral, muitos pesquisadores investem no uso de técnicas de processamento de imagens e aprendizado de máquina para desenvolver métodos computacionais que auxiliem especialistas na interpretação de imagens médicas (GWYNNE et al., 2012a; DINIZ et al., 2019; CRUZ et al., 2020). Essas técnicas computacionais são incorporadas em sistemas de Diagnóstico e de Detecção Auxiliados por Computador (*Computer Aided Diagnosis - CADx*) e (*Computer Aided Detection - CAD*) e são aplicadas a vários tipos de imagens médicas para fornecer suporte na análise de doenças (SELVANAYAKI; KARNAN, 2010; PADILLA et al., 2012; EL-DAHSHAN et al., 2014; GHAFOORIAN et al., 2017). Esses sistemas contribuem para reduzir a carga de trabalho do especialista e aumentar a eficiência da análise. Portanto, são úteis no processo de detecção e diagnóstico do câncer renal, por meio da pré-análise das informações fornecidas pelas imagens de TC.

As TCs são bastante usadas em métodos para segmentar os rins e detectar lesões renais, pois fornecem imagens de alta resolução com bons detalhes anatômicos. No entanto, essas imagens podem conter componentes distorcidos devido a artefatos encontrados na imagem, que incluem ruído, formato de anel, dispersão, pseudo-realce, artefatos metálicos, etc (BOAS; FLEISCHMANN, 2012). Um dos principais desafios da segmentação do rim na TC é incluir as regiões do campo renal recobertas por estruturas anormais. Em um cenário em que os pacientes são saudáveis ou, por exemplo, com pequenos nódulos, os métodos CAD geralmente são capazes de fornecer uma segmentação confiável, pois o contraste

entre as áreas renais e seus limites geralmente é mantido. No entanto, em um cenário mais complexo, os pacientes podem ter doenças que afetam seus rins com anormalidades, assim, a tarefa de segmentação do rim torna-se significativamente mais complexa (BERGERON, 2013; CANDEMIR et al., 2014). Na segmentação dos tumores renais, um dos principais desafios é delinear com precisão o limite da lesão renal entre os tecidos próximos, devido à presença de intensidade intrínseca na homogeneidade da textura do tecido ou limites que são inherentemente confusos (AGNELLO et al., 2020). Por essas razões, a tarefa de segmentar os rins e tumores renais torna-se significativamente complexa e desafiadora.

Dessa forma, métodos automáticos e robustos podem ser cruciais para superar essas dificuldades. Portanto, o problema abordado nesta tese consiste em propor um método automático capaz de segmentar com precisão rins e tumores renais em imagens tomográficas.

1.1 Objetivo Geral

O objetivo geral da tese proposta é desenvolver um método computacional para segmentar automaticamente os rins e tumores renais em TCs de pacientes doentes, usando principalmente os modelos ResUNet e DeepLabv3+ baseados em abordagem 2.5D, agrupamento de tumores renais acordo com os padrões identificados usando aprendizado de máquina e redução de falsos positivos com base em informações contextuais.

1.2 Objetivos Específicos

Para alcançar o objetivo geral deste trabalho, faz-se necessário atingir os seguintes objetivos específicos:

- Aplicar a técnica de aumento de dados em tempo real, tornando-a uma ferramenta poderosa para o método, com baixo consumo de recurso de máquina e maior capacidade de generalizar modelos de aprendizado profundo;
- Desenvolver um método automático que visa encontrar padrões de características existentes na base de imagens para organizá-la em grupos, de forma a garantir um modelo de aprendizagem profunda mais equilibrado e generalizado;
- Identificar os filtros mais efetivos para realçar os tumores renais e padronizar os rins;

- Compreender e aplicar as redes neurais convolucionais na tarefa de segmentar os rins e tumores renais em imagens de TC;
- Reconstruir regiões de tumores renais para recuperar tecido de tumores renais e rins não segmentados inicialmente, a fim de aumentar a precisão das segmentações dos rins e tumores renais;
- Validar o método proposto por meio de métricas comumente usadas em trabalhos de imagens médicas. Além disso, aplicar o teste de significância nas etapas do método para verificar se há diferença significativa entre elas;
- Avaliar o método proposto por meio da realização de experimentos, usando uma base de imagens públicas de TC;
- Comparar o método proposto com métodos já consolidados na literatura.

1.3 Organização do Trabalho

Os demais capítulos deste trabalho foram organizados em:

- O Capítulo 2 apresenta um resumo dos trabalhos relacionados à tarefa de segmentação de rins e tumores renais em imagens tomográficas.
- O Capítulo 3 trata dos conceitos fundamentais necessários para a construção desta pesquisa.
- O Capítulo 4 descreve todas as etapas do método proposto usado para a segmentação dos rins e tumores renais em imagens de TC.
- O Capítulo 5 mostra e discute os resultados alcançados no método proposto.
- O Capítulo 6 apresenta vários estudos de casos e um estudo comparativo com os trabalhos relacionados. Além disso, discute as vantagens e limitações encontradas no método proposto.
- Finalmente, o Capítulo 7 apresenta as considerações finais e sugestões de trabalhos futuros.

2 TRABALHOS RELACIONADOS

Na literatura existem estudos que tratam do mesmo problema abordado nesta proposta de tese, ou seja, métodos desenvolvidos para segmentação de rins e tumores renais em imagens de TC. As subseções a seguir apresentam resumos de alguns trabalhos relacionados, que se dividem em: segmentação de rins, segmentação de tumores renais, e segmentação de rins e tumores renais.

2.1 Segmentação de Rins

Nos últimos anos, a comunidade científica publicou vários métodos para segmentação de rins em TC. Khalifa et al. (2011) descrevem uma abordagem 3D para segmentação renal usando um modelo deformável baseado em *Level Set*. A abordagem proposta foi avaliada nos conjuntos de dados de TC de 29 pacientes, produzindo um Dice de 95% de rins. Os resultados apresentados indicam que combinar as características das imagens de TC na evolução do conjunto de níveis leva a resultados de segmentação mais precisos.

No trabalho de Wolz et al. (2013) foi apresentado um método automatizado para a segmentação de múltiplos órgãos de tomografias computadorizadas de abdômen. O método baseia-se em um esquema hierárquico de registro e ponderação de atlas. A segmentação final é obtida aplicando um modelo de intensidade aprendido, incorporando conhecimento espacial. Para avaliar o método, foram usados 150 exames de TC, resultando em um Dice de 92,5%.

No estudo de Zhao, Liang e Fan (2013) uma estrutura de segmentação baseada em fatias é construída para segmentar os rins automaticamente com base na continuidade contextual na sequência de imagens de TC. A estrutura tem quatro etapas principais: segmentação inicial, seleção da segmentação inicial mais confiável e localização e modificação de vazamentos. Um total de 10 pacientes são usados para avaliar a eficácia do método. O resultado final obtido foi de 94,7% de Dice.

Yang et al. (2014) propuseram um método automático baseado no registro de imagens multi-atlas. O método é baseado em duas etapas. A primeira etapa consiste no uso de atlas de baixa resolução para uma segmentação inicial. A partir da localização grosseira dos rins, a segunda etapa consiste em cortar os rins esquerdo e direito das imagens originais e alinhar com outro conjunto de imagens de atlas de alta resolução para

obter os resultados finais. Para avaliar o método, seis resultados de segmentação de rins com tumores renais grandes foram excluídos porque o tumor alterou significativamente a forma do rim. Finalmente, o método obteve 95,2% de Dice, usando um conjunto privado de 14 imagens.

No método de Khalifa et al. (2016) uma estrutura híbrida baseada na integração de modelos geométricos deformáveis e fatoração matricial não-negativa (*Nonnegative Matrix Factorization - NMF*) é usada para segmentação 3D dos rins. Além disso, um modelo de forma é construído usando o conjunto de treinamento e, durante a segmentação, um modelo baseado em aparência é atualizado, levando em consideração as localizações e aparências dos voxels. As interações espaciais são modeladas usando um modelo de campo aleatório de Potts Markov-Gibbs. Um conjunto de 36 exames de TC é usado para avaliar a abordagem proposta e apresentou 96,45% de Dice.

O método proposto por Skalski et al. (2017) é baseado em contorno ativo usando a estrutura *Level Set* com restrições de forma elipsoidal. A solução proposta leva em consideração as informações sobre a região e os termos de contorno, bem como as restrições dedicadas ao formato renal característico. O método proposto foi testado em 10 exames de pacientes com diagnóstico de câncer renal. A média em relação ao Dice foi igual a 86,2%.

O trabalho de Jackson et al. (2018) desenvolve uma ferramenta automatizada com base em redes neurais convolucionais (*Convolutional Neural Network - CNNs*) 3D capaz de segmentar o rim direito e esquerdo com precisão. Os autores optaram pela abordagem 3D por se mostrar mais adaptável a anormalidades estruturais mais complexas, como os cistos renais. Os experimentos foram realizados em TCs de 113 pacientes, obtendo-se resultados de 91% para o rim direito e 86% para o rim esquerdo, ou seja, 88,5% de Dice médio. O baixo desempenho foi observado em três pacientes com rins císticos.

Outro método é apresentado por Mehta et al. (2019). Os autores avaliam a precisão e a viabilidade de uso das segmentações renais por *crowdsourcing*. 42 TCs, foram rotuladas por 72 usuários na plataforma Robovision AI. As segmentações de *crowdsourcing* e as segmentações rotuladas por especialistas foram usadas individualmente e em conjunto como dados de treinamento para modelos separados da CNN. A validação primária foi realizada comparando as segmentações. O desempenho da CNN atingiu um resultado de 93,2% de Dice.

Les et al. (2021) apresentaram um sistema de apoio ao diagnóstico renal. O sistema permite a detecção e segmentação totalmente automática do rim usando um

conjunto de 270 imagens de TC. O método proposto tem duas etapas. Na primeira etapa, é realizada a segmentação semântica baseada na rede U-Net e *batch*. Na segunda etapa, um sistema de rastreamento (*middle-bottom, middle-top*) foi usado para preencher a imagem, o que permitiu encontrar o contorno do rim nas fatias restantes da TC. O método desenvolvido atingiu um Dice de 90,63%.

Ao analisar os trabalhos mencionados, observa-se que há desenvolvimento de métodos computacionais para segmentação de rins em imagens de TC. Além disso, é possível verificar que essa tarefa não é algo recente e trivial, ao longo dos anos vários métodos foram elaborados para a segmentação de rins. As abordagens baseadas em contorno ativo (SKALSKI et al., 2017) possuem dependência de pontos iniciais do contorno, parâmetros adequados, não seguem alterações topológicas de objetos e imagens de alta resolução requerem um longo tempo de execução. Além disso, são sensíveis aos estados mínimos locais, sua precisão depende da política de convergência (BAKOŠ, 2007).

Por sua vez, os métodos deformáveis baseiam-se principalmente no conhecimento ou usam atlas (WOLZ et al., 2013; YANG et al., 2014; YANG et al., 2018). A desvantagem desse tipo de abordagem é que geralmente possui um determinado tamanho para resolver o problema em um conjunto de imagens. Quando novas imagens de resoluções diferentes do atlas são testadas, o método não tem garantia de que funcionará (SOUZA et al., 2019). Além disso, técnicas que reduzem a interferência manual de médicos especialistas (MEHTA et al., 2019) para marcar a região de interesse não são triviais, pois podem ocorrer erros graves, que prejudicariam o modelo e, consequentemente, o diagnóstico médico.

Com o surgimento de *hardwares* cada vez mais sofisticados, começam a surgir trabalhos baseados em aprendizado profundo, que apresentam resultados expressivos em diferentes problemas. Trabalhos como Khalifa et al. (2011), Mehta et al. (2019) e Les et al. (2021) usam redes totalmente conectadas para segmentação de rins. A vantagem de usar essas técnicas é que não há necessidade de uma engenharia de características para determinar as melhores para que um classificador seja capaz de diferenciar classes. Ao usar camadas convolucionais, caberá à rede encontrar as melhores características para solucionar o problema.

2.2 Segmentação de Tumores Renais

Atualmente, vários estudos na literatura têm mostrado que o tumor renal é uma forma extremamente agressiva de câncer (MUGLIA; PRANDO, 2015; DREISIN; MATRANA, 2016; WARREN; HARRISON, 2018), e os métodos computacionais têm sido de grande importância para ajudar a acelerar a detecção precoce. No trabalho de Kaur, Juneja e Mandal (2019) uma técnica de segmentação híbrida foi proposta com base em dois métodos que incluem *Spatial Intuitionistic Fuzzy C-Means Clustering* (SIFCM) e *Distance Regularized Level-Sets Evolution* (DRLSE) para extração de lesões renais em imagens de TC. Um conjunto de 40 TCs foi usado para validar o método e atingiu um Dice de 88,2% para segmentação de lesões renais.

Yang et al. (2020) usaram um conjunto de CNNs fracamente supervisionada para segmentação de tumores renais. Para treinamento das CNNs, caixas delimitadoras de tumores renais foram fornecidas como anotações fracas. Com o intuito de melhorar o treinamento e a precisão da segmentação, propuseram duas fases de treinamento: em grupo e ponderado. No melhor cenário dos experimentos, apresentou resultado de 82,6% de Dice usando um conjunto de 200 TCs.

No método proposto por Ruan et al. (2020) um modelo de rede adversária gerativa de compartilhamento de recursos multilateral (*Multi-Branch Feature Sharing Generative Adversarial Network* - MB-FSGAN) foi proposto. Ele consiste em um extrator de características em várias escalas (*Multi-Scale Feature Extractor* - MSFE), um localizador da área de interesse e uma rede adversária geradora de compartilhamento de características (*Feature Sharing Generative Adversarial Network* - FSGAN). Os experimentos foram realizados em TCs de 113 pacientes com tumor renal. No melhor desempenho, obtiveram 85,9% de Dice.

Por sua vez, Yan et al. (2020) apresentaram um modelo denominado HybridNet 3D que consiste principalmente em duas partes: a rede de segmentação de primeiro plano e a rede de segmentação de *PointCloud Sparse*. A rede de segmentação de primeiro plano tem uma estrutura semelhante à V-Net (MILLETARI; NAVAB; AHMADI, 2016) e produz um mapa de segmentação de primeiro plano. A rede de segmentação de *PointCloud Sparse* consiste em um módulo *Voxel-to-Point* (VTP) e um módulo *Sparse Segmentation Network* (SSN), onde o módulo SSN contém uma U-Net construída pela *Submanifold Sparse Convolutional Networks* (SSCNs) em nuvens de pontos. A base de imagens usada

consiste em 300 TCs, tendo seu método atingido 79,7% de Dice para tumores renais.

Türk, Lüy e Yalcinkaya (2022) desenvolveram um sistema para auxiliar os médicos na segmentação de tumores renais. O método proposto consistiu em melhorias nas fases de codificação e decodificação do modelo V-Net. Com a estrutura do bloco gargalo de duplo estágio (*Two-Stage Bottleneck Block*), a arquitetura foi transformada em uma arquitetura única. Para avaliar o modelo, foram utilizadas 210 imagens de TC. O modelo desenvolvido obteve um Dice de 86,9%, apresentando os melhores resultados em relação às redes comparadas.

Tanimoto et al. (2022) apresentaram um método de aprendizado usando uma U-Net 3D, no qual erros em regiões implausíveis são suprimidos levando em consideração a estrutura global dos rins. No método proposto, adotou-se *patches* cuboides para cobrir toda a estrutura no plano axial (transversal), que tende a ter uma simetria rotacional quando a região dos rins está localizada no centro de cada *patch*. 213 imagens de TC foram usadas para avaliar o método proposto, obtendo-se 60,4% de Dice.

Como pode ser analisado nos estudos supracitados, a tarefa de segmentar tumores renais é extremamente difícil, até mesmo para especialistas, pois não há uma distinção clara entre o contraste do tumor renal e a textura do rim, entre outras anormalidades, como os cistos. As abordagens híbrida e fracamente supervisionada (KAUR; JUNEJA; MANDAL, 2019; YANG et al., 2020) minimizam a interferência manual de um médico especialista. Isso não é trivial, pois pode causar erros na marcação da região de interesse, uma vez que as técnicas propostas (SIFCM e *convolutional conditional random fields*) necessitam de parâmetros adequados para cada tipo de imagem (RUAN et al., 2020). Consequentemente, erros de segmentação podem ser gerados e interferir no diagnóstico do médico especialista. Por fim, estudos como Mehta et al. (2019), Yan et al. (2020), Türk, Lüy e Yalcinkaya (2022) e Tanimoto et al. (2022) usam as redes totalmente conectadas para segmentação dos tumores renais, o que acabaram tendo bom desempenho devido às suas arquiteturas de alto desempenho.

2.3 Segmentação de Rins e Tumores Renais

Existem vários métodos de aprendizado profundo bem-sucedidos para segmentar rins e tumores renais em imagens de TC. Esses métodos destacam-se devido sua arquitetura 3D, que apresenta alto desempenho. No trabalho de Yang et al. (2018) é proposta uma

rede 3D totalmente convolucional (*Fully Convolutional Network - FCN*) que combina um módulo de pooling de pirâmide (*Pyramid Pooling Module - PPM*) projetado para a segmentação de rins e tumores renais. Para extrair as regiões de interesse das imagens, usaram um método baseado em atlas. Os experimentos realizados em 140 pacientes de uma base de imagens privada, resultaram em um Dice de 93,1% para rins e 80,2% para tumores renais.

No estudo de Türk, Lüy e Barışçı (2020) foi proposto um modelo híbrido que usa recursos superiores dos modelos V-Nets existentes, considerando os estágios de codificação e decodificação separadamente. O codificador foi criado com base no modelo de fusão V-Net. O decodificador foi projetado com base na arquitetura ET-Net. Para validar o método proposto foi usado um conjunto de 210 tomografias. O modelo híbrido V-Net apresentou Dice médio de 97,7% e 86,5% para segmentação de rins e tumores renais, respectivamente.

Qayyum, Lalande e Meriaudeau (2020) propuseram uma rede residual 3D híbrida com um *Squeeze-and-Excitation (SE) block* para segmentação volumétrica de rins e tumores renais. A rede proposta usa blocos SE para capturar informações espaciais com base na função de reponderação da rede residual. Um esquema de bloco residual sucessivo foi usado em cada estágio de codificador e decodificador dos blocos residuais 3D. O método foi avaliado em um conjunto de 210 pacientes e obteve resultados de 97,8% de Dice para rins, e 86,8% para tumores renais.

Para lidar com os desafios da segmentação renal e tumoral em imagens de TC, Xie et al. (2020) propuseram um modelo denominado SE-ResNeXT U-Net (SERU), que combina as vantagens da SE-Net, ResNeXT e U-Net. Além disso, implementaram o modelo de maneira *coarse-to-fine*¹ para usar informações de contexto e fatias importantes dos rins esquerdo e direito. Um total de 300 tomografias foram usadas para realizar os experimentos. Os resultados apresentaram um Dice de 96,77% para rins e 74,32% para tumores renais.

Zhao et al. (2020) apresentaram uma U-Net 3D supervisionada em várias escalas, o MSS U-Net (*Multi-Scale Supervised 3D U-Net*) para segmentar rins e tumores renais. A arquitetura combina supervisão profunda com perda logarítmica exponencial para

¹ Segmentação *coarse* (grosseira) é uma segmentação que não tem muitos detalhes. Segmentação *fine* (fina) refere-se a uma segmentação com alto nível de detalhamento (ou seja, segmentação correta *pixel* a *pixel*).

aumentar a eficiência do treinamento 3D U-Net. Além disso, introduziram um método de pós-processamento baseado em componentes conectados para aprimorar o desempenho do processo geral. Essa arquitetura apresentou Dice de 96,9% para rins e 80,5% para tumores renais, usando uma base de imagens de 210 pacientes.

Geethanjali, Minavathi e Dinesh (2021) propuseram uma técnica automatizada para delinear os rins e o tumor em imagens de TC usando a arquitetura *Attention U-Net*. O modelo *Attention U-Net* foi usado para dar maior ênfase às regiões de interesse (rins e tumores) e menos ênfase nas áreas que não estão em foco. O modelo foi avaliado em 205 imagens de TC (cinco TCs foram excluídas inicialmente) e obteve um Dice de 95,65% e 93,86% para rins e tumores renais, respectivamente.

Outro trabalho que faz uso da abordagem 3D para segmentar rins e tumores renais foi apresentado por Heller et al. (2021). Os autores propuseram três arquiteturas inspiradas na U-Net 3D: uma simples, uma residual e uma residual de pré-ativação. Para medir a precisão do método proposto, um conjunto de 210 tomografias foi usado. Dentre todos os experimentos realizados, os melhores resultados foram obtidos usando a U-Net residual 3D, alcançando 97,4% de Dice para rins e 85,1% para tumores renais.

No estudo proposto por Lin et al. (2021) foram construídos dois modelos de segmentação 3D baseados em U-Net: um para segmentação automatizada do rim e outro para segmentação e detecção de tumor e cisto renal. O modelo de segmentação renal foi baseado em cascata de duas redes 3D U-Net. Para segmentar os tumores renais, uma única rede 3D U-Net foi construída. A avaliação do desempenho do modelo foi realizada por meio de uma base de imagens privada contendo 441 pacientes. O modelo proposto apresentou na segmentação de rins e tumores renais um Dice de 97,3% e 84,4%, respectivamente.

Yang et al. (2022) propuseram uma nova rede neural profunda, a 3D *Multi-Scale Residual Fully Convolutional Neural Network* (3D-MS-RFCNN) para melhorar a segmentação em tumores renais de tamanho grande. Para avaliar o desempenho do método proposto, uma base de imagens pública KiTS (210 imagens) foi usada para treinamento e um conjunto de dados hospitalares internos (270 imagens) para validação externa. Foram alcançados 91,62% de Dice para rins e 71,64% para tumores renais.

No trabalho de Kang et al. (2022) uma estrutura de segmentação em duas etapas: *coarse* e *fine*. Na etapa de segmentação *coarse*, um método de treinamento prévio assistido por contorno é proposto para melhorar a capacidade da rede de aprender os contornos das bordas e extrair completamente a imagem da região de interesse contendo o rim e

o tumor renal. No estágio de segmentação *fine*, é proposta uma rede U-Net melhorada (LC-Unet) fundindo ConvLSTM e CNN 3D. Vários experimentos foram feitos para verificar o desempenho do método proposto usando 300 imagens de TC. No melhor resultado foram obtidos 96,39% e 78,9% de Dice para rins e tumores renais, respectivamente.

Para segmentar os rins e tumores renais, observa-se que os métodos que utilizam abordagens 3D se destacaram devido ao seu alto desempenho. Em contrapartida, esse tipo de abordagem tem algumas desvantagens, como o consumo de memória da unidade de processamento gráfico (GPU - *Graphics Processing Unit*), alto custo computacional e necessidade de grande quantidade de conjuntos de dados totalmente anotados em 3D. Assim, as abordagens 3D nem sempre são opções viáveis.

Em geral, os trabalhos relacionados descritos neste capítulo, apresentam diferentes abordagens e têm alcançado resultados cada vez mais expressivos para solucionar o problema da segmentação de rins e tumores renais. Foram vistos desde estudos que trabalham com abordagens que usam atlas (WOLZ et al., 2013; YANG et al., 2014; YANG et al., 2018) a estudos que tem complexidade computacional, por meio de abordagens 3D (JACKSON et al., 2018; YANG et al., 2018; YAN et al., 2020; QAYYUM; LALANDE; MERIAUDEAU, 2020; ZHAO et al., 2020; HELLER et al., 2021; LIN et al., 2021; TANIMOTO et al., 2022; YANG et al., 2022; KANG et al., 2022). A Tabela 1 apresenta um resumo dos trabalhos relacionados descritos, na qual são apresentadas informações sobre as técnicas de segmentação, o número de pacientes e o desempenho Dice para rins e tumores.

Considerando as vantagens e limitações dos trabalhos relacionados, propõe-se um método eficiente para segmentar os rins e tumores renais em imagens de TC de pacientes doentes. De acordo com a literatura, os métodos que usam aprendizagem profunda superam as técnicas tradicionais em tarefas de visão computacional em geral (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), e na segmentação de imagens de TC em particular (CHRIST et al., 2017).

Desse modo, o método proposto aborda três problemas: segmentação inicial dos rins, o que consequentemente reduz o espaço de busca para o diagnóstico de doenças renais; segmentação de tumores renais, a partir do resultado da segmentação dos rins; e a reconstrução dos tumores renais para aumentar a precisão das segmentações dos rins e tumores renais. Para isso, propõe-se o uso de duas redes de aprendizado profundo: a ResUnet que é capaz de segmentar e recuperar os tumores renais; e a DeepLabv3+

que é capaz de segmentar os tumores renais com precisão. E, finalmente, técnicas de processamento de imagem são usadas para reduzir falsos positivos.

Tabela 1 – Resumo dos trabalhos relacionados.

	Trabalho	Técnica(s)	Base	Pacientes	Dice - Rim	Dice - Tumor
Rins	Khalifa et al. (2011)	Modelo deformável baseado em <i>Level Set</i>	Privada	29	95%	-
	Wolz et al. (2013)	Esquema hierárquico de registro e ponderação de atlas	Privada	150	92,5%	-
	Zhao, Liang e Fan (2013)	Continuidade contextual na sequência de imagens de TC	Privada	10	94,7%	-
	Yang et al. (2014)	Registro de imagens multi-atlas	Privada	14 fatias	95,2%	-
	Khalifa et al. (2016)	Estrutura híbrida baseada de modelos geométricos deformáveis e NMF	Privada	36	95%	-
	Skalski et al. (2017)	Contorno ativo usando a estrutura <i>Level Set</i>	Privada	10	86,2%	-
	Jackson et al. (2018)	CNNs 3D	Privada	113	88,5%	-
	Mehta et al. (2019)	Crowdsourcing e CNN	Privada	42	93,2%	-
	Les et al. (2021)	U-Net, <i>patch</i> e sistema de rastreamento	KiTS19	270	90,63%	-
Tumores	Kaur, Juneja e Mandal (2019)	Estrutura híbrida baseada em SIFCM e DRLSE	Privada	40	-	88,2%
	Yang et al. (2020)	CNNs fracamente supervisionada	Privada	200	-	82,6%
	Ruan et al. (2020)	MB-FSGAN	Privada	113	-	85,9%
	Yan et al. (2020)	HybridNet 3D	KiTS19	300	-	79,7%
	Türk, Lüy e Yalcinkaya (2022)	V-Net usando <i>Two-Stage Bottleneck Block</i>	KiTS19	210	-	86,9%
Rins e Tumores	Tanimoto et al. (2022)	3D U-Net preservando a simetria rotacional em cortes axiais	Privada	213	-	60,4%
	Yang et al. (2018)	Rede 3D com PPM + Atlas	Privada	140	93,1%	80,2%
	Türk, Lüy e Barışçı (2020)	Modelo híbrido V-Net	KiTS19	210	97,7%	86,5%
	Qayyum, Lalonde e Meriaudeau (2020)	Rede residual híbrida 3D com SE	KiTS19	210	97,8%	86,8%
	Xie et al. (2020)	SE-ResNeXT U-Net	KiTS19	300	96,77%	74,32%
	Zhao et al. (2020)	MSS U-Net 3D	KiTS19	210	96,9%	80,5%
	Geethanjali, Minavathi e Dinesh (2021)	<i>Attention</i> U-Net	KiTS19	205	95,65%	93,86%
	Heller et al. (2021)	U-Nets 3D: simples, residual e residual de pré-ativação	KiTS19	210	97,4%	85,1%
	Lin et al. (2021)	3D U-Net	Privada	441	97,3%	84,4%
	Yang et al. (2022)	3D-MS-RFCNN	KiTS19 + Privada	480	91,62%	71,64%
	Kang et al. (2022)	3D-CNN e ConvLSTM	KiTS19	300	96,39%	78,9%

2.4 Considerações Finais

Neste capítulo, foi feita uma revisão das produções científicas relevantes relacionadas à segmentação de rins e tumores renais em imagens de TC. Os métodos e os resultados obtidos nestes trabalhos foram brevemente apresentados. Além disso, foram feitas considerações sobre as diferenças entre as diferentes abordagens apresentadas, a fim de obter um panorama do que vem sendo produzido pela comunidade científica em

relação a esse tema.

No próximo capítulo, serão apresentados os conceitos teóricos fundamentais para o desenvolvimento deste trabalho.

3 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são detalhados os tópicos necessários para compreensão das técnicas usadas na elaboração do método proposto. As seções a seguir abordam conceitos sobre o exame de TC, o rim e tumores renais, aprendizado profundo e as métricas de desempenho para validar os resultados experimentais.

3.1 Rins e Tumores Renais

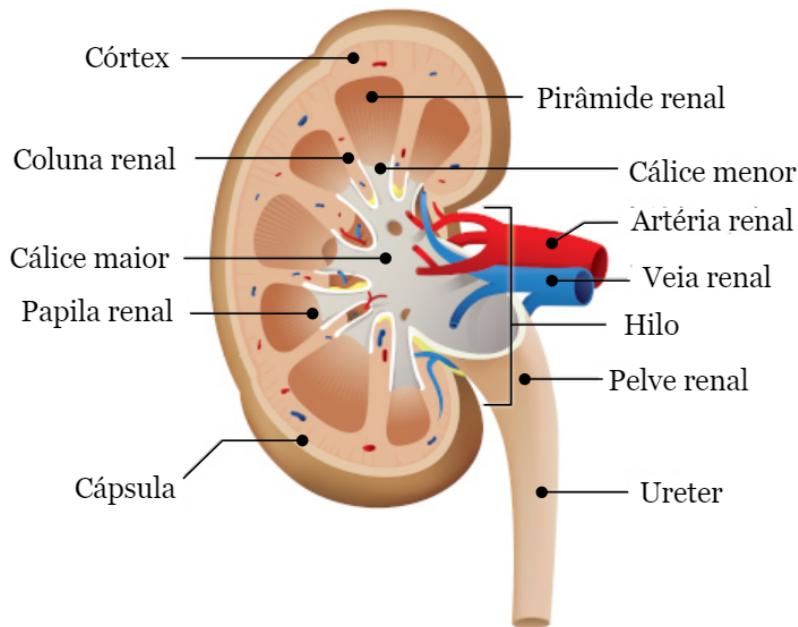
Os rins são órgãos bilaterais em forma de feijão, de cor marrom-avermelhada e localizados na região posterior do abdômen. Geralmente, o rim direito é ligeiramente menor e mais baixo que o esquerdo devido à presença do fígado. Cada rim pesa cerca de 125-170 gramas nos homens e 115-155 gramas nas mulheres, e são revestidos por uma cápsula renal fibrosa e resistente. Além disso, duas camadas de gordura servem como proteção (NEWMAN, 2019).

Internamente, os rins têm uma estrutura intrincada e única, conforme ilustrado na Figura 1. O rim apresenta uma camada central com uma cor mais escura (medula) e uma camada periférica mais clara (córTEX). Na medula renal, encontram-se as pirâmides renais, estruturas que têm o formato de um cone. O ápice de uma pirâmide renal está voltado para o córTEX, e termina em uma papila que se despeja no cálice menor. O córTEX se estende da cápsula até a base das pirâmides renais e também se encontra entre essas estruturas, formando as colunas renais (JONES, 2019; SANTOS, 2019).

O córTEX e a medula recebem sangue pela artéria renal e são drenados pela veia renal. Há também uma estrutura achatada e com formato de funil (pelve renal), responsável pela coleta da urina. A pelve se ramifica em direção à medula, formando cálices maiores e menores. Por fim, em cada rim, a margem medial é marcada por uma fissura profunda, conhecida como hilo renal. O hilo renal atua como uma porta de entrada para o rim, onde normalmente os vasos renais e o ureter entram e saem do rim (JONES, 2019; SANTOS, 2019; NEWMAN, 2019).

O principal papel dos rins é manter a homeostase. Isso significa que os rins gerenciam os níveis de fluidos, o equilíbrio eletrolítico e outros fatores que mantêm o ambiente interno do corpo consistente e confortável. Além disso, desempenham um amplo espectro de funções, tais como: remover uma série de resíduos a serem eliminados pela

Figura 1 – Estrutura interna do rim.



Fonte: (SANTOS, 2019).

urina, reabsorção de nutrientes do sangue, manter o pH estável e regular a osmolaridade e a pressão arterial (SOCIETY, 2017; NEWMAN, 2019; FUND, 2019).

Uma série de doenças pode afetar os rins, tais como: nefropatia diabética, insuficiência renal, hidronefrose renal, síndrome nefrótica, pedras nos rins e nefrite intersticial. Além disso, quando as células renais começam a se multiplicar de forma acelerada, podem surgir tumores renais. Os tumores renais podem ser benignos ou malignos. Os benignos não se espalham nem atacam os tecidos, mas os malignos podem ser agressivos a ponto de atingir outros órgãos por meio da corrente sanguínea. O câncer renal maligno mais comum é o carcinoma de células renais (SHUCH et al., 2015; NEWMAN, 2019).

A maioria dos casos de câncer renal tem cura com um diagnóstico precoce. Para iniciar o tratamento adequado do câncer renal, a segmentação acurada dos rins é necessária. Portanto, as informações estruturais do rim são analisadas por meio de exames de ultrassom, Ressonância Magnética (RM) ou TC de abdômen. Embora as imagens RM tenham sido usadas para segmentação renal em alguns estudos na literatura (GOCERI, 2011; GOCERI; GOCERI, 2015), a TC, dentre as modalidades disponíveis, é considerada o padrão ouro, especialmente para o diagnóstico de câncer precoce (KAUR; JUNEJA, 2016). Além disso, a TC é muito útil no estadiamento (verificação da extensão para outros

órgãos) e no planejamento terapêutico.

3.2 Tomografia Computadorizada

Para planejar o tratamento radioterápico, é necessário obter informações sobre os tumores renais, como tamanho e a localização exata. Para isso, é necessária uma imagem tridimensional detalhada da região do abdômen do paciente, geralmente realizada por meio do exame de TC. Essas informações são eventualmente usadas para orientar o radioterapeuta a produzir o melhor tratamento radioterápico para cada paciente.

A TC é um exame não invasivo, indolor e rápido, que combina uma série de imagens de raios-X obtidas de diferentes ângulos do corpo do paciente, produzindo sinais que são processados pelo computador para gerar imagens transversais (fatias) dos ossos, vasos sanguíneos e tecidos moles (BUZUG, 2011). Portanto, a TC funciona como outros exames de raios-X, mas fornece imagens mais detalhadas devido à emissão de vários feixes de raios-X ao mesmo tempo (GWYNNE et al., 2012b).

Os feixes de raios-X são emitidos por uma fonte e coletados por um conjunto de detectores que permitem avaliar a quantidade de radiação que é absorvida pelos diferentes tecidos. Os detectores são posicionados no lado diametralmente oposto para captar os feixes que passam pelos objetos observados. Finalmente, os dados coletados pelos detectores são processados e digitalizados em *pixels* que apresentam uma escala de cinza relacionada à atenuação sofrida pelos raios-X (medida da facilidade com que um material pode ser penetrado por um feixe de energia), expressa por uma escala numérica expressa em unidades Hounsfield (*Hounsfield Units - HU*) (BUZUG, 2011). A HU é calculada com base em uma transformação linear do coeficiente de atenuação linear da linha de base do feixe de raios X (ADAMS et al., 2012).

Para a realização do exame, o paciente é posicionado no tomógrafo e imobilizado com suportes e fitas de fixação. O tomógrafo é uma máquina com um túnel no centro onde, durante o exame, o paciente fica deitado sobre uma mesa estreita que desliza dentro do túnel. Cortes axiais são feitos para verificar o posicionamento e alinhamento do paciente tomando como referência o isocentro (ponto onde os feixes de radiação se cruzam). Em seguida, são realizadas radiografias digitais (topograma), que programam as sequências de imagens (fatias) e a espessura dos cortes (BUZUG, 2011; GWYNNE et al., 2012b). A Figura 2 ilustra um exemplo de equipamento de TC.

Figura 2 – Equipamento de TC.



Fonte: (TULLY, 2016).

Uma vez que várias fatias sucessivas são coletadas pelo computador da máquina, elas são “empilhadas” digitalmente para formar uma imagem tridimensional do paciente que permite a identificação e localização mais fácil das estruturas básicas do corpo, bem como possíveis tumores ou anormalidades. Para melhorar a interpretação dos resultados e garantir diagnósticos mais precisos, a administração de contraste intravenoso pode ser solicitada antes da aquisição das imagens. Estas substâncias são usadas para aumentar a visibilidade e a diferenciação entre tumores e estruturas circundantes normais. O produto para se obter esse contraste é geralmente administrado por via oral (engolido) ou por via intravenosa (injetado). Após a obtenção das imagens, o contraste é eliminado pela urina.

3.3 Processamento Digital de Imagens

O processamento digital de imagens é definido como um conjunto de técnicas computacionais que transformam uma imagem de entrada em uma saída desejada (GONZALEZ; WOODS, 2008). Esse conjunto de técnicas permite extrair e identificar informações de imagens e melhorar a qualidade visual de aspectos estruturais, facilitando a percepção humana e a interpretação automática por meio de máquinas (PEDRINI; SCHWARTZ, 2008).

Existem vários algoritmos com propósitos muito específicos, que juntos formam

a metodologia final. Portanto, uma metodologia difere de outra na forma como compõe suas etapas ou ferramentas, mas normalmente segue as etapas apresentadas por Gonzalez e Woods (2008):

1. Aquisição das imagens: as imagens são capturadas por meio de um dispositivo ou sensor e transformadas em um formato que pode ser processado posteriormente. Porém, como resultado desse processo, a imagem resultante pode apresentar falhas devido às condições de iluminação ou às características do dispositivo, por exemplo;
2. Pré-processamento: o objetivo é melhorar a qualidade da imagem por meio de técnicas para reduzir o ruído, realçar o contraste e suavizar certas estruturas da imagem;
3. Segmentação: consiste em localizar e extrair áreas de interesse contidas na imagem, com o objetivo de dividir a imagem isolando os diferentes objetos que a compõem;
4. Representação e descrição: denominada de extração de características, pois extrai informações que podem ser usadas para discriminar classes de objetos;
5. Reconhecimento e interpretação: reconhecimento é o processo que atribui um identificador aos objetos da imagem. A interpretação consiste em atribuir um significado ao conjunto de objetos reconhecidos.

A saída de uma etapa é usada como entrada para a próxima etapa, sendo que cada entrada e saída pode ser ou não uma imagem digital. Uma metodologia de processamento de imagem pode conter apenas um subconjunto de todas as etapas apresentadas (BRAZ, 2014).

3.4 Especificação do Histograma

O histograma básico de uma imagem digital é uma função que gera a distribuição dos valores de intensidade presentes em uma imagem, com níveis de intensidade variando no intervalo de $[0, L - 1]$. Ele é dado pela expressão $h(r_k) = n_k$, onde r_k é o k -ésimo nível de intensidade e n_k é a quantidade de *pixels* da imagem com intensidade r_k (GONZALEZ; WOODS, 2008). Portanto, o histograma é representado por um gráfico indicando o número de *pixels*, eixo vertical, para cada nível de intensidade, eixo horizontal (PEDRINI; SCHWARTZ, 2008).

A especificação do histograma (*histogram matching*) é o processo de transformar o histograma de uma imagem origem para corresponder ao histograma de uma imagem de referência, de modo que ambas tenham uma distribuição de *pixels* semelhante (GONZALEZ; WOODS, 2008). Para especificar o histograma de uma determinada imagem, $h_i(x)$, com o histograma de uma imagem de referência, $h_o(y)$, primeiro é necessário equalizar os níveis da imagem original $h_i(x)$ para obter um histograma intermediário $h^*(z)$. A expressão usada para calcular esse mapeamento é definida na Equação 3.1,

$$y(i) = \frac{L-1}{N} \sum_{i=0}^{L-1} n_i, \quad (3.1)$$

onde n_i representa o número de *pixels* com nível i , e N o total de *pixels* da imagem com L níveis de cinza.

A transformação $z = f(x)$ equaliza o histograma original $h_i(x)$ para o histograma intermediário $h^*(z)$, e $z = g(y)$ equaliza o histograma $h_o(y)$. Então, o mapeamento global dos níveis de intensidade necessários para produzir $h_o(y)$ a partir de $h_i(x)$ é obtido usando a transformação inversa $y = g^{-1}\{f(x)\}$. Isso representa o mapeamento dos valores da imagem de entrada equalizada para seus valores mais próximos no histograma especificado.

Neste trabalho, a técnica de especificação do histograma foi usada na etapa de pré-processamento para mapear as intensidades inter-exames de TC baseado em um exame de referência de TC.

3.5 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) são um ramo da Inteligência Artificial (IA), fundada na década de 1940, quando McCulloch e Pitts desenvolveram o primeiro modelo neural (MCCULLOCH; PITTS, 1943). O problema básico resolvido pelas RNAs é a aquisição indutiva de conceitos a partir de exemplos, a capacidade de aprender e generalizar a partir de dados, ou seja, imitar a capacidade humana de aprender com a experiência, torna as RNAs úteis na automatização do processo de aprendizado de regras de várias aplicações (MUKHOPADHYAY, 2011).

Em resumo, as RNAs são esquemas computacionais que representam parcialmente as redes neurais biológicas existentes nos cérebros humanos ou animais, expressas por nós conectados (neurônios artificiais) organizados adequadamente em camadas. Todos os neurônios artificiais são conectados e capazes de transmitir sinais, geralmente números reais,

por meio de suas conexões (sinapses artificiais) (MUKHOPADHYAY, 2011; BERSIMIS; VARLAMIS, 2019). Esses modelos são comumente usados na tarefa de reconhecimento de padrões, como detecção de objetos, identificação de nódulos de câncer, reconhecimento de voz e reconhecimento facial (ALANIS; ARANA-DANIEL; LOPEZ-FRANCO, 2019; BARONI, 2020).

3.5.1 Neurônio artificial

O neurônio artificial é a unidade básica de uma RNA, fundamental para a construção de modelos mais poderosos. Em termos matemáticos, o neurônio artificial fornece uma saída para um determinado conjunto de entradas, conforme expresso na Equação 3.2,

$$f(x) = \sigma \left(\sum_{i=1}^n x_i w_i + b \right), \quad (3.2)$$

em que x_i é a entrada i , w_i é o peso sináptico associado a entrada i , b é o termo *bias* e σ é a função de ativação. Portanto, cada neurônio combina suas entradas e, em seguida, passa por uma função de ativação, que pode ser um filtro linear ou não-linear.

Em relação aos filtros não lineares, a função sigmoide e a função tangente hiperbólica são citadas como exemplos. Além delas, existem outras funções não lineares propostas como as chamadas *Rectified Linear Units* (ReLU) e *Leaky Rectified Linear Unit* (Leaky ReLU), cuja velocidade de convergência em relação às supracitadas é até 6 vezes mais rápida (KRIZHEVSKY; SUTSKEVER; HINTON, 2012; MAAS; HANNUN; NG, 2013). Essas funções são expressas nas Equações 3.3 e 3.4, respectivamente,

$$\sigma(x) = \max(0, x), \quad (3.3)$$

$$\sigma(x) = \max(0.1, x). \quad (3.4)$$

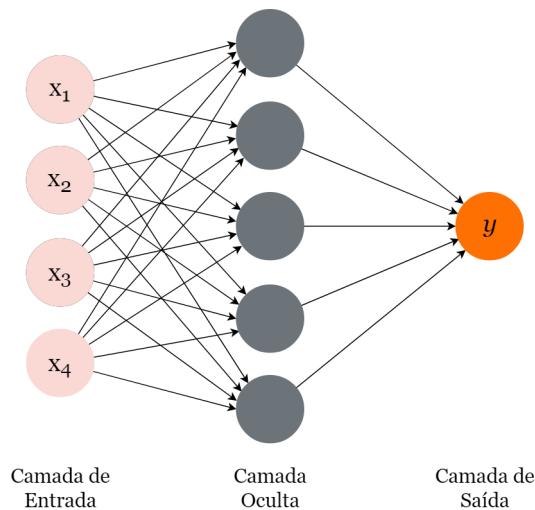
3.5.2 Redes *Multilayer Perceptron*

Embora existam inúmeras arquiteturas RNA, a arquitetura *Multilayer Perceptron* (MLP) é a mais frequentemente encontrada na literatura (PHAM et al., 2019). As MLPs são algoritmos de aprendizado supervisionado compostos por múltiplas camadas de

neurônios, todas completamente conectadas às suas subsequentes. Os neurônios, por sua vez, implementam as funções de ativação não-linear determinadas. De acordo com Hornik, Stinchcombe e White (1989), essas redes são capazes de aproximar qualquer função mensurável para qualquer grau de precisão desejado.

De modo geral, uma MLP é composta por uma camada inicial, que recebe os dados de entrada do modelo, seguida por uma ou mais camadas intermediárias (ocultas/escondidas) e, por fim, uma camada de saída (HORNIK; STINCHCOMBE; WHITE, 1989). A Figura 3 ilustra uma arquitetura básica de MLP com apenas uma camada oculta.

Figura 3 – Rede *Multilayer Perceptron*.



Fonte: Elaborado pela autora.

Os neurônios de cada camada são organizados e conectados uns aos outros em uma estrutura de múltiplas camadas *feedforward*. Em outras palavras, cada camada é totalmente conectada com sua camada adjacente e produz um vetor de saída de acordo com o vetor da camada anterior (BERSIMIS; VARLAMIS, 2019). A saída de cada camada é calculada aplicando a função de ativação de cada neurônio a todos os neurônios da camada. A Equação 3.5 descreve a saída de uma camada,

$$y^l = \sigma(W^l y^{l-1} + b^l), \quad (3.5)$$

em que y^l é o vetor de saída, W^l é a matriz dos pesos aplicados a cada par de neurônio da camada l e $l - 1$, e b^l é o vetor do termo *bias* de cada neurônio da camada l . Para treinar efetivamente as MLPs, o algoritmo de *backpropagation* é aplicado.

3.5.3 Backpropagation

Durante o treinamento de uma MLP, uma função de erro é definida para calcular o erro das previsões do modelo com base na saída esperada para os dados de entrada fornecidos. O objetivo do treinamento é minimizar o erro total da rede (soma das funções de erro aplicadas a todos os exemplos de uma base de imagens). Para isso, as MLPs usam um algoritmo denominado *backpropagation*.

O algoritmo *backpropagation* é caracterizado por dois passos consecutivos: primeiro, os dados são enviados para a camada de entrada da rede, e trafegam pela rede, camada por camada, sendo transformados por meio de cálculos realizados até que, na camada de saída, uma resposta seja produzida; segundo, a saída obtida é comparada com a saída desejada da entrada correspondente e um valor de erro é calculado. Em seguida, os erros derivados com relação aos parâmetros da rede são propagados novamente, da camada de saída até a camada de entrada, desta forma, os pesos das conexões nas camadas ocultas são modificados à medida que o erro é retropropagado. O processo continua várias vezes até que uma melhoria desejada na previsão do modelo seja alcançada.

Os passos do algoritmo *backpropagation* são brevemente apresentados a seguir. Os itens 1, 2 e 3 estão relacionados ao *feedforward* e o *backpropagation* inicia no item 4.

1. Inicializar os valores dos pesos sinápticos de cada neurônio aleatoriamente;
2. Apresentar as entradas da rede em um vetor x_1, x_2, \dots, x_n de características e especificar um vetor d_1, d_2, \dots, d_n de saídas desejadas;
3. Calcular as saídas da rede y_1, y_2, \dots, y_n conforme a Equação 3.2;
4. Reajustar os pesos começando pelos neurônios da camada de saída, retropropagando até a camada de entrada, de acordo com a Equação 3.6,

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_j x_i, \quad (3.6)$$

onde w_{ij} é o peso do neurônio j em uma iteração t , x_i corresponde a um neurônio de saída ou de entrada, η é a taxa de aprendizagem e δ_j é um erro de gradiente para o neurônio j . Se j for um neurônio de saída, então δ_j é definido pela Equação 3.7,

$$\delta_j = y_j(1 - y_j)(d_j - y_j), \quad (3.7)$$

onde d_j denota a saída desejada e y_j é a saída desejada da rede. Se o neurônio j for um neurônio oculto, então δ_j é definido pela Equação 3.8,

$$\delta_j = x_j(1 - x_j) \sum_k \delta_k w_{jk}, \quad (3.8)$$

onde k denota todos os neurônios da camada após a camada do neurônio j ;

5. Retornar ao passo 2 até que uma determinada condição seja satisfeita.

É importante ressaltar que a taxa de aprendizado (η) influencia a magnitude das mudanças dos pesos, desempenhando papel fundamental no aprendizado do modelo. Taxas de aprendizados muito pequenas implicam em pequenas variações, o que torna o modelo lento e aumentam as chances de parar em mínimos locais. No entanto, altas taxas de aprendizado tendem a produzir grandes oscilações, o que compromete o processo de aprendizado do modelo. A taxa de aprendizado é introduzida na rede para permitir uma convergência mais rápida ao valor ótimo desejado e ao mesmo tempo evitar que a rede oscile, diminuindo a taxa de aprendizado quando o erro tende a aumentar (SILVA, 2004; HAYKIN, 2007; SILVA; SPATTI; FLAUZINO, 2010).

3.6 Aprendizado Profundo

Durante o processo de aprendizagem, humanos e animais são inicialmente conduzidos a interpretar e compreender conceitos mais simples, a fim de aprender conceitos mais complexos a partir de conceitos previamente observados ao longo de suas vidas (FERNANDES, 2013). Esse processo de aprendizagem, denominado aprendizado profundo, sugere uma divisão em camadas hierárquicas do cérebro com diferentes responsabilidades (HUBEL; WIESEL, 1998).

O aprendizado profundo é uma subárea do aprendizado de máquina que usa arquiteturas hierárquicas para aprender abstrações de alto nível em um conjunto de imagens. É uma abordagem em evolução e tem sido amplamente usada em domínios tradicionais de inteligência artificial, como análise semântica, transferência de aprendizado, processamento de linguagem natural e visão computacional. Existem três fatores importantes para o crescimento do aprendizado profundo: o aumento da capacidade de processamento dos *chips* gráficos, os avanços consideráveis nos algoritmos de aprendizado

de máquina e o custo significativamente reduzido do *hardware* de computação (GUO et al., 2016).

As técnicas de aprendizagem profunda apresentam múltiplas camadas de processamento não linear para reconhecimento de padrões de forma semelhante ao cérebro (LECUN; BENGIO; HINTON, 2015). Diferente das RNAs tradicionais, essas técnicas permitem a extração automática de características do conjunto de treinamento, sem a necessidade de uma série de técnicas de processamento de imagens e reconhecimento de padrões (HUA et al., 2015). Como resultado, as etapas de extração, seleção e classificação de características são abstraídas no próprio modelo, com pouca intervenção humana (HUA et al., 2015; CHENG et al., 2016).

Atualmente, existem inúmeras técnicas de aprendizagem profundas disponíveis na literatura, como as redes neurais convolucionais, redes neurais recorrentes, redes de crença profunda, redes de memória de longo prazo, os auto-codificadores esparsos empilhados, entre outras. As diferentes arquiteturas têm diferentes vantagens com base na aplicação e nas características dos dados envolvidos. Por exemplo, em visão computacional, as redes neurais convolucionais são preferidas e, em sequências e modelagem de séries temporais, as redes neurais recorrentes. O aprendizado profundo é um campo em rápida evolução e arquiteturas mais novas com algoritmos de aprendizado mais recentes são frequentemente desenvolvidas para suportar a necessidade de desenvolver máquinas eficientes semelhantes a humanos em diferentes áreas de aplicação (SENGUPTA et al., 2020).

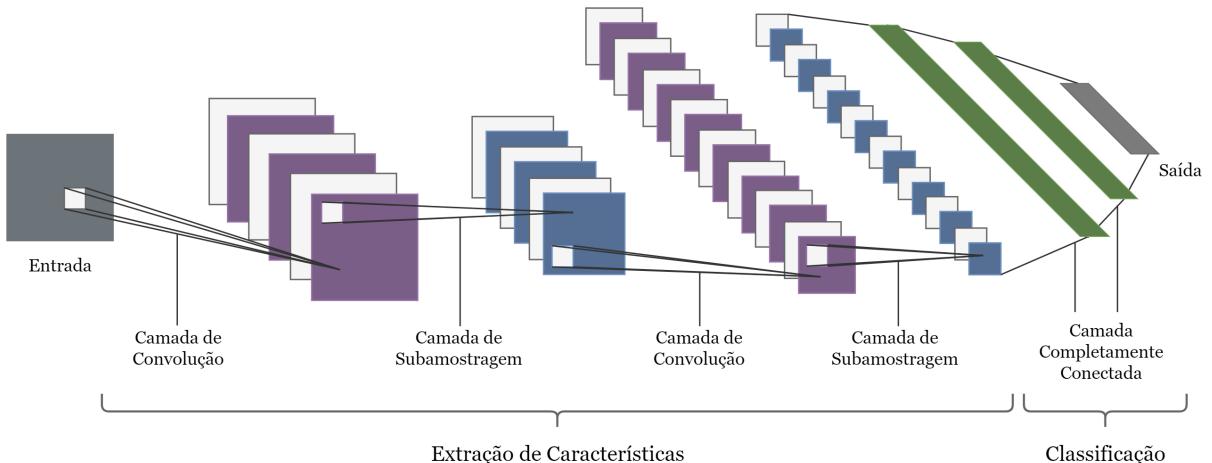
3.6.1 Redes Neurais Convolucionais

As redes neurais convolucionais (*Convolutional Neural Networks* - CNN) são modelos biologicamente inspirados capazes de construir um aprendizado hierárquico de características (LECUN et al., 1998; ZHANG et al., 2015). Essa arquitetura é bastante popular em tarefas de visão computacional, como detecção, segmentação e classificação de imagens e vídeos de vários domínios.

Geralmente, a CNN usa em sua arquitetura três tipos de camadas: convolução, subamostragem e completamente conectada (LECUN et al., 1998). As camadas de convolução extraem características de baixo nível importantes da imagem de entrada, como textura simples e bordas. À medida que mais camadas de convolução são adicionadas, características de alto nível são extraídas. As camadas de subamostragem reduzem a

resolução dos mapas de características, mantendo as informações de características. Finalmente, a camada totalmente conectada conecta a rede à camada discriminante (camada de saída), que fornece a saída desejada (LECUN; BENGIO; HINTON, 2015). A Figura 4 ilustra uma arquitetura básica de CNN, em seguida, os detalhes de cada camada são apresentados.

Figura 4 – Exemplo de uma arquitetura CNN.



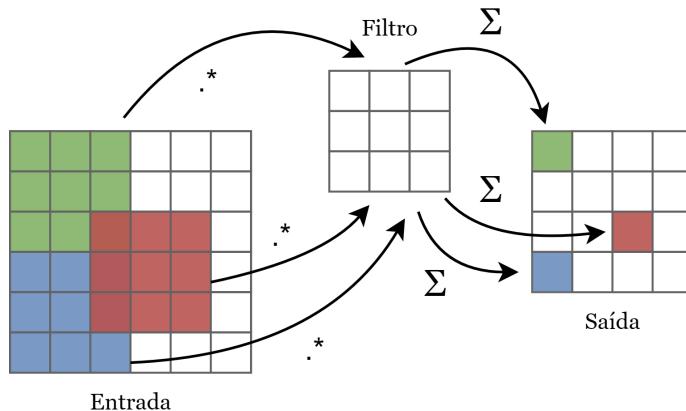
Fonte: Adaptado de (LIU et al., 2017).

3.6.1.1 Camada de Convolução

As camadas de convolução são compostas por filtros treináveis aplicados individualmente na imagem de entrada da rede, gerando diversos mapas de características. Os filtros são definidos por uma pequena área (3×3 , 5×5 , 7×7 pixels) e cada neurônio é ligado apenas aos neurônios nas proximidades da camada anterior. Portanto, para cada filtro aplicado a uma imagem de entrada, pode existir um neurônio conectado com a saída de um subconjunto de neurônios da camada anterior. Os pesos são compartilhados entre os neurônios, fazendo com que os filtros aprendam os padrões frequentes que ocorrem em qualquer parte da imagem (LECUN; KAVUKCUOGLU; FARABET, 2010; GUO et al., 2016). O processo de convolução em uma imagem é ilustrado na Figura 5.

Após gerar mapas de características, uma camada de ativação não linear é aplicada (WU; SAHOO; HOI, 2020). Ao fim do treinamento da rede, cada filtro será responsável por detectar uma característica específica da imagem, ou de parte dela (HAFEMANN, 2014).

Figura 5 – Ilustração de uma operação de convolução.



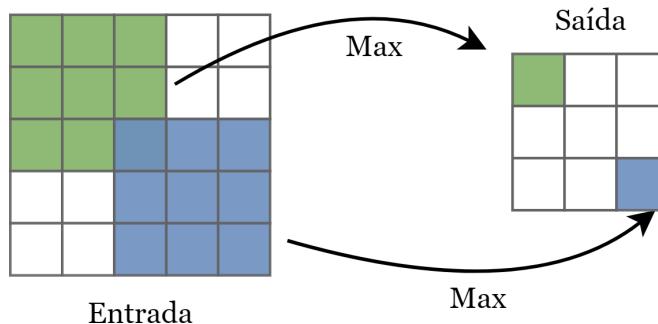
Fonte: Adaptado de (HAFEMANN, 2014).

3.6.1.2 Camada de Subamostragem

As camadas de subamostragem são geralmente subsequentes às camadas de convolução e são usadas para reduzir as resoluções dos mapas de características produzidos pelas camadas de convoluções anteriores e selecionar as características invariantes a deslocamentos e distorções (LECUN; KAVUKCUOGLU; FARABET, 2010; HAFEMANN, 2014).

A subamostragem consiste em aplicar filtros nos mapas de características e usar operações como média (*average pooling*) ou valor máximo (*max-pooling*) para extrair valores dessas sub-regiões. A Figura 6 ilustra a subamostragem com a aplicação da operação *max-pooling* deslizando uma janela de tamanho 3 com passo igual a 1, na qual apenas o valor máximo de *pixel* da janela de convolução é mantido. Ou seja, em uma máscara 3×3 com 9 valores, apenas o maior valor de *pixel* entre eles permanecerá.

Figura 6 – Ilustração da camada de subamostragem.



Fonte: Adaptado de (HAFEMANN, 2014).

Existem também as operações de *upsampling*. Basicamente, elas se comportam como o inverso das camadas de *max-pooling*, ou seja, dobram a resolução dos mapas de características da camada anterior por operações de interpolação (LONG; SHELHAMER; DARRELL, 2015). No entanto, é configurável.

3.6.1.3 Camada Completamente Conectada

Depois de extrair as características com as camadas de convolução e subamostragem, os mapas de características podem ser usados como entrada para as camadas totalmente conectadas, que por sua vez são responsáveis por classificar os padrões de entrada. Os neurônios em uma camada totalmente conectada têm conexões com todas as ativações na camada anterior, semelhante à estrutura do MLP (Seção 3.5.2).

Uma vez que a rede esteja treinada em um nível satisfatório, a rede pode ser usada como uma ferramenta para classificar novos dados. Neste trabalho, arquiteturas baseadas nos conceitos de redes neurais convolucionais são usadas para segmentar rins e tumores renais em imagens tomográficas.

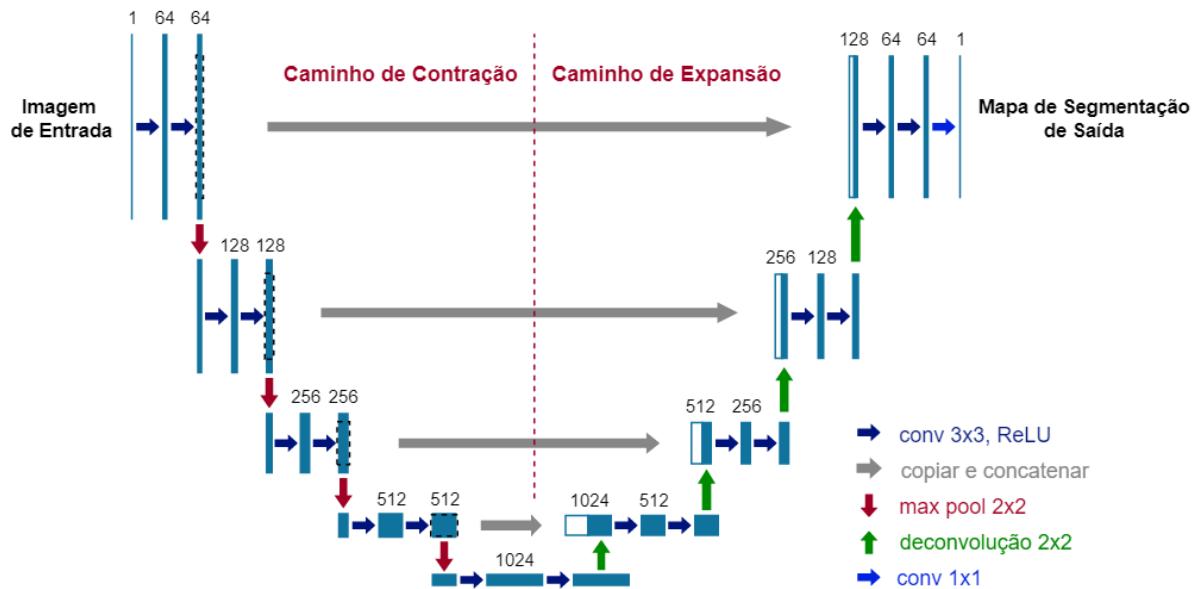
3.7 Arquitetura U-Net

Para realizar a segmentação dos rins e reconstrução dos tumores renais em imagens de TC, foi usada como base a arquitetura da U-Net (HE et al., 2016), que é uma rede capaz de generalizar mesmo em domínio de baixa variância.

A U-Net é uma RNA completamente convolucional proposta para segmentação de imagens biomédicas (RONNEBERGER; FISCHER; BROX, 2015), que desde 2015 vem superando os métodos existentes em diversos desafios biomédicos de segmentação de imagens. Essa rede é um tipo específico de rede profunda e avançada cuja arquitetura consiste em um caminho de contração e um caminho de expansão simétrico. A principal estratégia que diferencia a U-Net das outras arquiteturas totalmente convolucionais (*Fully Convolutional Network - FCN*) é a combinação entre os mapas de características no estágio de contração e seus correspondentes simétricos no estágio de expansão, o que permite a propagação de informações de contexto para os mapas de características de alta resolução. A Figura 7 ilustra a arquitetura e as camadas que constituem a U-Net.

O caminho de contração (lado esquerdo) captura o contexto e o caminho de expansão (lado direito) simétrico permite uma localização precisa. O caminho de contração

Figura 7 – Arquitetura U-Net.



Fonte: Adaptado de (RONNEBERGER; FISCHER; BROX, 2015).

é uma arquitetura CNN tradicional, várias aplicações repetidas de duas convoluções 3×3 , cada uma acompanhada por uma função de ativação ReLU e uma operação de subamostragem com operação de *max-pooling* 2×2 , reduzindo a largura e a altura da imagem, com passo (*stride*) 2 para subamostragem. Em cada etapa de subamostragem, o número de canais de características dobra. No caminho de contração as camadas de 1 a 5 são compostas por convoluções de 64, 128, 256, 512 e 1024 filtros, respectivamente.

O caminho de expansão consiste em um levantamento do mapa de características seguido de uma deconvolução (*upsampling*) 2×2 que faz a metade do número de canais de características. Posteriormente, uma junção é feita com o mapa de características correspondente cortado do caminho de contratação e duas convoluções 3×3 , cada uma seguida por uma função de ativação ReLU. Essa etapa do caminho expansivo é importante devido à perda de *pixels* de borda em cada convolução no caminho de contratação. Na última camada é usada uma convolução 1×1 para mapear cada vetor de características para o número desejado de classes.

Em resumo, a U-Net pode ser treinada do início ao fim com poucas imagens, onde a U-Net simplesmente concatena os mapas de características do codificador para mapear mapas de características do decodificador em todas as etapas para formar uma estrutura como escada. As conexões de concatenação nesta arquitetura permitem que o decodificador aprenda as características relevantes que são perdidas quando o codificador

é agrupado em cada fase.

3.8 Arquitetura DeepLabv3+

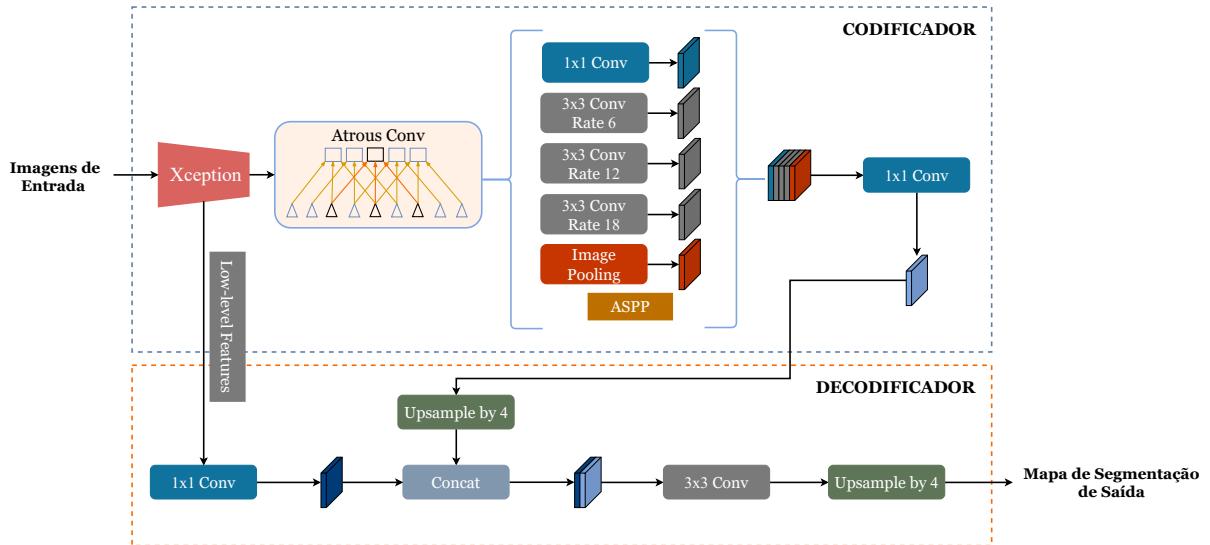
Para obter a segmentação de tumores renais em TC, foi usada como base a arquitetura de última geração DeepLabv3+ (CHEN et al., 2018), que é um modelo forte para codificar informações contextuais e refinar os resultados da segmentação, especialmente ao longo dos limites do objeto.

Como descrito anteriormente, as redes neurais profundas usam estruturas de codificador-decodificador para tarefas de segmentação semântica. O codificador extrai informações contextuais em várias escalas, enquanto o decodificador captura os limites dos objetos de forma mais clara, recuperando gradualmente as informações espaciais para reconstruir a saída (CHEN et al., 2018). Na arquitetura DeepLabv3+ o codificador padrão usado é o Xception (*Extreme Inception*) (CHOLLET, 2017). Em resumo, o Xception é uma extensão da arquitetura Inception que substitui os módulos Inception padrão por convoluções separáveis em profundidade (*depth wise separable convolution*).

A Figura 8 ilustra a arquitetura DeepLabv3+. As imagens de entrada da rede passam por uma estrutura de codificador-decodificador. Na fase de codificação, as características extraídas do codificador Xception são usadas como entrada para as operações de *Atrous Convolution* e *Atrous Spatial Pyramid Pooling* (ASPP). O *Atrous Convolution* captura características em escalas múltiplas ($rate = 6$, $rate = 12$ e $rate = 18$), e o ASPP é uma operação de *pooling* que ajuda a calcular objetos em escalas diferentes. Essas operações fazem a ponte entre a codificação e a decodificação (CHEN et al., 2018).

Na fase de decodificação, as informações extraídas na codificação são usadas para reconstruir a saída. Assim, após as operações de *Atrous Convolution* e ASPP, a operação de *upsampling* é aplicada usando um filtro bilinear de fator 4. Posteriormente, concatena-se com os *low-level features* (características de baixo nível) correspondentes da saída do codificador Xception, após passar por uma convolução 1×1 para reduzir o número de canais. Após a concatenação, algumas convoluções 3×3 são aplicadas para refinar as características, seguidas por outro *upsampling* bilinear simples de fator 4. Além disso, é usada a *depth wise separable convolution*, que é uma operação para reduzir o custo de computação e vários parâmetros, mantendo o desempenho (CHEN et al., 2018).

Figura 8 – Arquitetura DeepLabv3+.



Fonte: Adaptado de (CHEN et al., 2018).

3.9 Data Augmentation

Data augmentation é uma técnica comumente usada na literatura para aumentar o tamanho e a diversidade dos conjuntos de treinamento. Em visão computacional, o *data augmentation* se tornou uma técnica de regularização implícita comum para combater o *overfitting* em modelos de aprendizado profundo e é usado para melhorar o desempenho do modelo (KUKAČKA; GOLKOV; CREMERS, 2017; MIKOŁAJCZYK; GROCHOWSKI, 2018).

O aumento de dados é semelhante à imaginação ou sonho. Os seres humanos imaginam diferentes cenários com base na experiência. A imaginação ajuda a compreender melhor o mundo. Os métodos de *data augmentation*, podem “imaginar” alterações nas imagens, para que tenham um melhor entendimento delas (SHORTEN; KHOSHGOFTAAR, 2019). Portanto, técnicas são usadas para aumentar a diversidade da base de imagens aplicando transformações aleatórias, como transformações geométricas (rotação, translação, espelhamento, etc.) e filtros para realçar imagens (CLAHE, filtros de contraste, etc.), por exemplo.

No entanto, existem algumas desvantagens de usar o *data augmentation*, como memória e tempo de treinamento adicionais, e alto custo computacional (SHORTEN; KHOSHGOFTAAR, 2019). Estas desvantagens ocorrem quando o *data augmentation* é gerado *offline* (em disco) e adicionado ao conjunto de treinamento. Como não é prático

nem eficiente armazenar os dados aumentados na memória, técnicas de *data augmentation* em tempo real são usadas como uma solução para minimizar essas desvantagens. Como o nome sugere, o aumento é aplicado em tempo real, as transformações são realizadas em mini-lotes e, em seguida, inseridas ao treinamento do modelo. Dessa forma, as variações criadas artificialmente a partir de imagens existentes não precisam ser salvas na memória do disco, o que reduz o tempo de acesso ao disco para realizar operações de leitura e escrita.

No método proposto, o *data augmentation* foi realizado em tempo real (*online*) (SHORTEN; KHOSHGOFTAAR, 2019), não sendo necessário incluir novas imagens de *data augmentation* no conjunto de treinamento original, uma vez que foi aplicado diretamente ao conjunto de treinamento. Desta forma, novas imagens são geradas atualizando constantemente o próprio conjunto de treinamento. Essa técnica proporcionou maior poder de generalização e desempenho do modelo e baixo consumo de recursos de máquina. Na subseção 5.5.1 são apresentadas as operações de *data augmentation* em tempo real que foram aplicadas neste estudo.

3.10 Métricas de Desempenho

Para validar os resultados obtidos por um método proposto, quantificações de resultados são comumente adotadas. Esta atividade tem como objetivo avaliar o desempenho do método desenvolvido por meio da análise estatística dos resultados. Neste trabalho, as métricas usadas são comumente aplicadas em sistemas CAD/CADx para análise de imagens médicas: acurácia, sensibilidade, especificidade, coeficiente de similaridade Dice e índice de Jaccard (TAHA; HANBURY, 2015; BLAND, 2015).

O cálculo das métricas é baseado na matriz de confusão (Figura 9), que leva em consideração quatro variáveis: Verdadeiro Positivo (VP) indica a classificação correta dos *pixels* da classe positiva, ou seja, o alvo de segmentação (rins ou tumores renais, dependendo do modelo); Falso Positivo (FP) denota a classificação errônea dos *pixels* da classe negativa como classe positiva; Verdadeiro Negativo (VN) indica a classificação correta dos *pixels* da classe negativa; e Falso Negativo (FN) trata-se da classificação errônea dos *pixels* da classe positiva como classe negativa.

A acurácia (Acc) é definida como a razão entre o número de *pixels* classificados corretamente (classe positiva e negativa) e o número total de *pixels* na amostra do estudo.

Figura 9 – Matriz de confusão.

		Valor Verdadeiro (confirmado por análise)	
		Positivos	Negativos
Valor Previsto (predito pelo teste)	Positivos	VP Verdadeiro Positivo	FP Falso Positivo
	Negativos	FN Falso Negativo	VN Verdadeiro Negativo

Fonte: Adaptado de (DINIZ, 2021).

O cálculo da acurácia é expressa na Equação 3.9,

$$Acc = \frac{VP + VN}{VP + VN + FP + FN}. \quad (3.9)$$

A sensibilidade (Sen) é a capacidade do sistema em predizer corretamente a classe positiva, ou seja, indica a proporção de verdadeiros positivos. O cálculo da sensibilidade é definida na Equação 3.10,

$$Sen = \frac{VP}{VP + FN}. \quad (3.10)$$

A especificidade (Esp) é a capacidade do sistema em predizer corretamente a classe negativa, ou seja, a proporção de verdadeiros negativos. O cálculo da especificidade é definido na Equação 3.11,

$$Spe = \frac{VN}{VN + FP}. \quad (3.11)$$

O Dice, também conhecido como F-score e F-measure, é mais usada na validação da segmentação de imagens médicas (TAHA; HANBURY, 2015). É uma estatística usada para comparar a similaridade de duas amostras. Neste trabalho, é usado para comparar a marcação do especialista e a marcação produzida pelo método. A Equação 3.12 define o cálculo do Dice,

$$Dice = \frac{2VP}{2VP + FP + FN}. \quad (3.12)$$

O índice de Jaccard (Jacc), também conhecido como *Intersection over Union*

(IoU), busca apresentar de forma objetiva o nível de similaridade entre duas amostras. Pode ser calculado pela razão entre a intersecção de dois conjuntos (marcação obtida pelo método e marcação do especialista) e sua união. Em matrizes de confusão empregadas para classificação binária, o índice de Jaccard pode ser descrito pela Equação 3.13,

$$Jacc = \frac{VP}{VP + FP + FN}. \quad (3.13)$$

Além disso, foi realizado o teste de significância que é um procedimento estatístico que permite tomar uma decisão entre duas ou mais hipóteses (rejeitar ou não rejeitar), usando os dados observados de um determinado experimento (COX, 1982; ZOU et al., 2003). Assim, usamos o teste de significância entre os resultados das abordagens apresentados no Capítulo 5. Para isso, é necessário encontrar o *p-value*, que é uma espécie de teste de significância. Inicialmente, devem ser especificados os elementos para a formulação de um plano de análise, são eles: o método de teste e o nível de significância. O método de teste usado foi o teste z de duas proporções (ZOU et al., 2003), pois determina a diferença hipotética entre duas abordagens. O teste z é apropriado quando o tamanho da amostra é relativamente grande e possui uma distribuição normal. E o nível de significância adotado foi $\alpha = 0,05$, onde $p > 0,05$ não rejeita a hipótese nula, o que indica que não há diferença significativa entre as abordagens. O nível de significância aplicado é o mais frequentemente adotado na literatura (GAUVREAU; PAGANO, 1994).

Todas as métricas descritas visam mensurar o desempenho do método proposto como satisfatório ou não, bem como evidenciar aspectos positivos e negativos para trabalhos futuros.

3.11 Considerações Finais

Neste capítulo, foi apresentada a fundamentação teórica necessária para a compreensão das técnicas utilizadas e suas aplicações no método proposto. Foram abordados temas como: rins e tumores renais, exame de TC, técnicas de pré-processamento digital de imagens, aprendizado profundo e métricas de desempenhos.

4 MÉTODO PROPOSTO

Este capítulo descreve o método proposto para segmentar os rins e tumores renais em TCs de pacientes doentes. A base de imagens usada para validar o método proposto é descrita na Seção 5.1. Inicialmente, foi realizada a etapa de pré-processamento em todos os volumes de TC da base de imagens, seguindo fluxos diferentes para os modelos de segmentação. Na segunda etapa, a segmentação inicial dos rins foi realizada usando o modelo ResUNet 2.5D e a segmentação inicial a candidatos de tumores renais foi realizado em duas subetapas: na primeira, os candidatos a tumores renais foram segmentados dentro da região renal usando o modelo DeepLabv3+ 2.5D; na segunda, foram segmentados dentro da região abdominal usando o modelo ResUNet 2.5D. Na terceira etapa, a reconstrução das regiões tumorais é realizada por meio da segmentação de candidatos a tumores renais. Por fim, na etapa de redução de falsos positivos, os resultados finais das segmentações foram alcançados. A Figura 10 ilustra cada uma dessas etapas, e os detalhes são fornecidos nas seções a seguir.

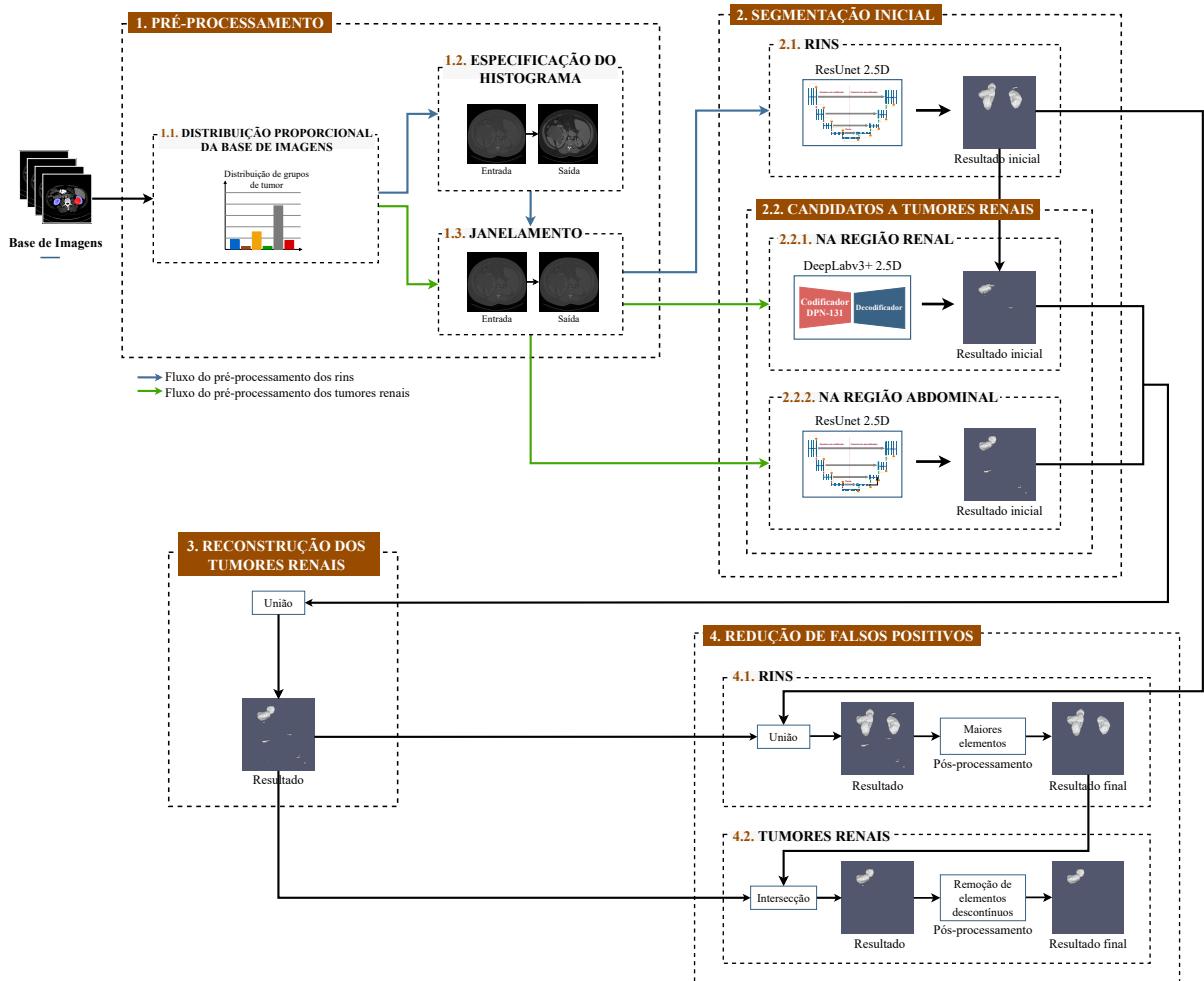
4.1 Pré-processamento

Na primeira etapa do método proposto, é aplicada a subetapa de distribuição proporcional da base da imagem. Esta subetapa é a base para os outros pré-processamentos. Em seguida, as imagens são submetidas às técnicas de melhoramento de imagens: especificação do histograma e janelamento. Essas técnicas visam aprimorar o desempenho dos modelos de segmentação e seguem diferentes fluxos, que são apresentados na Tabela 2. Na segmentação de rins, a especificação do histograma é inicialmente aplicada e, em seguida, o janelamento. Na segmentação a candidatos de tumores renais, apenas o janelamento é aplicado. Nas próximas subseções, cada subetapa dos modelos será explicada com mais detalhes.

Tabela 2 – Aplicação dos pré-processamentos para segmentação dos rins e candidatos a tumores renais.

Imagens	Pré-processamento
Segmentação de rins	<ul style="list-style-type: none"> - Especificação do histograma - Janelamento [-200, 500]
Segmentação de candidatos a tumores renais	<ul style="list-style-type: none"> - Janelamento [-200, 500]

Figura 10 – Etapas do método proposto.



Fonte: Elaborado pela autora.

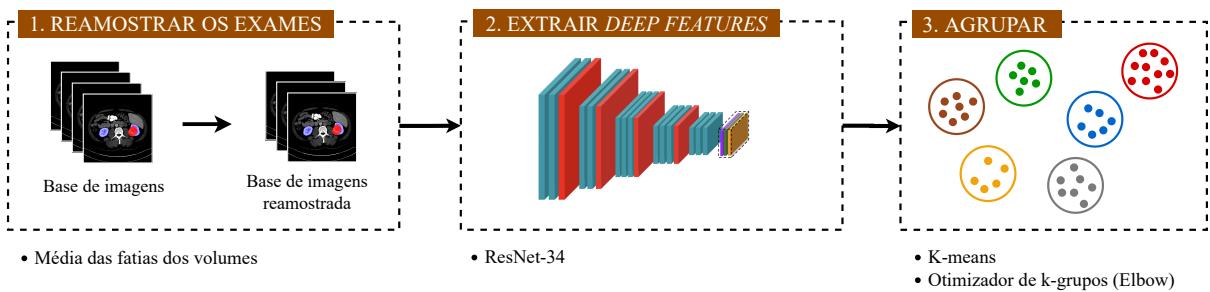
4.1.1 Distribuição Proporcional da Base de Imagens

No problema de segmentação de rins e tumores, sabe-se que eles podem ser heterogêneos em suas características, seja na forma ou na textura (SUN; JONASCH; LARA, 2015). Além disso, para que os modelos de aprendizado profundo sejam robustos, deve haver um bom equilíbrio entre os conjuntos de treinamento e validação para evitar problemas como *overfitting* (JOHNSON; KHOSHGOFTAAR, 2019).

Diante disso, foi desenvolvido um método automático com o objetivo de agrupar casos (exames) semelhantes de uma base de imagens, a fim de distribuí-los proporcionalmente entre os conjuntos de treinamento e validação, para garantir o bom desempenho dos modelos de segmentação. O método consiste em três etapas principais: (1) reamostrar os exames; (2) extrair *deep features*; e (3) agrupar. A Figura 11 ilustra as etapas descritas. Primeiramente, é importante destacar que para realizar o método discutido, são selecio-

nadas apenas as fatias dos exames de TC que contêm tumores do conjunto de dados de treino e validação da base de imagens escolhida. Isso é necessário porque as regiões mais relevantes para extraer características significativas são encontradas nas fatias que contém a classe positiva, ou seja, o alvo para a extração de *deep features* (tumores).

Figura 11 – Método proposto para agrupar os exames semelhantes.



Fonte: Elaborado pela autora.

Na primeira etapa, é realizada a técnica de reamostragem (DODGSON, 1992; LI; MUELLER; ERNST, 2004) das tomografias, que é uma técnica matemática usada para criar uma nova versão da imagem com diferentes larguras, alturas e/ou profundidades em *pixels* (DODGSON, 1992; SACHS, 2001). Essa etapa é necessária porque os exames de TC podem ter diferentes quantidades de fatias e precisam ser reamostrados para o mesmo número de fatias para que os dados de entrada das etapas consecutivas sejam normalizados. Essa transformação dos dados é uma prática comum para evitar que o algoritmo fique enviesado para as variáveis com maior ordem de grandeza (KIMBLE; MIOLIDAKIS, 2015).

Para isso, padroniza-se o número de fatias em cada exame de TC por meio da operação de média, que consiste em calcular a quantidade total de fatias de tumores contidos na base de imagens (conjunto de treino e validação) e dividir pela quantidade de exames totais. Como resultado, obtém-se o valor médio de fatias. Posteriormente, a reamostragem é aplicada aos exames usando a informação do valor médio e interpolação linear, que produz uma superfície de intensidade garantida de forma contínua (ANDREWS; PATTERSON, 1976; DODGSON, 1992). Por fim, com a técnica de reamostragem, todos os exames passam a ter um número n de fatias.

Na segunda etapa, são extraídas as *deep features* que servirão como dados de entrada para a próxima etapa. As *deep features* são extraídas pelos modelos ResNet-18,

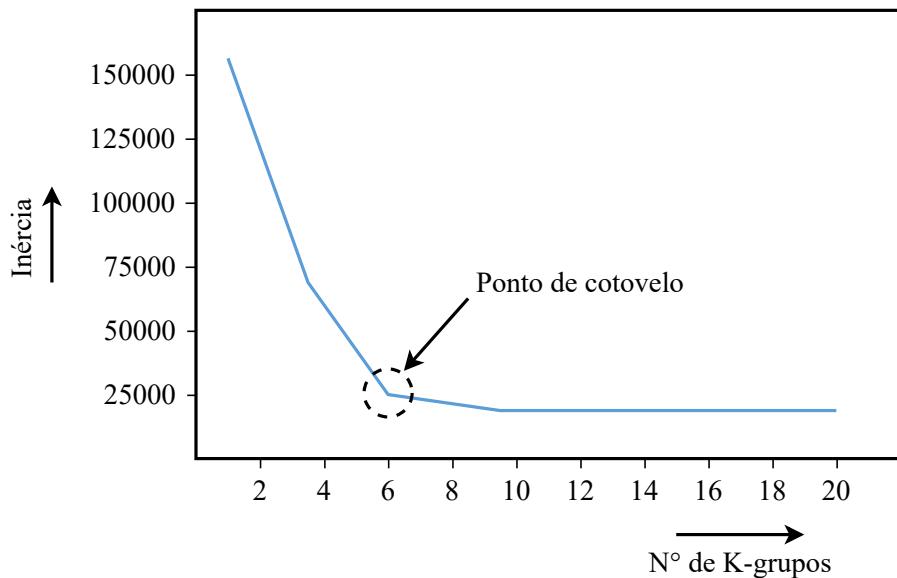
ResNet-34, ResNet-101 (HE et al., 2016), VGG-11, VGG-16 (SIMONYAN; ZISSERMAN, 2014), DPN-131 (CHEN et al., 2017) e Xception (CHOLLET, 2017). Esses modelos foram escolhidos por estarem consolidados na literatura para essa tarefa (TüRKOĞLU; HANBAY, 2019), apresentando um alto desempenho, além de já terem sido pré-treinados. Todos os modelos realizam implicitamente a extração e seleção de características da região dos tumores. As características obtidas são extraídas da última camada de convolução de cada arquitetura. Os modelos são inicializados usando pesos pré-treinados aplicados à base de imagens ImageNet (DENG et al., 2009).

Finalmente, na terceira etapa, as características extraídas por cada modelo são utilizadas individualmente para realizar experimentos relacionados ao agrupamento de dados. O agrupamento é o processo de dividir os dados inteiros em grupos (também conhecidos como *clusters*) com base nos padrões dos dados. Portanto, esta etapa visa agrupar os exames que possuam características semelhantes. Para isso, foi usado o K-means (MACQUEEN et al., 1967; HAMERLY; ELKAN, 2003), que é o algoritmo de agrupamento mais comumente usado na literatura, para dividir o conjunto de dados em um conjunto de K-grupos. No K-means, cada grupo é representado pelo seu centro (isto é, centroide) que corresponde à média dos pontos atribuídos ao grupo. Em geral, as características extraídas de cada exame na segunda etapa são classificadas em grupos, de modo que tenham alta similaridade intraclass e baixa similaridade interclasse.

Inicialmente, o modelo K-means é parametrizado em cada modelo de rede com diferentes números de K-grupos (de 2 a 20). Para cada valor de K, a inércia é calculada para medir o quanto bem um conjunto de dados foi agrupado pelo K-means. A inércia é calculada medindo a distância entre cada ponto de dados e seu centroide, elevando essa distância ao quadrado e somando esses quadrados em um grupo. Posteriormente, o hiperparâmetro K é otimizado com o método do cotovelo (Elbow) (JOSHI; NALWADE, 2013), que é frequentemente usado para encontrar o número ideal de grupos. Na Figura 12 o valor da inércia para cada K associado é ilustrado usando o método do cotovelo. Pode-se observar que à medida que o número de grupos aumenta, o valor da inércia começa a diminuir. Além disso, o gráfico muda rapidamente em um ponto, criando uma forma de cotovelo, a partir do qual o gráfico começa a se mover quase paralelo ao eixo X. O método do cotovelo seleciona esse ponto do cotovelo no gráfico de inércia, pois o valor K correspondente é o valor K ótimo ou um número ótimo de grupos.

Sabe-se que a ideia por trás de um bom agrupamento é ter um pequeno valor de

Figura 12 – Exemplo do método do cotovelo para seleção o número ideal de K-grupos.



Fonte: Elaborado pela autora.

inércia e um pequeno número de agrupamentos (JOSHI; NALWADE, 2013). Portanto, este mesmo processo descrito acima é realizado para cada modelo de rede a fim de encontrar o seu melhor grupo. Posteriormente, dentre os modelos, deve-se selecionar aquele com menor valor de inércia, pois representará o modelo com melhor número de grupos. Isso completa o método para agrupar exames semelhantes.

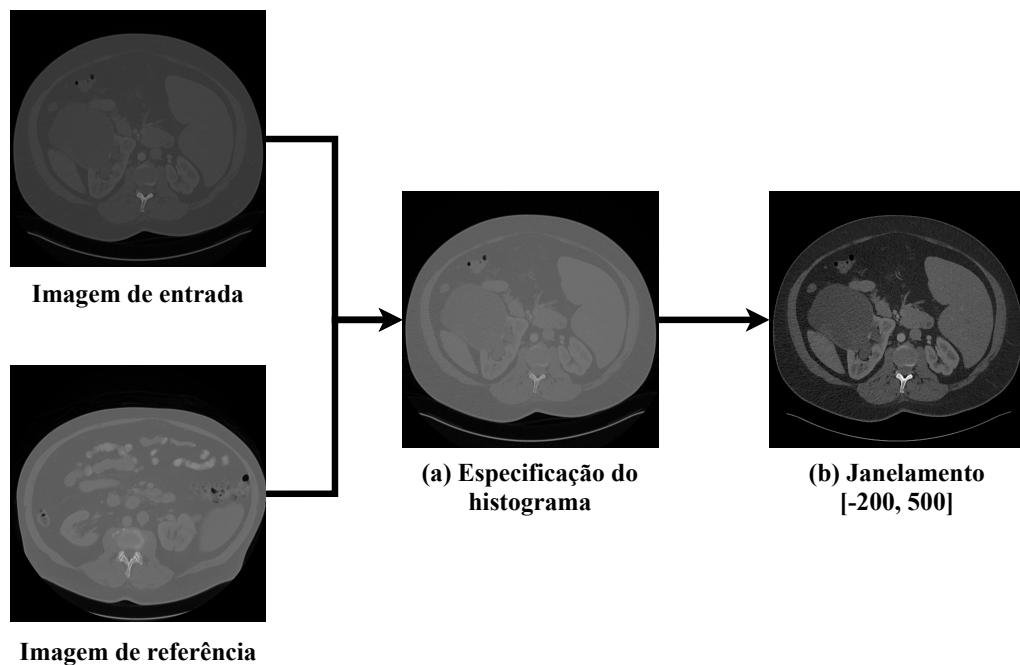
Vale ressaltar que antes de aplicar as etapas descritas, o conjunto de teste é escolhido aleatoriamente da base de imagens. Em sequência, os exames de cada um dos grupos gerados pelo método proposto são selecionados aleatoriamente e distribuídos proporcionalmente entre os conjuntos de dados de treinamento e validação. Isso garante um modelo mais equilibrado, pois deve haver exames de todos os grupos de tumores em ambos os conjuntos de dados. Consequentemente, a generalização do modelo é aumentada.

Depois de distribuir proporcionalmente a base de imagens, alguns outros pré-processamentos foram aplicados para aprimorar as imagens de TC. Conforme já mencionado, esses outros pré-processamentos foram realizados em diferentes fluxos (descrito nas próximas subseções) para cada modelo de segmentação. Posteriormente, as imagens pré-processadas foram aplicadas para segmentar os rins e candidatos a tumores renais.

4.1.2 Pré-processamentos Aplicados a Segmentação dos Rins

Antes de treinar o modelo de segmentação dos rins, os exames de TC são inicialmente submetidos a um processo de pré-processamento dividido em duas etapas. Essas etapas são ilustradas na Figura 13. A primeira etapa (Figura 13 (a)) é a normalização das intensidades dos *voxels* dos volumes de TC uma vez que as mesmas regiões dos rins podem ter intensidades muito diferentes em volumes diferentes, as quais podem dificultar uma eventual comparação das características de textura das regiões renais. Portanto, a especificação do histograma é aplicada para aproximar o histograma de um volume de TC ao histograma de um volume modelo escolhido aleatoriamente (fatia do volume modelo na Figura 13), de modo que ambos tenham uma distribuição dos *voxels* semelhante (GONZALEZ; WOODS, 2008).

Figura 13 – Pré-processamento: (a) especificação do histograma; (b) janelamento.



Fonte: Elaborado pela autora.

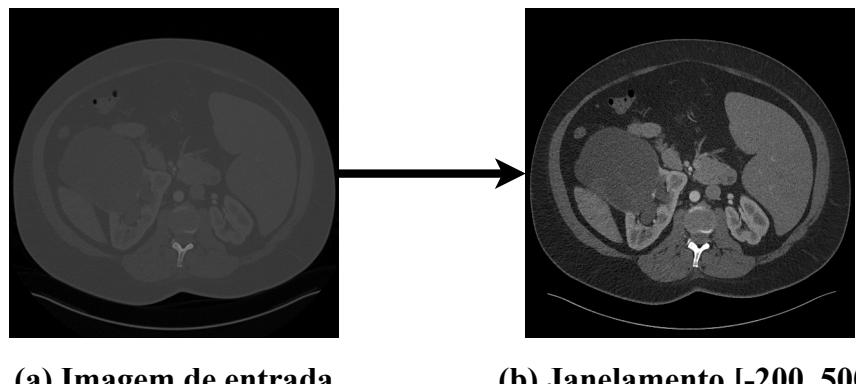
Devido a existência de ossos, ar no intestino e outros órgãos, a faixa dos valores de TC nas imagens pode variar de -10.240 a mais de 18.000 HU. Valores tomográficos de tecidos moles, como rins e tumores renais, têm a mesma distribuição, variando de aproximadamente -200 HU a 500 HU (BUZUG, 2011; ADAMS et al., 2012; YANG et al., 2018). Portanto, na segunda etapa, os valores de intensidades dos volumes foram limitados para a faixa de [-200, 500] HU (Figura 13 (b)). Este processo consiste em um janelamento, no qual se preserva os valores de *voxels* que estão no intervalo entre -200 e

500 HU, e os *voxels* com valores inferiores ou superiores a este intervalo são atribuídos os valores -200 e 500 HU, respectivamente. Isso melhora o contraste das imagens, facilitando a diferenciação das estruturas renais de outros órgãos (YANG et al., 2018; CRUZ et al., 2020). Em seguida, os valores de intensidade limitados foram normalizados entre 0 e 1. Esta transformação melhora a estabilidade da otimização, reduzindo a influência do problema de explosão de gradiente (BROWNLEE, 2019; BOHRA, 2021).

4.1.3 Pré-processamentos Aplicados à Segmentação Inicial de Candidatos a Tumores Renais

Para treinar os modelos de segmentação inicial a candidatos de tumores renais, as imagens de TC foram primeiramente submetidas ao pré-processamento de janelamento [-200, 500], que é o mesmo aplicado na segmentação de rins (Seção 4.1.2). A Figura 14 ilustra esse processo. Posteriormente, os valores de intensidade limitados com o janelamento também foram normalizados entre 0 e 1 (PREPROCESSING; AUGMENTATION, 2017).

Figura 14 – Pré-processamento: (a) imagem original; (b) janelamento.



Fonte: Elaborado pela autora.

4.2 Segmentação Inicial

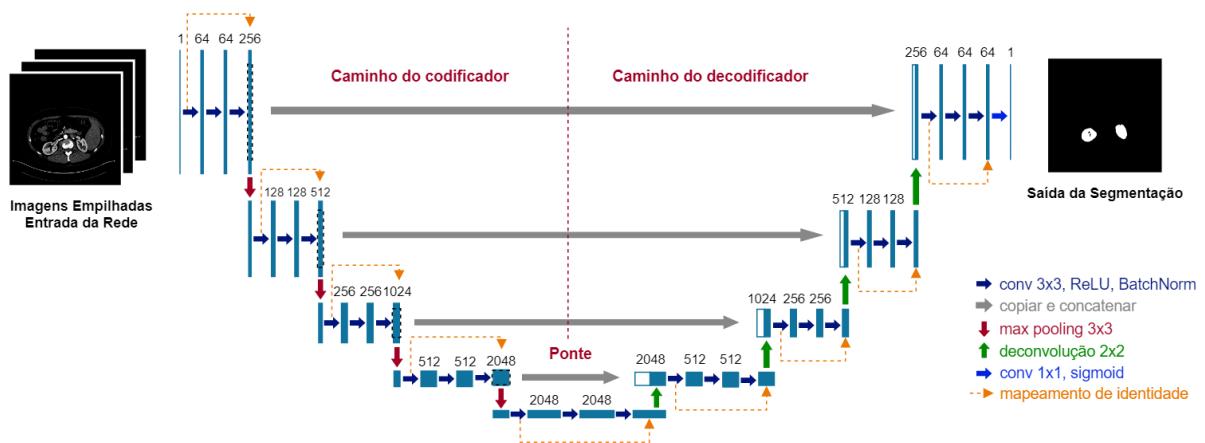
Esta etapa visa a segmentação inicial dos rins e candidatos a tumores renais. Inicialmente, o modelo ResUNet foi usado para segmentar os rins. Para segmentar os candidatos a tumores renais, foram abordados dois estágios. O primeiro estágio usa o modelo DeepLabv3+ e o segundo estágio usa o modelo ResUNet. Nas próximas subseções serão descritos os procedimentos para segmentar os rins e candidatos a tumores renais.

4.2.1 Segmentação dos rins usando a ResUNet

A ResUNet é uma arquitetura desenvolvida por Zhang, Liu e Wang (2018) para segmentação semântica. Foi inicialmente usada para extração de estradas a partir de imagens aéreas, mais tarde usada para várias outras aplicações, como segmentação de tumor cerebral e imagens humanas. Essa arquitetura é inspirada nos pontos fortes dos modelos de aprendizagem residual (HE et al., 2016) e U-Net (RONNEBERGER; FISCHER; BROX, 2015). Em resumo, a rede é construída com unidades residuais e tem arquitetura semelhante à da U-Net. As vantagens deste modelo são duplas: em primeiro lugar, as unidades residuais facilitam o treinamento de redes profundas; em segundo, as ricas conexões de salto dentro da rede podem facilitar a propagação de informações renais, permitindo projetar redes com menos parâmetros e mais desempenho.

A arquitetura ResUNet-101 usada neste trabalho é ilustrada na Figura 15. As imagens de entrada da rede são passadas por uma estrutura composta por três partes: codificação, ponte e decodificação. A primeira parte codifica a imagem de entrada em representações compactas. A última parte recupera as representações para uma segmentação semântica. A parte do meio serve como uma ponte conectando os caminhos de codificação e decodificação.

Figura 15 – Arquitetura ResUNet-101.



Fonte: Elaborado pela autora.

A ResUNet usa unidades residuais como bloco de construção básico em vez de bloco convolucional simples. O caminho de codificação e decodificação é composto por quatro unidades residuais. As três partes são construídas com unidades residuais que

consistem em dois blocos de convolução 3×3 e um mapeamento de identidade. Cada bloco de convolução inclui uma camada *batch normalization*, uma camada de ativação ReLU e uma camada convolucional. O mapeamento de identidade conecta (adição) a entrada e a saída da unidade. No caminho de decodificação, antes de cada unidade, há uma deconvolução de mapas de características de nível inferior e uma concatenação com os mapas de características do caminho de codificação correspondente. Após o último nível do caminho de decodificação, uma convolução 1×1 e uma camada de ativação sigmoide são usadas para projetar os mapas de características para a segmentação dos rins.

Resumidamente, a abordagem residual consiste em propagar informações sobre camadas, inserindo conexões de atalho entre as camadas de entrada e saída. Essas conexões de atalho simplesmente executam o mapeamento de identidade e suas saídas são adicionadas às saídas das camadas empilhadas. Assim, a inserção dos mapas de entrada na saída de cada camada evita que o conjunto de operações de *pooling* reduza as informações necessárias contidas nas imagens da base. Portanto, este modelo foi escolhido devido ao seu alto desempenho para segmentação renal, uma vez que as informações são propagadas com a abordagem residual, retendo e agregando características renais relevantes.

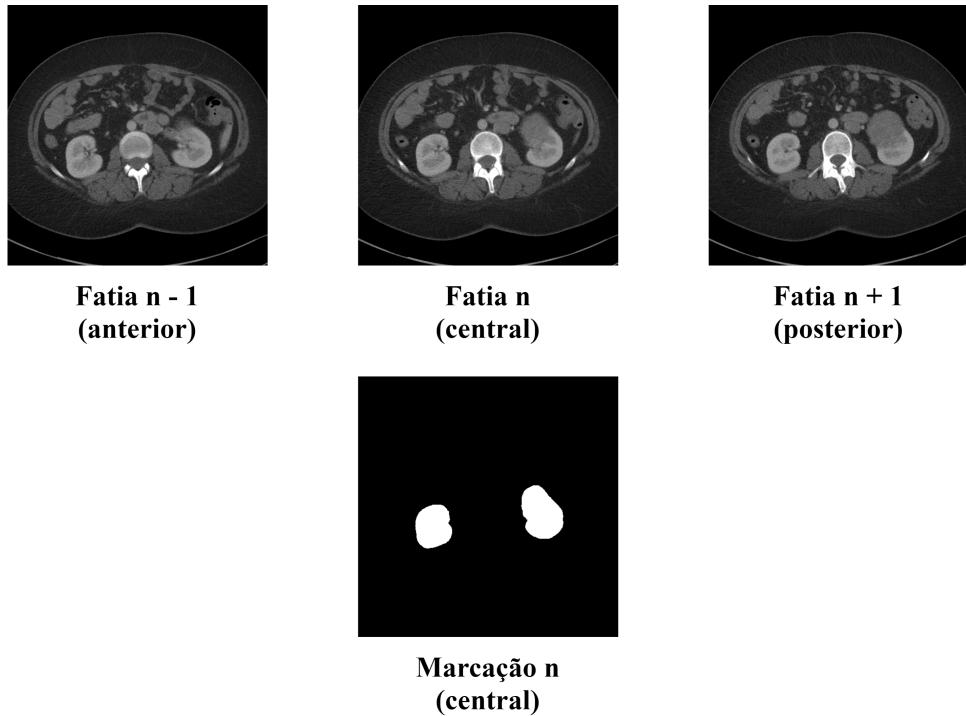
4.2.1.1 Treinamento da ResUNet 2.5D

A rede é treinada com imagens de TC no tamanho original de 512×512 pixels. O modelo usa uma abordagens 2.5D, que consiste em passar como entrada da rede uma pilha de três fatias consecutivas (anterior, central e posterior) e a máscara referente à fatia central com as marcações dos rins. Para analisar todo o volume de TC, o mesmo processo descrito foi realizado deslizando uma janela de tamanho 3×3 com um passo igual a 1 sobre as fatias. A saída da ResUNet é a máscara para a segmentação da fatia central da pilha. Um exemplo de fatias de entrada da rede usando a abordagem 2.5D é mostrado na Figura 16.

Uma das vantagens da abordagem 2.5D é que pode-se usar informações espaciais das fatias vizinhas para identificar o objeto de interesse, aumentando as chances de segmentação bem-sucedida e reduzindo os falsos positivos. Isso só foi possível devido à natureza da abordagem 2.5D, que possui conexões entre as fatias do volume de TC.

Além da abordagem 2.5D, um balanceamento de fatias de rins e não rins foi usado nos conjuntos de treinamento e validação. O balanceamento consistiu em usar as fatias

Figura 16 – Abordagem 2.5D: exemplo de imagens de entrada.



Fonte: Elaborado pela autora.

com rins e adicionar a mesma quantidade de fatias sem rins e depois retirar o excesso de fatias sem rins. Essa estratégia ajuda o modelo de aprendizado profundo a detectar se uma fatia tem rins ou não, a qual reduziu consideravelmente os falsos positivos.

Por fim, o treinamento da ResUNet 2.5D foi inicializado usando os pesos pré-treinados aplicados à base de imagens ImageNet (DENG et al., 2009). Essa prática ajuda a minimizar o tempo de treinamento e economizar recursos de *hardware*.

4.2.2 Segmentação Inicial de Candidatos a Tumores Renais

Esta etapa visa a segmentação inicial de candidatos a tumores renais. Para isso, foi dividida em dois estágios: o primeiro estágio usa os resultados da segmentação inicial dos rins como entrada para segmentar os candidatos a tumores renais usando o modelo DeepLabv3+; o segundo estágio usa outro modelo ResUNet para segmentar candidatos a tumores renais dentro da região abdominal (imagem completa). Mais detalhes de cada estágio são descritos nas próximas subseções.

4.2.2.1 Candidatos de Tumores Renais na Região Renal

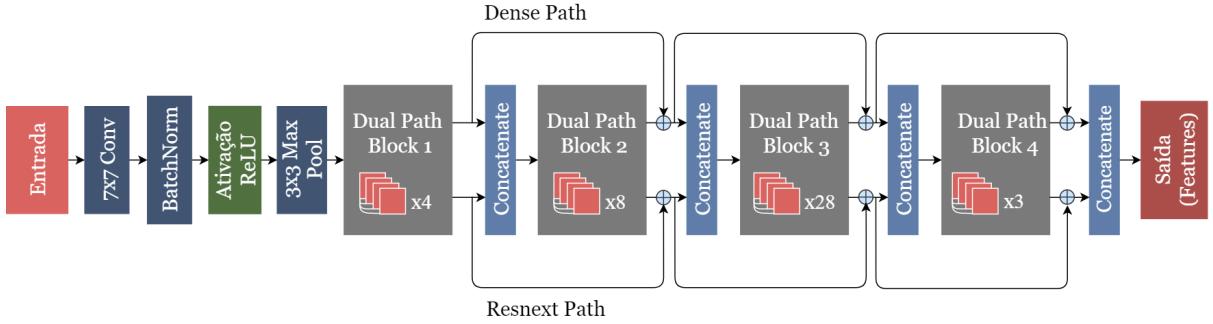
Como já mencionado, este primeiro estágio visa obter candidatos a tumores renais dentro da região renal. Para isso, os resultados adquiridos na etapa de segmentação de rins foram usados como entrada para segmentar os candidatos a tumores renais. Neste estágio, o modelo de aprendizado profundo DeepLabv3+ (Seção 3.8) foi adaptado para realizar a segmentação das regiões tumorais. Esta adaptação consiste em remover o codificador padrão (Xception (CHOLLET, 2017)) da DeepLabv3+, pela arquitetura de rede de caminho duplo (*Dual Path Network - DPN*)-131. Para isso, foi necessário remover a camada totalmente conectada da DPN-131 para funcionar como um extrator de características no modelo proposto.

A DPN é uma rede de classificação que apresenta uma topologia de caminhos de conexão interna. A DPN compartilha as vantagens da Rede Residual (*Residual Network - ResNet*) (HE et al., 2016), que permite a reutilização de recursos, e da Rede Densamente Convolucional (*Densely Convolutional Network - DenseNet*) (HUANG et al., 2017), que explora os novos recursos que são importantes para o aprendizado de boas representações (CHEN et al., 2017). A DPN-131 empilha vários blocos *dual path* modularizados, conforme mostrado na Figura 17. Inicialmente, uma convolução 7×7 é aplicada, seguida de *batch normalization*, função de ativação ReLU e *max-pooling* 3×3 . A estrutura de cada bloco é projetada com um estilo de gargalo (HE et al., 2016), que começa com uma camada convolucional 1×1 , seguida por uma 3×3 e finalizando com uma 1×1 .

Assim, a saída da última camada convolucional 1×1 é dividida em duas partes. Na primeira parte, é adicionado ao caminho residual e, na segunda parte, é concatenado com o caminho densamente conectado. Em outras palavras, as redes residuais adicionam os recursos de entrada aos recursos de saída por meio do caminho residual, e as redes densas usam um caminho densamente conectado para concatenar recursos de entrada com recursos de saída, permitindo que cada bloco receba informações de todos os blocos anteriores. A camada de convolução agrupada na segunda camada é usada como ResNeXt (XIE et al., 2017), aumentando a capacidade de aprendizado de cada bloco.

Como a arquitetura DPN-131 padrão foi usada como o codificador da DeepLabv3+, foi necessário remover a camada totalmente conectada do DPN-131 para funcionar como um extrator de características, conforme visto na Figura 17. Além disso, embora várias funções de ativação tenham sido aplicadas com redes profundas para obter alto

Figura 17 – Arquitetura DPN-131.



Fonte: Elaborado pela autora.

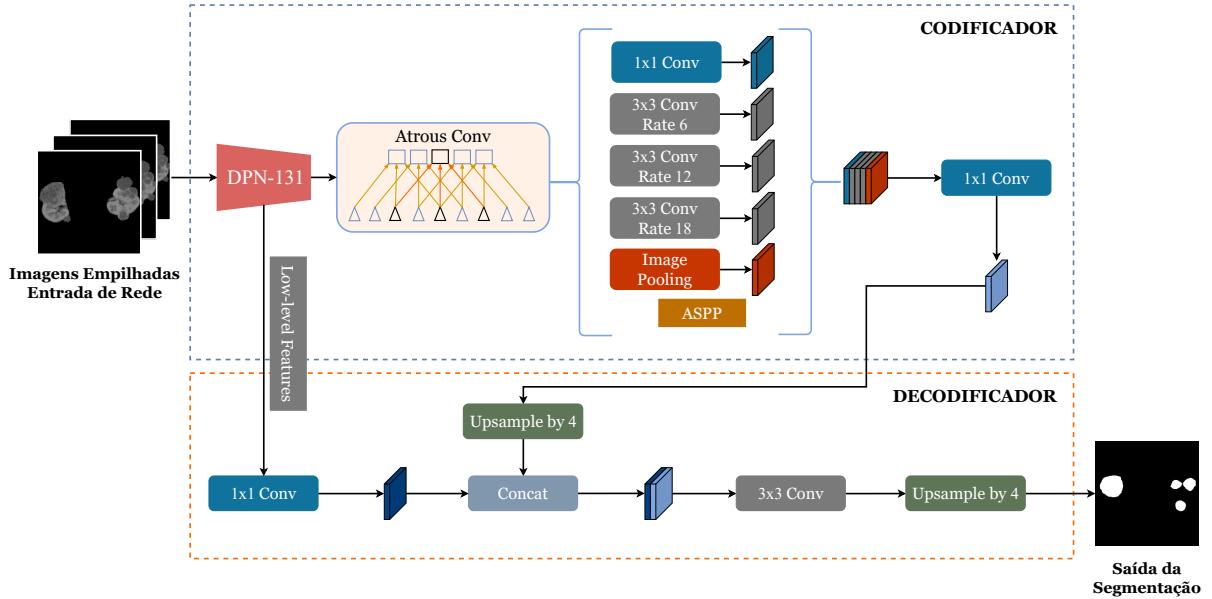
desempenho (GOCERI, 2019; GOCERI, 2021), foi usada a função ReLU devido à sua eficiência na segmentação da base de imagens e baixo custo computacional.

Finalmente, pode ser observado na Figura 18 a arquitetura DeepLabv3+ usada para a segmentação a candidatos de tumores renais. A escolha da DPN-131 foi feita por ser eficiente, introduzindo uma nova topologia de caminhos de conexão internamente. Além disso, vale ressaltar que a DeepLabv3+ é um modelo de aprendizado profundo de última geração capaz de refinar os resultados da segmentação com foco especial nos limites do objeto. Outro ponto importante é que a rede extraí mapas de características densos para capturar contextos de longo alcance (informações contextuais de pixels mais distantes), melhorando a tarefa de segmentar os diferentes tamanhos e formas dos tumores encontrados na base de imagens. Portanto, a combinação resultou em um codificador-decodificador robusto, com alto desempenho para segmentação a candidatos de tumores renais.

4.2.2.1.1 Treinamento da DeepLabv3+ 2.5D

Antes de realizar o treinamento, as regiões renais foram inicialmente recortadas automaticamente nas imagens de TC. Os rins foram encontrados usando a marcação do especialista e aplicando uma caixa delimitadora 3D que envolvesse os dois rins (Figura 19 (a)). Então, o volume é cortado para o valor correspondente à caixa delimitadora 3D (Figura 19 (b)). Posteriormente, um redimensionamento proporcional para 256×256 pixels foi aplicado às imagens cortadas maiores que 256 pixels de altura ou largura (Figura 19 (c)). Finalmente, as imagens foram centralizadas com um tamanho final de 256×256 pixels e apenas a textura da região do rim é mantida (Figura 19 (d)) para a segmentação

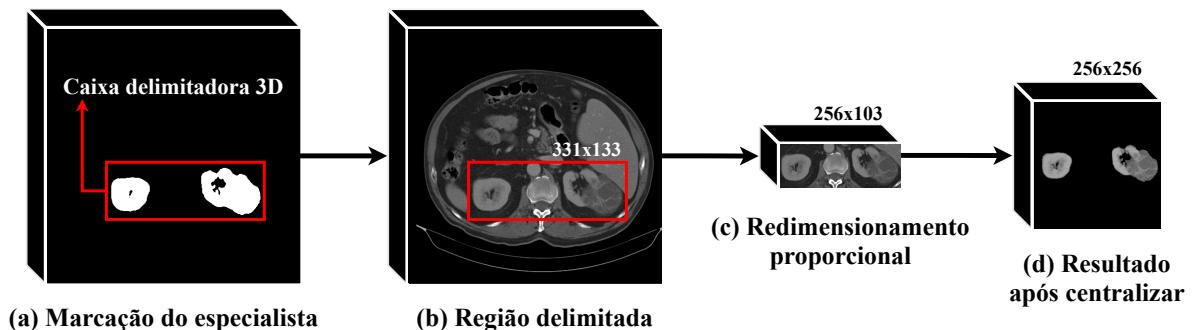
Figura 18 – Arquitetura DeepLabv3+.



Fonte: Adaptado de (CHEN et al., 2018).

da região dos tumores. Todo esse processo foi realizado porque havia muitas regiões irrelevantes (outros órgãos) para capturar características. Também foi necessário devido às limitações de *hardware*, já que a arquitetura é profunda e requer memória considerável para representar todos os mapas de características.

Figura 19 – Delimitação da região de interesse.



Fonte: Elaborado pela autora.

Finalmente, o modelo DeepLabv3+ foi treinado usando abordagem 2.5D, conforme descrito na Seção 4.2.1.1. No entanto, foram passadas para entrada da rede três fatias consecutivas da região dos tumores e a máscara referente a fatia central com as marcações dos tumores.

Nos experimentos realizados também aplicou-se o balanceamento de fatias (Se-

ção 4.2.1.1) nos conjuntos de treinamento e validação. Entretanto, o balanceamento de fatias para o modelo de segmentação a candidatos de tumores renais na região renal consistiu em usar as fatias com tumores e somar a mesma quantidade de fatias sem tumores e, posteriormente, remover as fatias excedentes sem tumores. Isso reduziu bastante os falsos positivos, devido às regiões que são muito semelhantes a tumores renais, como cistos.

A inicialização do treinamento da DeepLabv3+ 2.5D também utilizou os pesos pré-treinados aplicados à base de imagens ImageNet (DENG et al., 2009).

4.2.2.2 Candidatos de Tumores Renais na Região Abdominal

O segundo estágio também visa segmentar candidatos a tumores renais. No entanto, toda a região abdominal é analisada. Isso é feito porque no primeiro estágio o modelo está limitado a segmentar candidatos a tumores renais apenas dentro da região renal resultante da segmentação inicial dos rins. Esse procedimento pode acabar não trazendo os resultados desejados por falta de informações contextuais e porque algumas regiões renais não foram obtidas na segmentação inicial dos rins. Logo, essas regiões poderiam ter tumores renais e acabaram não tendo a possibilidade de segmentá-los.

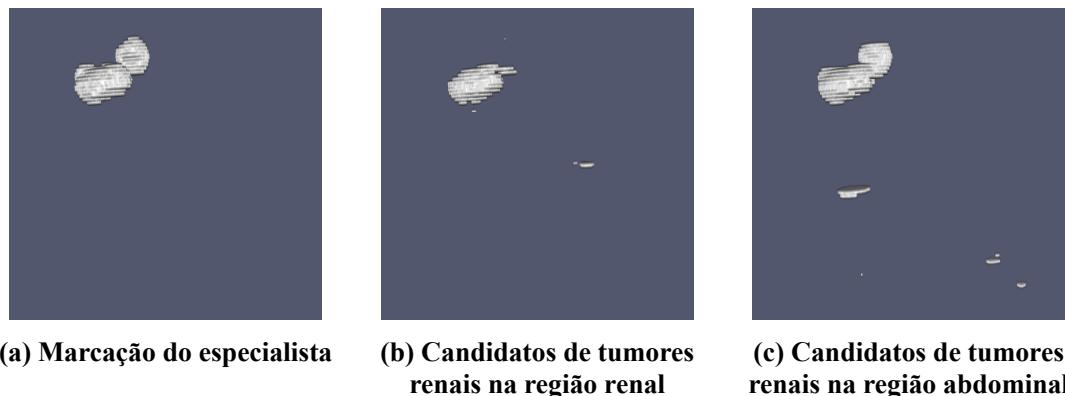
Portanto, o segundo estágio tem o objetivo de segmentar os candidatos a tumores renais na região abdominal, analisando mais informações contextuais em uma região mais ampla. Vale ressaltar que essa técnica é mais suscetível a falsos positivos, pois o tumor renal não tem um formato padronizado, o que pode causar confusão com outras estruturas (órgãos). Entretanto, o modelo traz resultados interessantes, pois é capaz de segmentar tumores renais que não haviam sido segmentados no estágio anterior.

Dessa forma, um segundo modelo de aprendizado profundo é treinado para realizar a segmentação de candidatos a tumores renais na região abdominal. Para isso, é usado o modelo ResUNet que é a mesma arquitetura descrita na Seção 4.2.1. No entanto, o treinamento do modelo é realizado usando apenas as marcações dos tumores renais, mas com informações mais ampla (regiões abdominais). Ou seja, o modelo é especializado nas segmentações dos tumores nas imagens de TC. Além disso, o modelo também usa a abordagem 2.5D e balanceamento de fatias (Seção 4.2.2.1.1).

A Figura 20 ilustra a segmentação inicial dos candidatos a tumores renais em uma imagem de TC aplicada à entrada dos modelos (DeepLabv3+ e ResUNet). Na Figura 20 (a) apresenta a marcação do especialista. O resultado da etapa de candidatos a tumores

renais na região renal é mostrado na Figura 20 (b). Finalmente, os candidatos a tumores renais na região abdominal que podem não ter sido segmentados no primeiro estágio são ilustrados na Figura 20 (c). No entanto, nota-se que alguns fragmentos semelhantes a tumores renais também foram segmentados.

Figura 20 – Candidatos a tumores renais: (a) marcação do especialista; (b) candidatos de tumores renais na região renal; (c) candidatos de tumores renais na região abdominal.

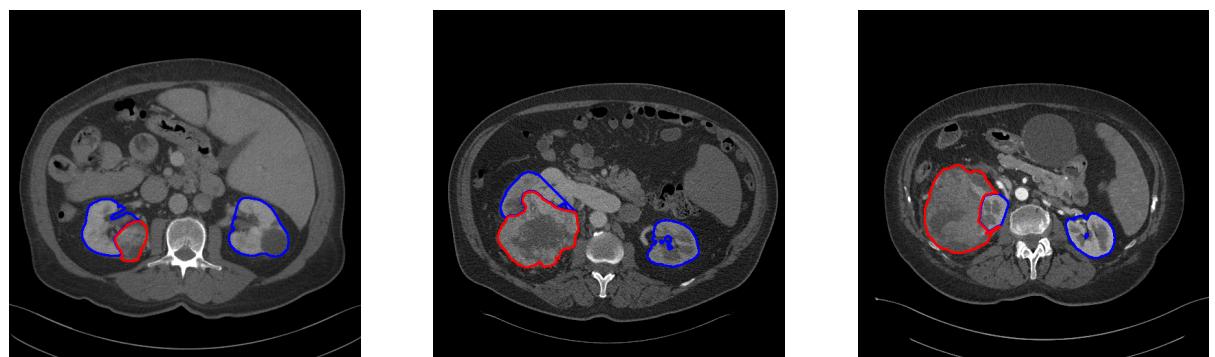


Fonte: Elaborado pela autora.

4.3 Reconstrução dos Tumores Renais

Em bons cenários, onde a segmentação inicial dos rins foi capaz de obter grande parte da região renal, e em que os pacientes apresentam tumores bem definidos, apenas o primeiro estágio é suficiente para alcançar bons resultados. Porém, em casos difíceis como surgimento de cistos, texturas homogêneas, tumores em estágios avançados (Figura 21) e perda de regiões renais na segmentação inicial dos rins, o modelo acaba não segmentando bem os tumores.

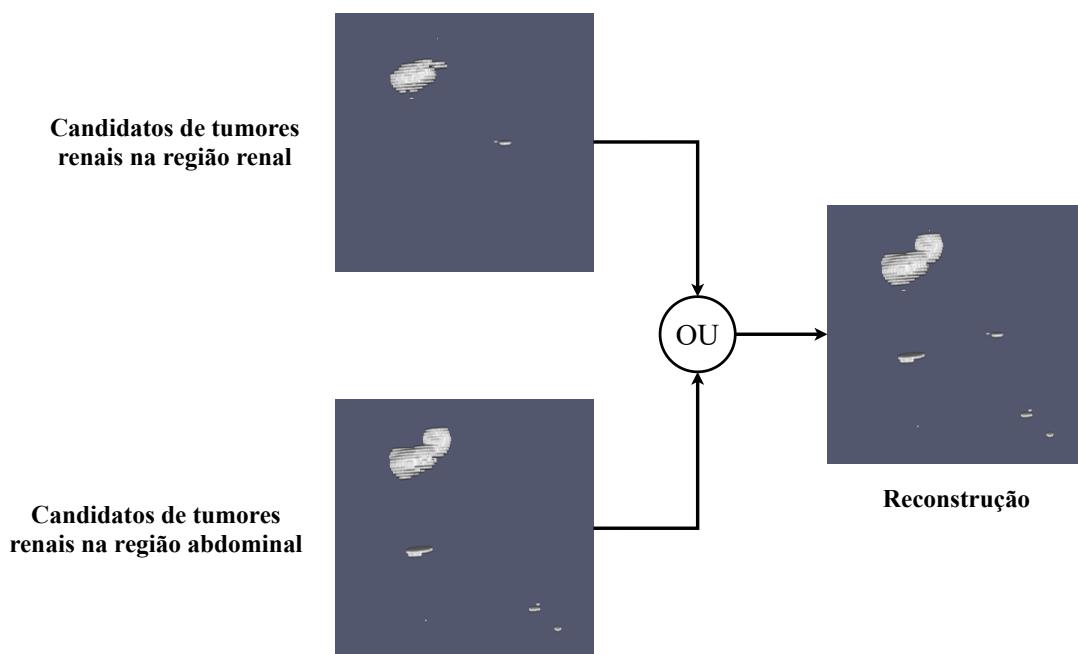
Figura 21 – Exemplos de casos difíceis. Rins marcados em azul e tumores renais em vermelho.



Fonte: Elaborado pela autora.

Dessa forma, o segundo estágio é um fator importante para segmentar regiões que não puderam ser segmentadas no primeiro estágio. No entanto, afim de desenvolver um modelo mais robusto de segmentação de candidatos a tumores, foi realizada uma reconstrução de tumores renais. Esta etapa faz a união entre o primeiro e o segundo estágio usando a operação binária “OU” nas duas máscaras de segmentação. Isso leva a uma segmentação mais precisa da região tumoral, pois partes que não foram segmentadas no primeiro estágio serão segmentadas no segundo estágio, e posteriormente unidas nesta etapa. A Figura 22 mostra um exemplo desse processo.

Figura 22 – Etapas da reconstrução de tumores renais.



Fonte: Elaborado pela autora.

O principal benefício da reconstrução dos tumores renais é que esta etapa é capaz de unir porções consideráveis das regiões tumorais, o que resultou em uma melhor definição dos contornos e regiões tumorais. No entanto, também pode segmentar várias regiões que não são tumores. Porém, na etapa final de segmentação, essas regiões segmentadas extras são removidas usando o pós-processamento.

4.4 Redução de Falsos Positivos

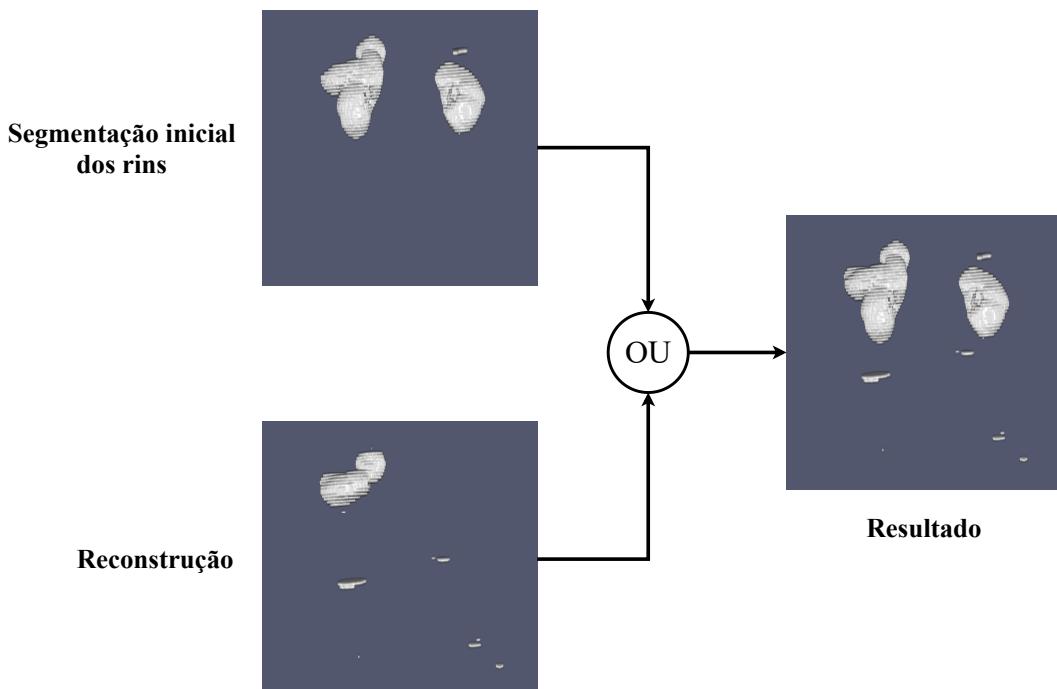
Nesta etapa são obtidos os resultados finais das segmentações dos rins e tumores. Para isso, o resultado da etapa de reconstrução dos tumores renais é fundamental para

se obter o resultado final dos rins e tumores. Mais detalhes são fornecidos nas próximas subseções.

4.4.1 Redução de Falsos Positivos para os Rins

Para obter a segmentação final dos rins, foi feito inicialmente uma união da segmentação inicial dos rins com o resultado da etapa de reconstrução dos tumores renais, usando a operação binária “OU” (Figura 23). O objetivo é melhorar os contornos que não foram obtidos na etapa de segmentação inicial dos rins por meio da reconstrução dos tumores renais, uma vez que os tumores renais também fazem parte da região renal.

Figura 23 – União da segmentação inicial dos rins e reconstrução dos tumores renais.



Fonte: Elaborado pela autora.

No entanto, isso também pode afetar a segmentação inicial dos rins. Portanto, uma etapa de pós-processamento é necessária para remover as regiões que não fazem parte dos rins que foram segmentadas na etapa de união descrita anteriormente. Essa etapa consiste em manter os dois maiores elementos do volume que corresponde aos rins e remover o restante dos fragmentos segmentados (falsos positivos). A Figura 24 (a) mostra o resultado em 3D da união descrita anteriormente e a Figura 24 (b) a marcação do especialista. Na Figura 24 (c) encontra-se o resultado visual após a etapa de pós-

processamento. Pode-se observar que alguns fragmentos não renais foram removidos e, consequentemente, a precisão da segmentação dos rins também aumentou.

Figura 24 – Resultado final da segmentação de rins após pós-processamento.



Fonte: Elaborado pela autora.

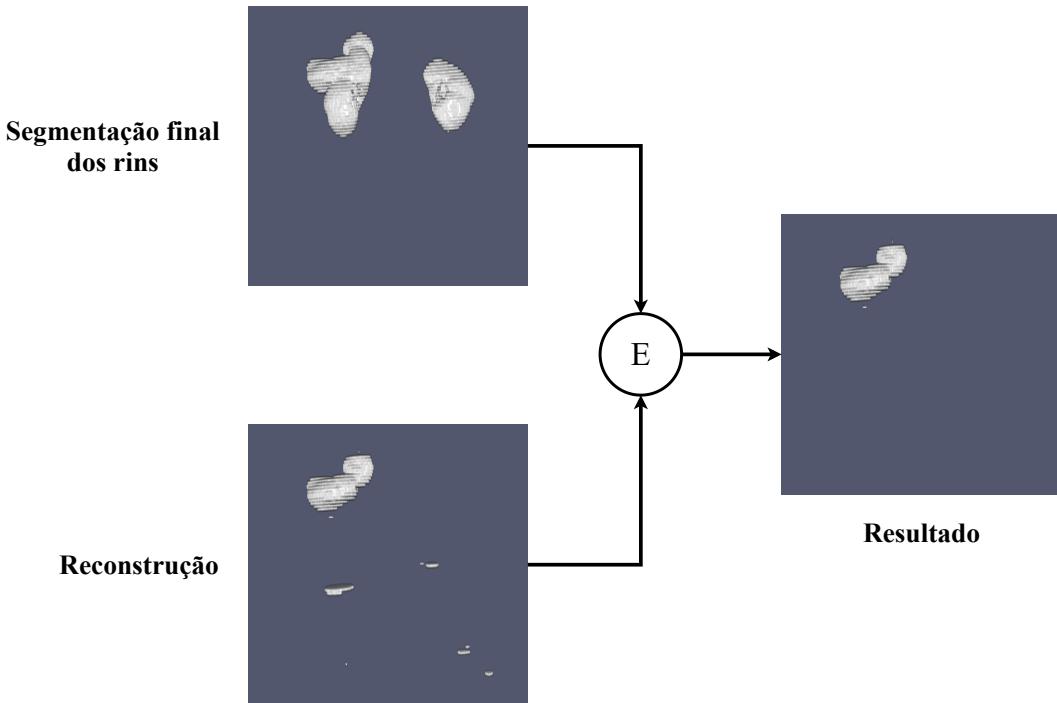
4.4.2 Redução de Falsos Positivos para os Tumores Renais

Nesta subseção são apresentados os procedimentos para obtenção da segmentação final dos tumores renais. Inicialmente, foi aplicada a operação binária “E” para fazer a intersecção da segmentação final dos rins com a etapa de reconstrução dos tumores renais (Figura 25). Essa etapa inicial foi feita para extrair apenas a região que contém os tumores, removendo assim parte dos falsos positivos que foram adquiridos na etapa de reconstrução dos tumores renais. Posteriormente, uma etapa de pós-processamento foi aplicada para remover alguns elementos descontínuos que possivelmente eram falsos positivos que permaneceram.

Os tumores renais são estruturas contínuas que aparecem em várias fatias de um volume de TC (HEALTH, 2021; CARE, 2021). Analisando as informações contextuais vizinhas, o modelo DeepLabv3+ 2.5D resultou na segmentação da fatia central. Portanto, é fundamental verificar se os elementos segmentados pertencem à mesma região. Assim, a etapa de pós-processamento desenvolvida remove as estruturas segmentadas sem informações contextuais suficientes para representar os tumores renais, ou seja, elementos que apresentavam menos de três fatias contínuas.

A Figura 26 ilustra a etapa de pós-processamento aplicada. Na Figura 26 (a), tem-se o resultado em 3D da intersecção descrita acima e a Figura 26 (b) mostra a máscara do especialista. Por fim, a Figura 26 (c) apresenta o resultado visual após a remoção

Figura 25 – Intersecção da segmentação final dos rins e reconstrução dos tumores renais.



Fonte: Elaborado pela autora.

de elementos descontínuos. É possível verificar que alguns falsos positivos restantes semelhantes aos tumores foram removidos. Assim, a etapa de pós-processamento foi capaz de melhorar a segmentação dos tumores.

Figura 26 – Resultado final da segmentação de tumores renais após a aplicação do pós-processamento.



Fonte: Elaborado pela autora.

4.5 Considerações Finais

Este capítulo apresentou e descreveu em detalhes o método proposto para segmentação de rins e tumores em imagens de TC. Foram detalhadas cada uma das etapas que

compõem o método, e apresentadas as adaptações empregadas para a segmentação de rins e tumores nesta tese. Entre as principais contribuições estão: a distribuição proporcional da base de imagens; balanceamento de fatias do alvo de segmentação; segmentação de rins e candidatos a tumores renais na região abdominal usando o modelo ResUNet 2.5D; combinação do modelo DeepLabv3+ com o codificador DPN-131 para segmentação de candidatos a tumores renais na região renal; reconstrução dos tumores renais usando operações binárias; e técnicas de processamento de imagens para redução de falsos positivos com base em informações contextuais.

No próximo capítulo, serão apresentados os resultados obtidos em cada uma das etapas do método proposto. Além disso, são apresentadas a base de imagens aplicada, a configuração experimental das redes usadas e alguns experimentos para validar as etapas do método proposto.

5 RESULTADOS

Este capítulo apresenta os resultados experimentais de cada etapa do método proposto. Primeiramente, é apresentada a base de imagens que foi usada. Em seguida, é mostrada a avaliação dos resultados, com os seguintes procedimentos: 1) configuração experimental; 2) resultados gerais do método proposto; 3) resultados iniciais da segmentação de rins; 4) resultados iniciais de candidatos a tumores renais; 5) reconstrução dos tumores renais; e 6) resultados finais da segmentação de rins.

5.1 Base de Imagens

Neste trabalho, a base de imagens usada para realizar os experimentos foi de um desafio denominado KiTS19. Este desafio teve como objetivo acelerar a pesquisa e desenvolver novos recursos nefrométricos para auxiliar no prognóstico do tumor renal e no planejamento do tratamento; e também a criação de métodos confiáveis de segmentação semântica para rim e tumor renal (HELLER et al., 2019).

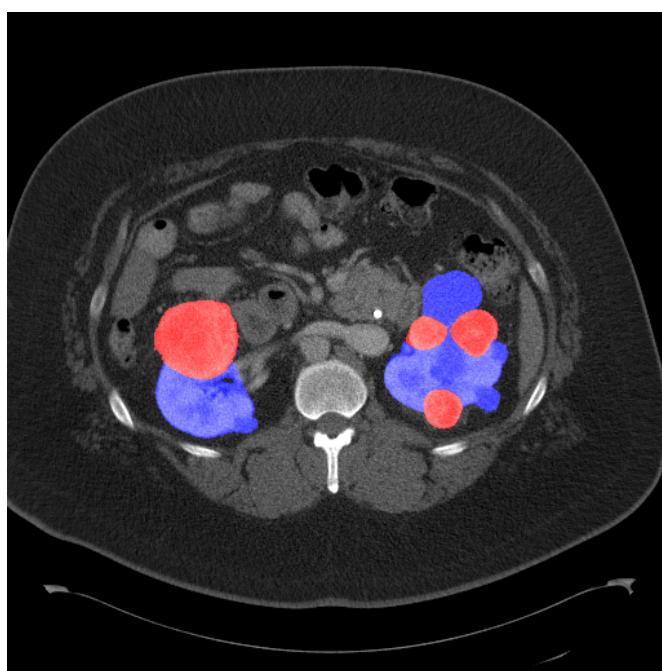
A base de imagens KiTS19 consiste em TCs de pacientes que foram submetidos a nefrectomia parcial ou radical para um ou mais tumores renais. Os exames foram coletados pelo Centro Médico da Universidade de Minnesota (*University of Minnesota Medical Center*) entre 2010 e 2018. De acordo com o desafio, a base de imagens contém 300 TCs, divididos em 210 TCs para treinamento e validação e os 90 restantes para avaliar o desempenho do modelo. Para o conjunto de treinamento e validação, o desafio disponibilizou máscaras de segmentação semântica (anotações) e foram esses conjuntos de imagens que foram usados para realizar os experimentos desta tese. Essas anotações foram fornecidas por estudantes de medicina que realizaram segmentações manuais sob a supervisão dos radiologistas e patologistas cirúrgicos (Drs. Christopher Weight e Niranjan Sathianathan). Isso ajudou na localização precisa dos tumores para evitar classificação incorreta como cistos que também podem estar presentes (HELLER et al., 2019).

Todos os pacientes foram tratados por urologistas em um único centro de atendimento e mais da metade dos exames foram adquiridos em mais de 50 instituições de referência. Para cada instituição, uma variedade de protocolos de aquisição é aplicada. Isso torna a base de imagens KiTS19 diversa em termos de dimensões de *voxel*, tempo de contraste, assinatura de tabela e campo de visão do scanner. Além disso, ajuda a

minimizar algumas preocupações sobre a validade externa de um coorte retrospectivo¹ de uma única instituição, como o KiTS19 (HELLER et al., 2021).

Cada volume de TC (imagens e máscaras) foi fornecido no formato NIFTI (INITIATIVE, 2000). Todos os volumes têm resolução 512×512 em escala de cinza, e o número de fatias varia de 29 a 1059. O método proposto tem como objetivo segmentar os rins e também os tumores renais entre os cistos presentes, o que é um desafio devido às semelhanças. A Figura 27 ilustra um exemplo de fatia de um volume, que contém as marcações: rins (azul) e tumores (vermelho).

Figura 27 – Informações da base de imagens: fatia com marcações de rins (azul) e tumores (vermelho).



Fonte: Elaborado pela autora.

5.1.1 Distribuição Proporcional da Base de Imagens

Esta seção apresenta o resultado da aplicação do método para distribuição proporcional da base de imagens. Para realizar os experimentos, foi usada a base de imagens KiTS19, na qual o conjunto de teste (31 TCs) foi inicialmente selecionado aleatoriamente entre os 210 TCs disponíveis. Os demais exames de TC foram usados como entrada para o método, seguindo as três etapas principais (Seção 4.1.1): (1) reamostrar os exames; (2) extrair *deep features*; e (3) agrupar.

¹ O pesquisador colhe informação pregressa dos fatores de exposição e acompanha por um período de tempo os indivíduos (CAMARGO; SILVA; MENEGUETTI, 2019).

É importante mencionar que antes de reamostrar os exames, a quantidade de fatias de tumores renais em cada exame variou de 3 a 256. A primeira etapa, foi realizada a técnica de reamostragem dos exames de TC. Na qual, obteve-se o valor médio de fatias de exames de TC por meio da operação média, que consistiu em calcular a quantidade total de fatias de tumores (4.927) contidos no conjunto de treino e validação, e dividir pela quantidade de exames totais (179). Como resultado, o valor médio foi de 27 fatias. Por fim, foi aplicada a técnica de reamostragem e todos os exames passaram a ter 27 fatias.

Na segunda etapa, são extraídas as *deep features* dos exames reamostrados. A Tabela 3 mostra o número de características extraídas dos exames para cada modelo descrito na Seção 4.1.1.

Tabela 3 – Total de características extraídas em cada modelo de CNN.

Modelo	Nº de características por fatia	Total (27 fatias)
ResNet-18	512	13.824
ResNet-34	512	13.824
ResNet-101	2.048	55.296
VGG-11	1.000	27.000
VGG-16	1.000	27.000
DPN-131	2.688	72.576
Xception	2.048	55.296

Na terceira etapa, as características extraídas por cada modelo são usadas para realizar o agrupamento de dados. A Tabela 4 apresenta os modelos, o melhor grupo encontrado entre os K-grupos (2 a 20) e o valor da inércia. Dentre eles, o modelo ResNet-34, que tem o valor seis representando o melhor grupo, foi o escolhido por apresentar a menor inércia (49.158).

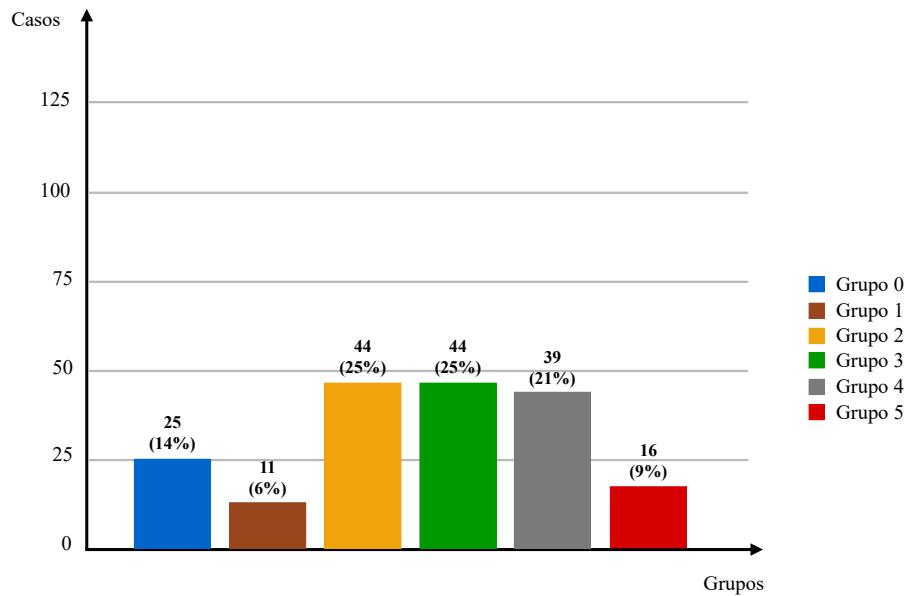
Tabela 4 – Melhores grupos e suas inéncias encontrados para os modelos usando o método do cotovelo.

Modelo	Nº ótimo de grupos	Inércia
ResNet-18	6	55.260
ResNet-34	6	49.158
ResNet-101	5	83.542
VGG-11	6	248.324.688
VGG-16	8	134.276.912
DPN-131	7	5.828.450.304
Xception	6	2.972.436

A Figura 28 apresenta o gráfico com o número de exames para cada um dos grupos encontrados. Finalmente, os casos (exames) de cada um dos seis grupos gerados

pelo método proposto foram selecionados aleatoriamente e distribuídos proporcionalmente entre os conjuntos de dados de treinamento (148 exames) e validação (31 exames).

Figura 28 – Grupos de tumores encontrados com a aplicação do método a partir dos conjuntos de treinamento e validação.



Fonte: Elaborado pela autora.

5.1.2 Preparação da Base de Imagens

Inicialmente, os volumes de TC foram pré-processados (Seção 4.1) antes de serem usados como entrada para a ResUNet 2.5D (Seções 4.2.1 e 4.2.2.2) e DeepLabv3+ 2.5D (Seção 4.2.2.1). Para fins de teste e validação, foram separados aproximadamente 148 (70%) volumes de TC para treinamento, 31 (15%) para validação e 31 (15%) para teste do conjunto de 210 TCs disponíveis. Inicialmente, o conjunto de teste foi selecionado aleatoriamente (31 casos), e com os casos restantes (179 casos), uma distribuição proporcional foi feita para separar os casos entre os conjuntos de treinamento e validação, conforme descrito na Seção 5.1.1. Além disso, foi realizado um balanceamento de fatias para cada modelo de aprendizado profundo (Seções 4.2.1.1, 4.2.2.1.1 e 4.2.2.2).

5.2 Configuração Experimental

O método proposto foi implementado usando a linguagem de programação Python e principalmente a biblioteca de aprendizado profundo Pytorch (PASZKE et al., 2019). O computador usado nos experimentos consistia em uma CPU Intel Core i7-7700K 4.20GH,

16 GB de RAM e placa de vídeo Nvidia GeForce GTX 1080-Ti, executado no sistema operacional Windows 10.

O critério de parada nos modelos de treinamento foi usando a técnica early stop (BROWNLEE, 2018), que consiste em parar o treinamento da rede quando a métrica monitorada (o valor da função de perda nos dados de validação) parasse de melhorar em 15 épocas. Os hiperparâmetros usados nos modelos tinham um tamanho de *batch* igual a 8, função de perda Dice (DICE, 1945), otimizador estocástico Adam (KINGMA; BA, 2014) com uma taxa de aprendizagem inicial de 0,0001 com decaimento de 10% se o valor da função de perda nos dados de validação não cair em 7 épocas, beta-1 0,9 e beta-2 0,999 (taxas de decaimento iniciais usadas ao estimar o primeiro e o segundo momentos do gradiente). Após a aplicar o algoritmo de otimização *Tree of Parzen Estimators* (TPE) usando a biblioteca Hyperopt (BERGSTRA; YAMINS; COX, 2013), este foi o conjunto de hiperparâmetros que produziu os melhores resultados. O tempo de execução total de otimização foi de aproximadamente 4 semanas.

As métricas (Seção 3.10) usadas para avaliar o desempenho do método proposto foram: coeficiente de similaridade Dice (Dice), índice de Jaccard (Jacc), acurácia (Acc), sensibilidade (Sen) e especificidade (Esp). Além disso, foram realizados testes de significância² das etapas do método.

5.3 Resultados Gerais do Método Proposto

Nesta seção, são apresentados os resultados obtidos em cada etapa do método proposto. Para que o método fosse validado, os conjuntos de treino, validação e teste foram distribuídos 5 vezes usando os mesmos procedimentos descritos na Seção 5.1.2. Posteriormente, em cada etapa do método, a CNN foi executada 5 vezes (*5-folds*) usando os conjuntos de dados mencionados. Essa prática tem a finalidade de verificar a eficiência do método independentemente do conjunto de dados avaliado, garantindo que o método não seja tendencioso (BURMAN, 1989).

A Tabela 5 apresenta os resultados dos *5-folds* para cada uma das etapas do método proposto, seguido da média e desvio padrão. Por fim, é mostrado o tempo de treinamento para cada época das abordagens. Analisando os resultados individuais, pode-

² Teste de significância é um procedimento estatístico que permite tomar uma decisão entre duas ou mais hipóteses (rejeitar ou não rejeitar), usando os dados observados de um determinado experimento (COX, 1982; ZOU et al., 2003).

se perceber que a segmentação dos rins apresentou bons resultados iniciais, atingindo uma média de Dice igual a 96,37% com desvio padrão de 1,20. Nos candidatos a tumores renais, obtiveram-se resultados médios de 78,49% de Dice e desvio de 2,06, levando em consideração que os candidatos a tumores renais foram segmentados dentro da região renal obtidas pela segmentação inicial dos rins.

Tabela 5 – Resultados iniciais das etapas do método proposto.

Segmentação Inicial	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)
Rins	96,77	93,80	99,93	97,58	99,96
	96,54	93,46	99,90	96,08	99,97
	94,40	91,10	99,90	93,31	99,97
	97,65	95,43	99,96	97,55	99,98
	96,52	93,74	99,94	95,82	99,98
	96,37 ± 1,20	93,51 ± 1,55	99,93 ± 0,03	96,07 ± 1,74	99,97 ± 0,01
Candidatos a Tumores Renais na Região Renal	81,02	69,66	99,93	80,52	99,97
	80,10	68,65	99,90	81,66	99,97
	77,72	67,36	99,89	82,02	99,96
	77,68	67,55	99,95	78,10	99,97
	75,91	63,91	99,95	74,98	99,98
	78,49 ± 2,06	67,43 ± 2,17	99,93 ± 0,03	79,46 ± 2,94	99,97 ± 0,01
Candidatos a Tumores Renais na Região Abdominal	68,88	58,45	99,91	85,57	99,93
	60,95	47,11	99,82	75,54	99,91
	58,54	47,45	99,82	84,01	99,85
	57,40	45,57	99,91	76,34	99,93
	59,55	47,41	99,86	75,86	99,89
	61,06 ± 4,56	49,20 ± 5,23	99,86 ± 0,05	79,46 ± 4,90	99,90 ± 0,04
Reconstrução de Tumores Renais	68,20	57,17	99,90	90,33	99,91
	63,77	49,82	99,86	86,80	99,90
	57,99	45,93	99,80	89,82	99,82
	60,10	47,79	99,91	85,95	99,92
	61,90	49,32	99,87	86,00	99,88
	62,39 ± 3,89	50,01 ± 4,28	99,87 ± 0,04	87,78 ± 2,13	99,88 ± 0,04

Na segmentação inicial de candidatos a tumores renais na região abdominal, foram obtidos resultados médios de 61,06% de Dice com desvio padrão de 4,56. Porém, nesta etapa, a informação mais relevante é obtida pela métrica de sensibilidade, pois informa a quantidade percentual de tumores segmentados. Portanto, para a métrica sensibilidade, obtiveram-se resultados médios de 79,46% com desvio padrão de 4,90. Por fim, a etapa de reconstrução de tumores renais apresentou resultados médios de 87,78% de sensibilidade com desvio padrão de 2,13. Este resultado permite observar a importância desta etapa para a reconstrução de tumores renais, uma vez que porções consideráveis de tumores foram obtidas unindo os estágios a candidatos a tumores renais, adquirindo assim resultados de sensibilidade superiores aos estágios de candidatos a tumores renais.

Na Tabela 6 os resultados finais da segmentação de rins e tumores são mostrados. Pode-se observar que na segmentação final dos rins foram obtidos resultados superiores

em todas as métricas em relação aos resultados iniciais dos rins. Na segmentação dos tumores também foram obtidos resultados superiores aos resultados iniciais dos candidatos a tumores renais. Para ambas as segmentações, a obtenção dos resultados superiores em relação aos resultados iniciais se deu devido a técnica de reconstrução robusta que acabou delimitando partes de regiões tumorais que também fazem partes de regiões renais. Além disso, as técnicas de pós-processamento potencializaram as segmentações por meio de seu refinamento.

Tabela 6 – Resultados finais do método proposto.

Segmentação Final	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)
Rins	97,45	95,05	99,95	98,44	99,96
	97,47	95,11	99,93	97,27	99,98
	95,45	92,69	99,93	94,83	99,97
	97,80	95,70	100,00	97,90	99,98
	96,81	94,29	99,95	96,24	99,98
	97,00 ± 0,94	94,57 ± 1,16	99,95 ± 0,03	96,94 ± 1,43	99,97 ± 0,01
Tumores Renais	84,06	75,04	99,94	88,33	99,95
	81,70	70,47	99,92	86,80	99,96
	81,64	70,76	99,90	89,58	99,92
	81,95	71,29	99,95	85,71	99,97
	82,61	71,70	99,96	85,20	99,98
	82,39 ± 1,01	71,85 ± 1,84	99,94 ± 0,02	87,12 ± 1,82	99,96 ± 0,02

Por fim, é importante mencionar que os resultados médios e seus desvios padrões comprovam que os resultados individuais fornecem fortes evidências que o método proposto é robusto, independentemente do conjunto de dados usado. Portanto, essa abordagem usando *k-fold* permitiu fazer uma análise mais detalhada do método proposto em geral. No entanto, deve-se sempre ter muita cautela ao aplicar essa abordagem. Isso porque, dependendo do conjunto de dados, o custo computacional pode ser muito grande, pois o mesmo modelo é treinado em várias subdivisões.

Portanto, nas próximas seções, os resultados de cada etapa são apresentados com mais detalhes, além da realização de outros experimentos e análises usando outras abordagens. Para isso, foi selecionado o melhor *fold* do método proposto considerando aquele que obteve o melhor resultado para os tumores, visto que a segmentação tumoral é mais delicada e propensa a ruídos. Portanto, o melhor *fold* é o primeiro, e este foi usado como referência para descrever os resultados individuais, estudos de caso e comparação com trabalhos relacionados.

5.4 Segmentação Inicial dos Rins

Conforme descrito na Seção 4.2.1, o modelo ResUNet 2.5D foi usado para a segmentação inicial dos rins em imagens de TC. Para construir o modelo, foram usadas 23.670 fatias do conjunto de treino, 5.062 fatias do conjunto de validação e 5.506 fatias do conjunto de teste.

Para verificar a eficácia da abordagem 2.5D escolhida e a capacidade de uma melhor generalização, as abordagens 2D e 2.5D foram comparadas. Basicamente, o modelo ResUNet 2.5D foi substituído pelo modelo ResUNet 2D, usando as mesmas configurações descritas na Seção 5.2 e o balanceamento de fatias (Seção 4.2.1.1). Em outras palavras, foram mostrados os resultados do método completo usando a abordagem 2D e a 2.5D em cada etapa do método. Os resultados iniciais da segmentação de rins usando essas abordagens são mostrados na Tabela 7.

Tabela 7 – Resultados da etapa de segmentação inicial dos rins.

Modelo	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)	T. por época
ResUNet 2D	92,67	86,93	99,85	93,79	99,91	32m 47s
ResUNet 2.5D	96,77	93,80	99,93	97,58	99,96	33m 29s
<i>P-value</i>	0,0000	0,0000	0,2054	0,0000	0,3033	-

A Tabela 7 apresenta as porcentagens de validação, teste de significância e o tempo de treinamento para cada época das abordagens. O modelo ResUNet 2D obteve 92,67% de Dice, 86,93% de Jaccard, 99,85% de acurácia, 93,79% de sensibilidade e 99,91% de especificidade e o tempo de execução por época aproximadamente 32m e 47s. No entanto, o modelo ResUNet 2.5D produziu resultados mais significativos, alcançando 96,77% de Dice, 93,80% de Jaccard, 99,93% de acurácia, 97,58% de sensibilidade e 99,96% de especificidade e o tempo de execução por época foi de aproximadamente 33m e 29s.

Em termos estatísticos, as abordagens 2D e 2.5D, apresentaram um *p-value* igual a 0 para as métricas Dice, Jaccard e sensibilidade (destacado em cinza). Portanto, a hipótese nula³ é rejeitada, demonstrando que existe diferença significativa entre as abordagens 2D e 2.5D. É importante mencionar que em imagens médicas, as métricas descritas são de grande importância para verificar a eficiência do método de segmentação em termos de similaridade volumétrica (TAHA; HANBURY, 2015).

³ Adotando um nível de significância de $\alpha = 0,05$, se $p\text{-value} > 0,05$ não rejeita-se a hipótese nula, o que indica que não há diferença significativa entre as abordagens analisadas.

Ressalta-se que, na abordagem 2.5D, são levadas em consideração não apenas as informações locais de uma fatia, mas também as informações espaciais considerando as fatias anterior e posterior, que consequentemente produzem resultados mais expressivos. Assim, acredita-se que a opção pela abordagem 2.5D fornece informações cruciais para a eficácia do modelo proposto para segmentar os rins, conseguindo atingir métricas robustas, destacando-se o Dice de 96,77%.

5.5 Segmentação Inicial de Candidatos a Tumores Renais

Nesta seção, são apresentados os resultados iniciais da segmentação de candidatos a tumores renais. Nas subseções 5.5.1 e 5.5.2 os resultados iniciais dos dois estágios (candidatos a tumores renais na região renal e abdominal) são descritos.

5.5.1 Candidatos a Tumores Renais na Região Renal

Nesta subseção, são apresentados os resultados alcançados do primeiro estágio de segmentação de candidatos a tumores renais. De acordo com a Seção 4.2.2.1.1, para o treinamento e validação do modelo DeepLabv3+ (Seção 4.2.2.1) foram usadas 7.870 e 1.944 fatias, respectivamente. Para testar o modelo, foram aplicadas as 1.990 fatias que contém rins.

Vale ressaltar que, assim como na segmentação inicial dos rins, nesta etapa também foi verificada a eficácia do modelo DeepLabv3+ usando as abordagens 2D e 2.5D para a segmentação inicial dos candidatos a tumores renais. Para isso, as diferentes abordagens tiveram as mesmas configurações descritas na Seção 5.2 e o balanceamento de fatias (Seção 4.2.2.1.1).

Além disso, foram aplicadas seis operações de *data augmentation* em tempo real (de execução), duas das quais foram baseadas em probabilidades. Em seguida, combinações aleatórias foram feitas nas operações. Foram elas: inversão horizontal, escalas entre 50% e 120% do tamanho total da imagem (fatia), rotações entre -15° e +15° e transformação elástica (cisalhamento) (MIKOŁAJCZYK; GROCHOWSKI, 2018). Além disso, operações de probabilidade foram aplicadas às imagens, nas quais cada imagem tinha 10% de probabilidade de aplicação do filtro *gaussian blur* e/ou contraste linear (GONZALEZ; WOODS, 2008). Os parâmetros escolhidos modificam ligeiramente a textura da imagem, gerando novas representações dos rins e tumores encontrados em cenários reais, tornando-os

úteis em tarefa de segmentação de tumores. As operações de *data augmentation* reduziram o *overfitting* e, consequentemente, tornaram o modelo mais generalista, uma vez que os pesos estavam mais adequados à realidade do problema devido à diversidade gerada no conjunto de treinamento.

Os resultados obtidos na segmentação inicial dos rins (Seção 5.4) são usados como entrada para segmentar os candidatos a tumores renais na fase de teste. Os resultados estão descritos na Tabela 8. O modelo DeepLabv3+ 2D atingiu um Dice de 72,75%, Jaccard de 61,60%, acurácia de 99,48%, sensibilidade 78,85% e especificidade de 99,59% e o tempo de execução por época de 11m e 34s. No entanto, o modelo DeepLabv3+ 2.5D forneceu resultados superiores de Dice e da maioria das outras métricas, alcançando 81,02%, 69,66%, 99,93%, 80,52% e 99,97% de Dice, Jaccard, acurácia, sensibilidade e especificidade, respectivamente, e tempo de execução por época de aproximadamente 13m e 11s. Além disso, em termos estatísticos, os modelos DeepLabv3+ 2D e 2.5D obtiveram um *p-value* abaixo de 0,0103 para as métricas Dice, Jaccard, acurácia e especificidade (destacado em cinza). Portanto, a hipótese nula é rejeitada, demonstrando que as abordagens são significativamente diferentes.

Tabela 8 – Resultados da etapa de segmentação de candidatos a tumores renais na região renal.

Modelo	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)	T. por época
DeepLabv3+ 2D	72,75	61,60	99,48	78,85	99,59	11m 34s
DeepLabv3+ 2.5D	81,02	69,66	99,93	80,52	99,97	13m 11s
<i>P-value</i>	0,0000	0,0000	0,0092	0,1908	0,0103	-

Pode-se observar que mais uma vez a escolha da abordagem 2.5D é fundamental para a obtenção de resultados expressivos. Além disso, vale ressaltar que embora a abordagem 2.5D leve em consideração mais fatias, e consequentemente a análise de mais informações, o tempo de execução das abordagens 2D e 2.5D foi próximo. Portanto, a DeepLabv3+ 2.5D foi capaz de obter bons resultados com baixo custo computacional.

5.5.2 Candidatos a Tumores Renais na Região Abdominal

Neste segundo estágio, são apresentados os resultados iniciais para segmentação de candidatos a tumores renais na região abdominal. De acordo com a Seção 4.2.2.2, um segundo modelo ResUNet foi usado para segmentar regiões de tumores renais. Foram aplicadas as mesmas configurações descritas na Seção 5.2 e o balanceamento de fatias (Seção 4.2.2.1.1). Na construção do modelo, foram usadas 7.870 fatias do conjunto de treino

e 1.944 fatias do conjunto de validação, aplicando-se também o mesmo *data augmentation* usado no primeiro estágio. Na fase de teste, foram aplicadas as 5.506 fatias do conjunto de teste.

As diferentes abordagens (2D e 2.5D) na arquitetura ResUNet também foram analisadas a fim de verificar a eficácia na segmentação de candidatos a tumores renais na região abdominal. A Tabela 9 mostra os resultados obtidos no segundo estágio. No modelo ResUNet 2D, os resultados obtidos para o Dice, Jaccard, acurácia, sensibilidade e especificidade foram 50,01%, 38,22%, 99,79%, 76,30% e 99,82%, respectivamente, e aproximadamente 10m e 43s de tempo de execução. Para o modelo ResUNet 2.5D foram alcançados 68,88% de Dice, 58,45% de Jaccard, 99,91% de acurácia, 85,57% de sensibilidade e 99,93% de especificidade, e o tempo de execução total foi de aproximadamente 11m e 22s.

Tabela 9 – Resultados da etapa de segmentação de candidatos a tumores renais na região abdominal.

Modelo	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)	T. por época
ResUNet 2D	50,01	38,22	99,79	76,30	99,82	10m 43s
ResUNet 2.5D	68,88	58,45	99,91	85,57	99,93	11m 22s
<i>P-value</i>	0,0000	0,0000	0,1038	0,0000	0,1024	-

Em relação às métricas, observa-se que os resultados não foram tão elevados. Porém, é importante enfatizar que, no segundo estágio, o objetivo é segmentar os candidatos a tumores renais em regiões mais ampla, analisando mais informações contextuais, embora também segmente falsos positivos. Portanto, para este estágio, a métrica mais importante é a sensibilidade, pois representa o percentual de segmentação na classe positiva, ou seja, a quantidade de tumores segmentados. Assim, analisando a Tabela 9, pode-se concluir que para a segmentação inicial a candidatos a tumores renais na região abdominal, a abordagem 2.5D obteve melhor desempenho, alcançando resultado superior a 12,57% de sensibilidade em relação à abordagem 2D. Estatisticamente, os modelos 2D e 2.5D são significativamente diferentes, pois obtiveram um *p-value* próximo a 0 para as métricas Dice, Jaccard e sensibilidade. Portanto, foi rejeitada a hipótese nula.

5.6 Reconstrução dos Tumores Renais

Nesta etapa, são apresentados os resultados obtidos na reconstrução dos tumores renais. Para isso, foi aplicada a união da segmentação inicial dos candidatos a tumores

renais na região renal e abdominal nas abordagens 2D e 2.5D. Assim, é possível verificar a eficácia de ambas abordagens na etapa de reconstrução dos tumores renais.

A Tabela 10 apresenta os resultados da reconstrução dos tumores renais aplicando as abordagens 2D e 2.5D dos candidatos a tumores renais. Vale ressaltar que, assim como o segundo estágio da etapa de segmentação inicial de candidatos a tumores renais, a métrica mais importante para o objetivo da reconstrução é a sensibilidade, pois mostra com precisão o percentual de tumores reconstruídos após a união dos dois estágios dos candidatos a tumores renais. Portanto, analisando a métrica de sensibilidade, a abordagem 2D alcançou 83,40% e a abordagem 2.5D 90,33%. Isso comprova que a abordagem 2.5D foi capaz de ajustar com mais eficiência os valores da classe positiva, recuperando partes consideráveis das regiões tumorais por meio da união entre os estágios. Em termos estatísticos, as abordagens 2D e 2.5D obtiveram um *p-value* igual a 0 para as métricas Dice, Jaccard e sensibilidade, portanto, a hipótese nula foi rejeitada, comprovando que as abordagens são significativamente diferentes.

Tabela 10 – Resultados da etapa de reconstrução dos tumores renais.

Abordagem	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)
2D	49,94	38,16	99,77	83,40	99,79
2.5D	68,20	57,17	99,90	90,33	99,91
<i>P-value</i>	0,0000	0,0000	0,0928	0,0000	0,1038

Além disso, vale ressaltar que a principal funcionalidade da etapa de reconstrução dos tumores renais é unir os estágios de candidatos a tumores renais a fim de obter uma região segmentada maior de tumores renais. Observando a sensibilidade, percebe-se que isso foi alcançado, no entanto, a métrica Dice apresenta resultados inferiores em relação às segmentações iniciais dos candidatos a tumores renais. Isso significa que mais falsos positivos foram adicionados. Como já mencionado, esta etapa é suscetível a falsos positivos, mas na última etapa do método proposto (redução de falsos positivos), essas regiões extras adicionadas são removidas por meio do pós-processamento.

5.6.1 Redução de Falsos Positivos para os Rins

Nesta etapa, o objetivo é delimitar as regiões dos rins segmentadas pelas etapas anteriores. Inicialmente, é feita uma união da segmentação inicial dos rins com o resultado da etapa de reconstrução dos tumores renais para melhorar os contornos dos rins. Em seguida, uma etapa de pós-processamento é necessária para remover as regiões (falsos posi-

tivos) que não pertencem aos rins. Conforme descrito na Seção 4.4.1, o pós-processamento projetado é baseado em componentes conectados que mantêm os dois maiores elementos (rins) segmentados de cada volume de TC.

A Tabela 11 apresenta os resultados obtidos após a etapa de pós-processamento, concluindo o método proposto para segmentação de rins. O desempenho final usando a abordagem 2.5D atingiu 97,45% de Dice, 95,05% de Jaccard, 99,95% de acurácia, 98,44% de sensibilidade e 99,96% de especificidade, apresentando resultados mais significativos do que o desempenho final da abordagem 2D. Estatisticamente, as abordagens obtiveram um *p-value* igual a 0 nas métricas Dice, Jaccard e sensibilidade, rejeitando a hipótese nula.

É importante ressaltar que as métricas Dice e Jaccard foram consideravelmente maiores que na segmentação inicial dos rins (Seção 5.4), o que implica em segmentações aprimoradas, pois essas métricas indicam semelhança e sobreposição entre os objetos. Além disso, houve melhora na sensibilidade, demonstrando um impacto positivo na etapa de reconstrução de tumores para obtenção de mais regiões renais. Portanto, esta etapa final oferece melhorias para a segmentação dos rins, tornando-a mais precisa.

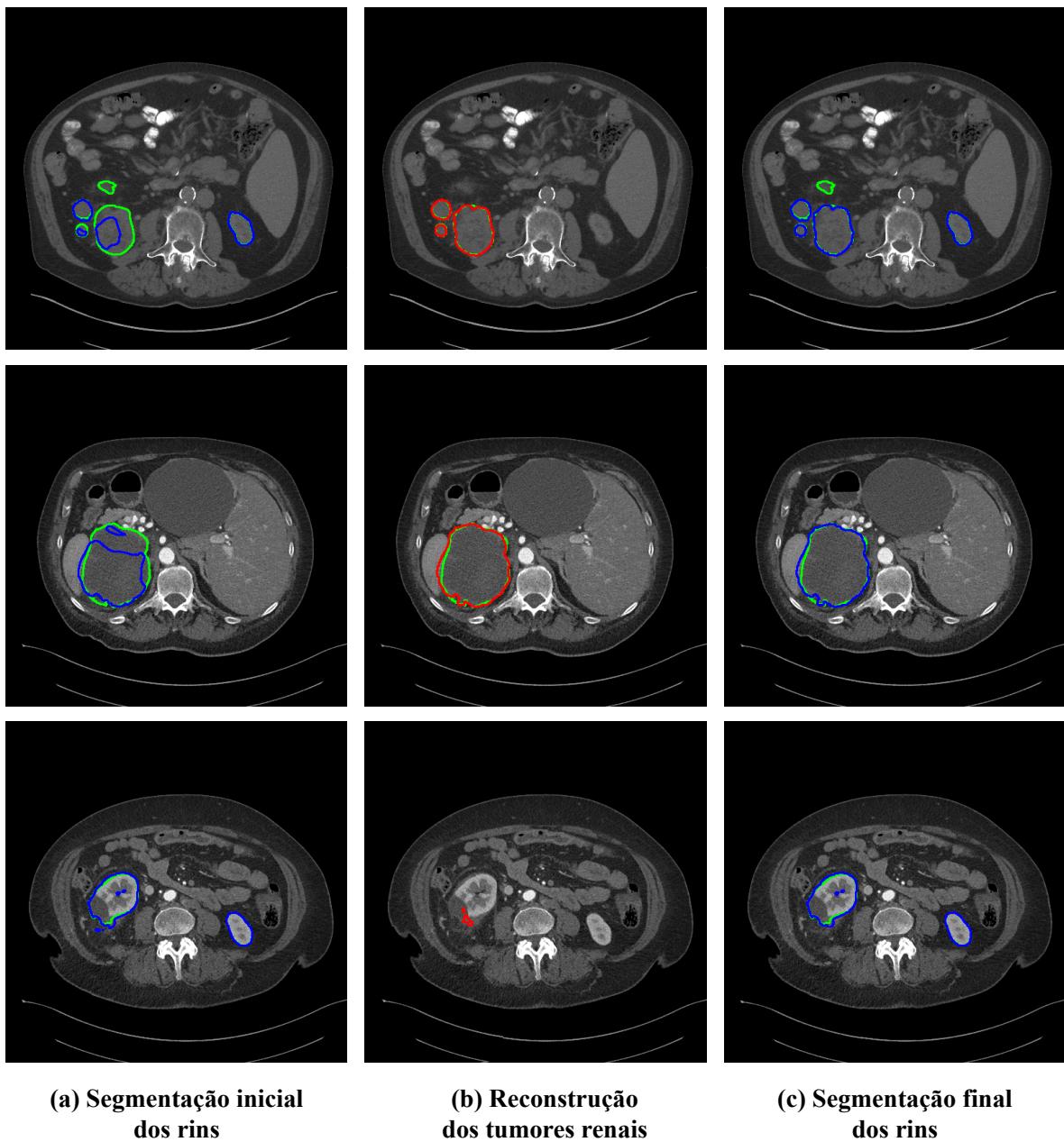
Tabela 11 – Resultados da etapa de segmentação final dos rins.

Resultado Final	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)
2D	93,54	88,86	99,87	94,02	99,92
2.5D	97,45	95,05	99,95	98,44	99,96
<i>P-value</i>	0,0000	0,0000	0,1616	0,0000	0,3914

A Figura 29 ilustra três casos aplicando as etapas para obter o resultado final da segmentação de rins. A marcação do especialista está em verde e as segmentações finais dos rins e tumores em azul e vermelho, respectivamente. É possível observar que na segmentação inicial (Figura 29 (a)) alguns tecidos de rins (rins + tumores renais) não foram segmentados. Com a aplicação da etapa de reconstrução de tumores renais (Figura 29 (b)) mais algumas regiões dos tumores foram segmentadas. Isso significa que regiões dos rins também foram recuperadas, já que regiões tumorais também fazem parte dos rins. Por fim, a Figura 29 (c) apresenta o resultado final da segmentação dos rins, na qual foram incluídos mais tecidos renais devido a etapa de reconstrução tumoral.

Vale destacar o último caso (imagens inferiores), em que alguns falsos positivos segmentados foram observados na etapa inicial de segmentação renal (Figura 29 (a) inferior). No entanto, foram removidos após a aplicação do pós-processamento (manter os

Figura 29 – Etapas da segmentação dos rins. Marcação do especialista (verde) e segmentações do método (azul - rins, vermelho - tumores renais).



Fonte: Elaborado pela autora.

dois maiores elementos). Portanto, acredita-se que este conjunto de etapas seja capaz de contribuir para melhorar o resultado da segmentação de rins.

5.6.2 Redução de Falsos Positivos para os Tumores Renais

Esta etapa completa o método proposto para segmentação de tumores renais. Para isso, são usadas as segmentações das etapas anteriores obtidas, conforme descrito na Seção 5.6.1. Com as segmentações obtidas, foi feita a intersecção da segmentação

final dos rins com a etapa de reconstrução dos tumores renais para extrair apenas a região que contém os tumores, removendo assim parte dos falsos positivos adquiridos na reconstrução dos tumores. Finalmente, os resultados dessa etapa são passados pela etapa de pós-processamento para reduzir os falsos positivos que não foram removidos na etapa anterior.

O pós-processamento consistiu na remoção de estruturas segmentadas sem informação contextual suficiente para representar as regiões dos tumores renais, ou seja, elementos com menos de três fatias contínuas (Seção 4.4.2). Os resultados após o pós-processamento estão descritos na Tabela 12. A segmentação final atingiu 84,06% de Dice, 75,04% de Jaccard, 99,94% de acurácia, 88,33% de sensibilidade e 99,95% de especificidade.

Tabela 12 – Resultados da etapa de segmentação final dos tumores renais.

Resultado Final	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)
2D	73,62	61,21	99,89	81,03	99,91
2.5D	84,06	75,04	99,94	88,33	99,95
<i>P-value</i>	0,0000	0,0000	0,3680	0,0000	0,4275

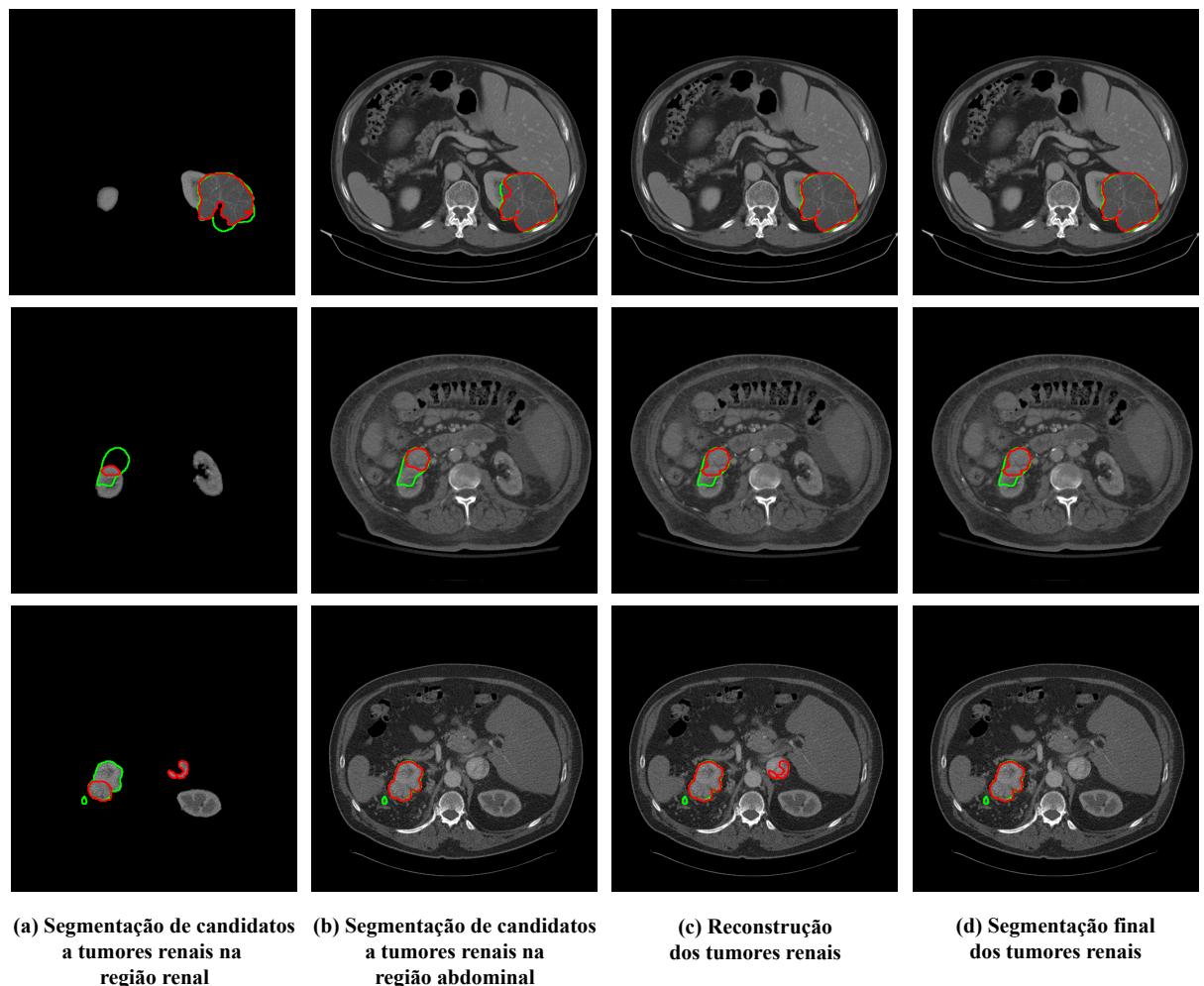
Em termos estatísticos, a segmentação final dos tumores renais usando as abordagens 2D e 2.5D apresentou um *p-value* igual a 0 para as métricas Dice, Jaccard e sensibilidade. Portanto, comprova que existe uma diferença significativa entre as segmentações finais, visto que a hipótese nula foi rejeitada. Isso demonstra que as segmentações finais usando a abordagem 2.5D são mais precisas, com menos falsos positivos e mais regiões de tumores. Uma vez que, as métricas descritas (Dice, Jaccard e sensibilidade) indicam a capacidade de identificar regiões de tumores e a semelhança entre os objetos em estudo.

Além disso, pode-se observar que houve melhorias consideráveis nas métricas Dice, Jaccard e sensibilidade em relação às segmentações iniciais dos candidatos a tumores renais (Seção 5.5.1 e 5.5.2). Isso comprova a importância das etapas anteriores para a obtenção de resultados expressivos. Portanto, devido à melhora significativa nos resultados finais, o trabalho dos especialistas para analisar todos os elementos segmentados é reduzido.

Três casos aplicando as etapas para obter a segmentação final dos tumores renais são apresentados na Figura 30. Primeiramente, pode-se observar que a segmentação de candidatos a tumores renais na região renal (Figura 30 (a)) não foi capaz de segmentar com precisão as regiões dos tumores. Além disso, muitas regiões renais não foram segmentadas na etapa de segmentação inicial dos rins, o que acabou comprometendo o desempenho

do modelo, uma vez que ficou impossibilitado de segmentar mais regiões tumorais. No entanto, a etapa de segmentação de candidatos a tumores renais na região abdominal (Figura 30 (b)) compensa a deficiência da segmentação inicial dos rins e candidatos a tumores renais na região renal, permitindo a segmentação de mais regiões tumorais. Isso é possível porque esta etapa analisa mais informações contextuais em uma região mais ampla (abdominal).

Figura 30 – Etapas da segmentação dos tumores renais. Marcação do especialista (verde) e segmentação do método (vermelho).



Fonte: Elaborado pela autora.

Posteriormente, a etapa de reconstrução de tumores renais (Figura 30 (c)) combiou as etapas de segmentação de candidatos a tumores renais na região renal e abdominal. Isso levou a uma segmentação mais precisa da região tumoral, pois regiões tumorais não segmentadas no primeiro estágio foram segmentadas no segundo estágio, agregando os resultados. Finalmente, a segmentação final dos tumores é mostrada na Figura 30 (d).

Pode-se analisar que mais tecidos das regiões tumorais foram incluídos devido à etapa de reconstrução do tumor.

Por fim, vale destacar o último caso (imagens inferiores) da Figura 30. No qual pode-se observar que na etapa de segmentação dos candidatos a tumores renais na região renal (Figura 30 (a) inferior), as regiões tumorais foram segmentadas, mas também falsos positivos. Entretanto, foram eliminados após a aplicação do pós-processamento (remoção de elementos descontínuos). Portanto, acredita-se que todas as etapas foram importantes para melhorar a segmentação dos tumores.

5.6.3 Método Proposto com/sem Balanceamento de Fatias

Após os experimentos usando abordagens 2D e 2.5D, foram realizados mais alguns experimentos para validar as demais etapas do método proposto. Primeiramente, foi verificada a eficácia da técnica de balanceamento de fatias no conjunto de treinamento e validação. Para isso, foi realizado um experimento com e sem balanceamento de fatias no método proposto, no qual, ao invés de usar a estratégia de balanceamento de fatias, os modelos foram treinados e validados usando todas as fatias dos volumes de TCs. Posteriormente, todas as outras etapas do método proposto foram aplicadas com base neste experimento. Os resultados finais do método proposto com e sem balanceamento proporcional são apresentados na Tabela 13. Além disso, é mostrado o tempo de treinamento para cada época dos modelos (ResUNet 2.5D e DeepLabv3+ 2.5D). Na segmentação dos tumores renais é apresentado o resultado da soma dos tempos de treinamento dos modelos para segmentar candidatos a tumores renais na região renal e abdominal.

Tabela 13 – Resultado do método proposto com/sem balanceamento de fatias.

Segmentação	Técnica	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)	T. por época
Reconstrução dos Tumores Renais	Todas as fatias	43,79	32,11	99,63	77,00	99,65	-
	Balanceado	68,20	57,17	99,90	90,33	99,91	-
Rins	Todas as fatias	92,96	87,38	99,84	93,05	99,92	1h 01m 52s
	Balanceado	97,45	95,05	99,95	98,44	99,96	33m 29s
Tumores Renais	Apenas fatias com rins	62,03	49,15	99,82	75,99	99,85	46m 03s
	Balanceado	84,06	75,04	99,94	88,33	99,95	24m 33s
<i>P-value</i>		0,0000	0,0000	0,0690	0,0000	0,0969	-

Pode-se observar que a abordagem proposta com balanceamento de fatias produz

melhorias significativas em todas as colunas em comparação com a abordagem sem balanceamento de fatias. Inclusive o tempo com balanceamento é inferior, pois tem menos fatias que a abordagem sem balanceamento. Analisando a métrica sensibilidade entre as abordagens, pode-se verificar que sofreu um impacto negativo em relação à abordagem com balanceamento de fatias. Isso acontece porque, sem aplicar o balanceamento de fatias, os modelos terão como entrada mais fatias de fundo (sem rim, sem tumores, dependendo do modelo) do que fatias com o objeto de interesse (rins ou tumores). Dessa forma, o modelo tende a aprender mais fatias de fundo e não consegue segmentar com precisão o objeto de interesse, portanto, o modelo tem baixa sensibilidade e, consequentemente, resultados baixos para Dice e Jaccard.

Além disso, estatisticamente, obteve-se um *p-value* de 0,0034 na acurácia em relação à reconstrução de tumores e um *p-value* igual a 0 para as métricas Dice, Jaccard e sensibilidade em relação a todas as abordagens com/sem balanceamento de fatias e, portanto, a hipótese nula foi rejeitada. Portanto, com os resultados experimentais foi possível comprovar que a técnica de balanceamento de fatias é fundamental para a obtenção de resultados mais consistentes na segmentação dos rins e segmentação e reconstrução de tumores renais. Essa melhora se deve à capacidade do balanceamento detectar fatias com e sem tumores, possibilitando segmentar mais regiões de interesse e minimizar falsos positivos.

5.6.4 Método Proposto com Distribuição Aleatória e Proporcional da Base de Imagens

O próximo experimento realizado foi relacionado a etapa da distribuição proporcional da base de imagens. Para validar essa etapa, o conjunto de teste permaneceu o mesmo e os modelos de segmentação e reconstrução foram treinados com outro conjunto de treino e validação distribuídos aleatoriamente. As demais etapas do método proposto foram aplicadas posteriormente. Os resultados finais do método proposto com distribuição aleatória e proporcional são apresentados na Tabela 14. Também são apresentados o tempo de treinamento para cada época dos modelos (ResUNet 2.5D e DeepLabv3+ 2.5D). Na segmentação dos tumores renais, é apresentado o resultado da soma dos tempos de treinamento dos modelos para segmentar candidatos a tumores renais na região renal e abdominal.

Tabela 14 – Resultado do método proposto com distribuição aleatória e proporcional.

Segmentação	Técnica	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)	T. por época
Reconstrução dos Tumores Renais	Aleatório	53,39	40,78	99,78	82,22	99,81	-
	Distribuição proporcional	68,20	57,17	99,90	90,33	99,91	-
Rins	Aleatório	94,18	90,06	99,90	94,44	99,95	34m 33s
	Distribuição proporcional	97,45	95,05	99,95	98,44	99,96	33m 29s
Tumores Renais	Aleatório	71,29	59,17	99,89	80,64	99,91	27m 14s
	Distribuição proporcional	84,06	75,04	99,94	88,33	99,95	24m 33s
<i>P-value</i>		0,0000	0,0000	0,1152	0,0000	0,1605	-

A partir dos resultados obtidos, é possível ver a importância da etapa de distribuição proporcional, não só por ter obtido resultados melhores do que com a aplicação da distribuição aleatória, mas também por atingir um *p-value* igual a 0 para várias métricas (Dice, Jaccard e sensibilidade), revelando-se significativamente diferentes. Portanto, esta etapa garante que os modelos sejam equilibrados e obtenham um melhor desempenho, além de ser importante para o sucesso das demais etapas do método proposto.

5.6.5 Segmentação de Rins sem/com a Etapa de Reconstrução de Tumores Renais

Neste experimento, o foco foi avaliar a etapa de reconstrução de tumores renais do método proposto para segmentação de rins. Basicamente, a etapa de reconstrução tumoral foi excluída para verificar o impacto nos resultados da segmentação renal. Vale ressaltar que todas as configurações experimentais e demais etapas foram aplicadas. Os resultados são apresentados na Tabela 15.

Tabela 15 – Resultados da segmentação de rins com e sem a etapa de reconstrução tumoral.

Segmentação de Rins	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)
Sem Reconstrução de Tumores Renais	97,11	94,44	99,94	97,57	99,97
Com Reconstrução de Tumores Renais	97,45	95,05	99,95	98,44	99,96
<i>P-value</i>	0,2794	0,1550	0,8767	0,0011	0,9375

Analizando os resultados das métricas de avaliação, pode-se notar que os melhores resultados da segmentação de rins é aplicando também a etapa de reconstrução tumoral. Essa etapa é capaz de recuperar regiões consideráveis de tumores, que por sua vez também são regiões renais e possivelmente não foram inicialmente segmentadas. Portanto, ao

aplicar a reconstrução tumoral, as regiões renais também são reconstruídas, delimitando melhor os contornos renais. Estatisticamente, obteve-se um *p-value* de 0,0011 para a métrica de sensibilidade em relação às abordagens comparadas. Isso significa que o uso da etapa de reconstrução tumoral tem impacto positivo na obtenção de mais regiões da classe positiva (rins). Portanto, acredita-se que essa etapa é válida para obter resultados mais precisos para a segmentação renal.

5.6.6 Segmentação de Tumores Renais aplicando a Máscara do Método/Especialista

No último experimento, a segmentação de tumores renais foi analisada usando diferentes abordagens de imagens de entrada para o método. Para validar isso, apenas as formas de entrada para a segmentação de candidatos a tumores renais na região renal foram usadas para estudo. As demais etapas não foram utilizadas para não causar impacto no estudo avaliado em questão. Portanto, o objetivo deste experimento é investigar o impacto da segmentação de tumores quando a máscara renal marcada pelo especialista é aplicada como entrada da rede e como as demais etapas são refletidas.

Tabela 16 – Resultados da segmentação de tumores renais aplicando a máscara do método e do especialista.

Segmentação de Tumores Renais	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)
Máscara do Especialista	84,95	75,06	99,68	84,44	99,80
Máscara do Método	81,02	69,66	99,93	80,52	99,97
<i>P-value</i>	0,0010	0,0001	0,0738	0,0011	0,1136

As duas abordagens avaliadas foram: aplicar a máscara proveniente da segmentação inicial dos rins (etapa anterior) como entrada da rede; e usar a máscara renal marcada pelo especialista como entrada da rede. Os resultados dessas abordagens são apresentados na Tabela 16. É perceptível que os resultados usando como entrada da rede a máscara do especialista apresentaram resultados superiores. Além disso, em termos estatísticos também comprovam que as abordagens são significativamente diferentes.

A partir disto, é possível tirar conclusões sobre a segmentação inicial dos rins obtida pelo método proposto. Em que, embora 96,37% de Dice seja obtido na segmentação inicial dos rins (Tabela 7), partes de regiões renais que também são tumores não são segmentadas. Portanto, usando a máscara dos rins da segmentação inicial dos rins, não foi possível segmentar maiores regiões tumorais devido à perda de regiões renais. No entanto,

resultados superiores foram produzidos aplicando a máscara do especialista ao mesmo modelo de segmentação (DeepLabv3+ 2.5D). Portanto, é possível concluir que o modelo de segmentação de tumores apresenta bom desempenho, pois o que causa resultados inferiores é somente quando as regiões renais que possuem tumores não são visíveis como áreas para segmentação.

Entretanto, é importante destacar que a deficiência da segmentação de rins é suprida com as demais etapas do método proposto. Em que as regiões não segmentadas inicialmente são reconstruídas usando a etapa de reconstrução dos tumores renais. Mediante isso, os resultados finais para a segmentação de tumores (84,06% de Dice, na Tabela 12) são equiparáveis aos resultados quando não há perda de regiões renais/tumorais. Mais uma vez, é notável como as etapas do método proposto são essenciais e robustas, o que acaba impactando nos bons resultados para a segmentação de rins e tumores.

5.7 Diversos Experimentos Testados

Finalmente, esta última seção descreve um resumo dos principais experimentos realizados na elaboração do método proposto. A Tabela 17 apresenta o resumo dos principais pré-processamentos, arquiteturas, codificadores e abordagens aplicadas. Os diversos experimentos foram aplicados em diferentes etapas do método e foram feitas combinações entre eles.

Tabela 17 – Resumo dos diversos experimentos testados.

Pré-processamentos	Arquiteturas	Codificadores	Abordagens
CLAHE, bilateral, média, equalização do histograma e unsharp masking	UNet, UNet++, DeepLabv3, DenseUNet e Pyramid Attention Network	AlexNet, DenseNet, ResNet-18, ResNet-34, ResNet-50, ResNet-152, ResNeXt, SqueezeNet, Xception, Inception, DPN-98, EfficientNet-B3 e EfficientNet-B7	2D, 2.5D (com 5 e 7 fatias)

Na primeira coluna, podem ser vistos os principais pré-processamentos que foram aplicados na etapa inicial do método proposto. Todos eles foram testados individualmente e em conjunto com o objetivo de melhorar as imagens usadas no método, visando aprimorar o desempenho de segmentação. A segunda coluna contém as arquiteturas de rede que foram testadas para segmentação de rins e candidatos a tumores renais, nas quais pode-se observar que foram usadas desde arquiteturas clássicas como a U-Net até arquiteturas de última geração como DeepLabv3 e *Pyramid Attention Network* (LI et al., 2018). A terceira coluna mostra os codificadores que foram testados nas arquiteturas descritas.

A última coluna, mostra as abordagens 2D e 2.5D. No caso da abordagem 2.5D, foram testados com 5 e 7 fatias. *O data augmentation offline* também foi testado, mas desvantagens como memória e tempo de treinamento adicionais e alto custo computacional foram encontrados. De qualquer forma, *data augmentation online* apresentou resultados superiores com baixo custo computacional em relação aos outros. Por fim, é importante destacar que para cada um dos experimentos foram aplicadas diversas combinações entre eles. Em geral, como pode ser visto, durante o desenvolvimento do método proposto atual uma série de testes foram realizados. No entanto, os experimentos descritos não superaram os resultados atuais, apresentando resultados inferiores a 90,12% e 66,7% de Dice na segmentação de rins e tumores, respectivamente, levando-se em consideração a melhor combinação.

5.8 Considerações Finais

Neste capítulo, os resultados obtidos pelos experimentos realizados no desenvolvimento do método proposto foram apresentados e discutidos. Além disso, mais alguns experimentos foram realizados a fim de validar as etapas do método proposto.

No próximo capítulo serão apresentados os estudos de casos, também será feita uma comparação com os trabalhos descritos na literatura, visando contextualizar a relevância da pesquisa desenvolvida neste trabalho. Por fim, serão descritos os principais aspectos e limitações encontrados após a análise das etapas propostas.

6 DISCUSSÃO

Este capítulo apresenta vários estudos de caso, bem como uma comparação do método proposto com os trabalhos relacionados encontrados na literatura (Capítulo 2). Por fim, são discutidas as vantagens e limitações do método.

6.1 Estudos de Caso – Segmentação de Rins

Esta seção apresenta um conjunto de estudos de caso para analisar a influência das etapas propostas na segmentação de rins. Os casos observados foram: (1) a segmentação inicial apresentou uma segmentação precisa e as demais etapas não foram necessárias; (2) as etapas de reconstrução dos tumores renais e pós-processamento reduziram ligeiramente o desempenho da segmentação inicial; e (3) as etapas de reconstrução de tumores renais e pós-processamento foram necessárias e capazes de melhorar a qualidade da segmentação inicial.

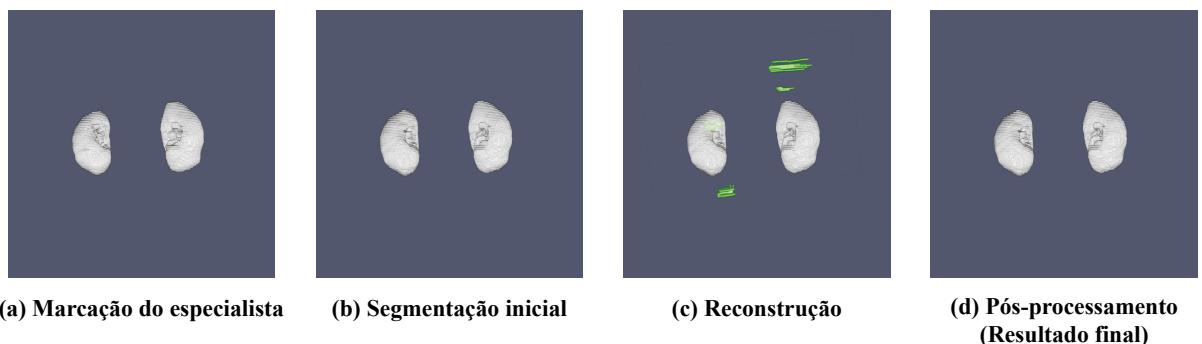
6.1.1 Estudo de Caso 1 – Paciente case_00018

O primeiro estudo de caso trata de um exame em que as etapas de reconstrução e pós-processamento dos tumores foram desnecessárias. Este caso é ilustrado na Figura 31, onde pode-se ver que nenhum fragmento segmentado está presente, exceto os rins. Além disso, é notável que a etapa de reconstrução dos tumores renais (Figura 31 (c), em cinza a segmentação inicial dos rins e em verde a reconstrução dos tumores renais) não contribuiu para uma maior segmentação na região renal, uma vez que as regiões reconstruídas dos tumores renais já haviam sido segmentadas na etapa inicial (dentro do rim). A segmentação dos rins obteve 94,67% Dice, 89,88% Jaccard, 99,91% de acurácia, 99,88% de sensibilidade e 99,91% de especificidade. Após as etapas de reconstrução e pós-processamento (Figura 31 (d)) dos tumores, os resultados permaneceram os mesmos. Portanto, a segmentação inicial foi capaz de determinar contornos renais muito precisos. Esse tipo de comportamento foi encontrado em 4 pacientes.

6.1.2 Estudo de Caso 2 – Paciente case_00149

Neste estudo de caso, as etapas de reconstrução e pós-processamento dos tumores afetaram sutilmente a segmentação inicial. Pode-se ver na Figura 32 que a combinação das

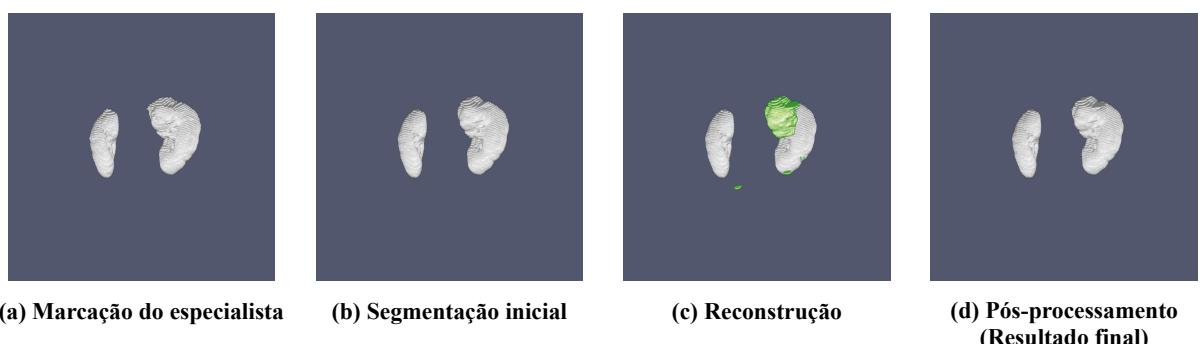
Figura 31 – Estudo de caso 1. Segmentações dos rins (cinza) e reconstrução dos tumores renais (verde).



Fonte: Elaborado pela autora.

etapas descritas não foi capaz de melhorar os resultados da segmentação inicial. A etapa mais crítica foi a reconstrução de tumores renais (Figura 32 (c)), na qual é possível observar que comprometeu a região renal com mais fragmentos de falsos positivos. A segmentação inicial obteve 98,48% Dice, 97,00% Jaccard, 99,95% de acurácia, 98,37% de sensibilidade e 99,98% de especificidade e a segmentação final (reconstrução + pós-processamento) obteve 98,37% Dice, 96,79% Jaccard, 99,94% de acurácia, 98,50% de sensibilidade e 99,97% de especificidade. Apesar disso, as taxas de desempenho da maioria dos quantificadores não foram tão afetadas. Esse tipo de comportamento foi encontrado em apenas 2 pacientes.

Figura 32 – Estudo de caso 2. Segmentações dos rins (cinza) e reconstrução dos tumores renais (verde).



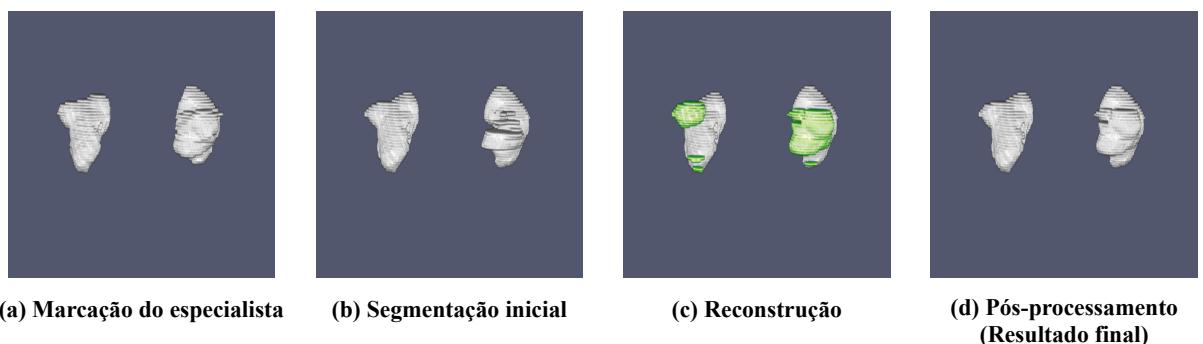
Fonte: Elaborado pela autora.

6.1.3 Estudo de Caso 3 – Paciente case_00045

No último estudo de caso dos rins, as etapas de reconstrução e pós-processamento dos tumores funcionaram bem. É possível ver na Figura 33 que a segmentação inicial dos rins gerou alguns fragmentos errôneos (falsos positivos). Além disso, algumas regiões renais

não foram segmentadas, porém a reconstrução dos tumores renais foi capaz de suprir essa deficiência, pois conseguiu recuperar regiões tumorais que consequentemente também fazem parte dos rins (Figura 33 (c)). Finalmente, com a aplicação do pós-processamento (Figura 33 (d)), pode-se ver como a marcação do método proposto é muito semelhante à marcação do especialista. Nesse caso, a segmentação inicial atingiu 89,93% de Dice, 81,71% de Jaccard, 99,81% de acurácia, 86,63% de sensibilidade e 99,94% de especificidade. Após as etapas de reconstrução e pós-processamento dos tumores, os resultados foram 94,78% Dice, 90,07% Jaccard, 99,89% de acurácia, 96,31% de sensibilidade e 99,92% de especificidade, basicamente uma melhoria de 4,85% de Dice em relação à segmentação inicial. Foram encontrados 25 pacientes semelhantes a este comportamento.

Figura 33 – Estudo de caso 3. Segmentação dos rins (cinza) e reconstrução dos tumores renais (verde).



Fonte: Elaborado pela autora.

6.2 Estudos de Caso – Segmentação de Tumores Renais

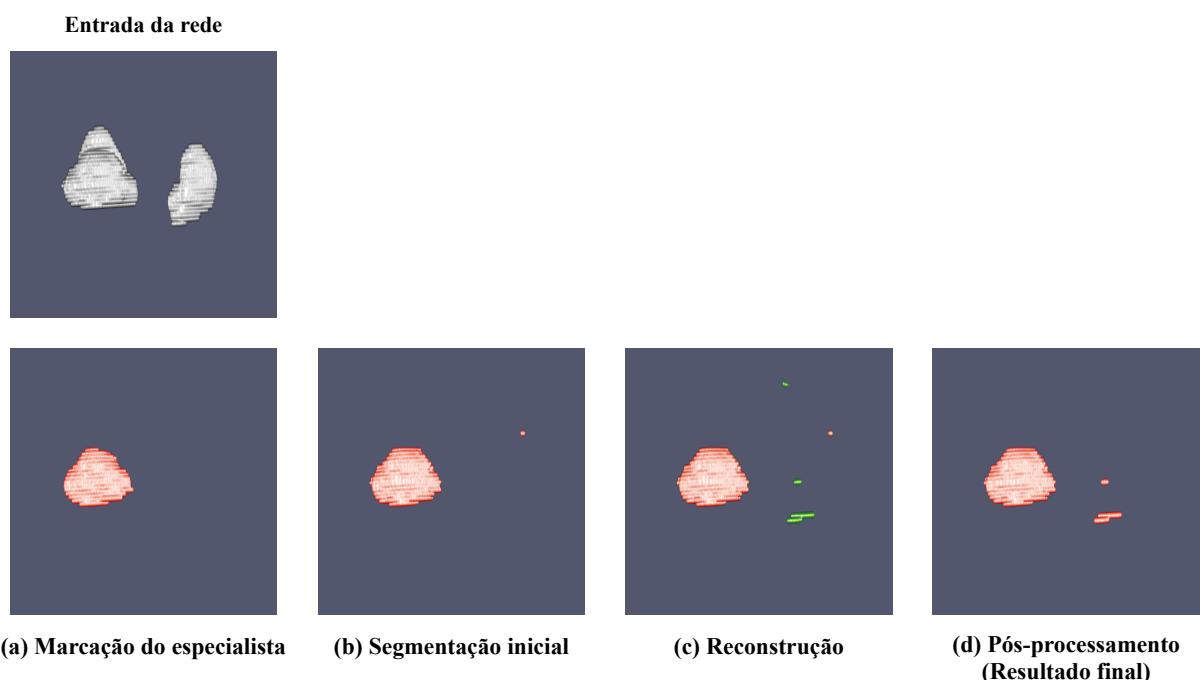
Nesta seção, são mostradas e discutidas as situações mais comuns observadas na aplicação das etapas do método proposto para a segmentação de tumores. Os cenários encontrados foram os seguintes: (1) as etapas de reconstrução de tumores renais e pós-processamento afetaram a qualidade da segmentação inicial dos candidatos a tumores renais na região renal; e (2) as etapas de reconstrução de tumores renais e pós-processamento melhoraram a segmentação inicial dos candidatos a tumores renais na região renal.

6.2.1 Estudo de Caso 1 – Paciente case_00012

O primeiro estudo de caso explora uma circunstância em que as etapas de reconstrução e pós-processamento dos tumores afetaram a segmentação inicial de candidatos

a tumores renais na região renal. Observa-se na Figura 34 a entrada da rede de segmentação (segmentação final dos rins), a marcação do especialista (b) e as etapas do método proposto (c) e (d). Após as duas últimas etapas (Figura 34 (d)), houve uma deterioração no desempenho em comparação com a segmentação inicial. Neste caso, os resultados finais foram 89,85% de Dice, 81,57% de Jaccard, 99,92% de acurácia 98,71% de sensibilidade e 99,92% especificidade. No entanto, a segmentação inicial atingiu 90,69%, 82,97%, 99,92%, 98,04% e 99,93% de Dice, Jaccard, acurácia, sensibilidade e especificidade, respectivamente. Dessa forma, as etapas de reconstrução e pós-processamento dos tumores não foram capazes de definir com maior precisão as regiões dos tumores. Este cenário foi observado em 4 pacientes.

Figura 34 – Estudo de caso 1. Segmentação dos rins (cinza), segmentação dos tumores renais (vermelho) e reconstrução dos tumores renais (verde).



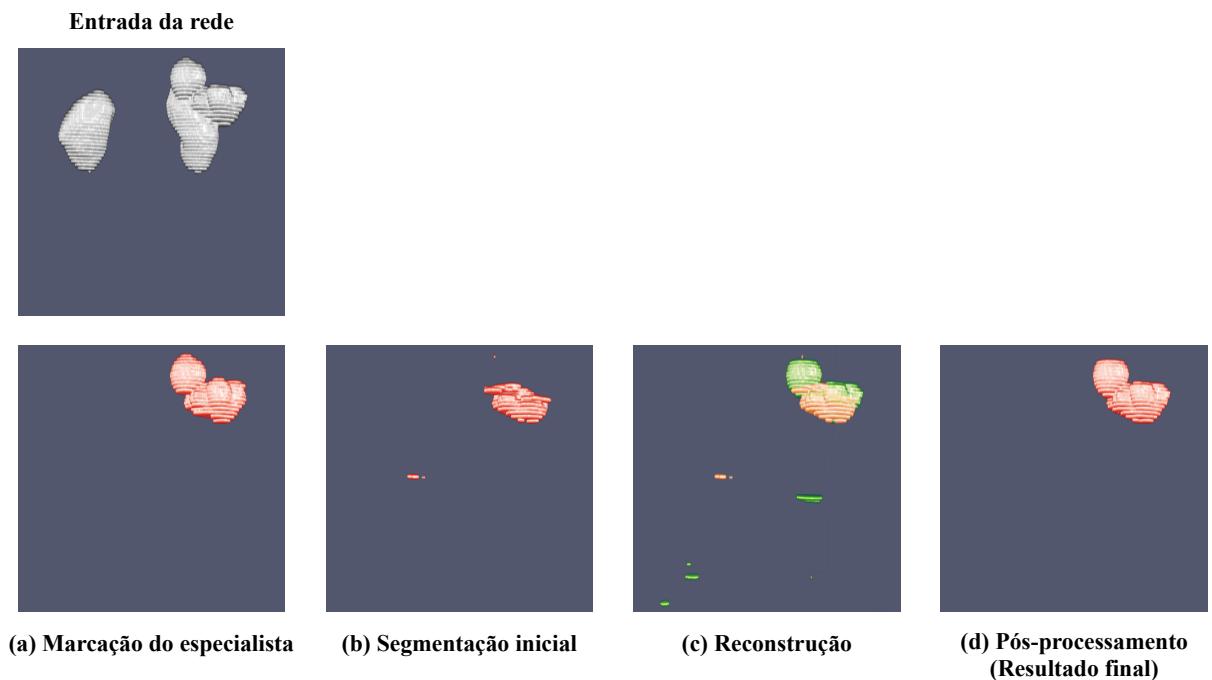
Fonte: Elaborado pela autora.

6.2.2 Estudo de Caso 2 – Paciente case_00028

Neste estudo de caso, as etapas de reconstrução e pós-processamento dos tumores compensaram algumas deficiências da segmentação inicial. Na Figura 35 (b), observa-se que a segmentação inicial gerou vários fragmentos errôneos. Porém, após as etapas de reconstrução e pós-processamento (Figura 35 (c) e Figura 35 (d), respectivamente), a

segmentação pelo método proposto eliminou um grande número de falsos positivos. Neste caso, foram obtidos 94,50% de Dice, 89,57% Jaccard, 99,97% de acurácia, 93,85% de sensibilidade e 99,98% de especificidade na segmentação final, basicamente uma melhoria significativa de 27,67% na métrica Dice em comparação com a segmentação inicial, onde os resultados foram 66,83% de Dice, 50,18% de Jaccard, 99,84% de acurácia, 51,52% de sensibilidade e 99,99% de especificidade. Este cenário foi observado em 27 casos de teste.

Figura 35 – Estudo de caso 2. Segmentação dos rins (cinza), segmentação dos tumores renais (vermelho) e reconstrução dos tumores renais (verde).



Fonte: Elaborado pela autora.

6.3 Comparação do Método Proposto com os Trabalhos Relacionados

Nesta seção é apresentada uma análise comparativa dos resultados obtidos pelo método proposto em relação aos trabalhos relacionados (Capítulo 2). As subseções a seguir apresentam uma comparação com os trabalhos relacionados à segmentação de rins, segmentação de tumores, e segmentação de rins e tumores renais.

6.3.1 Segmentação de Rins

Esta subseção mostra os resultados dos trabalhos relacionados à segmentação de rins. A Tabela 18 apresenta um resumo desta comparação, fornecendo informações das

técnicas, quantidade de pacientes e as métricas de desempenho usadas em seus métodos. Por fim, são apresentados os resultados do método proposto.

Tabela 18 – Comparação dos resultados do método com trabalhos relacionados à segmentação de rins.

Técnica(s)	Base	Pacientes	Dice (%)	Jacc (%)	Acc (%)	Sen (%)	Esp (%)
Modelo deformável baseado em <i>Level Set</i> (KHALIFA et al., 2011)	Privada	29	95	-	96,36	93,8	-
Esquema hierárquico de registro e ponderação de atlas (WOLZ et al., 2013)	Privada	150	92,5	86,8	-	98,3	-
Continuidade contextual na sequência de imagens de TC (ZHAO; LIANG; FAN, 2013)	Privada	10	94,7	-	-	-	-
Registro de imagens multi-atlas (YANG et al., 2014)	Privada	14 fatias	95,2	-	-	-	-
Estrutura híbrida baseada de modelos geométricos deformáveis e NMF (KHALIFA et al., 2016)	Privada	36	96,45	-	-	-	-
Contorno ativo usando a estrutura <i>Level Set</i> (SKALSKI et al., 2017)	Privada	10	86,2	-	-	-	-
CNNs 3D (JACKSON et al., 2018)	Privada	113	88,5	-	-	-	-
Crowdsourcing e CNN (MEHTA et al., 2019)	Privada	42	93,2	-	-	-	-
U-Net, <i>patch</i> e sistema de rastreamento (LES et al., 2021)	KiTS19	270	90,63	-	-	-	-
Método proposto com ResUNet 2.5D, DeepLabv3+ 2.5D, agrupador de tumores e técnicas para remover falsos positivos	KiTS19	210	97,45	95,05	99,95	98,44	99,96

Nos últimos anos, muitas técnicas de segmentação renal foram desenvolvidas e aplicadas em diversos estudos. Por exemplo, Khalifa et al. (2011) usaram modelos deformáveis baseados em conjuntos de níveis geométricos para segmentação renal em TC. Khalifa et al. (2016) também fizeram uso de modelos deformáveis. Os modelos deformáveis são controlados por informações sobre a intensidade e a forma estatística dos rins. Em casos de ruído, imagens de baixa resolução e limites difusos, pode aumentar a instabilidade quanto à identificação dos contornos renais e resultar em falhas no método. Khalifa et al. (2011) atingiram um Dice de 95%, acurácia de 96,36% e sensibilidade de 93,8%, considerando apenas 29 pacientes. Em Khalifa et al. (2016) 95% de Dice foram alcançados, considerando 36 pacientes. Os dois estudos fazem uso de TCs de pacientes saudáveis, ou seja, sem presença de tumores ou cistos. O método proposto nesta tese usa TCs de pacientes com presença de anormalidades, o que dificulta a segmentação dos rins. Ainda assim, este método superou os resultados dos dois trabalhos descritos.

Alguns trabalhos que usaram atlas e registros foram propostos por Wolz et al. (2013) e Yang et al. (2014) alcançando métricas promissoras na tarefa de segmentação renal. A desvantagem de suas abordagens é que geralmente elas tem um certo tamanho para resolver o problema em um conjunto de imagens e, ao testar novas imagens de

tamanhos diferentes do atlas, o método não tem garantia de funcionar. Observe que em Wolz et al. (2013) o conjunto de imagens usado para validar o método contém 150 pacientes saudáveis e em Yang et al. (2014) apenas 14 fatias são usadas, das quais seis resultados de segmentação renal com grandes tumores foram excluídos porque o tumor alterou significativamente a forma do rim. O método aqui proposto não usa atlas e registros, o que o torna mais robusto na tarefa de segmentação renal. Além disso, todas as imagens de TC disponíveis na base de imagens foram utilizadas, independentemente dos resultados. O índice de Dice alcançado em nosso método proposto foi de 97,45%.

O método proposto por Zhao, Liang e Fan (2013) faz uso da continuidade contextual de fatias adjacentes em imagens de TC, então uma estrutura de segmentação baseada em fatias é construída para segmentar automaticamente os rins. Para este propósito, quatro parâmetros contextuais, incluindo distância e sobreposição, são selecionados para estimar a continuidade entre fatias adjacentes. Para isso, utilizam-se algoritmos convencionais de segmentação e modificação (modelo deformável tradicional e limiar iterativo local), chegando a 94,7% de Dice, utilizando um banco de dados privado composto por 10 pacientes com doenças renais. No método em estudo, propõe-se o uso de duas redes neurais convolucionais para atingir uma segmentação precisa para os rins, assim, são inseridas abordagens capazes de segmentar sem a necessidade de um ponto de partida.

A segmentação dos rins baseada no método Contorno Ativo usando uma estrutura *Level Set* com restrições elipsoidais foi proposta em Skalski et al. (2017). A solução proposta leva em consideração as informações sobre a região e os termos de contorno, bem como as restrições dedicadas à característica de formato do rim, o que torna o método bastante restrito e, não é garantido que o método funcione bem com o surgimento de novas imagens com formatos ligeiramente diferentes. O método foi testado em 10 tomografias de pacientes com câncer renal, obtendo resultados de 86,2% de Dice. Em nossa proposta não há necessidade de informações anatômicas, e alcançou resultados superiores a 97% de Dice usando 31 pacientes com tumores. Jackson et al. (2018) desenvolveram uma ferramenta automatizada com base em CNNs 3D e alcançaram 88,5% de Dice usando 113 TCs e *data augmentation offline*.

Em Mehta et al. (2019), os autores usaram uma rede neural convolucional e uma ferramenta de *crowdsourcing* para segmentação de órgãos em larga escala. As validações foram realizadas comparando as segmentações de *crowdsourcing* e segmentações marcadas por especialistas, e verificou que o desempenho não foi significativamente diferente. Porém,

a realização de marcações renais por usuários sem o conhecimento especializado na área é imprudente, pois isso pode ser feita de forma errada e afetar a segmentação dos rins e, consequentemente, gerar resultados errados. Os resultados foram 93,2% de Dice usando para validar uma base de 42 pacientes saudáveis. Les et al. (2021) apresentaram um método composto por duas etapas. Na primeira etapa, é realizada a segmentação semântica baseada na rede U-Net e *batch*. Na segunda etapa, um sistema de rastreamento (*middle-bottom, middle-top*) foi usado para encontrar contornos do rim nas fatias de TC restantes. O método desenvolvido atingiu um Dice de 90,63%.

6.3.2 Segmentação de Tumores Renais

Esta segunda subseção apresentada uma análise comparativa do método proposto com os trabalhos relacionados à segmentação de tumores renais. O resumo desses trabalhos é apresentado na Tabela 19, com informações sobre as técnicas usadas, quantidade de pacientes e as métricas de desempenho aplicadas. Além disso, a última linha da tabela apresenta os resultados obtidos pelo método proposto neste estudo.

Tabela 19 – Comparação dos resultados do método com trabalhos relacionados à segmentação de tumores renais.

Técnica(s)	Base	Pacientes	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)
Estrutura híbrida baseada em SIFCM e DRLSE (KAUR; JUNEJA; MANDAL, 2019)	Privada	40	88,2	88,5	86,51	-	-
CNNs fracamente supervisionada (YANG et al., 2020)	Privada	200	82,6	-	-	-	-
MB-FSGAN (RUAN et al., 2020)	Privada	113	85,9	-	95,7	86,2	89,4
HybridNet 3D (YAN et al., 2020)	KiTS19	300	79,7	-	-	-	-
V-Net usando <i>Two-Stage Bottleneck Block</i> (TÜRK; LÜY; YALCINKAYA, 2022)	KiTS19	210	86,9	76,8	-	-	-
3D U-Net preservando a simetria rotacional em cortes axiais (TANIMOTO et al., 2022)	Privada	213	60,4	-	-	-	-
Método proposto com ResUNet 2.5D, DeepLabv3+ 2.5D, agrupador de tumores e técnicas para remover falsos positivos	KiTS19	210	84,06	75,04	99,94	88,33	99,95

É importante mencionar que todos os trabalhos relacionados nesta subseção desenvolveram métodos semiautomáticos que segmentam os tumores a partir da marcação dos rins feita por especialistas. Isso torna o método dependente do especialista humano, pois não houve uma etapa para segmentação dos rins por técnicas computacionais. Kaur, Juneja e Mandal (2019) sugerem uma técnica de segmentação híbrida baseada em dois métodos que incluem SIFCM (*Spatial Intuitionistic Fuzzy C-Means Clustering*) que integra

detalhes de imagem espacial e DRLSE (*Distance Regularized Level-Sets Evolution*) para extração de lesão renal. Esta abordagem requer uma estimativa grosseira da região de interesse fornecida pelo SIFCM dentro do rim. Assim, o resultado correto da segmentação depende das informações preliminares e da posição da função de definição de nível. Portanto, a inicialização adequada do ajuste de nível próximo ao limite do tumor renal é essencial para a demarcação correta. O método foi validado em 40 TCs de pacientes com câncer renal, obtendo resultados de 88,2% Dice, 88,5% Jaccard e acurácia de 86,51%. Em nosso método foram testados mais exames, e não há necessidade de aproximar a região de interesse, além disso, não foram usadas marcações dos rins feitas por especialistas, mas sim por técnica computacional.

Yang et al. (2020) usaram um conjunto de CNNs fracamente supervisionada para segmentação de tumores. Uma estrutura de três estágios foi introduzida para treinar as CNNs com as anotações fracas de tumores (caixas delimitadoras). Para treinamento das CNNs, cada imagem é aplicada *data augmentation offline*, totalizando 14.400 novas imagens. O resultado do trabalho foi de 82,6% de Dice usando TCs de 200 pacientes. Nosso método obteve resultados acima de 82% de Dice. Além disso, foi utilizado *data augmentation* em tempo real, o que apresenta um método com melhor desempenho e baixo consumo de recursos de máquina.

No estudo de Ruan et al. (2020) é proposto o MB-FSGAN, que consiste em um extrator de características em várias escalas, e alcançaram 85,9% de Dice, 95,7% de acurácia, 86,2% de sensibilidade e 89,4% de especificidade, usando 113 TCs. Neste estudo, levaram em consideração apenas 20 pacientes para testar o método. No estudo recente de Tanimoto et al. (2022), um modelo U-Net 3D foi usado para segmentar tumores levando em consideração a estrutura global dos rins. O método obteve resultados promissores de 60,4% de Dice utilizando 213 TCs.

Trabalhos como os de Yan et al. (2020), Türk, Lüy e Yalcinkaya (2022) também usaram abordagens 3D para segmentar tumores. Esses métodos apresentam desempenho superior a 79,7% de Dice, consumindo alto poder computacional. Vale ressaltar que essas abordagens que usam como entrada regiões renais completas, podem apresentar resultados mais expressivos, já que para segmentar os tumores usam as marcações manuais dos rins feitas por especialistas. Em nosso estudo, foram alcançados 84,06% de Dice para os tumores usando uma abordagem completa (segmentação dos rins e tumores) com baixo custo computacional comparado aos trabalhos citados.

6.3.3 Segmentação de Rins e Tumores Renais

Nesta última subseção, são apresentados os trabalhos relacionados a segmentação de rins e tumores renais. Na Tabela 20 é apresentado o resumo dos trabalhos relacionados, da mesma forma que nas subseções anteriores, no qual contém as informações sobre as técnicas aplicadas, quantidade de pacientes e as métricas de desempenho. Os resultados para rins são expostos na linha superior e os tumores renais na linha inferior. Finalmente, são mostrados os resultados obtidos pelo método proposto.

Tabela 20 – Comparação dos resultados do método proposto com os trabalhos relacionados à segmentação de rins e tumores renais.

Técnica(s)	Base	Pacientes	Dice(%)	Jacc(%)	Acc(%)	Sen(%)	Esp(%)
Rede 3D com PPM + Atlas (YANG et al., 2018)	Privada	140	93,1 80,2	-	-	-	-
Modelo híbrido V-Net (TÜRK; LÜY; BARIŞÇI, 2020)	KiTS19	210	97,7 86,5	-	-	-	-
Rede residual híbrida 3D com SE (QAYYUM; LALANDE; MERIAUDEAU, 2020)	KiTS19	210	97,8 86,8	-	-	95,6 91,31	99,6 91,45
SE-ResNeXT U-Net (XIE et al., 2020)	KiTS19	300	96,77 74,32	-	-	-	-
MSS U-Net 3D (ZHAO et al., 2020)	KiTS19	210	96,9 80,5	-	-	-	-
Attention U-Net (GEETHANJALI; MINAVATHI; DINESH, 2021)	KiTS19	205	95,65 93,86	-	-	-	-
U-Nets 3D: simples, residual e residual de pré-ativação (HELLER et al., 2021)	KiTS19	210	97,4 85,1	-	-	-	-
3D U-Net (LIN et al., 2021)	Privada	441	97,3 84,4	-	-	-	-
3D-MS-RFCNN (YANG et al., 2022)	KiTS19 + Privada	480	91,62 71,64	-	-	-	-
3D-CNN e ConvLSTM (KANG et al., 2022)	KiTS19	300	96,39 78,9	-	-	-	-
Método proposto com ResUNet 2.5D, DeepLabv3+ 2.5D, agrupador de tumores e técnicas para remover falsos positivos	KiTS19	210	97,45 84,06	95,05 75,04	99,95 99,94	98,44 88,33	99,96 99,95

Vale mencionar que nesta subseção são considerados os trabalhos que fazem o processo completo, segmentando os rins e tumores automaticamente. Esses trabalhos são propostos por Yang et al. (2018), Türk, Lüy e Barışçı (2020), Qayyum, Lalande e Meriaudeau (2020), Xie et al. (2020), Zhao et al. (2020), Geethanjali, Minavathi e Dinesh (2021), Heller et al. (2021), Lin et al. (2021), Yang et al. (2022), Kang et al. (2022). Türk, Lüy e Barışçı (2020) apresentam um novo modelo híbrido usando recursos aprimorados

em modelos V-Net existentes. Usando um conjunto de 20 TCs para testar o modelo, alcançaram resultados de 97,7% e 86,8% de Dice para os rins e tumores, respectivamente. Em nosso estudo, obtivemos resultados semelhantes usando um conjunto de 31 TCs para teste.

Nos trabalhos de Xie et al. (2020) e Geethanjali, Minavathi e Dinesh (2021) foram feita combinações de arquiteturas. Xie et al. (2020) combinaram as vantagens da SE-Net, ResNeXT e U-Net, e obtiveram 96,77% e 74,32% de Dice para rins e tumores, respectivamente. Geethanjali, Minavathi e Dinesh (2021) implementaram um modelo U-Net com camadas *Attention Gate*, alcançando em seu método um Dice de 95,65% para rins e 93,86% para tumores, usando 20 TCs para teste. Em nosso estudo, também realizamos combinações de arquiteturas, como a DeepLabv3+ ajustada com o codificador DPN-131, e obtivemos resultados semelhantes, 97,45% de Dice para rins e 84,06% para tumores usando 31 TCs para teste.

Finalmente, são apresentados vários outros trabalhos que usaram aprendizado profundo para segmentar os rins e tumores. Esses métodos apresentaram resultados bem-sucedidos e se destacaram pela arquitetura 3D, que apresenta alto desempenho, como no trabalho de Qayyum, Lalande e Meriaudeau (2020) que propuseram uma rede residual 3D híbrida com *Squeeze-and-Excitation* (SE). A rede foi treinada em 4.000 épocas, com alto poder de processamento. Foram obtidos resultados de 97,8% de Dice para rins e 86,8% de Dice para tumores, usando uma base de imagens de 210 pacientes, dentre eles, 10% para testar o método. Zhao et al. (2020) implementaram uma U-Net 3D supervisionada em várias escalas, MSS U-Net 3D U-Net, e também tiveram um bom desempenho, alcançando um Dice de 96,9% e 80,5% para rins e tumores, respectivamente.

No estudo de Kang et al. (2022) é proposta uma rede combinando as características da ConvLSTM e U-Net 3D e são obtidos resultados expressivos de 96,39% de Dice para rins e 78,9% para tumores. Lin et al. (2021) construíram dois modelos de segmentação 3D baseados em U-Net. No trabalho de Yang et al. (2022) propuseram uma nova rede neural profunda, a 3D-MS-RFCNN para melhorar a segmentação em tumores de tamanho extremamente grande. Ambos os trabalhos tiveram bons desempenhos para segmentação de rins e tumores, usando base de imagens extensas, porém privadas. Isso acaba limitando a realização de experimentos com essas bases de imagens.

Heller et al. (2021) e Yang et al. (2018) também propuseram abordagens 3D. Heller et al. (2021) usaram três modelos 3D da UNet e alcançaram o primeiro lugar no

desafio KiTS19. Para treinar a rede usaram 1.000 épocas. Além disso, os autores afirmam que não mostraram nenhum benefício significativo na construção do modelo. O poder da abordagem 3D foi suficiente para alcançar os melhores resultados, no entanto, alto poder computacional é consumido. Essa abordagem atingiu 97,4% de Dice para rins e 85,1% de Dice para os tumores. Yang et al. (2018) apresentaram um método que usa redes 3D combinadas com PPM e atlas. Além disso, foi aplicado o *data augmentation offline*, gerando 90.000 novas imagens, cerca de 1.000 vezes mais que o número de imagens originais. Como resultado, obtiveram 93,1% e 80,2% de Dice para rins e tumores, respectivamente, usando 140 TCs. Nosso método obteve excelente desempenho nas segmentações dos rins e tumores em comparação aos modelos 3D existentes com baixo custo computacional.

De acordo com as Tabelas 18, 19 e 20, pode-se observar que existem vários métodos na literatura que investigam a segmentação dos rins e tumores. Além disso, as técnicas se tornaram cada vez mais robustas ao longo dos anos e quase sempre alcançaram um desempenho superior a 86% e 80% em todas as métricas de validação para segmentação de rins e tumores, respectivamente. Ademais, é possível verificar que os melhores resultados fazem uso de abordagens 3D em seu método. As técnicas de aprendizado profundo são promissoras, mas apresentam várias desvantagens, como o alto custo computacional e grande quantidade de conjunto de imagens necessário. Portanto, as abordagens 3D nem sempre são opções viáveis.

Em contraste com os métodos descritos na literatura, destaca-se que o trabalho proposto usa modelos de última geração em conjunto com abordagens 2.5D de baixo custo computacional. Esses modelos são capazes de segmentar rins e tumores com alta precisão, garantindo melhor os limites (bordas) dos objetos e removendo o máximo de previsões erradas. Isso diminui o trabalho do médico especialista em verificar todas as regiões consideradas rins e tumores. Além disso, o método proposto obteve resultados comparáveis aos trabalhos da literatura, mostrando-se promissor e robusto. Por fim, vale ressaltar que diversas métricas são utilizadas para validar o método proposto, ao contrário dos trabalhos relacionados, o que impossibilitam outras comparações. Dessa forma, é possível demonstrar a viabilidade da utilização do método proposto para segmentação de rins e tumores em TC.

6.4 Aspectos Importantes do Método Proposto

A segmentação de rins e tumores em exames tomográficos não é uma tarefa trivial. Desenvolver um método capaz de contornar todas as adversidades desses órgãos e atingir uma boa taxa de acerto ainda é um desafio. Nesta tese, as principais etapas de um sistema CAD (aquisição de imagens, pré-processamento, segmentação e pós-processamento) foram aplicadas. O método se mostrou robusto e preciso na segmentação de rins e tumores. Ressalta-se a importância dos aspectos e limitações encontrados após a análise das etapas propostas.

1. O estudo em questão é um método completo para resolver o problema de segmentação dos rins e tumores. Nesta área de pesquisa, a dificuldade desta tarefa é amplamente reconhecida. Por esse motivo, os resultados obtidos ganham lugar de destaque entre os métodos de última geração encontrados na literatura.
2. Tendo conhecimento sobre a anatomia dos rins e tumores renais, sabe-se que podem ser heterogêneos em suas características (forma e textura). Dessa forma, é possível ver a importância de uma etapa adicional para encontrar padrões de características existentes e agrupar os casos (exames) semelhantes de acordo com os padrões identificados. Com a formação dos grupos de tumores e distribuindo-os proporcionalmente entre o conjunto de treinamento e validação, garantiu-se que os modelos ficassem equilibrados e tivessem melhor desempenho, pois existem diferentes tipos de tumores nos conjuntos de treinamento e validação. Essa etapa foi fundamental para o sucesso das demais etapas.
3. Em relação ao uso de aumento de dados em tempo real, é uma técnica de regularização implícita para combater o *overfitting* de modelos de aprendizado profundo (KUKAČKA; GOLKOV; CREMERS, 2017). O que não foi diferente nesse estudo, a combinação das operações aplicadas em tempo de execução obteve uma maior diversidade dos tumores renais no conjunto de treinamento, resultando em um maior poder de generalização. Consequentemente, houve um impacto positivo nas métricas de validação para segmentação dos tumores. Esses aumentos mostram que é uma abordagem poderosa para melhorar a generalização e robustez de modelos usando baixo consumo de recursos de máquina.

4. O método proposto é um processo automatizado executado por dois modelos CNN, que são ferramentas muito robustas que realizam implicitamente a extração e seleção de características. Este é um aspecto positivo, pois elimina a necessidade de extrair empiricamente o conjunto de características a serem utilizadas no processo de aprendizagem e de definir as técnicas a serem usadas na seleção das características.
5. Outro aspecto importante foi o uso da abordagem 2.5D. Uma das vantagens dessa abordagem é que as informações espaciais de fatias vizinhas são usadas para identificar tumores, sem comprometer a complexidade computacional. Isso aumentou as chances de segmentação bem-sucedida e reduziu os falsos positivos. Além disso, foi feito um balanceamento das fatias de rins e tumores. Isso contribuiu para que a rede também aprendesse as fatias sem rins/tumores, o que ajudou a reduzir ainda mais os falsos positivos, que é um dos maiores desafio.
6. É importante notar que a etapa inicial de segmentação do rim por si só foi capaz de fornecer bons resultados, comparáveis a outros estudos relevantes na literatura. Acredita-se que isso tenha sido possível devido ao desempenho robusto da arquitetura ResUNet, visto que as informações sobre as camadas são propagadas com a abordagem residual, mantendo e agregando características renais relevantes, o que consequentemente resultou em melhores resultados.
7. Para segmentar os candidatos a tumores na região renal, foi usado o modelo DeepLabv3+ 2.5D com o codificador DPN-131. Esse codificador compartilha as vantagens de redes robustas (ResNet e DenseNet) (CHEN et al., 2017) e ao combinar com a DeepLabv3+ 2.5D, permitiu a exploração de novos recursos, possibilitando o aprendizado de representações tumorais a partir da captura de suas informações contextuais, como diferentes formas e tamanhos. Além disso, o decodificador DeepLabv3+ recupera gradualmente as informações espaciais para reconstruir a saída, capturando melhor os limites dos objetos (CHEN et al., 2018). Essa combinação resultou em uma rede codificador-decodificador eficiente com alto desempenho para segmentação de candidatos a tumores.
8. Ademais, para segmentar os candidatos a tumores renais na região renal, uma abordagem em cascata é aplicada. Essa abordagem consiste em usar os resultados da segmentação dos rins como entrada para segmentar os tumores. Esse processo reduz

o escopo do problema e os erros de segmentação inicial dos tumores. No entanto, pode acontecer de regiões que não foram segmentadas nos rins sejam possíveis regiões tumorais. Caso isso aconteça, não é possível predizer essas regiões e isso acaba reduzindo a chance de obter melhores resultados. Pensando nisso, incluiu-se a etapa de segmentar os candidatos a tumores renais na região abdominal, o que ocasionou melhor segmentação, pois foram recuperadas partes das regiões tumorais.

9. A aplicação da etapa de reconstrução dos tumores renais foi capaz de unir porções consideráveis das regiões tumorais, melhorando os resultados da segmentação dos tumores e também dos rins. Isso foi possível porque à medida que as regiões tumorais foram recuperadas, as regiões renais também foram, pois as regiões tumorais fazem parte dos rins. Portanto, a reconstrução dos tumores renais também possibilitou melhorar a segmentação dos rins.
10. Com a etapa de reconstrução dos tumores renais, vários falsos positivos foram adicionados. Entretanto, ao selecionar apenas os dois maiores elementos na etapa de pós-processamento dos rins, foi possível eliminar não só os falsos positivos encontrados na segmentação inicial, mas também na reconstrução dos tumores renais. Esse pós-processamento acabou impactando positivamente no resultado final da segmentação dos rins e tumores.
11. Na etapa de pós-processamento dos tumores renais, os elementos segmentados com informações contextuais insuficientes para representar os tumores foram removidos, preservando-se apenas os elementos contínuos. Assim, o pós-processamento melhorou a precisão da segmentação dos tumores, removendo muitos falsos positivos. Essa etapa resultou em uma melhoria considerável na maioria dos casos.
12. Finalmente, a combinação de todas as técnicas neste estudo proporcionou uma melhor segmentação dos rins e tumores. Até onde sabe-se, este é o primeiro método que combinou explicitamente todas essas técnicas. Comparado com trabalhos relacionados, o método proposto apresentou resultados expressivos e ganhou destaque, alcançando na segmentação de rins um Dice de 97,45%, Jaccard de 95,05%, acurácia de 99,95%, sensibilidade de 98,44% e especificidade de 99,96%. E na segmentação dos tumores atingiu 84,06%, 75,04%, 99,94%, 88,33% e 99,95% de Dice, Jaccard, acurácia, sensibilidade e especificidade, respectivamente.

Embora o método proposto apresente vários fatores positivos, também apresenta algumas limitações, nas quais destacam-se:

1. A segmentação dos rins e tumores é promissora e significativa, sendo comparável a trabalhos relevantes da literatura. No entanto, a abordagem 2.5D é limitada em termos espaciais do volume de TC. Acredita-se que para obter um melhor desempenho é necessário explorar recursos 3D. Entretanto, não foram explorados devido às limitações de consumo de *hardware* e memória.
2. Com a etapa de reconstrução dos tumores renais, muitas regiões de rins e tumores que não foram segmentadas acabaram sendo recuperadas. Entretanto, algumas regiões não foram totalmente recuperadas. Portanto, outra técnica (por exemplo, *ensemble*) poderia ser adicionada para segmentar mais regiões tumorais, a fim de melhorar a sensibilidade.
3. Por fim, vale destacar que o pós-processamento usado na segmentação de rins e tumores apresenta deficiências, em que ora o pós-processamento contribui efetivamente para a remoção de falsos positivos e ora afeta a segmentação. Portanto, é importante explorar novas técnicas de pós-processamento que sejam mais estáveis para a etapa de remoção de falsos positivos.

Os aspectos positivos discutidos nas etapas do método proposto contribuíram para que os resultados obtidos fossem comparáveis aos trabalhos encontrados na literatura. É perceptível que existem diversos métodos na literatura que propõem soluções para o problema da segmentação dos rins e tumores utilizando abordagens 3D. No entanto, essas abordagens requerem um alto custo/esforço computacional. O método proposto, apesar de algumas limitações, foi capaz de obter resultados satisfatórios com baixo custo computacional. Portanto, este estudo apresenta contribuições para o meio científico e é de fundamental importância na demarcação dos rins e tumores pelo especialista, contribuindo para aumentar a produtividade e melhorar os índices diagnósticos.

7 CONCLUSÃO

As altas taxas de incidência do câncer de rim em todo o mundo mostram a importância do desenvolvimento de pesquisas que forneçam suporte ao diagnóstico precoce da doença. Essa é uma etapa fundamental para viabilizar o tratamento adequado aos pacientes. Nesse contexto, este trabalho apresentou um método automático para a tarefa de segmentação de rins e tumores renais baseado em aprendizado profundo e técnicas de processamento de imagens para reduzir os falsos positivos.

O método proposto apresenta quatro etapas principais. Na primeira etapa, utiliza técnicas de processamento de imagens, como realce da imagem e normalização das intensidades do *voxel*. Além de aplicar técnicas para agrupar tumores renais a fim de distribuir proporcionalmente o conjunto de treinamento e validação. Na segunda etapa, as segmentações iniciais de rins e tumores são obtidas usando ResUNet 2.5D e DeepLabv3+ 2.5D, respectivamente. Na terceira etapa, a reconstrução dos tumores renais usa outro modelo ResUNet 2.5D. Finalmente, na quarta etapa, as segmentações finais dos rins e tumores renais são obtidas por meio do pós-processamento para remoção de elementos que não fazem parte dos rins e tumores.

Para validar o método proposto, foi usada a base de imagem pública KiTS19. Essa base de imagens consiste em tomografias de pacientes em diferentes estágios da doença, tornando-a altamente heterogênea e complexa. Dessa forma, a segmentação dos rins e tumores é uma tarefa desafiadora em cenários realistas. Os resultados experimentais revelaram um desempenho promissor para a tarefa proposta, atingindo 97,47% de Dice, 95,09% de Jaccard, 99,95% de acurácia, 97,86% de sensibilidade e 99,97% de especificidade para segmentação de rins. Para a segmentação de tumores foi obtido 82,94% de Dice, 72,54% de Jaccard, 99,95% de acurácia, 86,51% de sensibilidade e 99,96% de especificidade. De maneira geral, os resultados fornecem fortes evidências de que o método proposto é uma ferramenta poderosa que pode ser incorporada aos sistemas CAD para auxiliar no diagnóstico da doença.

7.1 Contribuições

As principais contribuições do método proposto nesta tese são descritas a seguir:

1. Desenvolvimento de um método automático para agrupar casos (exames) e distribuí-

los proporcionalmente no conjunto de treinamento e validação para garantir que os modelos fossem equilibrados e obtivessem melhor desempenho;

2. Adaptação da arquitetura DeepLabv3+, com o uso da DPN-31 como codificador para a segmentação de tumores renais, fornecendo resultados precisos;
3. Aplicação da técnica de balanceamento de fatias nos conjuntos de treinamento e validação juntamente com a abordagem 2.5D, contribuindo para reduzir consideravelmente os falsos positivos e detectando se uma fatia apresenta rins/tumores ou não;
4. A aplicação da etapa de reconstrução de tumores renais por meio da combinação de duas CNNs recuperou partes consideráveis de regiões de tumores e, consequentemente, proporcionou melhores resultados de segmentação renal;
5. Aplicação de técnicas de processamento de imagens baseadas em informações contextuais para reduzir falsos positivos em segmentações de rins e tumores.

7.2 Trabalhos Futuros

Apesar dos bons resultados obtidos para a segmentação de rins e tumores renais em imagens de TCs, melhorias ainda podem ser feitas no método proposto visando incrementar a sua eficiência. A seguir são destacadas algumas sugestões:

1. Considerando a diversidade de aquisições da base de imagens, padronizar as entradas da rede usando uma rede *autoencoder* como um pré-processamento poderia aumentar a eficiência dos modelos de segmentação;
2. Assim como na abordagem 2.5D usada no método proposto, as fatias são visualizadas sequencialmente. Acredita-se que o uso de recursos recorrentes da rede neural pode melhorar o desempenho, pois essas redes podem “lembrar” informações de fatias anteriores, adicionando recursos ao modelo;
3. Aprimorar ou construir um *ensemble* na etapa de reconstrução dos tumores renais a fim de melhorar a sensibilidade;
4. Construir um método para fazer uma triagem de TCs com e sem tumores renais. Este novo método seria integrado ao método proposto como uma etapa inicial para

identificar quais TCs têm tumores renais. Posteriormente, TCs com tumores seriam usadas como entrada para segmentar as regiões com tumores.

7.3 Produções Científicas

A Tabela 21 apresenta os artigos diretamente relacionados ao método proposto para a segmentação de rins e tumores renais em imagens de TC. Além disso, a Tabela 22 lista os artigos científicos publicados e submetidos em outras aplicações de processamento de imagens e visão computacional desde o início do doutorado.

Tabela 21 – Produções científicas em relação ao método proposto para segmentação de rins e tumores renais.

Artigo	_tipo	Qualis	Status
Kidney Segmentation from Computed Tomography Images using Deep Neural Network. Em: Computers in Biology and Medicine. Ano: 2020.	Periódico	A1	Publicado
Kidney Tumor Segmentation from Computed Tomography Images using DeepLabv3+ 2.5D Model. Em: Expert Systems with Applications. Ano: 2021.	Periódico	A1	Publicado

Tabela 22 – Produções científicas em outras aplicações de processamento de imagens e visão computacional.

Artigo	Tipos	Qualis	Status
Interferometer Eye Image Classification for Dry Eye Categorization using Phylogenetic Diversity Indexes for Texture Analysis. Em: Computer Methods and Programs in Biomedicine. Ano: 2019.	Periódico	A1	Publicado
Tear Film Classification in Interferometry Eye Images using Phylogenetic Diversity Indexes and Ripley's K Function. Em: IEEE Journal of Biomedical and Health Informatics. Ano: 2020.	Periódico	A1	Publicado
Planejamento Cirúrgico de Estrabismo Horizontal Utilizando Árvore de Regressão de Múltiplas Saídas. Em: Simpósio Brasileiro de Computação Aplicada à Saúde. Ano: 2020.	Simpósio	B3	Publicado
An Automatic Method for Segmentation of Liver Lesions in Computed Tomography Images Using Deep Neural Networks. Em: Expert Systems with Applications. Ano: 2021.	Periódico	A1	Publicado
Surgical Planning of Horizontal Strabismus Using Multiple Output Regression Tree. Em: Computers in Biology and Medicine. Ano: 2021.	Periódico	A1	Publicado
Automatic Method for Classifying COVID-19 Patients Based on Chest X-ray Images, using Deep Features and PSO-optimized XGBoost. Em: Expert Systems with Applications. Ano: 2021.	Periódico	A1	Publicado
Segmentation and Quantification of COVID-19 Infections in CT using Pulmonary Vessels Extraction and Deep Learning. Em: Multimedia Tools and Applications. Ano: 2021.	Periódico	A2	Publicado
Classificação Automática de Glóbulos Brancos usando Desritores de Forma e Textura e eXtreme Gradient Boosting. Em: Simpósio Brasileiro de Computação Aplicada à Saúde. Ano: 2020.	Simpósio	B3	Publicado
Liver Segmentation from Computed Tomography Images using Cascade Deep Learning. Em: Computers in Biology and Medicine. Ano: 2021.	Periódico	A1	Publicado
Heart Segmentation in Planning CT using 2.5D U-Net++ with Attention Gate. Em: Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization. Ano: 2021.	Periódico	A3	Publicado

REFERÊNCIAS

- ADAMS, J. E.; MUGHAL, Z.; DAMILAKIS, J.; OFFIAH, A. C. Chapter 12 - radiology. In: GLORIEUX, F. H.; PETTIFOR, J. M.; JÜPPNER, H. (Ed.). **Pediatric Bone (Second Edition)**. Second edition. San Diego: Academic Press, 2012. p. 277–307. ISBN 978-0-12-382040-2.
- AGNELLO, F.; ALBANO, D.; MICCI, G.; BUONO, G. D.; AGRUSA, A.; SALVAGGIO, G.; PARDO, S.; SPARACIA, G.; BARTOLOTTA, T. V.; MIDIRI, M.; LAGALLA, R.; GALIA, M. Ct and mr imaging of cystic renal lesions. **Insights into imaging**, Springer, v. 11, n. 1, p. 1–11, 2020.
- ALANIS, A. Y.; ARANA-DANIEL, N.; LOPEZ-FRANCO, C. **Artificial neural networks for engineering applications**. [S.l.]: Academic Press, 2019.
- ANDREWS, H. C.; PATTERSON, C. L. Digital interpolation of discrete images. **IEEE Transactions on Computers**, IEEE, v. 100, n. 2, p. 196–202, 1976.
- BAKOŠ, M. Active contours and their utilization at image segmentation. In: **5th Slovakian-Hungarian Joint symposium on applied machine intelligence and informatics, Poprad, Slovakia**. [S.l.: s.n.], 2007. p. 313–317.
- BARONI, M. Linguistic generalization and compositionality in modern artificial neural networks. **Philosophical Transactions of the Royal Society B**, The Royal Society, v. 375, n. 1791, p. 20190307, 2020.
- BERGERON, A. The pulmonologist's point of view on lung infiltrates in haematological malignancies. **Diagnostic and Interventional Imaging**, v. 94, n. 2, p. 216 – 220, 2013. ISSN 2211-5684. Haematological malignancies.
- BERGSTRA, J.; YAMINS, D.; COX, D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In: PMLR. **International conference on machine learning**. [S.l.], 2013. p. 115–123.
- BERSIMIS, F. G.; VARLAMIS, I. Chapter 2 - use of health-related indices and classification methods in medical data. In: DEY, N. (Ed.). **Classification Techniques for Medical Image Analysis and Computer Aided Diagnosis**. [S.l.]: Academic Press, 2019, (Advances in ubiquitous sensing applications for healthcare). p. 31–66.
- BLAND, M. **An introduction to medical statistics**. [S.l.]: Oxford University Press (UK), 2015. 446 p.
- BOAS, F. E.; FLEISCHMANN, D. Ct artifacts: causes and reduction techniques. **Imaging Med**, v. 4, n. 2, p. 229–240, 2012.
- BOHRA, Y. **The Challenge of Vanishing/Exploding Gradients in Deep Neural Networks**. 2021. Disponível em: <<https://www.analyticsvidhya.com/blog/2021/06/the-challenge-of-vanishing-exploding-gradients-in-deep-neural-networks/>>. Acesso em: 03 de março de 2019.
- BRAY, F.; FERLAY, J.; SOERJOMATARAM, I.; SIEGEL, R. L.; TORRE, L. A.; JEMAL, A. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: a cancer journal for clinicians**, Wiley Online Library, v. 68, n. 6, p. 394–424, 2018.

BRAZ, G. J. **Detecção de regiões de massas em mamografias usando índices de diversidade, geoestatística e geometria côncava.** Tese (Doutorado) — Universidade Federal do Maranhão, Programa de Pós-Graduação em Engenharia de Eletricidade. São Luis - MA, 2014.

BROWNLEE, J. **Better deep learning: train faster, reduce overfitting, and make better predictions.** [S.l.]: Machine Learning Mastery, 2018.

_____. **How to Avoid Exploding Gradients With Gradient Clipping.** 2019. Disponível em: <<https://machinelearningmastery.com/how-to-avoid-exploding-gradients-in-neural-networks-with-gradient-clipping/>>. Acesso em: 03 de março de 2019.

BURMAN, P. A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. **Biometrika**, Oxford University Press, v. 76, n. 3, p. 503–514, 1989.

BUZUG, T. M. Computed tomography. In: _____. **Springer Handbook of Medical Technology.** Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 311–342. ISBN 978-3-540-74658-4.

CAMARGO, L. A.-s. M. A.; SILVA, R. P. M.; MENEGUETTI, D. U. d. O. Research methodology topics: Cohort studies or prospective and retrospective cohort studies. **Journal of Human Growth and Development**, scieloepsic, v. 29, p. 433 – 436, 12 2019. ISSN 0104-1282. Disponível em: <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0104-12822019000300016&nrm=iso>.

CANDEMIR, S.; JAEGER, S.; PALANIAPPAN, K.; MUSCO, J. P.; SINGH, R. K.; XUE, Z.; KARARGYRIS, A.; ANTANI, S.; THOMA, G.; MCDONALD, C. J. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. **IEEE Transactions on Medical Imaging**, v. 33, n. 2, p. 577–590, Feb 2014. ISSN 1558-254X.

CARE, U. **What is a Renal Mass and What is a Localized Renal Tumor?** 2021. Disponível em: <<https://www.urologyhealth.org/urology-a-z/r/renal-mass-and-localized-renal-tumors>>. Acesso em: 28 de abril de 2021.

CHEN, L.-C.; ZHU, Y.; PAPANDREOU, G.; SCHROFF, F.; ADAM, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: **Proceedings of the European conference on computer vision (ECCV).** [S.l.: s.n.], 2018. p. 801–818.

CHEN, Y.; LI, J.; XIAO, H.; JIN, X.; YAN, S.; FENG, J. Dual path networks. **Advances in neural information processing systems**, v. 30, p. 4467–4475, 2017.

CHENG, J.-Z.; NI, D.; CHOU, Y.-H.; QIN, J.; TIU, C.-M.; CHANG, Y.-C.; HUANG, C.-S.; SHEN, D.; CHEN, C.-M. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in us images and pulmonary nodules in ct scans. **Scientific reports**, Nature Publishing Group, v. 6, n. 1, p. 1–13, 2016.

CHOLLET, F. Xception: Deep learning with depthwise separable convolutions. In: **Proceedings of the IEEE conference on computer vision and pattern recognition.** [S.l.: s.n.], 2017. p. 1251–1258.

CHRIST, P. F.; ETTLINGER, F.; GRÜN, F.; ELSHAERA, M. E. A.; LIPKOVA, J.; SCHLECHT, S.; AHMADDY, F.; TATAVARTY, S.; BICKEL, M.; BILIC, P.; REMPFLER, M.; HOFMANN, F.; ANASTASI, M. D.; AHMADI, S.-A.; KAISSIS, G.; HOLCH, J.; SOMMER, W.; BRAREN, R.; HEINEMANN, V.; MENZE, B. Automatic liver and tumor segmentation of ct and mri volumes using cascaded fully convolutional neural networks. **arXiv preprint arXiv:1702.05970**, 2017.

COX, D. Statistical significance tests. **British Journal of Clinical Pharmacology**, v. 14, n. 3, p. 325–331, 1982.

CRUZ, L. B. da; ARAÚJO, J. D. L.; FERREIRA, J. L.; DINIZ, J. O. B.; SILVA, A. C.; ALMEIDA, J. D. S. de; PAIVA, A. C. de; GATTASS, M. Kidney segmentation from computed tomography images using deep neural network. **Computers in Biology and Medicine**, Elsevier, p. 103906, 2020. ISSN 0010-4825.

DALLAL, A. H.; AGARWAL, C.; ARBABSHIRANI, M. R.; PATEL, A.; MOORE, G. Automatic estimation of heart boundaries and cardiothoracic ratio from chest x-ray images. In: III, S. G. A.; PETRICK, N. A. (Ed.). **Medical Imaging 2017: Computer-Aided Diagnosis**. [S.l.]: SPIE, 2017. v. 10134, p. 134 – 143.

DENG, J.; DONG, W.; SOCHER, R.; LI, L.-J.; LI, K.; FEI-FEI, L. Imagenet: A large-scale hierarchical image database. In: IEEE. **2009 IEEE conference on computer vision and pattern recognition**. [S.l.], 2009. p. 248–255.

DICE, L. R. Measures of the amount of ecologic association between species. **Ecology**, Ecological Society of America, v. 26, n. 3, p. 297–302, 1945. ISSN 00129658, 19399170.

DINIZ, J. O. B. **Métodos para segmentação de medula espinhal e esôfago em tomografia computadorizada de planejamento à radioterapia**. Tese (Doutorado) — Universidade Federal do Maranhão, Programa de Pós-Graduação em Engenharia Elétrica. São Luís - MA, 2021.

DINIZ, J. O. B.; DINIZ, P. H. B.; VALENTE, T. L. A.; SILVA, A. C.; PAIVA, A. C. Spinal cord detection in planning ct for radiotherapy through adaptive template matching, imslic and convolutional neural networks. **Computer methods and programs in biomedicine**, Elsevier, v. 170, p. 53–67, 2019.

DODGSON, N. A. **Image resampling**. [S.l.], 1992.

DREISIN, A.; MATRANA, M. R. Treating renal cell carcinoma in young adults: challenges and solutions. **Clinical Oncology in Adolescents and Young Adults**, Dove Press, v. 6, p. 31–38, 2016.

EL-DAHSHAN, E.-S.; MOHSEN, H.; REVETT, K.; SALEM, A.-B. Computer-aided diagnosis of human brain tumor through mri: A survey and a new algorithm. **Expert Systems with Applications**, v. 41, n. 11, p. 5526–5545, 2014.

FERLAY, J.; SOERJOMATARAM, I.; DIKSHIT, R.; ESER, S.; MATHERS, C.; REBELO, M.; PARKIN, D. M.; FORMAN, D.; BRAY, F. Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2012. **International journal of cancer**, Wiley Online Library, v. 136, n. 5, p. E359–E386, 2015.

FERNANDES, B. J. T. **Redes neurais com extração implícita de características para reconhecimento de padrões visuais.** Tese (Doutorado) — Universidade Federal de Pernambuco, Recife, 2013.

FUND, W. C. R. **Kidney Cancer Statistics.** 2019. Disponível em: <<https://www.wcrf.org/dietandcancer/cancer-trends/kidney-cancer-statistics>>. Acesso em: 27 de nov. de 2020.

GAUVREAU, K.; PAGANO, M. **WHY 5-PERCENT.** [S.l.]: ELSEVIER SCIENCE INC 655 AVENUE OF THE AMERICAS, NEW YORK, NY 10010, 1994. 93–94 p.

GEETHANJALI, T. M.; MINAVATHI; DINESH, M. S. Semantic segmentation of tumors in kidneys using attention u-net models. In: **2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT).** [S.l.: s.n.], 2021. p. 286–290.

GHAFOORIAN, M.; MEHRTASH, A.; KAPUR, T.; KARSSEMEIJER, N.; MARCHIORI, E.; PESTEIE, M.; GUTTMANN, C. R.; LEEUW, F.-E. de; TEMPANY, C. M.; GINNEKEN, B. V. et al. Transfer learning for domain adaptation in mri: Application in brain lesion segmentation. In: SPRINGER INTERNATIONAL PUBLISHING. **International conference on medical image computing and computer-assisted intervention.** [S.l.], 2017. p. 516–524.

GHOSN, M.; EID, R.; HAMADA, E.; AZIM, H. A.; ZEKRI, J.; AL-MANSOUR, M.; JALOUDI, M.; NASR, F.; ERRIHANI, H.; BOUNEDJAR, A.; MEZLINI, A.; BOUSSEN, H.; KATTAN, J.; KARAK, F. E.; FARHAT, F. a. Ossmar: An observational study to describe the use of sunitinib in real-life practice for the treatment of metastatic renal cell carcinoma. **Journal of Global Oncology**, American Society of Clinical Oncology, n. 5, p. 1–10, 2019. PMID: 31657977.

GOCERI, E. Automatic kidney segmentation using gaussian mixture model on mri sequences. In: **Electrical Power Systems and Computers.** [S.l.]: Springer, 2011. p. 23–29.

_____. Analysis of deep networks with residual blocks and different activation functions: Classification of skin diseases. In: **2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA).** [S.l.: s.n.], 2019. p. 1–6.

_____. Deep learning based classification of facial dermatological disorders. **Computers in Biology and Medicine**, v. 128, p. 104118, 2021. ISSN 0010-4825. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0010482520304492>>.

GOCERI, N.; GOCERI, E. A neural network based kidney segmentation from mr images. In: **2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA).** [S.l.: s.n.], 2015. p. 1195–1198.

GONZALEZ, R.; WOODS, R. **Digital image processing.** [S.l.]: Pearson, Prentice Hall, 2008. 954 p.

GUO, Y.; LIU, Y.; OERLEMANS, A.; LAO, S.; WU, S.; LEW, M. S. Deep learning for visual understanding: A review. **Neurocomputing**, v. 187, p. 27 – 48, 2016. ISSN 0925-2312. Recent Developments on Deep Big Vision.

- GWYNNE, S.; WEBSTER, R.; ADAMS, R.; MUKHERJEE, S.; COLES, B.; STAFFURTH, J. Image-guided radiotherapy for rectal cancer — a systematic review. **Clinical Oncology**, v. 24, n. 4, p. 250 – 260, 2012. ISSN 0936-6555.
- _____. Image-guided radiotherapy for rectal cancer—a systematic review. **Clinical Oncology**, Elsevier, v. 24, n. 4, p. 250–260, 2012.
- HAFEMANN, L. G. **An analysis of deep neural networks for texture classification**. Dissertação (Mestrado) — Universidade Federal do Paraná, Curitiba - PR, 2014.
- HAMERLY, G.; ELKAN, C. Learning the k in k-means. **Advances in neural information processing systems**, v. 16, 2003.
- HAYKIN, S. **Redes neurais: princípios e prática**. [S.l.]: Bookman Editora, 2007.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. 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.
- HEALTH, U. **What is kidney cancer?** 2021. Disponível em: <<https://www.urologyhealth.org/urology-a-z/k/kidney-cancer>>. Acesso em: 11 de maio de 2021.
- HELLER, N.; ISENSEE, F.; MAIER-HEIN, K. H.; HOU, X.; XIE, C.; LI, F.; NAN, Y.; MU, G.; LIN, Z.; HAN, M.; YAO, G.; GAO, Y.; ZHANG, Y.; WANG, Y.; HOU, F.; YANG, J.; XIONG, G.; TIAN, J.; ZHONG, C.; MA, J.; RICKMAN, J.; DEAN, J.; STAI, B.; TEJPAUL, R.; OESTREICH, M.; BLAKE, P.; KALUZNIAK, H.; RAZA, S.; ROSENBERG, J.; MOORE, K.; WALCZAK, E.; RENGEL, Z.; EDGERTON, Z.; VASDEV, R.; PETERSON, M.; MCSWEENEY, S.; PETERSON, S.; KALAPARA, A.; SATHIANATHEN, N.; PAPANIKOLOPOULOS, N.; WEIGHT, C. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced ct imaging: Results of the kits19 challenge. **Medical Image Analysis**, v. 67, p. 101821, 2021. ISSN 1361-8415.
- HELLER, N.; SATHIANATHEN, N.; KALAPARA, A.; WALCZAK, E.; MOORE, K.; KALUZNIAK, H.; ROSENBERG, J.; BLAKE, P.; RENGEL, Z.; OESTREICH, M.; DEAN, J.; TRADEWELL, M.; SHAH, A.; TEJPAUL, R.; EDGERTON, Z.; PETERSON, M.; RAZA, S.; REGMI, S.; PAPANIKOLOPOULOS, N.; WEIGHT, C. **The KiTS19 Challenge Data: 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes**. 2019.
- HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural networks**, Elsevier, v. 2, n. 5, p. 359–366, 1989.
- HUA, K.-L.; HSU, C.-H.; HIDAYATI, S. C.; CHENG, W.-H.; CHEN, Y.-J. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. **OncoTargets and therapy**, Dove Press, v. 8, 2015.
- HUANG, G.; LIU, Z.; MAATEN, L. V. D.; WEINBERGER, K. Q. Densely connected convolutional networks. In: **2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2017. p. 2261–2269.
- HUBEL, D. H.; WIESEL, T. N. Early exploration of the visual cortex. **Neuron**, Elsevier, v. 20, n. 3, p. 401–412, 1998.

INITIATIVE, N. I. T. **NIfTI**. 2000. Disponível em: <<https://nifti.nimh.nih.gov/>>. Acesso em: 27 de nov. de 2020.

JACKSON, P.; HARDCASTLE, N.; DAWE, N.; KRON, T.; HOFMAN, M. S.; HICKS, R. J. Deep learning renal segmentation for fully automated radiation dose estimation in unsealed source therapy. **Frontiers in Oncology**, v. 8, p. 215, 2018. ISSN 2234-943X.

JOHNSON, J. M.; KHOSHGOFTAAR, T. M. Survey on deep learning with class imbalance. **Journal of Big Data**, Springer, v. 6, n. 1, p. 1–54, 2019.

JONES, O. **The Kidneys**. 2019. Disponível em: <<https://teachmeanatomy.info/abdomen/viscera/kidney/>>. Acessado 03 Maio 2021.

JOSHI, K. D.; NALWADE, P. Modified k-means for better initial cluster centres. **International Journal of Computer Science and Mobile Computing**, v. 2, n. 7, p. 219–223, 2013.

KANG, L.; ZHOU, Z.; HUANG, J.; HAN, W. Renal tumors segmentation in abdomen ct images using 3d-cnn and convlstm. **Biomedical Signal Processing and Control**, v. 72, p. 103334, 2022. ISSN 1746-8094. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1746809421009319>>.

KAUR, R.; JUNEJA, M. A survey of different imaging modalities for renal cancer. **Indian Journal of Science and Technology**, v. 9, n. 44, p. 1–6, 2016.

KAUR, R.; JUNEJA, M.; MANDAL, A. K. A hybrid edge-based technique for segmentation of renal lesions in ct images. **Multimedia Tools and Applications**, Springer, v. 78, n. 10, p. 12917–12937, 2019.

KHALIFA, F.; GIMEL'FARB, G.; EL-GHAR, M. A.; SOKHADZE, G.; MANNING, S.; MCCLURE, P.; OUSEPH, R.; EL-BAZ, A. A new deformable model-based segmentation approach for accurate extraction of the kidney from abdominal ct images. In: **2011 18th IEEE International Conference on Image Processing**. [S.l.: s.n.], 2011. p. 3393–3396. ISSN 1522-4880.

KHALIFA, F.; SOLIMAN, A.; TAKIELDEEN, A.; SHEHATA, M.; MOSTAPHA, M.; SHAFFIE, A.; OUSEPH, R.; ELMAGHRABY, A.; EL-BAZ, A. Kidney segmentation from ct images using a 3d nmf-guided active contour model. In: **2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)**. [S.l.: s.n.], 2016. p. 432–435. ISSN 1945-8452.

KIMBLE, C.; MIOLIDAKIS, G. Big data and business intelligence: Debunking the myths. **Global Business and Organizational Excellence**, Wiley Online Library, v. 35, n. 1, p. 23–34, 2015.

KINGMA, D.; BA, J. Adam: A method for stochastic optimization. **International Conference on Learning Representations**, 12 2014.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Advances in neural information processing systems**, v. 25, p. 1097–1105, 2012.

KUKAČKA, J.; GOLKOV, V.; CREMERS, D. Regularization for deep learning: A taxonomy. **arXiv preprint arXiv:1710.10686**, 2017.

- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, p. 436–44, 05 2015.
- LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, v. 86, n. 11, p. 2278–2324, 1998.
- LECUN, Y.; KAVUKCUOGLU, K.; FARABET, C. Convolutional networks and applications in vision. In: **Proceedings of 2010 IEEE International Symposium on Circuits and Systems**. [S.l.: s.n.], 2010. p. 253–256.
- LES, T.; MARKIEWCZ, T.; DZIEKIEWICZ, M.; LORENT, M. Kidney segmentation from computed tomography images using u-net and batch-based synthesis. In: **2021 International Joint Conference on Neural Networks (IJCNN)**. [S.l.: s.n.], 2021. p. 1–8.
- LI, A.; MUELLER, K.; ERNST, T. Methods for efficient, high quality volume resampling in the frequency domain. In: **IEEE Visualization 2004**. [S.l.: s.n.], 2004. p. 3–10.
- LI, H.; XIONG, P.; AN, J.; WANG, L. **Pyramid Attention Network for Semantic Segmentation**. 2018.
- LIN, Z.; CUI, Y.; LIU, J.; SUN, Z.; MA, S.; ZHANG, X.; WANG, X.-y. Automated segmentation of kidney and renal mass and automated detection of renal mass in ct urography using 3d u-net-based deep convolutional neural network. **European Radiology**, v. 31, 01 2021.
- LIU, W.; WANG, Z.; LIU, X.; ZENG, N.; LIU, Y.; ALSAADI, F. E. A survey of deep neural network architectures and their applications. **Neurocomputing**, v. 234, p. 11 – 26, 2017. ISSN 0925-2312.
- LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2015. p. 3431–3440.
- MAAS, A. L.; HANNUN, A. Y.; NG, A. Y. Rectifier nonlinearities improve neural network acoustic models. In: CITESEER. **Proceedings of the 30th International Conference on Machine Learning**. [S.l.], 2013. v. 30, n. 1, p. 3.
- MACQUEEN, J. et al. Some methods for classification and analysis of multivariate observations. In: OAKLAND, CA, USA. **Proceedings of the fifth Berkeley symposium on mathematical statistics and probability**. [S.l.], 1967. v. 1, n. 14, p. 281–297.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, n. 4, p. 115–133, 1943.
- MEHTA, P.; SANDFORT, V.; GHEYSENS, D.; BRAECKEVELT, G.; BERTE, J.; SUMMERS, R. M. Segmenting the kidney on ct scans via crowdsourcing. In: **2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)**. [S.l.: s.n.], 2019. p. 829–832. ISSN 1945-7928.

- MIKOŁAJCZYK, A.; GROCHOWSKI, M. Data augmentation for improving deep learning in image classification problem. In: IEEE. **2018 international interdisciplinary PhD workshop (IIPhDW)**. [S.l.], 2018. p. 117–122.
- MILLETARI, F.; NAVAB, N.; AHMADI, S.-A. **V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation**. arXiv, 2016. Disponível em: <<https://arxiv.org/abs/1606.04797>>.
- MUGLIA, V. F.; PRANDO, A. Renal cell carcinoma: histological classification and correlation with imaging findings. **Radiologia brasileira**, SciELO Brasil, v. 48, n. 3, p. 166–174, 2015.
- MUKHOPADHYAY, S. 13 - artificial neural network applications in textile composites. In: MAJUMDAR, A. (Ed.). **Soft Computing in Textile Engineering**. [S.l.]: Woodhead Publishing, 2011, (Woodhead Publishing Series in Textiles). p. 329–349. ISBN 978-1-84569-663-4.
- NEWMAN, T. **What do the kidneys do?** 2019. Disponível em: <<https://www.medicalnewstoday.com/articles/305488>>. Acesso em: 03 de maio de 2021.
- PADILLA, P.; LOPEZ, M.; GORRIZ, J. M.; RAMIREZ, J.; SALAS-GONZALEZ, D.; ALVAREZ, I. Nmf-svm based cad tool applied to functional brain images for the diagnosis of alzheimer's disease. **IEEE Transactions on Medical Imaging**, v. 31, n. 2, p. 207–216, Feb 2012. ISSN 1558-254X.
- PASZKE, A.; GROSS, S.; MASSA, F.; LERER, A.; BRADBURY, J.; CHANAN, G.; KILLEEN, T.; LIN, Z.; GIMELSHEIN, N.; ANTIGA, L.; DESMAISON, A.; KOPF, A.; YANG, E.; DEVITO, Z.; RAISON, M.; TEJANI, A.; CHILAMKURTHY, S.; STEINER, B.; FANG, L.; BAI, J.; CHINTALA, S. Pytorch: An imperative style, high-performance deep learning library. In: WALLACH, H.; LAROCHELLE, H.; BEYGELZIMER, A.; ALCHÉ-BUC, F. d'; FOX, E.; GARNETT, R. (Ed.). **Advances in Neural Information Processing Systems 32**. [S.l.]: Curran Associates, Inc., 2019. p. 8024–8035.
- PEDRINI, H.; SCHWARTZ, W. R. **Análise de imagens digitais: princípios, algoritmos e aplicações**. [S.l.]: Thomson Learning, 2008.
- PHAM, B. T.; NGUYEN, M. D.; BUI, K.-T. T.; PRAKASH, I.; CHAPI, K.; BUI, D. T. A novel artificial intelligence approach based on multi-layer perceptron neural network and biogeography-based optimization for predicting coefficient of consolidation of soil. **CATENA**, v. 173, p. 302–311, 2019. ISSN 0341-8162.
- PREPROCESSING, H.-O. A. P. . I. D.; AUGMENTATION. **A Tutorial Series for Software Developers, Data Scientists, and Data Center Managers**. 2017. Disponível em: <<https://software.intel.com/en-us/articles/hands-on-ai-part-14-image-data-preprocessing-and-augmentation>>. Acesso em: 03 de março de 2019.
- QAYYUM, A.; LALANDE, A.; MERIAUDEAU, F. Automatic segmentation of tumors and affected organs in the abdomen using a 3d hybrid model for computed tomography imaging. **Computers in Biology and Medicine**, v. 127, p. 104097, 2020. ISSN 0010-4825.

- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. **International Conference on Medical image computing and computer-assisted intervention**. [S.l.], 2015. v. 9351, p. 234–241. ISBN 978-3-319-24573-7.
- ROSSI, S. H.; KLATTE, T.; USHER-SMITH, J.; STEWART, G. D. Epidemiology and screening for renal cancer. **World Journal of Urology**, Springer, v. 36, n. 9, p. 1341–1353, Sep 2018. ISSN 1433-8726.
- RUAN, Y.; LI, D.; MARSHALL, H.; MIAO, T.; COSSETTO, T.; CHAN, I.; DAHER, O.; ACCORSI, F.; GOELA, A.; LI, S. Mb-fsgan: Joint segmentation and quantification of kidney tumor on ct by the multi-branch feature sharing generative adversarial network. **Medical Image Analysis**, v. 64, p. 101721, 2020. ISSN 1361-8415. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1361841520300852>>.
- SACHS, J. Image resampling. **Digital Light & Color**, v. 14, 2001.
- SANTOS, V. S. dos. **Rins**. 2019. Disponível em: <<https://brasilescola.uol.com.br/biologia/os-rins.htm>>. Acesso em: 03 de maio de 2021.
- SELVANAYAKI, K.; KARNAN, M. Cad system for automatic detection of brain tumor through magnetic resonance image-a review. **International Journal of Engineering Science and Technology**, v. 2, n. 10, p. 5890–5901, 2010.
- SENGUPTA, S.; BASAK, S.; SAIKIA, P.; PAUL, S.; TSALAVOUTIS, V.; ATIAH, F.; RAVI, V.; PETERS, A. A review of deep learning with special emphasis on architectures, applications and recent trends. **Knowledge-Based Systems**, p. 105596, 2020. ISSN 0950-7051.
- SEYFRIED, T. N.; HUYSENTRUYT, L. C. On the origin of cancer metastasis. **Critical reviews in oncogenesis**, NIH Public Access, v. 18, n. 1-2, p. 43–73, 2013.
- SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. **Journal of Big Data**, Springer, v. 6, n. 1, p. 60, 2019.
- SHUCH, B.; AMIN, A.; ARMSTRONG, A. J.; EBLE, J. N.; FICARRA, V.; LOPEZ-BELTRAN, A.; MARTIGNONI, G.; RINI, B. I.; KUTIKOV, A. Understanding pathologic variants of renal cell carcinoma: Distilling therapeutic opportunities from biologic complexity. **European Urology**, v. 67, n. 1, p. 85 – 97, 2015. ISSN 0302-2838.
- SILVA, A. C. **Algoritmos para Diagnóstico Assistido de nódulos Pulmonares Solitários em Imagens de Tomografia Computadorizada**. Tese (Doutorado) — Pontifícia Universidade Católica do Rio de Janeiro, Brasil, 2004.
- SILVA, I. d.; SPATTI, D. H.; FLAUZINO, R. A. Redes neurais artificiais para engenharia e ciências aplicadas. **São Paulo: Artliber**, v. 23, n. 5, p. 33–111, 2010.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014.
- SKALSKI, A.; HERYAN, K.; JAKUBOWSKI, J.; DREWNIAK, T. Kidney segmentation in ct data using hybrid level-set method with ellipsoidal shape constraints. **Metrology and Measurement Systems**, Polish Academy of Sciences Committee on Metrology and Scientific Instrumentation, v. 24, n. 1, p. 101–112, 2017.

SOCIETY, A. C. **What is kidney cancer?** 2017. Disponível em: <<https://www.cancer.org/cancer/kidney-cancer/about/what-is-kidney-cancer>>. Acesso em: 27 de nov. de 2020.

SOUZA, J. C.; DINIZ, J. O. B.; FERREIRA, J. L.; SILVA, G. L. F. da; SILVA, A. C.; PAIVA, A. C. de. An automatic method for lung segmentation and reconstruction in chest x-ray using deep neural networks. **Computer Methods and Programs in Biomedicine**, v. 177, p. 285 – 296, 2019. ISSN 0169-2607.

SUN, D.; JONASCH, E.; LARA, P. N. Genetic heterogeneity of kidney cancer. In: _____. **Kidney Cancer: Principles and Practice**. Cham: Springer International Publishing, 2015. p. 83–88. ISBN 978-3-319-17903-2.

TAHA, A. A.; HANBURY, A. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. **BMC medical imaging**, BioMed Central, v. 15, n. 1, p. 1–28, 2015.

TANIMOTO, R.; HIGUCHI, K.; ISHIGURO, T.; MORI, K.; KOJO, K.; KOJIMA, T.; KAKEYA, H. Segmentation of renal tumors in ct images by 3d u-net preserving rotational symmetry in axial slices. **Opt. Continuum**, OSA, v. 1, n. 2, p. 297–305, Feb 2022.

TULLY, E. **CT Scanners, Diagnostic Imaging Equipment**. 2016. Disponível em: <<https://www.oncologysystems.com/blog/everybody-know-buying-used-ge-ct-scanner>>. Acesso em: 10 de jun. de 2021.

TÜRK, F.; LÜY, M.; BARİŞÇİ, N. Kidney and renal tumor segmentation using a hybrid v-net-based model. **Mathematics**, Multidisciplinary Digital Publishing Institute, v. 8, n. 10, p. 1772, 2020.

TüRK, F.; LüY, M.; YALCINKAYA, F. Kidney tumor segmentation using two-stage bottleneck block architecture. **Intelligent Automation & Soft Computing**, v. 33, p. 349–363, 01 2022.

TÜRKOĞLU, M.; HANBAY, D. Combination of deep features and knn algorithm for classification of leaf-based plant species. In: **2019 International Artificial Intelligence and Data Processing Symposium (IDAP)**. [S.l.: s.n.], 2019. p. 1–5.

UK, K. C. **Risk Factors for Kidney Cancer**. 2019. Disponível em: <<https://www.kcuk.org.uk/kidneycancer/what-is-kidney-cancer/causes-of-kidney-cancer/>>. Acesso em: 11 de maio de 2021.

WARREN, A. Y.; HARRISON, D. Who/isup classification, grading and pathological staging of renal cell carcinoma: standards and controversies. **World journal of urology**, Springer, v. 36, n. 12, p. 1913–1926, 2018.

WOLZ, R.; CHU, C.; MISAWA, K.; FUJIWARA, M.; MORI, K.; RUECKERT, D. Automated abdominal multi-organ segmentation with subject-specific atlas generation. **IEEE Transactions on Medical Imaging**, v. 32, n. 9, p. 1723–1730, Sep. 2013. ISSN 1558-254X.

WU, X.; SAHOO, D.; HOI, S. C. Recent advances in deep learning for object detection. **Neurocomputing**, 2020. ISSN 0925-2312. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0925231220301430>>.

- XIE, S.; GIRSHICK, R.; DOLLÁR, P.; TU, Z.; HE, K. Aggregated residual transformations for deep neural networks. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2017. p. 1492–1500.
- XIE, X.; LI, L.; LIAN, S.; CHEN, S.; LUO, Z. Seru: A cascaded se-resnext u-net for kidney and tumor segmentation. **Concurrency and Computation: Practice and Experience**, v. 32, n. 14, p. e5738, 2020. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/cpe.5738>>.
- XU, W.; ATKINS, M. B.; MCDERMOTT, D. F. Checkpoint inhibitor immunotherapy in kidney cancer. **Nature Reviews Urology**, Nature Publishing Group, p. 1–14, 2020.
- YAN, X.; YUAN, K.; ZHAO, W.; WANG, S.; LI, Z.; CUI, S. An efficient hybrid model for kidney tumor segmentation in ct images. In: **2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)**. [S.l.: s.n.], 2020. p. 333–336.
- YANG, E.; KIM, C. K.; GUAN, Y.; KOO, B.-B.; KIM, J.-H. 3d multi-scale residual fully convolutional neural network for segmentation of extremely large-sized kidney tumor. **Computer Methods and Programs in Biomedicine**, v. 215, p. 106616, 2022. ISSN 0169-2607. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169260722000013>>.
- YANG, G.; GU, J.; CHEN, Y.; LIU, W.; TANG, L.; SHU, H.; TOUMOULIN, C. Automatic kidney segmentation in ct images based on multi-atlas image registration. In: **2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society**. [S.l.: s.n.], 2014. p. 5538–5541. ISSN 1558-4615.
- YANG, G.; LI, G.; PAN, T.; KONG, Y.; WU, J.; SHU, H.; LUO, L.; DILLENSEGER, J.; COATRIEUX, J.; TANG, L.; ZHU, X. Automatic segmentation of kidney and renal tumor in ct images based on 3d fully convolutional neural network with pyramid pooling module. In: **2018 24th International Conference on Pattern Recognition (ICPR)**. [S.l.: s.n.], 2018. p. 3790–3795. ISSN 1051-4651.
- YANG, G.; WANG, C.; YANG, J.; CHEN, Y.; TANG, L.; SHAO, P.; DILLENSEGER, J.-L.; SHU, H.; LUO, L. Weakly-supervised convolutional neural networks of renal tumor segmentation in abdominal cta images. **BMC Medical Imaging**, Springer, v. 20, p. 1–12, 2020.
- ZHANG, K.; LIU, Q.; SONG, H.; LI, X. A variational approach to simultaneous image segmentation and bias correction. **IEEE Transactions on Cybernetics**, v. 45, n. 8, p. 1426–1437, 2015.
- ZHANG, Z.; LIU, Q.; WANG, Y. Road extraction by deep residual u-net. **IEEE Geoscience and Remote Sensing Letters**, v. 15, n. 5, p. 749–753, 2018.
- ZHAO, E.; LIANG, Y.; FAN, H. Contextual information-aided kidney segmentation in ct sequences. **Optics Communications**, v. 290, p. 55–62, 03 2013. ISSN 0030-4018.
- ZHAO, W.; JIANG, D.; QUERALTA, J. P.; WESTERLUND, T. Mss u-net: 3d segmentation of kidneys and tumors from ct images with a multi-scale supervised u-net. **Informatics in Medicine Unlocked**, Elsevier, v. 19, p. 100357, 2020.

ZNAOR, A.; LORTET-TIEULENT, J.; LAVERSANNE, M.; JEMAL, A.; BRAY, F. International variations and trends in renal cell carcinoma incidence and mortality. **European urology**, Elsevier, v. 67, n. 3, p. 519–530, 2015.

ZOU, K. H.; FIELDING, J. R.; SILVERMAN, S. G.; TEMPANY, C. M. C. Hypothesis testing i: Proportions. **Radiology**, v. 226, n. 3, p. 609–613, 2003.