

CENTRO UNIVERSITÁRIO FEI
MURILLO FREITAS BOUZON

**PLANEJAMENTO SEMI-AUTOMÁTICO DE TRAJETÓRIA DE URETEROSCOPIA
BASEADO NA CLASSIFICAÇÃO E SEGMENTAÇÃO DE CÁLCULOS URINÁRIOS
POR INTELIGÊNCIA ARTIFICIAL**

São Bernardo do Campo

2026

MURILLO FREITAS BOUZON

**PLANEJAMENTO SEMI-AUTOMÁTICO DE TRAJETÓRIA DE URETEROSCOPIA
BASEADO NA CLASSIFICAÇÃO E SEGMENTAÇÃO DE CÁLCULOS URINÁRIOS
POR INTELIGÊNCIA ARTIFICIAL**

Tese de Doutorado apresentada ao Centro Universitário FEI para obtenção do título de Doutor em Engenharia Elétrica. Orientado pelo Prof. Dr. Carlos Eduardo Thomaz.

São Bernardo do Campo

2026

Freitas Bouzon, Murillo.

Planejamento Semi-Automático de Trajetória de Ureteroscopia Baseado na Classificação e Segmentação de Cálculos Urinários por Inteligência Artificial / Murillo Freitas Bouzon. São Bernardo do Campo, 2025.

119 p. : il.

Tese - Centro Universitário FEI.

Orientador: Prof. Dr. Carlos Eduardo Thomaz.

1. Planejamento de trajetória de Ureteroscopia. 2. Inteligência artificial. 3. Detecção de cálculos urinários. 4. Segmentação de imagens médicas. 5. Tomografia Computadorizada. I. Eduardo Thomaz, Carlos, orient. II. Título.

Elaborada pelo sistema de geração automática de ficha catalográfica da FEI com os dados fornecidos pelo(a) autor(a).



APRESENTAÇÃO DE TESE
ATA DA BANCA EXAMINADORA
Programa de Pós-Graduação Stricto Sensu em Engenharia Elétrica

Doutorado

PGE-10

Aluno(a): Murillo Freitas Bouzon

Matrícula: 521.104-0

Título do Trabalho: PLANEJAMENTO SEMI-AUTOMÁTICO DE TRAJETÓRIA DE URETEROSCOPIA BASEADO NA CLASSIFICAÇÃO E SEGMENTAÇÃO DE CÁLCULOS URINÁRIOS POR INTELIGÊNCIA ARTIFICIAL

Área de Concentração: Processamento de Sinais e Imagens

Orientador(a): Prof. Dr. Carlos Eduardo Thomaz

Data da realização da defesa: 28/11/2025

ORIGINAL ASSINADA

Avaliação da Banca Examinadora:

O aluno realizou sua apresentação em 45 minutos e, posteriormente, foi questionado pelos 4 examinadores individualmente, em sessão virtual. A banca considerou a tese com contribuições experimentais relevantes, mas com a necessidade de aperfeiçoamentos na redação científica relacionados à integração entre as 3 metodologias propostas (classificação, segmentação e planejamento de trajetória). Há também a necessidade de rever a Introdução e a Conclusão neste contexto integrador. A banca parabeniza o aluno pelo trabalho e pelo esforço empenhado, aprovando-o, ao final, por unanimidade.

A Banca Julgadora abaixo-assinada atribuiu ao aluno o seguinte resultado:

APROVADO ☒

REPROVADO ☐

MEMBROS DA BANCA EXAMINADORA

Prof. Dr. Carlos Eduardo Thomaz

Prof. Dr. Leonardo Anjoletto Ferreira

Prof^a Dr^a Maria Claudia Ferrari de Castro

Prof. Dr. Leandro Augusto da Silva

Prof. Dr. Luiz Antônio Pereira Neves

Aprovação do Coordenador do Programa de Pós-graduação

Prof. Dr. Carlos Eduardo Thomaz

Dedico esta tese ao Prof. Dr. Paulo Sérgio Silva Rodrigues, meu orientador inicial e mentor em toda a minha jornada acadêmica, iniciada há 11 anos. Não teria chegado até aqui sem o seu apoio. Muito obrigado por todos esses anos de colaboração e inspiração.

AGRADECIMENTOS

A Deus, por ser meu refúgio e amparo em todos os momentos desta caminhada. Pela força, fé e sabedoria que me concedeu, e por me agraciar com o dom da inteligência e da memória. Por ter colocado tantas pessoas generosas em minha vida que me ajudaram tanto durante este período. E, acima de tudo, pelo Amor que me moveu até aqui. Sem o Amor, nenhuma palavra desta tese teria sido escrita. Aos meus pais e à minha avó, pelo amor incondicional, apoio constante e pela compreensão diante das renúncias necessárias ao longo desta jornada. Ao restante de minha família, por compreender minhas faltas e celebrar comigo cada conquista, sendo o pilar que me sustentou em todos os momentos.

Ao Prof. Dr. Paulo Sérgio Silva Rodrigues, meu primeiro orientador e verdadeiro mentor, que me acompanhou desde o início da minha trajetória acadêmica, há mais de uma década. Agradeço por todos os ensinamentos, pela paciência, por ter me impulsionado a superar meus limites todos os dias e por sempre acreditar no meu potencial, mesmo nos momentos mais desafiadores. Ao Prof. Dr. Carlos Thomaz, por sua orientação na etapa final deste projeto, contribuindo de forma essencial para a conclusão desta tese, com conselhos assertivos, dedicação e grande generosidade intelectual. Ao Dr. Oscar Fugita, pela contribuição fundamental para a validação clínica deste trabalho. Como urologista, sua visão médica foi essencial para adaptar a proposta desenvolvida a um problema real enfrentado na prática clínica da urologia. Sua generosidade e disponibilidade em compartilhar conhecimento, esclarecer dúvidas e oferecer orientações práticas foram indispensáveis para que esta pesquisa mantivesse relevância e aplicabilidade no contexto cirúrgico real. Ao Prof. Dr. Gilson Giraldi, pela orientação próxima e constante durante o período em que estive no LNCC, em Petrópolis, e também posteriormente, à distância. Sua dedicação ultrapassou os limites institucionais, acompanhando de perto o desenvolvimento dos experimentos, a redação dos artigos e a validação dos resultados. Agradeço pela paciência, pelas discussões sempre construtivas e pelo exemplo de rigor científico e humildade acadêmica, que tanto contribuíram para meu crescimento como pesquisador.

Aos meus amigos do Immersive Lab, pelo apoio técnico e humano durante todo o desenvolvimento deste trabalho, em especial aos alunos Fernando Pujaico Riviera e Gabriel Nunes Missima, cuja parceria, dedicação e amizade foram fundamentais. Aos meus amigos de São Bernardo do Campo, que foram um suporte essencial para que eu seguisse em frente. Obrigado por me inspirarem, ajudarem a aliviar a pressão desta jornada e fazerem parte da minha vida pessoal de forma tão significativa. Em especial, agradeço a Pedro Lagosta, Rodolpho Frodl, Igor Savordelli, Kaue Guisani, Giovanni Vilella, Matheus Centurion e Victor Abreu, pela amizade, incentivo constante e por estarem ao meu lado nos momentos que mais precisei.

À Carla Ferreira, que me acolheu com generosidade em um momento crucial, oferecendo um espaço e um ambiente de paz para que eu pudesse concluir a escrita da minha qualificação.

Aos amigos de Petrópolis, pelo acolhimento e carinho durante minha estadia, especialmente ao Emerson Ribeiro, que me motivou a não desistir no momento mais decisivo desta

caminhada. Aos colegas do LNCC, especialmente a Jalil, Bruna Sozzo, Letícia Assis, Pedro Luga, Frank Acasiete e Eliaquim Monteiro, pela convivência agradável, pela troca de experiências e pelo companheirismo diário. E à toda cidade de Petrópolis, que me acolheu com tanto amor, me proporcionou experiências memoráveis e marcou profundamente o meu amadurecimento.

Aos meus irmãozinhos da Paróquia Santa Teresinha, cuja juventude, força, amor e gentileza me inspiram a ser uma pessoa melhor e a nunca esquecer o verdadeiro propósito da vida. Ao Padre Alexandre Mosquelli, pelas orações e palavras encorajadoras que me reergueram para a finalização desta tese.

Ao Hospital Universitário da Universidade de São Paulo (USP), pelo suporte na obtenção e compreensão das imagens radiológicas, que foram fundamentais para a etapa clínica desta pesquisa. Ao Laboratório Nacional de Computação Científica (LNCC), pela acolhida, infraestrutura e oportunidade de aprimoramento técnico e científico que tanto contribuíram para o meu crescimento profissional. Ao Centro Universitário FEI, pela bolsa de estudos, pela estrutura e pelo laboratório que me proporcionaram um ambiente adequado ao desenvolvimento desta pesquisa. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo apoio financeiro concedido durante esta investigação.

E, por fim, a todas que, de alguma forma, fizeram parte desta trajetória, deixando marcas de afeto, aprendizado e inspiração ao longo do caminho, deixo o meu mais profundo e sincero agradecimento.

“Há pessoas que desejam saber só por saber, e isso é curiosidade; outras, para alcançarem fama, e isso é vaidade; outras, para enriquecerem com a sua ciência, e isso é um negócio mesquinho; outras, para serem edificadas, e isso é prudência; outras, para edificarem os outros, e isso é amor.”

Santo Agostinho

RESUMO

Cirurgias de ureterosopia são procedimentos minimamente invasivos realizados para o tratamento de cálculos urinários, nos quais o cirurgião utiliza um ureteroscópio flexível para acessar e remover o cálculo presente no sistema urinário. O sucesso desse tipo de cirurgia depende fortemente da experiência do cirurgião em navegar por um ambiente anatômico estreito e complexo, com múltiplas bifurcações e variações geométricas. O uso de sistemas de navegação assistida ou de cirurgias robóticas pode reduzir riscos e aumentar a precisão, mas exige, como etapa preliminar, o planejamento automático da trajetória tridimensional (3D) que conecta a bexiga ao cálculo. Esta tese propõe uma abordagem que combina diagnóstico assistido por Inteligência Artificial (IA) e planejamento semi-automático da trajetória para cirurgias de ureterosopia. O método é composto por três módulos principais: (i) classificação automática de imagens de tomografia computadorizada (TC) para detecção de cálculos urinários, (ii) segmentação automática de cálculos urinários, e (iii) planejamento de trajetória 3D utilizando algoritmos de busca e suavização. Na etapa de classificação, foi proposto um modelo baseado em *ensemble learning* e fusão tardia de múltiplas bases de dados, alcançando acurácia média superior a 99% na detecção de cálculos urinários em imagens coronais de TC, com baixo custo computacional e boa capacidade de generalização. Na etapa de segmentação, foi introduzido o conjunto de dados público *KSSD2025*, contendo 838 imagens axiais de TC com anotações manuais, e avaliadas arquiteturas profundas como U-Net, U-Net++, U-Net3+ e TransU-Net, todas com coeficiente Dice médio superior a 95%, destacando-se a U-Net pelo equilíbrio entre desempenho e simplicidade. No planejamento de trajetórias, o método proposto combinou uma versão modificada do algoritmo A* com filtros de suavização, sendo as curvas B-Spline as que apresentaram melhores resultados em termos de suavidade, coerência anatômica e compatibilidade com a flexibilidade de ureteroscópios reais. Embora os módulos tenham sido avaliados de forma parcialmente independente, o método proposto permite uma futura integração completa em um sistema unificado de suporte ao planejamento cirúrgico. Os resultados obtidos indicam que a combinação entre aprendizado profundo e modelagem geométrica constitui uma solução promissora para o desenvolvimento de sistemas de apoio clínico, contribuindo para cirurgias mais seguras, eficientes e personalizadas.

Palavras-chave: Planejamento de trajetória de Ureterosopia. Inteligência artificial. Detecção de cálculos urinários. Segmentação de imagens médicas. Tomografia Computadorizada.

ABSTRACT

Ureteroscopy surgeries are minimally invasive procedures performed to treat urinary stones, in which the surgeon uses a flexible ureteroscope to access and remove the stone present in the urinary system. The success of this type of surgery depends heavily on the surgeon's experience in navigating a narrow and complex anatomical environment with multiple bifurcations and geometric variations. The use of assisted navigation systems or robotic surgery can reduce risks and increase accuracy, but requires, as a preliminary step, the automatic planning of the three-dimensional (3D) trajectory connecting the bladder to the stone. This thesis proposes an approach that combines Artificial Intelligence (AI)-assisted diagnosis and semi-automatic trajectory planning for ureteroscopy surgeries. The method consists of three main modules: (i) automatic classification of computed tomography (CT) images for the detection of urinary stones, (ii) automatic segmentation of urinary stones, and (iii) 3D trajectory planning using search and smoothing algorithms. In the classification stage, a model based on ensemble learning and late fusion of multiple databases was proposed, achieving an average accuracy of over 99% in the detection of urinary calculi in coronal CT images, with low computational cost and good generalization capacity. In the segmentation stage, the public dataset KSSD2025 was introduced, containing 838 manually annotated axial CT images, and deep architectures such as U-Net, U-Net++, U-Net3+, and TransU-Net were evaluated, all with an average Dice coefficient greater than 95%, with U-Net standing out for its balance between performance and simplicity. In trajectory planning, the proposed method combined a modified version of the A* algorithm with smoothing filters, with B-Spline curves showing the best results in terms of smoothness, anatomical consistency, and compatibility with the flexibility of real ureteroscopes. Although the modules were evaluated partially independently, the proposed architecture allows for their future complete integration into a unified surgical planning support system. The results indicate that the combination of deep learning and geometric modeling offers a promising solution for developing clinical support systems, contributing to safer, more efficient, and personalized surgeries.

Keywords: Ureteroscopy. Urinary stones. Kidney stone detection. Computed tomography. Artificial intelligence. Deep learning. Medical image segmentation. 3D trajectory planning. Computer-assisted surgery.

LISTA DE ILUSTRAÇÕES

Ilustração 1 – Ilustração de um sistema urinário.	20
Ilustração 2 – Ilustração da doença de ureterolitíase.	21
Ilustração 3 – Exemplo de um ureteroscópio flexível do modelo Flex-X2S.	21
Ilustração 4 – Ilustração de uma cirurgia de ureterosopia.	22
Ilustração 5 – Medição de um cálculo em uma tomografia para planejar o melhor tratamento.	23
Ilustração 6 – Exemplo de fatias de uma TC empilhadas em um espaço 3D.	23
Ilustração 7 – Ilustração de curvas com diferentes encurvamentos.	27
Ilustração 8 – Ilustração de um círculo osculador e seu raio de curvatura.	28
Ilustração 9 – Ilustração de uma rede neural <i>perceptron</i> multicamadas.	31
Ilustração 10 – Função ReLU	33
Ilustração 11 – Exemplo de uma CNN com duas camadas de convolução e uma camada de <i>pooling</i>	33
Ilustração 12 – Exemplo de uma convolução 2D em uma matriz	34
Ilustração 13 – Exemplo de uma operação de <i>max pooling</i> e <i>average pooling</i> , considerando um filtro de <i>pooling</i> de tamanho 2×2	35
Ilustração 14 – Exemplo de uma operação de <i>flatten</i> , transformando uma matriz M em um vetor F	35
Ilustração 15 – Exemplo de imagens de entrada da rede <i>U-Net</i> . A imagem da esquerda é a imagem original e a da direita o seu mapa de segmentação esperado. . .	36
Ilustração 16 – Arquitetura da <i>U-Net</i>	37
Ilustração 17 – Ilustração da presença de cálculos em diferentes partes de um sistema urinário humano.	43
Ilustração 18 – Exemplo de fatias de uma tomografia computadorizada empilhadas em um espaço 3D.	43
Ilustração 19 – Fluxograma da metodologia proposta neste trabalho.	47
Ilustração 20 – Fonte: Autor	47
Ilustração 21 – Exemplos de amostras da Base de Dados A.	48
Ilustração 22 – Exemplos de amostras da Base de Dados B.	48
Ilustração 23 – Exemplo de amostras após o aumento de dados.	50
Ilustração 24 – Arquitetura Conv4.	51
Ilustração 25 – Fluxograma da metodologia da arquitetura DE-MLP proposta neste trabalho. . .	52
Ilustração 26 – Matrizes de confusão geradas. 26a apresenta a matriz do modelo DE-MLP treinado com modelos que não passaram por aumento de dados. 26b apresenta a matriz do modelo DE-MLP treinado com modelos que passaram por aumento de dados.	57

Ilustração 27 – Gráficos de convergência gerados. (a) e (b) apresentam gráficos de modelos da base A. (c) e (d) apresentam gráficos de modelos da base B. (e) e (f) apresentam gráficos de modelos da base conjunta A e B.	64
Ilustração 28 – Mapas de calor Grad-CAM das regiões mais ativadas pelo modelo proposto nas classificações de amostras que possuem cálculos renais.	65
Ilustração 29 – Mapas de calor Grad-CAM das regiões mais ativadas pelo modelo proposto nas classificações de amostras que não possuem cálculos renais.	65
Ilustração 30 – Comparação da amostra que o DE-MLP classificou incorretamente com os mapas de ativação do extrator de características dos modelos <i>Conv4A</i> e <i>Conv4B</i>	66
Ilustração 31 – Diagrama da construção do KSSD2025 e da validação dos modelos baseados em U-Net.	70
Ilustração 32 – Ilustração do pipeline de anotação do dataset.	72
Ilustração 33 – Exemplos de comparação de máscaras anotadas entre anotadores.	73
Ilustração 34 – Arquitetura original da U-Net.	74
Ilustração 35 – Arquitetura U-Net modificada.	76
Ilustração 36 – Resultados de segmentação obtidos por cada modelo baseado em U-Net treinado com KSSD2025. A primeira coluna mostra as amostras originais. A segunda coluna mostra as amostras de referência manualmente anotadas. A terceira coluna apresenta a previsão obtida pelo modelo U-Net. A quarta coluna apresenta a previsão obtida pelo modelo U-Net++. A quinta coluna apresenta a previsão obtida pelo modelo U-Net3+. A sexta coluna apresenta a previsão obtida pelo modelo TransU-Net.	80
Ilustração 37 – Valores de p do teste t pareado unilateral ($A > B$). Valores de p menores indicam superioridade estatisticamente significativa da Arquitetura A sobre B.	81
Ilustração 38 – Os resultados de segmentação revelaram que os modelos baseados em U-Net apresentaram erros. A primeira coluna exibe as amostras originais. A segunda coluna exibe as amostras originais ampliadas. A terceira coluna exibe a amostra padrão-ouro ampliada. A quarta coluna mostra a imagem de predição obtida pelo modelo U-Net. A quinta coluna mostra a imagem de predição obtida pelo modelo U-Net++. A sexta coluna mostra a imagem de predição obtida pelo modelo U-Net3+. A sétima coluna mostra a imagem de predição obtida pelo modelo TransU-Net.	82

Ilustração 39 – Resultados de segmentação em amostras de dados externos. A primeira linha mostra os resultados obtidos em uma amostra de cálculo no rim esquerdo. A segunda linha mostra os resultados obtidos em uma amostra de cálculo no rim direito. A terceira linha mostra os resultados obtidos em uma amostra de cálculo no ureter. Todos os modelos foram treinados exclusivamente no conjunto de dados KSSD2025.	84
Ilustração 40 – Resultados de segmentação em amostras sem cálculos renais. A primeira coluna mostra as fatias originais de TC. A segunda coluna apresenta as máscaras padrão-ouro anotadas manualmente. Da terceira à sexta colunas estão as máscaras de segmentação preditas pelos modelos U-Net, U-Net++, U-Net3+ e TransU-Net, respectivamente. Todos os modelos foram treinados exclusivamente no conjunto de dados KSSD2025.	85
Ilustração 41 – Fluxograma da metodologia proposta para o planejamento automático da trajetória 3D em ureteroscopia.	102
Ilustração 42 – Visualização 3D das trajetórias suavizadas: comparação entre Savitzky-Golay, B-Spline e Laplacian.	103
Ilustração 43 – Trajetória otimizada com <i>B-Spline</i> (azul) sobreposta à reconstrução 3D do trato urinário.	104

LISTA DE TABELAS

Tabela 1 – Hiperparâmetros utilizados para o treinamento dos modelos de aprendizado profundo.	54
Tabela 2 – Comparação de desempenho do DE-MLP com os modelos Conv4.	56
Tabela 3 – Tabela comparativa do método proposto com métodos anteriores que utilizaram a Base de Dados A para detecção de cálculos renais	60
Tabela 4 – Tabela comparativa do método proposto com métodos anteriores que utilizaram a Base de Dados B para detecção de cálculos renais	61
Tabela 5 – Comparação entre a U-Net modificada e a original em termos de custo computacional e uso de recursos.	75
Tabela 6 – Hiperparâmetros utilizados para o treinamento das arquiteturas de deep learning.	77
Tabela 7 – Resultados da validação cruzada 5-fold para modelos baseados em U-Net treinados com KSSD2025.	78
Tabela 8 – Coeficiente Dice médio, intervalos de confiança e teste de normalidade para cada arquitetura.	79
Tabela 9 – Coeficientes de Dice (%) obtidos por cada modelo quando testados em um conjunto de dados externo, segmentado por região anatômica.	83
Tabela 10 – Informações de complexidade de cada modelo baseado em U-Net utilizado neste trabalho.	85
Tabela 11 – Comparação entre o estudo proposto e abordagens do estado da arte.	86
Tabela 12 – Pontuações da função objetivo para cada técnica em diferentes resoluções.	99
Tabela 13 – Comparação quantitativa entre as melhores trajetórias.	99

LISTA DE ALGORITMOS

Algoritmo 1 – Pseudocódigo do algoritmo A*	25
Algoritmo 2 – Algoritmo Bresenham 3D	26

SUMÁRIO

1	Introdução	17
1.1	Objetivo	18
1.2	Estrutura do Trabalho	18
2	Conceitos Fundamentais	20
2.1	Urolitíase	20
2.2	Ureteroscópio Flexível	20
2.3	Ureteroscopia	22
2.4	Tomografia Computadorizada	22
2.5	Convolução Discreta em 2 dimensões	24
2.6	Algoritmo A*	24
2.6.1	Algoritmo de Bresenham 3D	24
2.6.2	Campos de Repulsão	25
2.7	Curvatura de Função Vetorial Paramétrica	27
2.7.1	Raio de uma Curvatura	28
2.8	Filtro Savitzky-Golay	28
2.9	B-Spline	29
2.10	Filtro Laplaciano	29
2.11	Redes Neurais Artificiais	30
2.11.1	<i>Perceptron</i>	30
2.11.2	Rede Neural <i>Perceptron</i> Multicamadas	30
2.11.3	Retropropagação	31
2.12	Função de ativação	32
2.12.1	Unidade Linear Retificada	32
2.13	Redes Neurais Convolucionais	32
2.13.1	Camada convolucional	34
2.13.2	Camada de <i>pooling</i>	34
2.13.3	Camada de <i>flatten</i>	35
2.13.4	Camada da rede totalmente conectada	36
2.14	U-Net	36
2.14.1	Arquitetura da <i>U-Net</i>	36
2.15	Métricas de avaliação	37
2.15.1	Entropia Cruzada	38
2.15.2	Distância Euclidiana	38
2.15.3	Acurácia	38
2.15.4	Precisão	38
2.15.5	<i>Recall</i>	39
2.15.6	F1-Score	39

2.15.7	Coeficiente de Dice	39
2.15.8	Coeficiente de Jaccard	39
2.16	Validação Cruzada <i>k-Fold</i>	40
2.17	Validação Cruzada de Monte Carlo	40
3	Classificação de imagens para detecção automática de cálculos urinários em TC	42
3.1	Contextualização	42
3.2	Trabalhos Relacionados	44
3.3	Materiais e Métodos	46
3.3.1	Base de Dados Combinada	46
3.3.2	Pré-processamento	48
3.3.3	Modelo único (Conv4)	50
3.3.4	Modelo proposto DE-MLP	51
3.3.5	Configuração do treinamento e hiperparâmetros	53
3.3.6	Métricas de avaliação	54
3.4	Resultados e Discussão	55
3.4.1	Métricas de desempenho do modelo proposto	55
3.4.2	Matriz de confusão: DE-MLP	56
3.4.3	Taxa de convergência	56
3.4.4	Mapas de ativação	58
3.4.5	Análise do caso de amostra classificada incorretamente	59
3.4.6	Contextualização com o estado-da-arte	59
3.4.7	Aplicações clínicas do modelo DE-MLP proposto	61
3.4.7.1	<i>Assistente de diagnóstico de cálculos renais para radiologistas</i>	61
3.4.7.2	<i>Ferramenta de diagnóstico inicial</i>	61
3.4.7.3	<i>Integração automatizada de radiologia</i>	61
3.4.8	Limitações do modelo	62
3.4.9	Direções futuras	62
4	Segmentação Automática de Cálculos Urinários em TCs	67
4.1	Contextualização	67
4.2	Trabalhos Relacionados	68
4.3	Materiais e Métodos	70
4.3.1	Imagens de TC de Cálculos Renais	71
4.3.2	Anotação dos dados	71
4.3.3	Modelos U-Net Modificados	74
4.3.4	Configuração de Treinamento e Hiperparâmetros	76
4.4	Resultados e Discussão	77
4.4.1	Análise de Desempenho com Validação Cruzada 5-Fold	78
4.4.2	Análise Visual dos Resultados	78

4.4.3	Análise Estatística	79
4.4.4	Análise dos Casos de Erro	81
4.4.5	Avaliação dos Modelos em Dados Externos	82
4.4.6	Avaliação dos Modelos em Amostras Sem Cálculos Renais	84
4.4.7	Análise de Complexidade	85
4.4.8	Paralelo com o estado da arte	86
4.4.9	Aplicações Clínicas	87
4.4.10	Limitações do Estudo	88
4.4.11	Direções Futuras	89
5	Planejamento Automático de Trajetória 3D de Ureteroscopia	90
5.1	Contextualização	90
5.2	Trabalhos Relacionados	91
5.3	Materiais e Métodos	96
5.3.1	Aquisição e Pré-processamento das Imagens	96
5.3.2	Segmentação do Cálculo Ureteral e do Trato Urinário	96
5.3.3	Reconstrução Tridimensional	96
5.3.4	Planejamento da Trajetória Inicial	97
5.3.5	Suavização da Trajetória	97
5.3.6	Função de Avaliação Multiobjetivo	98
5.4	Experimentos e Resultados	99
5.4.1	Resultados Quantitativos	99
5.4.2	Resultados Visuais	100
5.5	Discussão	100
5.5.1	Análise do resultados	100
5.5.2	Potenciais Clínicos	100
5.5.3	Limitações	101
5.5.4	Trabalhos Futuros	101
6	Conclusão	105
	REFERÊNCIAS	107

1 INTRODUÇÃO

A formação de cálculos urinários, ou urolitíase, resulta da cristalização e agregação de minerais presentes na urina, originando estruturas sólidas que podem se localizar em diferentes regiões do sistema urinário. Essa condição apresenta elevada prevalência e incidência crescente em nível mundial, configurando-se como um problema relevante de saúde pública, em função do impacto direto sobre a qualidade de vida dos pacientes e do aumento da demanda por atendimentos médicos e intervenções cirúrgicas (Sorokin *et al.*, 2017; Jha *et al.*, 2013; Alelign; Petros, 2018). Os cálculos urinários podem provocar impactos mecânicos nas paredes do trato urinário ou causar obstrução parcial ou total do fluxo urinário, resultando em quadros de dor intensa, processos inflamatórios e, em casos mais graves, comprometimento funcional dos rins e evolução para doença renal crônica, quando não tratados de forma adequada (Shokeir, 2002; Ingimarsson; Krambeck; Pais, 2016).

A intervenção cirúrgica mais frequente para tratar cálculos renais é a ureteroscopia (Ordon *et al.*, 2014), um procedimento minimamente invasivo em que o cirurgião realiza uma trajetória até a localização do cálculo renal para tratá-lo. Apesar de ser um procedimento que evoluiu nas últimas décadas, a ocorrência de complicações durante a cirurgia não é incomum, sendo um dos principais fatores a falta de experiência do cirurgião responsável (Schuster *et al.*, 2001; El-Nahas *et al.*, 2009; Mandal *et al.*, 2012; Komori *et al.*, 2015). No entanto, o treinamento de cirurgiões novatos em pacientes reais tornou-se uma preocupação por questões éticas e legais (Gallagher; Traynor, 2008). Por esse motivo, o uso de simuladores se tornou comum para o treinamento de cirurgias ureteroscópicas (Beverini *et al.*, 2021). Da mesma forma, tecnologias de realidade virtual têm se tornado uma tendência em simuladores de urologia (Hameed *et al.*, 2022), permitindo que cirurgiões visualizem e naveguem por dentro do sistema urinário do paciente de forma imersiva.

Com o avanço da Inteligência Artificial (IA), especialmente dos métodos de aprendizado profundo e sua aplicação em imagens médicas, tornou-se possível realizar etapas importantes do diagnóstico pré-operatório de forma automática, como a classificação e segmentação de estruturas anatômicas e de cálculos urinários em tomografias computadorizadas (TC) (Chan; Hadjiiski; Samala, 2020; Hesamian *et al.*, 2019). A TC sem contraste é o padrão-ouro para detecção de cálculos urinários (Saw *et al.*, 2000; Kambadakone *et al.*, 2010), fornecendo informações detalhadas sobre tamanho, localização e composição, essenciais para o planejamento cirúrgico. Métodos automáticos de segmentação baseados em redes neurais convolucionais têm mostrado resultados promissores para identificar a posição exata dos cálculos e delimitar as estruturas do sistema urinário (Cui *et al.*, 2020; Li, D. *et al.*, 2022; Kim, Y. I. *et al.*, 2023).

A partir dessas informações diagnósticas, torna-se possível realizar o planejamento automático da trajetória que o ureteroscópio deverá percorrer até alcançar o cálculo. Essa etapa é fundamental tanto para o suporte à decisão cirúrgica quanto para aplicações futuras em sistemas de navegação ou assistência robótica. A geração automática de trajetórias tridimensionais

(3D) consiste em definir, dentro do volume reconstruído do trato urinário, um caminho suave e anatomicamente viável que conecte a bexiga ao cálculo, respeitando a geometria do ureter e evitando regiões de risco.

Trabalhos anteriores abordaram o planejamento de trajetórias para ureteroscopia de forma limitada. Fan Yang *et al.* (2019) consideraram apenas um segmento do ureter, sem levar em conta a posição real do cálculo. A dissertação de Scarzanella (2020) tratou apenas da navegação da uretra até a entrada do ureter, também sem integrar informações de diagnóstico. Além disso, nenhum dos dois estudos explorou a integração entre segmentação automática baseada em aprendizado profundo e algoritmos de busca para planejamento completo do trajeto.

Assim, neste trabalho é proposta uma abordagem integrada de diagnóstico assistido por IA e planejamento de trajetória para ureteroscopia. O método é composto por três etapas principais: (i) classificação de imagem para detecção automática de cálculos urinários a partir de exames de TC, utilizando redes neurais profundas; (ii) segmentação automática de cálculos urinários em TCs a partir de redes neurais baseadas na arquitetura U-Net e (iii) geração semi-automática da trajetória 3D até o cálculo, empregando algoritmos de busca combinados com filtros que simulam a flexibilidade do ureteroscópio, garantindo que o caminho seja suavizado, contínuo e compatível com a anatomia do paciente. Essa integração permite transformar dados de imagem pré-operatórios em informações diretamente utilizáveis no planejamento cirúrgico, potencialmente aumentando a precisão e a segurança dos procedimentos.

1.1 OBJETIVO

Este trabalho tem como objetivo desenvolver um método de planejamento semi-automático de cirurgias ureteroscópicas para o tratamento de cálculos urinários baseado em diagnóstico de exames de TCs assistido por IA. O método combina técnicas de aprendizado profundo para classificação e segmentação automáticas com algoritmos de busca e filtros digitais para simular a navegação de um ureteroscópio flexível, considerando informações anatômicas e a localização do cálculo urinário obtidas na etapa de diagnóstico.

1.2 ESTRUTURA DO TRABALHO

O restante deste trabalho é dividido da seguinte maneira:

O Capítulo 2 apresenta os Conceitos Fundamentais necessários para a compreensão do restante da tese. São abordados tópicos relacionados à formação de cálculos urinários, tomografia computadorizada, fundamentos de redes neurais convolucionais e algoritmos de busca aplicados para o planejamento de trajetórias tridimensionais.

No Capítulo 3, é apresentada uma nova abordagem para classificação de imagens de TC para detecção de cálculos urinários, explorando modelos de aprendizado profundo treinados com bases de dados de fontes distintas. São descritos os modelos utilizados, os conjuntos de

dados empregados, os procedimentos de treinamento e os resultados obtidos para a tarefa de detecção.

O Capítulo 4 aborda um novo método para segmentação automática de cálculos urinários em TCs, detalhando a construção de uma nova base de dados, os métodos de segmentação baseados em U-Net, os experimentos realizados e os resultados obtidos.

No Capítulo 5, é apresentado o método proposto para planejamento semi-automático de trajetória 3D de ureteroscopia, integrando as informações anatômicas e de localização dos cálculos obtidas nas etapas de diagnóstico assistido por IA. São descritos os algoritmos de busca e filtros utilizados para simular a navegação de um ureteroscópio flexível, bem como a função de avaliação multiobjetivo empregada para selecionar a trajetória final.

Por fim, o Capítulo 6 apresenta as conclusões gerais do trabalho, destacando as principais contribuições científicas e analisando como as abordagens propostas podem se integrar em um sistema único e seu potencial de aplicação no suporte ao planejamento cirúrgico de ureteroscopia

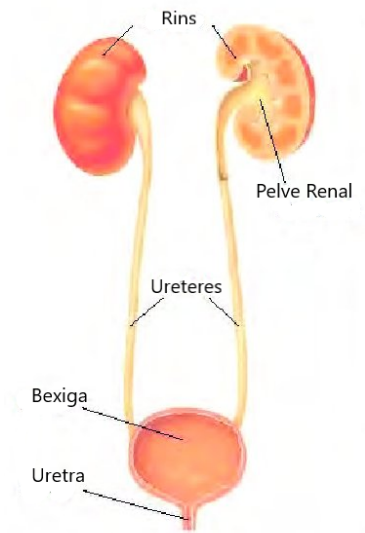
2 CONCEITOS FUNDAMENTAIS

Neste capítulo serão apresentados os conceitos fundamentais relacionados ao problema de planejamento de trajetória em ureteroscopia para o tratamento de urolitíase e a explicação do funcionamento das principais técnicas utilizadas na metodologia deste trabalho.

2.1 UROLITÍASE

Urolitíase é uma doença que afeta o sistema urinário, composto principalmente pela uretra, a bexiga, um par de rins, que possuem uma pelve renal, e um par de ureteres, ilustrado na Figura 1. Essa doença ocorre quando existe a formação de cálculos urinários, que são substâncias sólidas formadas a partir da agregação de minerais da urina.

Figura 1 – Ilustração de um sistema urinário.



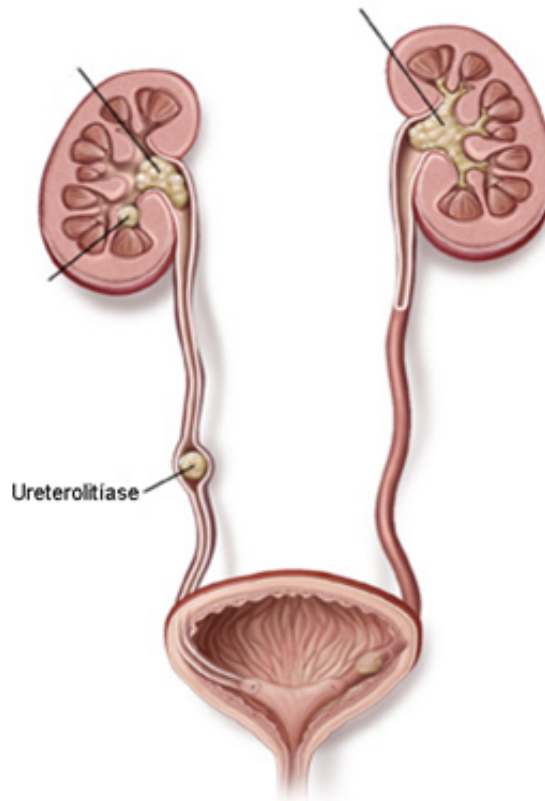
Fonte: Cauberg *et al.* (2009).

A ureterolitíase, especificamente, ocorre quando esses cálculos estão localizados no ureter, obstruindo a passagem da urina, como mostrado na Figura 2. Essa condição pode causar dores agudas, além de outras complicações no restante do sistema urinário, podendo causar dano permanente nos ureteres e nos rins. Uma das formas de tratamento escolhida é a ureteroscopia, dependendo das circunstâncias do paciente, do tamanho e da localização do cálculo (Glazer; Brea; Vaitla, 2022), onde é utilizado um equipamento chamado de ureteroscópio flexível.

2.2 URETEROSCÓPIO FLEXÍVEL

Um ureteroscópio flexível é um equipamento constituído de um cabo longo, fino e flexível, com uma fonte de luz e uma câmera em sua ponta (Gupta *et al.*, 2020). É uma das principais ferramentas utilizadas para cirurgias de ureteroscopia no tratamento de cálculos urinários e de

Figura 2 – Ilustração da doença de ureterolítase.



Fonte: Autor.

outras doenças que afetam o sistema urinário (Zhao *et al.*, 2021). A Figura 3 apresenta um exemplo de um ureteroscópio flexível. Os modelos de ureteroscópio possuem geralmente um cabo de tamanho entre 64,5 e 70cm, aproximadamente, e um ângulo de deflexão que pode chegar até 280°, dependendo do modelo de ureteroscópio utilizado (Paffen *et al.*, 2008; Mazzucchi *et al.*, 2022).

Figura 3 – Exemplo de um ureteroscópio flexível do modelo Flex-X2S.

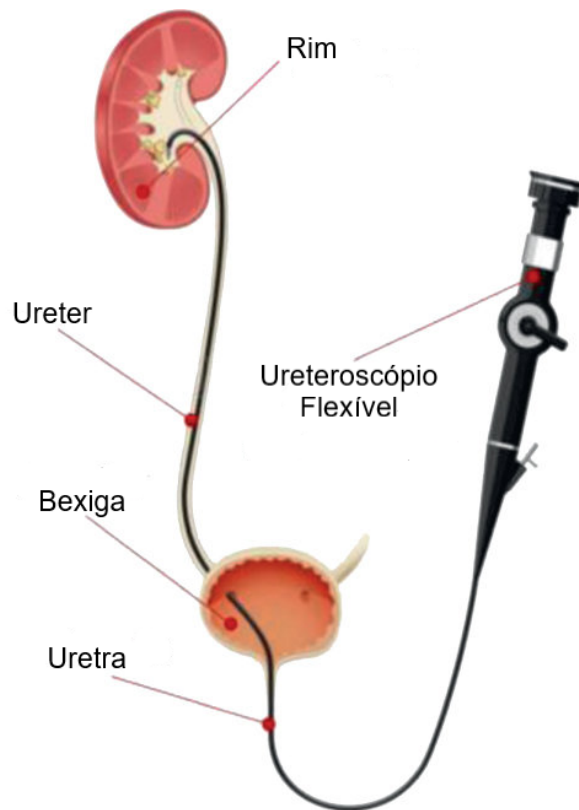


Fonte: Mager *et al.* (2018).

2.3 URETEROSCOPIA

A cirurgia de ureteroscopia é um procedimento minimamente invasivo, onde o interior do paciente é acessado através de um pequeno orifício, realizado para o tratamento de ureterolitíase. É o tipo de intervenção mais comum para cálculos renais e uma das cirurgias mais realizadas na área de urologia (Ordon *et al.*, 2014). Nesse procedimento, é utilizado um ureteroscópio, sendo inserido o seu cabo dentro paciente pela sua uretra, passando pelo seu sistema urinário até a localização do cálculo ser encontrada. Em seguida, é feita a remoção do cálculo, puxando ele pelo caminho de volta, ou quebrando-o, dependendo do seu tamanho. A Figura 4 ilustra um exemplo de uma cirurgia de ureteroscopia.

Figura 4 – Ilustração de uma cirurgia de ureteroscopia.



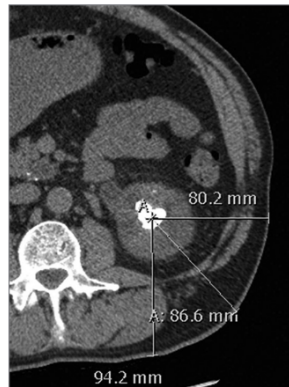
Fonte: Zhao *et al.* (2021).

2.4 TOMOGRAFIA COMPUTADORIZADA

A Tomografia Computadorizada (TC) é um exame de coleta de imagem não invasivo, apresentado por Hounsfield (1973). Por meio de um equipamento de raio-X, múltiplas imagens de radiografia do paciente são combinadas para criar um exame mais detalhado em relação à uma radiografia comum. Por permitir a análise do interior do paciente através da coleta de diversas fatias de imagens, esse exame é adotado para diagnosticar doenças como acidentes vasculares cerebrais, pneumonia, embolia pulmonar, aneurisma, entre outros.

Esse é um dos principais exames utilizados para o diagnóstico e planejamento de tratamento de pacientes com suspeita de cálculo no sistema urinário, sendo realizado na região abdominal. Tem sido preferido ao invés de outros tipos de exames de imagens por sua velocidade, facilidade de coleta de imagem, disponibilidade e por permitir diagnosticar outras anormalidades no sistema urinário (Andrabi *et al.*, 2015). A Figura 5 apresenta um exemplo da medição de um cálculo urinário a partir de um exame de TC.

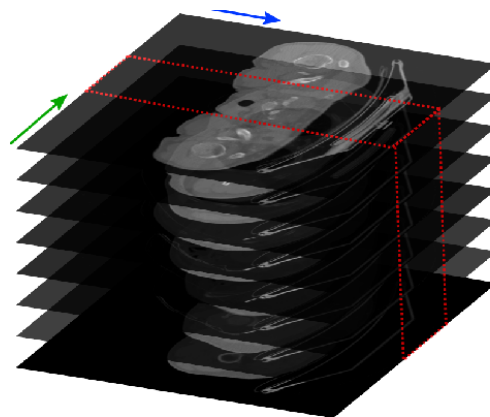
Figura 5 – Medição de um cálculo em uma tomografia para planejar o melhor tratamento.



Fonte: Andrabi *et al.* (2015).

Por coletar diversas fatias de imagem, o exame de TC permite a manipulação e visualização 3D do interior do paciente, expandindo as formas de diagnóstico e planejamento de tratamento. A Figura 6 apresenta um exemplo de um exame de tomografia visualizado em um espaço 3D.

Figura 6 – Exemplo de fatias de uma TC empilhadas em um espaço 3D.



Fonte: Kanavati *et al.* (2018a).

2.5 CONVOLUÇÃO DISCRETA EM 2 DIMENSÕES

Considerando uma matriz w de tamanho $m \times n$ e uma matriz $f(x,y)$, a convolução de w em $f(x,y)$, denotada como $(w ** f)(x,y)$ pode ser definida pela Equação (1) (Gonzales; Wintz, 1987).

$$(w ** f)(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s,t) f(x - s, y - t), \quad (1)$$

onde $a = \frac{(m-1)}{2}$, $b = \frac{(n-1)}{2}$, assumindo que m e n são números inteiros e ímpares. Com isso, é obtida a soma do produto de cada elemento de w percorrendo cada elemento de f . Esse processo também é chamado de filtragem espacial.

2.6 ALGORITMO A*

O algoritmo A* (A-estrela) é um algoritmo de busca informada apresentado por Hart, Nilsson e Raphael (1968) para encontrar o caminho ótimo dentro de um grafo ponderado. O caminho é encontrado expandindo o menor número de nós por priorizar os nós que possuem a melhor solução atual. Dado um vértice inicial s , ele encontra uma sequência de nós até o vértice objetivo t de forma que o custo total seja mínimo. Assim, em cada iteração do algoritmo, o vértice a ser expandido é selecionado de forma que minimize a função de custo, dada pela Equação (2).

$$f(n) = g(n) + h(n), \quad (2)$$

sendo $g(n)$ o custo do caminho do vértice inicial s até n , $h(n)$ é uma heurística que mede o custo do caminho de n até o vértice objetivo t (Russell, 2010). O algoritmo finaliza em dois casos. Quando o nó atual a ser expandido é o nó t , retornando a lista de nós de s até o nó objetivo t , ou quando não existe mais nós a serem expandidos, retornando uma lista vazia, pois a solução não foi encontrada. O Algoritmo 1 apresenta um pseudocódigo do A*.

2.6.1 Algoritmo de Bresenham 3D

O algoritmo de Bresenham é um método originalmente desenvolvido para rasterização de linhas em sistemas gráficos bidimensionais (Bresenham, 1965). Sua principal característica é a eficiência na determinação dos pixels que melhor aproximam uma linha contínua em uma grade discreta, utilizando apenas operações inteiras e incrementais.

Para aplicações em ambientes tridimensionais, o algoritmo foi estendido para o caso 3D, permitindo determinar todos os voxels que compõem o trajeto entre dois pontos no espaço. O Algoritmo 2 apresenta um pseudocódigo do algoritmo Bresenham 3D:

Algoritmo 1 – Pseudocódigo do algoritmo A*

```

1 Função A*(Grafo G, Pesos w, Vértice inicial s, Vértice objetivo t)
2  $g[s] = 0$ 
3  $f[s] = g[s] + h(s,t)$ 
4  $Q[0] = s$ 
5 while existir vértices não explorados do
6    $u = \text{ExtrairMinCusto}(Q)$ 
7   for  $v \in \text{lista de adjacência de } u$  do
8      $g_{temp} = g[u] + w(u,v)$ 
9     if  $v$  ainda não foi explorado or  $g_{temp} < g[v]$  then
10        $g[v] = g_{temp}$ 
11        $f[v] = g[v] + h(v,t)$ 
12       if  $v == t$  then
13         retornar caminho ótimo de  $s$  até  $t$ 
14       end
15       if  $v$  não foi explorado then
16          $Q.push(v)$ 
17       end
18     end
19   end
20 end
21 retornar caminho não encontrado

```

2.6.2 Campos de Repulsão

Embora o algoritmo A* produza um caminho de custo mínimo entre dois pontos, ele não garante que a trajetória esteja adequadamente afastada das paredes do ambiente. Para contornar esse problema, foi implementado um campo de repulsão no espaço tridimensional do volume segmentado, inspirado em técnicas de planejamento de movimento em robótica (Khatib, 1986).

O campo de repulsão é construído a partir de uma *função potencial*, gerada por uma convolução de um filtro gaussiano. Essa operação cria um gradiente espacial onde regiões próximas às paredes apresentam valores de repulsão mais altos, enquanto as regiões centrais permanecem neutras. Durante a execução do A*, a função de custo $f(n)$ é modificada para incluir um termo adicional proporcional à intensidade do campo de repulsão no voxel correspondente ao nó n , conforme a Equação 3.

$$f'(n) = g(n) + h(n) + \lambda \cdot R(n), \quad (3)$$

onde $R(n)$ representa a intensidade do campo de repulsão no ponto n e λ é um fator de ponderação que controla o peso dessa penalização.

Algoritmo 2 – Algoritmo Bresenham 3D

```

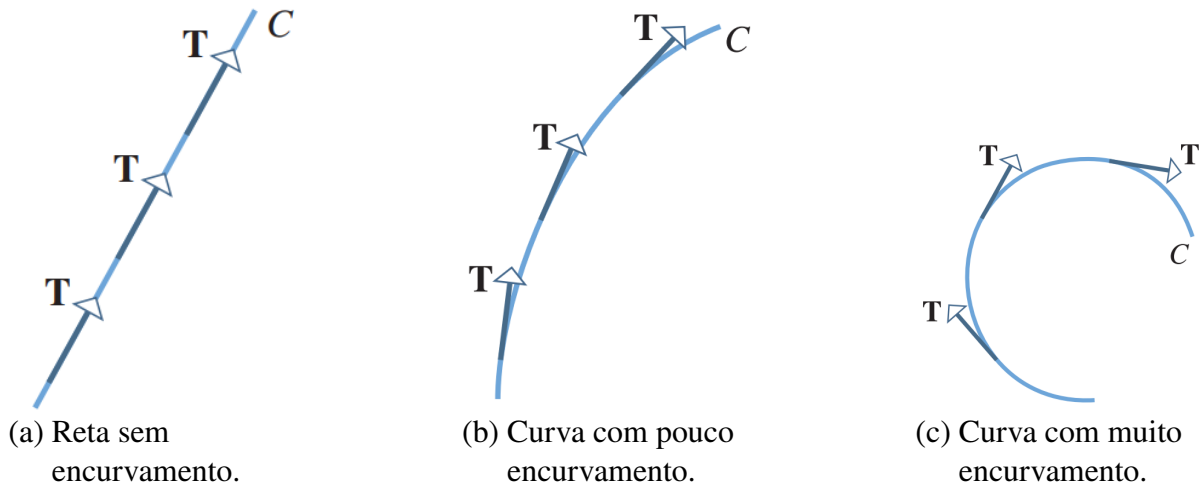
1 Input: Pontos inicial e final  $(x_0, y_0, z_0)$  e  $(x_1, y_1, z_1)$ 
2 Output: Lista de voxels da linha 3D
3  $\Delta x \leftarrow |x_1 - x_0|$ ,  $\Delta y \leftarrow |y_1 - y_0|$ ,  $\Delta z \leftarrow |z_1 - z_0|$ 
4  $s_x \leftarrow \text{sign}(x_1 - x_0)$ ,  $s_y \leftarrow \text{sign}(y_1 - y_0)$ ,  $s_z \leftarrow \text{sign}(z_1 - z_0)$ 
5 if  $\Delta x \geq \Delta y$  and  $\Delta x \geq \Delta z$  then
6    $p_y \leftarrow 2\Delta y - \Delta x$ ;  $p_z \leftarrow 2\Delta z - \Delta x$ 
7   for  $i = 0$  to  $\Delta x$  do
8      $\text{plot}(x_0, y_0, z_0)$ 
9     if  $p_y \geq 0$  then
10       $y_0 \leftarrow y_0 + s_y$ ;  $p_y \leftarrow p_y - 2\Delta x$ 
11    end
12    if  $p_z \geq 0$  then
13       $z_0 \leftarrow z_0 + s_z$ ;  $p_z \leftarrow p_z - 2\Delta x$ 
14    end
15     $p_y \leftarrow p_y + 2\Delta y$ ;  $p_z \leftarrow p_z + 2\Delta z$ ;  $x_0 \leftarrow x_0 + s_x$ 
16  end
17 end
18 else if  $\Delta y \geq \Delta x$  and  $\Delta y \geq \Delta z$  then
19    $p_x \leftarrow 2\Delta x - \Delta y$ ;  $p_z \leftarrow 2\Delta z - \Delta y$ 
20   for  $i = 0$  to  $\Delta y$  do
21      $\text{plot}(x_0, y_0, z_0)$ 
22     if  $p_x \geq 0$  then
23       $x_0 \leftarrow x_0 + s_x$ ;  $p_x \leftarrow p_x - 2\Delta y$ 
24    end
25    if  $p_z \geq 0$  then
26       $z_0 \leftarrow z_0 + s_z$ ;  $p_z \leftarrow p_z - 2\Delta y$ 
27    end
28     $p_x \leftarrow p_x + 2\Delta x$ ;  $p_z \leftarrow p_z + 2\Delta z$ ;  $y_0 \leftarrow y_0 + s_y$ 
29  end
30 end
31 else
32    $p_x \leftarrow 2\Delta x - \Delta z$ ;  $p_y \leftarrow 2\Delta y - \Delta z$ 
33   for  $i = 0$  to  $\Delta z$  do
34      $\text{plot}(x_0, y_0, z_0)$ 
35     if  $p_x \geq 0$  then
36       $x_0 \leftarrow x_0 + s_x$ ;  $p_x \leftarrow p_x - 2\Delta z$ 
37    end
38    if  $p_y \geq 0$  then
39       $y_0 \leftarrow y_0 + s_y$ ;  $p_y \leftarrow p_y - 2\Delta z$ 
40    end
41     $p_x \leftarrow p_x + 2\Delta x$ ;  $p_y \leftarrow p_y + 2\Delta y$ ;  $z_0 \leftarrow z_0 + s_z$ 
42  end
43 end
44 return lista de voxels

```

2.7 CURVATURA DE FUNÇÃO VETORIAL PARAMÉTRICA

Pode-se dizer que a curvatura de um ponto em uma curva paramétrica representa a taxa de variação de sua direção ao longo do comprimento do arco da curva. De acordo com Anton, Bivens e Davis (2014), considerando o gráfico de uma curva parametrizada C , representada por uma função vetorial, a sua curvatura está relacionada com $\frac{\partial T}{\partial s}$, sendo a taxa de variação de direção do vetor tangente unitário T em relação ao comprimento do arco s . A Figura 7 ilustra curvas com diferentes curvaturas. Quando C é uma reta, como na Figura 7a, a direção de T se mantém constante e a curvatura é nula. Na Figura 7b, a direção de T muda gradualmente e C se encurva levemente. Quando C possui um grande encurvamento, a direção de T varia com grande intensidade, como na Figura 7c.

Figura 7 – Ilustração de curvas com diferentes encurvamentos.



Fonte: Anton, Bivens e Davis (2014).

Sendo $r(t)$ uma função vetorial no espaço bi ou tri-dimensional, dado qualquer valor de t onde $T'(t)$ e $r''(t)$ existam, a curvatura de $r(t)$, denotada por $\kappa = \kappa(t)$, pode ser definida pela Equação(4). Considerando que $T(t)$ é definida pela Equação(5), $\kappa(t)$ pode ser reescrita pela Equação(6). Considerando que $r(t) = (x(t), y(t), z(t))$ é uma função vetorial tri-dimensional, a curvatura $k(t)$ é definida pela Equação(7).

$$\kappa(t) = \frac{||T'(t)||}{||r'(t)||} \quad (4)$$

$$T(t) = \frac{r'(t)}{||r'(t)||} \quad (5)$$

$$\kappa(t) = \frac{||r'(t) \times r''(t)||}{||r'(t)||^3} \quad (6)$$

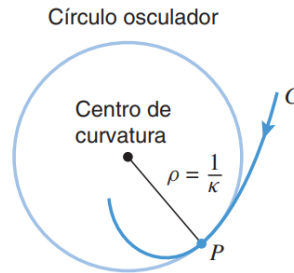
$$\kappa(t) = \frac{\sqrt{(z''(t)y'(t) - y''(t)z'(t))^2 + (x''(t)z'(t) - z''(t)x'(t))^2 + (y''(t)x'(t) - x''(t)y'(t))^2}}{(x'(t)^2 + y'(t)^2 + z'(t)^2)^{\frac{3}{2}}} \quad (7)$$

2.7.1 Raio de uma Curvatura

Geralmente, para uma curva C que possui curvatura κ em um ponto P , existe um círculo que possui uma tangente em comum com C em P e tem como centro o lado côncavo da curva em P (Anton; Bivens; Davis, 2014). Esse círculo é chamado de círculo osculador e seu raio ρ é definido pela Equação (8), sendo o círculo que melhor aproxima a curva C na vizinhança de P . A Figura 8 ilustra um círculo osculador de um ponto P em uma curva C . O raio ρ de um círculo osculador em P é chamado de raio de curvatura e o centro do círculo é chamado de centro de curvatura em P .

$$\rho = \frac{1}{\kappa} \quad (8)$$

Figura 8 – Ilustração de um círculo osculador e seu raio de curvatura.



Fonte: Anton, Bivens e Davis (2014).

O raio de uma curvatura ρ também pode ser calculado em função do comprimento do arco da curva A , e do seu ângulo de curvatura, G_C , seguindo a Equação(9) (Punmia, 2005).

$$\rho = \frac{180^\circ A}{\pi G_c} \quad (9)$$

2.8 FILTRO SAVITZKY-GOLAY

O filtro Savitzky-Golay foi proposto por Savitzky e Golay (1964) com o objetivo de suavizar sinais e preservar suas características de forma, como picos e curvaturas, sem introduzir atrasos de fase significativos. O método baseia-se no ajuste de um polinômio de baixa ordem a uma janela móvel de pontos sucessivos, utilizando o método dos mínimos quadrados.

Ao invés de aplicar uma média simples, o filtro ajusta localmente um polinômio de grau k sobre $2m+1$ amostras, minimizando o erro quadrático entre o polinômio e os valores observados. O valor suavizado $\hat{y}_i$ em cada ponto i é obtido conforme a Equação 10:

$$\hat{y}_i = \sum_{j=-m}^m c_j y_{i+j}, \quad (10)$$

em que c_j são os coeficientes calculados a partir da matriz de Vandermonde do polinômio de ajuste, e y_{i+j} são as amostras da série original.

2.9 B-SPLINE

As curvas B-Spline (*Basis Splines*) são amplamente utilizadas para modelagem de trajetórias suaves e contínuas em aplicações de computação gráfica e planejamento de movimento (Schoenberg, 1946). Uma B-Spline é uma combinação linear de funções base polinomiais ponderadas por pontos de controle, sendo definida de forma paramétrica pela Equação 11:

$$C(u) = \sum_{i=0}^n N_{i,p}(u) P_i, \quad u \in [0,1], \quad (11)$$

onde P_i são os pontos de controle, $N_{i,p}(u)$ são as funções base de B-Spline de grau p , e u é o parâmetro da curva. As funções base são definidas recursivamente conforme o algoritmo de De Boor (De Boor; De Boor, 1978).

A interpolação B-Spline garante continuidade de primeira e segunda derivada (C^1 e C^2), o que a torna ideal para representar trajetórias suaves em espaços tridimensionais.

2.10 FILTRO LAPLACIANO

O filtro Laplaciano é uma técnica baseada em operadores diferenciais de segunda ordem, tradicionalmente utilizada em processamento de imagens para realce de bordas e detecção de variações abruptas de intensidade. No entanto, quando aplicado de forma inversa (com sinal negativo ou suavização iterativa), o operador pode ser usado para eliminar flutuações locais e suavizar trajetórias discretas.

O operador Laplaciano discreto unidimensional é definido conforme a Equação 12:

$$\nabla^2 y_i = y_{i-1} - 2y_i + y_{i+1}. \quad (12)$$

A suavização Laplaciana é então obtida atualizando iterativamente cada ponto y_i em direção à média de seus vizinhos, conforme a Equação 13:

$$y_i^{(t+1)} = y_i^{(t)} + \alpha \nabla^2 y_i^{(t)}, \quad (13)$$

onde α é um parâmetro de relaxação que controla a intensidade da suavização.

2.11 REDES NEURAIS ARTIFICIAIS

As Redes Neurais Artificiais (RNAs) são técnicas de aprendizado de máquina inspiradas em modelos de neurônios biológicos, simulando a conexão e troca de sinais entre eles. Essa inspiração apareceu inicialmente no trabalho de McCulloch e Pitts (1943), sendo proposto um modelo de neurônio artificial que possuía apenas uma saída gerada pela soma das entradas multiplicadas por um conjunto de pesos, com o intuito de simular o processo de sinapse.

2.11.1 *Perceptron*

Baseando-se no trabalho de McCulloch e Pitts (1943), outro modelo de neurônio, chamado de *Perceptron* foi proposto por Rosenblatt (1958), sendo aplicado para classificação binária. Dado um vetor de entrada x , é feita a soma do produto de x por um vetor de pesos w , e então o resultado é passado para uma função de ativação $f(x)$, definida pela Equação (14).

$$f(x) = \begin{cases} 1 & \text{se } w \cdot x + b \geq 0 \\ 0 & \text{caso contrário} \end{cases}, \quad (14)$$

sendo b um valor constante chamado de *bias*, adicionando um viés no aprendizado do *perceptron*.

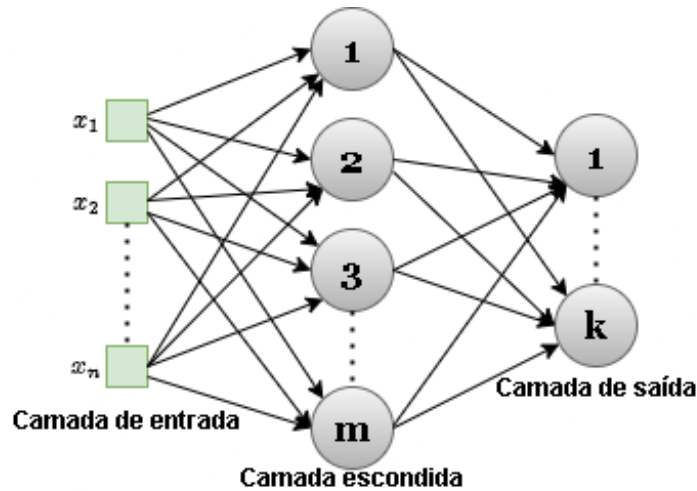
Foi apontado por Minsky e Papert (1969) que o *perceptron* tinha limitações em resolver problemas de classificação que não são linearmente separáveis, além de não conseguir trabalhar com classificação não-binária. Para resolver esses problemas, surgiram as redes neurais *perceptron* multicamadas.

2.11.2 Rede Neural *Perceptron* Multicamadas

A rede neural *perceptron* multicamadas foi proposta por Rumelhart, Hinton e Williams (1986) aprimorando o modelo neural de *perceptron*, apresentado na Seção 2.11.1. A rede *perceptron* multicamadas possui 1 camada de entrada, N camadas escondidas e 1 camada de saída. Também é chamada de rede totalmente conectada por ter todos os neurônios de uma camada conectados com os da próxima camada. Além disso, é uma rede do tipo *feed-forward* por não possuir ciclos e propagar as informações para uma única direção. A Figura 9 apresenta um exemplo de uma rede neural *perceptron* multicamadas com n entradas, m neurônios na camada escondida e k saídas.

As conexões entre os neurônios possuem um peso que pondera os dados conforme eles são propagados. A rede recebe dados numéricos na camada de entrada, sendo somados, multiplicados pelos pesos das conexões e aplicada a função de ativação antes de passar o resultado

Figura 9 – Ilustração de uma rede neural *perceptron* multicamadas.



Fonte: Autor.

para os neurônios das próximas camadas. O resultado final obtido pela rede são os dados após propagarem pela camada de saída.

Para a rede aprender a classificar os dados corretamente, é necessário encontrar o conjunto de pesos que melhor ajuste a saída de acordo com o valor real esperado. Um dos algoritmos mais populares para realizar essa otimização é o algoritmo de Retropropagação, também conhecido como *Backpropagation*.

2.11.3 Retropropagação

O algoritmo de Retropropagação foi apresentado por Rumelhart, Hinton e Williams (1986) para treinar redes neurais *perceptron* multicamadas através da retro-propagação dos erros da camada de saída para as camadas intermediárias. Dessa forma, a rede aprende de forma iterativa a ajustar os pesos que resultam na saída mais próxima do resultado esperado.

Esse processo de aprendizado possui dois estágios:

- Feed-Forward*: Os dados de entrada são propagados pela rede até a camada de saída.
- Feed-Backward*: Com a saída da rede, é calculado o erro e retro propagado da camada de saída até a primeira camada escondida, ajustando os pesos entre os neurônios.

A saída da camada de saída da rede $\hat{y}$ é definida pela Equação (15):

$$\hat{y}(x) = W^{(N)T} h^{(N-1)} + b^{(N)}, \quad (15)$$

sendo que $h^{(N-1)}$ a saída da última camada escondida, $g^{(N-1)}$ a função de ativação da camada de saída, $W^{(N-1)}$ é a matriz de pesos da camada de saída, x é o vetor de entrada da rede neural, N o número de camadas escondidas da rede e $b^{(N-1)}$ é a matriz de *bias* da camada de saída.

Por ser um aprendizado supervisionado, após a entrada ser propagada pela rede neural e a saída $\hat{y}$ ser gerada, ela é comparada com um vetor de respostas y para calcular uma função de custo $C(y, \hat{y})$ que representa a qualidade da rede em realizar previsões. O objetivo do treinamento é minimizar função de custo, ajustando os pesos da rede para encontrar o ponto mínimo de C . Ou seja, calcular derivada de C em relação ao conjunto de pesos e *bias* da rede neural $\frac{\partial C}{\partial \theta}$.

O algoritmo de retropropagação calcula primeiro a derivada na camada de saída da rede. Esse resultado é passado para o cálculo da derivada da penúltima camada, que é armazenado e reutilizado para antepenúltima rede e assim sucessivamente. Assim, o erro é retropropagado da última camada até a primeira camada da rede, utilizando o último cálculo de derivada para calcular a derivada da camada atual e atualizando os seus pesos por sucessivas iterações, seguindo a Equação (16):

$$\theta^{t+1} = \theta^t - \alpha \frac{\partial C}{\partial \theta}, \quad (16)$$

onde α é a taxa de aprendizado que define o quanto a função de custo é variada, θ^t o conjunto de pesos e *obias* na iteração t .

2.12 FUNÇÃO DE ATIVAÇÃO

Uma função de ativação determina a saída dos neurônios de uma rede neural. Também ajudam com a normalização de valores em intervalos de 0 a 1 ou -1 a 1. Outro ponto importante é a capacidade representativa que é dada para as RNAs por meio de um componente não linear, permitindo o aprendizado além de relações lineares entre variáveis dependentes e independentes. Uma das funções de ativação utilizadas é a Unidade Linear Retificada, mais conhecida como ReLU, descrita a seguir.

2.12.1 Unidade Linear Retificada

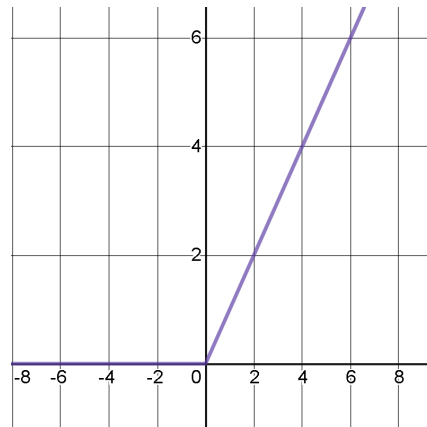
A Unidade Linear Retificada, ou ReLU, é uma função de ativação que ganhou destaque com o seu uso em redes neurais profundas, aumentando a velocidade de convergência da rede (Nair; Hinton, 2010). A função ReLU é definida pela Equação (17) e ilustrada na Figura 10.

$$\text{ReLU}(x) = \max(0, x) \quad (17)$$

2.13 REDES NEURAIS CONVOLUCIONAIS

As redes neurais convolucionais, conhecidas como *Convolutional Neural Network* (CNN), são tipos de RNAs propostas por LeCun *et al.* (1998) para aplicações na área de visão computacional possibilitando o aprendizado de aspectos importantes da imagem e diferenciá-los entre

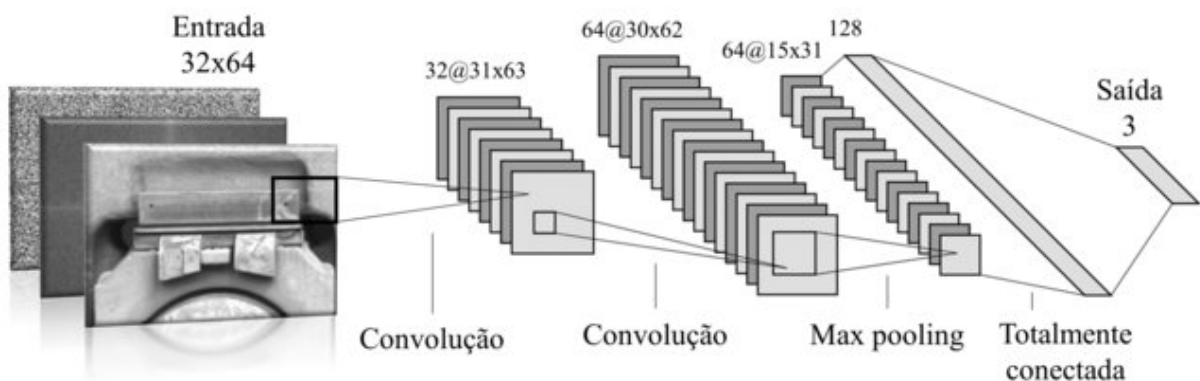
Figura 10 – Função ReLU



Fonte: Autor.

as imagens de classes diferentes. Isso ocorre porque a CNN tem a capacidade de otimizar os filtros e características que melhor auxilia na classificação das imagens.

A CNN também possui múltiplas camadas com múltiplos neurônios, assim como a rede *perceptron* multicamadas, apresentada na Seção 2.11.2, porém, ela recebe como entrada matrizes tridimensionais que podem representar imagens. Uma CNN tradicional pode ser dividida em duas partes: a primeira parte é responsável por aprender a extrair características da imagem de entrada e a segunda aprende a classificar corretamente, baseando-se nas características extraídas. A primeira parte é normalmente composta de camadas de convolução, *pooling* e *flatten*, que serão melhor explicadas posteriormente (Goodfellow; Bengio; Courville, 2016). A segunda parte é feita através de uma rede totalmente conectada, conforme explicada na Seção 2.11.2. A Figura 11 apresenta um exemplo da arquitetura de uma rede do tipo CNN.

Figura 11 – Exemplo de uma CNN com duas camadas de convolução e uma camada de *pooling*

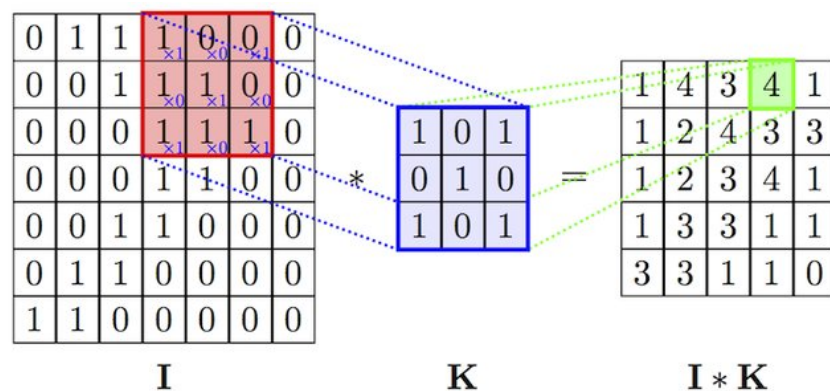
Fonte: Rocha *et al.* (2019).

2.13.1 Camada convolucional

Na camada convolucional é feita a convolução 2D sobre as imagens de entrada, seguindo o mesmo conceito explicado na Seção 2.5. Assim, nessa camada são aplicados filtros espaciais com o objetivo de extrair características, como as bordas, da imagem. Os filtros percorrem toda a matriz de entrada, gerando uma matriz de saída chamada de mapa de características. Esse mapa de características passa por uma função de ativação e o resultado é passado para as próximas camadas.

Nesse procedimento, são definidos dois parâmetros importantes: *stride*, sendo o tamanho do passo em píxel que o filtro percorre a imagem. Outro parâmetro é o *padding*, que define o preenchimento da matriz com o valor zero durante a operação de convolução, podendo resultar na diminuição do mapa de características. A Figura 12 ilustra o processo de convolução 2D em uma matriz.

Figura 12 – Exemplo de uma convolução 2D em uma matriz



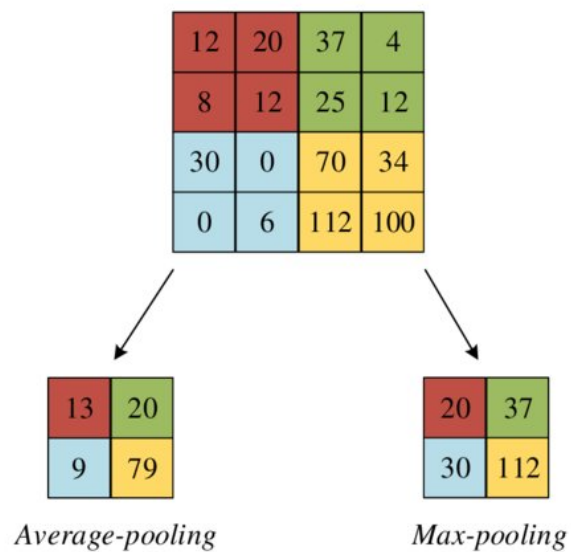
Fonte: Mohamed (2017).

2.13.2 Camada de *pooling*

A camada de *pooling* normalmente é responsável por reduzir o tamanho do mapa de características recebido da camada de convolução. Essa é uma etapa importante para reduzir a dimensionalidade dos dados, reduzindo o poder computacional necessário, enquanto mantém características dominantes que estão presentes nos dados. O *pooling* realiza um processo semelhante à convolução, percorrendo o mapa de características com uma matriz de tamanho pré-definido, filtrando os dados de acordo com os critérios de *pooling* adotados na arquitetura da rede.

Existem dois tipos tradicionais de *pooling*. O *max pooling*, onde é selecionado o maior valor dentro da matriz de *pooling*, e o *average pooling*, onde é realizada a média de todos os elementos da matriz. A Figura 13 apresenta um exemplo da aplicação do filtro de *max pooling* e *average pooling*.

Figura 13 – Exemplo de uma operação de *max pooling* e *average pooling*, considerando um filtro de *pooling* de tamanho 2×2 .

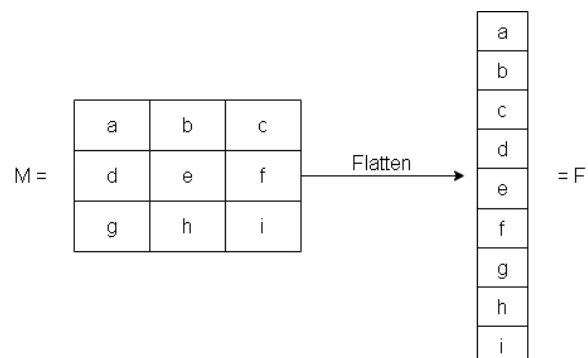


Fonte: Mohamed (2017).

2.13.3 Camada de *flatten*

A camada de *flatten* consiste em transformar a saída da camada de seleção de característica em forma matricial para um vetor. A Figura 14 apresenta um exemplo da operação realizada na camada de *flatten*. A saída vetorial da camada *flatten* é então utilizada na etapa de classificação, servindo como um vetor de características de entrada para uma rede totalmente conectada.

Figura 14 – Exemplo de uma operação de *flatten*, transformando uma matriz M em um vetor F .



Fonte: Autor.

2.13.4 Camada da rede totalmente conectada

Geralmente, a última camada da CNN é composta por uma rede totalmente conectada para realizar a classificação dos dados de entrada. Nesta camada é utilizada uma rede perceptron multicamadas, conforme explicada na Seção 2.11.2, que é alimentada com o vetor de características gerado na camada *flatten*. Assim, a CNN termina de processar os dados, obtendo como saída a classificação pela rede totalmente conectada.

2.14 U-NET

U-Net é uma arquitetura de CNN (Seção 2.13) desenvolvida por Ronneberger, Fischer e Brox (2015) para segmentação de imagens médicas. Ela obteve ótimo desempenho em desafios de segmentação de imagem e é uma das arquiteturas mais populares para segmentação semântica.

Essa rede recebe como entrada uma imagem e o seu mapa de segmentação, com as regiões rotuladas de acordo com a segmentação desejada. Os rótulos normalmente são atribuídos por valores de píxel, onde os pixels de mesmo valor fazem parte da mesma classe. A Figura 15 apresenta um exemplo de uma imagem original e a sua segmentação semântica esperada. Dessa forma, o objetivo da rede *U-Net* é prever um mapa de segmentação dada uma imagem de entrada.

Figura 15 – Exemplo de imagens de entrada da rede *U-Net*. A imagem da esquerda é a imagem original e a da direita o seu mapa de segmentação esperado.



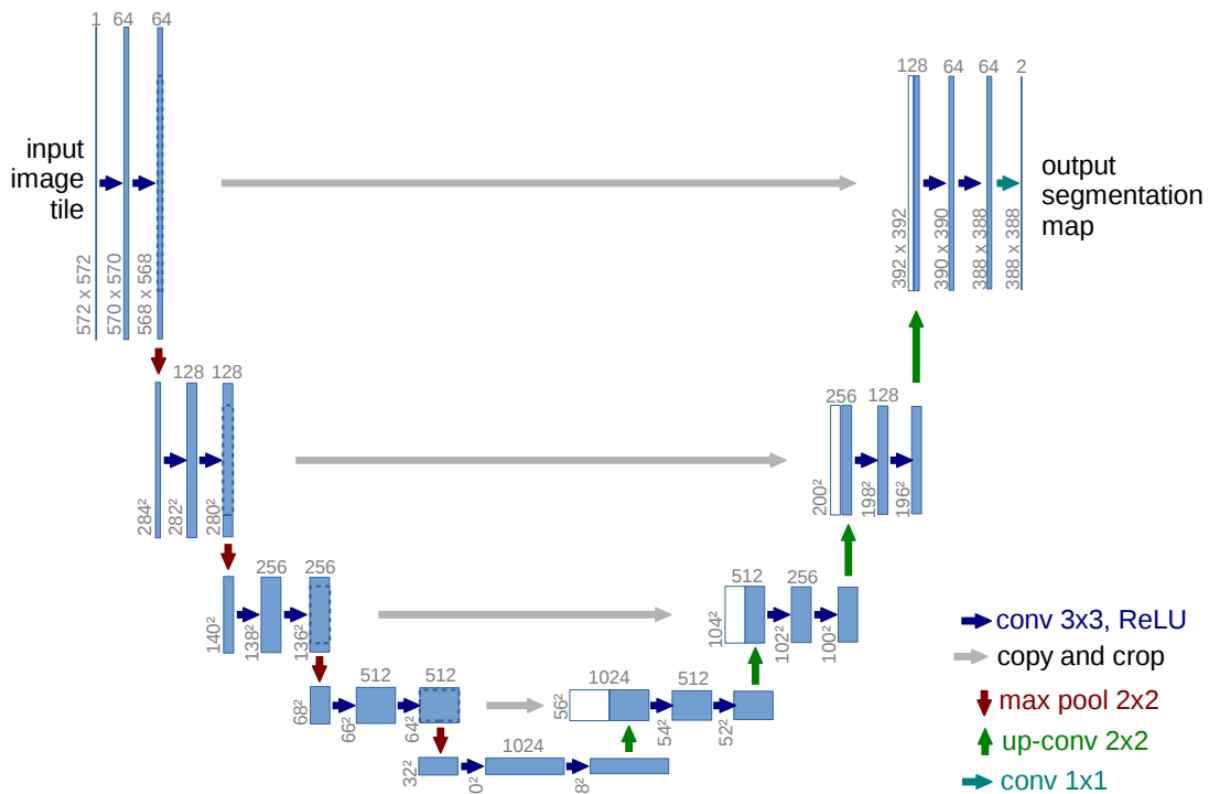
Fonte: Aston Zhang *et al.* (2021).

2.14.1 Arquitetura da *U-Net*

A Figura 34 apresenta a arquitetura da *U-Net*. Ela é composta por um caminho de contração e outro de expansão. O caminho de contração consiste na repetição da aplicação de duas camadas de convolução de tamanho 3×3 , sem *padding* e utilizando função ReLU como função de ativação, seguida de uma camada de *max pooling* de tamanho 2×2 com *stride* de tamanho 2 para reduzir o tamanho dos dados. Cada vez que os dados passam pela camada de *max pooling*,

o número de canais de características é dobrado. No caminho de expansão, é feito um aumento de amostras do mapa de características seguido por uma convolução de tamanho 2×2 que divide pela metade o número de canais de características, uma concatenação com um corte correspondente ao mapa de característica do caminho de contração para recuperar a perda das bordas em cada convolução, duas camadas de convolução 3×3 , seguida de uma função ReLU. Na última camada é aplicada uma convolução de tamanho 1×1 para mapear os 64 componentes do vetor de características para o número desejado de classes. Assim, a *U-Net* possui um total de 23 camadas de convolução.

Figura 16 – Arquitetura da *U-Net*.



2.15.1 Entropia Cruzada

A entropia cruzada, é uma métrica que mede a entropia relativa entre duas distribuições de probabilidade. Ela pode ser utilizada para definir uma função de custo em algoritmos de aprendizado de máquina, sendo definida pela Equação (18) (Bishop, 2006).

$$CE = - \sum_i^N y \log \hat{y} + (1 - y) \log(1 - \hat{y}), \quad (18)$$

onde y é a saída esperada e $\hat{y}$ é a saída estimada.

2.15.2 Distância Euclidiana

A distância euclidiana mede a distância entre dois pontos $X = (x_1, x_2, \dots, x_n)$ e $Y = (y_1, y_2, \dots, y_n) \in R^n$. Ela pode ser representada pela Equação (19).

$$\sqrt{\sum_{i=0}^n (x_i - y_i)^2} \quad (19)$$

2.15.3 Acurácia

A acurácia avalia o desempenho geral de um modelo de classificação, medindo sua capacidade de prever corretamente amostras de ambas as classes (positiva e negativa na classificação binária). A métrica de acurácia é representada pela Equação (20):

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (20)$$

em que TP (Verdadeiros Positivos) são o número de amostras positivas em que o modelo previu corretamente, TN (Verdadeiros Negativos) são o número de amostras negativas que o modelo previu corretamente, FP (Falsos Positivos) são o número de amostras em que o modelo previu incorretamente, e FN (Falsos Negativos) são o número de amostras em que o modelo previu incorretamente.

2.15.4 Precisão

A precisão mede a proporção de previsões positivas corretas em relação a todas as previsões positivas produzidas pelo modelo. Essa métrica considera o custo das classificações falsas positivas e quantifica o quanto o modelo produziu de previsões positivas precisas. A métrica de precisão é definida pela Equação (21):

$$\text{Precisão} = \frac{TP}{TP + FP} \quad (21)$$

2.15.5 Recall

O *recall*, ou sensibilidade, estima a taxa de previsão positiva verdadeira do modelo. Representa o número de amostras de positivas do conjunto de dados que o modelo classifica corretamente. O *recall* é definido pela Equação(22):

$$Recall = \frac{TP}{TP + FN} \quad (22)$$

2.15.6 F1-Score

Uma métrica útil para avaliar modelos de classificação é o *F1-Score*. Essa métrica é uma média harmônica entre precisão e recall e fornece uma medida equilibrada entre as duas métricas. É uma métrica para avaliar o desempenho do modelo considerando um equilíbrio entre falsos negativos e falsos positivos. Na análise realizada, o *F1-Score* indica a capacidade do modelo de detectar corretamente cálculos renais nas amostras de dados. O F1-score é definido pela Equação (23):

$$F1-score = \frac{2 * Precisão * Recall}{Precisão + Recall} \quad (23)$$

2.15.7 Coeficiente de Dice

O coeficiente de Dice foi proposto por Dice (1945) para medir a similaridade entre duas amostras. Dadas duas amostras representadas pelos conjuntos *A* e *B*, o coeficiente de Dice é o cálculo da área de interseção entre *A* e *B* dividida pela soma dos dois conjuntos, sendo representada na Equação (24). É uma métrica popular para avaliar segmentação de imagens utilizando uma máscara binária como padrão-ouro.

$$Dice(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (24)$$

2.15.8 Coeficiente de Jaccard

Outro coeficiente para medir a similaridade entre dois conjuntos e utilizado para avaliar segmentação de imagens é o de Jaccard, proposto por Jaccard (1912). Dado dois conjuntos *A* e *B*, o coeficiente de Jaccard é representado pela Equação (25), sendo a razão entre a interseção e a união dos dois conjuntos.

$$Jaccard(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (25)$$

2.16 VALIDAÇÃO CRUZADA *K-FOLD*

A validação cruzada *k-Fold* é uma técnica utilizada em aprendizado de máquina para avaliar a capacidade de generalização de modelos preditivos, ou seja, sua habilidade em produzir resultados precisos para dados não utilizados durante o treinamento (Kohavi *et al.*, 1995).

O princípio das técnicas de validação cruzada consiste em particionar a base de dados em subconjuntos mutuamente exclusivos, de forma que uma parte seja utilizada para o treinamento do modelo e outra para sua validação. Essa abordagem reduz o viés associado a uma única divisão de dados e proporciona uma estimativa mais generalizada do desempenho médio do modelo.

O método *k-Fold* destaca-se por sua simplicidade, eficiência computacional e por produzir estimativas de erro menos enviesadas quando comparado a técnicas como o método *hold-out*. O procedimento é composto pelas seguintes etapas:

- a) Embaralhar aleatoriamente a base de dados.
- b) Dividir o conjunto de dados em k subconjuntos de tamanhos aproximadamente iguais.
- c) Para cada subconjunto:
 - Definir o subconjunto atual como base de validação.
 - Utilizar os demais subconjuntos como base de treinamento.
 - Treinar o modelo com os dados de treinamento e avaliá-lo com os dados de validação.
 - Armazenar as métricas de desempenho obtidas.
- d) Calcular o desempenho médio do modelo a partir das métricas obtidas em todas as iterações.

2.17 VALIDAÇÃO CRUZADA DE MONTE CARLO

A validação cruzada de Monte Carlo é uma variação flexível dos métodos tradicionais de particionamento, utilizada para avaliar a estabilidade e a capacidade de generalização de modelos de aprendizado de máquina (Xu; Liang, 2001).

Diferentemente da validação *k-Fold*, que realiza divisões fixas e mutuamente exclusivas, o método de Monte Carlo gera múltiplas divisões aleatórias e independentes da base de dados em subconjuntos de treinamento e validação. Cada divisão é usada para treinar e testar o modelo, e o processo é repetido diversas vezes. Essa repetição permite obter uma estimativa mais estável e confiável do desempenho médio do modelo, especialmente em bases de dados de pequeno porte.

O procedimento da validação cruzada de Monte Carlo pode ser descrito pelas seguintes etapas:

- a) Definir a proporção dos dados destinados ao treinamento e à validação (por exemplo, 80% e 20%).

- b) Repetir N vezes o processo:
 - Embaralhar aleatoriamente a base de dados.
 - Selecionar aleatoriamente os conjuntos de treinamento e validação conforme a proporção definida.
 - Treinar o modelo no conjunto de treinamento e avaliá-lo no conjunto de validação.
 - Armazenar as métricas de desempenho obtidas.
- c) Calcular a média e o desvio padrão das métricas ao longo das N repetições.

A principal vantagem desse método é a redução da dependência em uma única divisão de dados, tornando a avaliação mais robusta a variações no conjunto de dados.

3 CLASSIFICAÇÃO DE IMAGENS PARA DETECÇÃO AUTOMÁTICA DE CÁLCULOS URINÁRIOS EM TC

Neste capítulo é apresentado um novo método para classificação de cálculos renais em exames de TC baseado na combinação de bases de dados coletadas de fontes distintas. O método se baseia em Redes Neurais Convolucionais, utilizando uma abordagem híbrida que combina dois modelos treinados com bases de dados diferentes, chamado de DE-MLP (*Dataset Ensemble MLP*). O modelo proposto foi validado e comparado com outros modelos semelhantes do estado-da-arte, alcançando resultados promissores.

3.1 CONTEXTUALIZAÇÃO

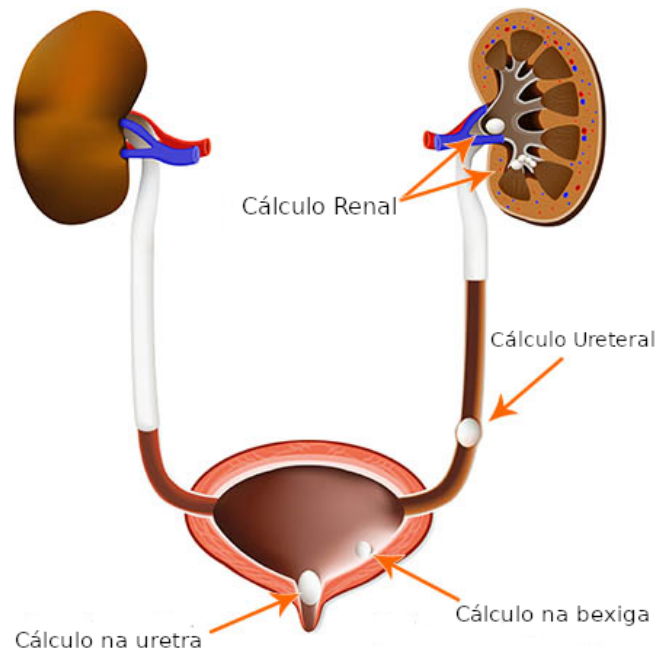
A urolitíase, conhecida popularmente como cálculos renais ou pedras nos rins, é uma doença comum com aumento em sua taxa de incidência nas últimas décadas pelo mundo inteiro (Sorokin *et al.*, 2017). Essa doença é caracterizada pela agregação de minerais da urina, formando substâncias sólidas chamadas de cálculos, localizados no sistema urinário. A Fig. 17 ilustra exemplos de cálculos urinários. Os cálculos podem causar obstrução do fluxo urinário, levando a dores severas e danos permanentes aos órgãos do paciente (Ingimarsson; Krambeck; Pais, 2016). A rápida detecção e seu diagnóstico precoce são importantes para prevenir que a formação destes cálculos prejudique gravemente os pacientes, além de permitir a elaboração de uma estratégia de tratamento eficiente que possa ser iniciada o mais breve possível (Rasulova; Aminova; Ismailova, 2023).

A utilização de exames de imagens se tornou comum para diagnosticar pacientes com suspeitas de cálculos urinários. O exame de TC (Hounsfield, 1973) é o principal exame de coleta de imagens para diagnóstico de cálculos. Através da combinação de múltiplas imagens de radiografia capturadas por um equipamento de raio-x, obtêm-se imagens que fornecem informações detalhadas dos órgãos do paciente, permitindo a manipulação e visualização 3D do interior do paciente, além de aplicações importantes para melhorar o diagnóstico e planejamento de tratamento. A Fig. 18 apresenta um exemplo de um exame de tomografia visualizado em um espaço 3D. O exame de TC é considerado o padrão-ouro no diagnóstico de urolitíase por sua velocidade, facilidade de coleta de imagem, detalhamento dos órgãos internos do paciente e por permitir o diagnóstico de outras anormalidades no sistema urinário (Andrabi *et al.*, 2015).

A partir da TC, o médico realiza o diagnóstico de urolitíase detectando a presença do cálculo no paciente. Baseando-se nessas informações, ele decide a melhor estratégia de tratamento. No entanto, esse processo pode demorar pelo excesso de exames que precisam ser avaliados pelo médico. Além disso, erros no diagnóstico podem ocorrer devido à falta de experiência do médico responsável e à variação das características dos exames utilizados.

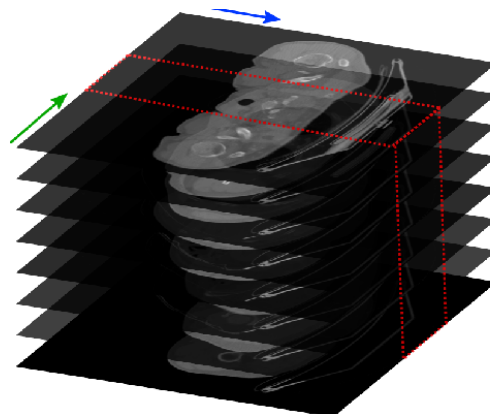
A IA tornou-se uma ferramenta benéfica para aplicações médicas, reduzindo os encargos clínicos, melhorando a precisão do diagnóstico e acelerando os processos de tratamento. Téc-

Figura 17 – Ilustração da presença de cálculos em diferentes partes de um sistema urinário humano.



Fonte: Adaptado de *Urologic Surgeons of Washington* ¹.

Figura 18 – Exemplo de fatias de uma tomografia computadorizada empilhadas em um espaço 3D.



Fonte: (Kanavati *et al.*, 2018b).

nicas baseadas em IA estão sendo incorporadas em sistemas assistidos por computador e são utilizadas em muitas aplicações, tais como assistência ao diagnóstico, educação médica (Younis *et al.*, 2024) e segmentação de imagens médicas (Ma *et al.*, 2024). Na urologia, a IA tem sido recentemente explorada em aplicações de cirurgia robótica e análise de imagens (Abid; Hussein; Guru, 2024), previsão de resultados cirúrgicos (Suarez-Ibarrola *et al.*, 2020) e, especialmente, detecção automática de cálculos renais usando aprendizado de máquina com base em exames de imagem (Panthier *et al.*, 2024). O desenvolvimento de um sistema automatizado de diag-

nóstico de cálculos renais baseado em técnicas de IA ganhou atenção para melhorar a eficiência diagnóstica.

Esse tipo de sistema poderia reduzir a carga de trabalho dos radiologistas, a variabilidade entre observadores e o tempo decorrido antes do início do tratamento adequado. Entre as abordagens baseadas em IA, o aprendizado de máquina surgiu como uma das abordagens mais eficazes para essa tarefa. Por meio do aprendizado supervisionado, esses modelos podem identificar automaticamente a presença de cálculos renais, melhorando a precisão do diagnóstico e auxiliando na tomada de decisões clínicas.

3.2 TRABALHOS RELACIONADOS

O desenvolvimento recente de técnicas de aprendizado profundo aplicadas à visão computacional tem beneficiado a área da medicina, obtendo sucesso em aplicações de sistemas de diagnóstico inteligente (Chan; Hadjiiski; Samala, 2020), (Hesamian *et al.*, 2019). Essas técnicas também estão sendo utilizadas para resolver problemas na área urológica, como o diagnóstico automático de cálculos urinários a partir de exames de imagens (Hameed *et al.*, 2021), análise da composição do cálculo (Kim, U. S. *et al.*, 2022), predição de tratamentos e de consequências cirúrgicas (Suarez-Ibarrola *et al.*, 2020).

Trabalhos anteriores exploraram técnicas de aprendizado de máquina para a detecção de cálculos renais, empregando diferentes algoritmos de classificação e métodos de extração de características para aumentar a robustez e a generalização. Soni e Rai (2020) aplicaram equalização de histograma e aprimoramento de bordas como etapas de pré-processamento antes de classificar cálculos renais usando uma Máquina de Vetor de Suporte (SVM) (Cortes; Vapnik, 1995), alcançando 98,71% de acurácia em um conjunto de dados de 156 imagens. Kolli *et al.* (2022) treinaram modelos de aprendizagem supervisionada em 12.446 imagens de TC, comparando Random Forest (Breiman, 2001), Regressão Logística e uma abordagem combinando as duas técnicas, alcançando 97,00% de acurácia. Mais recentemente, Hama, Majeed e Nari-man (2025) desenvolveram um modelo baseado em SVM usando o Padrão Binário Local (LBP) (Ojala; Pietikäinen; Harwood, 1996) para extração de características de textura, obtendo 96,31% de acurácia em um conjunto de dados de 6.454 imagens. Outros estudos exploraram redes neurais para a detecção de cálculos renais, como o de Akshaya *et al.* (2020), que aplicaram uma MLP para classificar cálculos renais em tomografias computadorizadas, usando a Transformada Discreta de Wavelet (Mallat, 1989) para pré-processamento e a Matriz de Coocorrência de Níveis de Cinza (GLCM) para extração de características, alcançando 98,8% de acurácia. Da mesma forma, Antony *et al.* (2020) (Antony *et al.*, 2020) aplicaram GLCM com um classificador MLP e aplicaram filtragem gaussiana e bilateral para suavização de imagens, obtendo 96,0% de acurácia. Esses estudos utilizaram ML clássico para a detecção de cálculos renais, demonstrando o potencial de diferentes técnicas de classificação e métodos de extração de características para a detecção de cálculos renais com base em exames de TC.

Técnicas tradicionais de aprendizado de máquina, como MLP, SVM e Random Forest, mostraram resultados promissores na detecção de cálculos renais. No entanto, técnicas de aprendizado profundo, particularmente CNNs, tornaram-se a abordagem mais amplamente adotada para tarefas de imagem médica, demonstrando desempenho excepcional no reconhecimento e análise de imagens médicas. As abordagens baseadas em CNN ganharam atenção como uma técnica de ponta para o diagnóstico automático de cálculos renais, dada sua capacidade de generalizar previsões em várias condições de imagem e alcançar alta precisão de classificação (Sulaksono; Adi *et al.*, 2023), (N *et al.*, 2024).

Vários estudos propuseram técnicas baseadas em CNN para essa tarefa. Yildirim *et al.* (2021a) aplicaram o XResNet-50 a 1799 imagens de TC adquiridas, alcançando 96,82% de acurácia. Da mesma forma, Baygin *et al.* (2022) introduziram o modelo ExDark19, uma arquitetura DarkNet19 modificada, que atingiu 99,71% de acurácia quando aplicada ao mesmo conjunto de dados de 1799 imagens de TC. Asif *et al.* (2023) (Asif *et al.*, 2023) desenvolveram o StoneNet, um modelo leve baseado em convoluções separáveis (MobileNet (Howard *et al.*, 2017)), alcançando 97,98% de acurácia. Patro *et al.* (2023) (Patro *et al.*, 2023) propuseram uma CNN com convoluções de Kronecker, que melhorou a redução da redundância de recursos e superou o modelo de aprendizado profundo ResNet-50, alcançando 98,50% de acurácia. Manudhas *et al.* (2024) (Manudhas; Manoranjitham, 2024) usaram o XResNet152 para classificar imagens de TC como normais ou com cálculos renais, alcançando 99,80% de acurácia.

Expandindo o escopo dos trabalhos anteriores, Islam *et al.* (2022) propuseram um sistema automatizado para detectar cistos, cálculos e tumores renais usando Vision Transformers e transferência de aprendizado. Eles coletaram 12.446 imagens de TC (3.709 cistos, 5.077 normais, 1.377 cálculos e 2.283 tumores) e testaram seis modelos, incluindo Vision Transformers (EANet, CCT, Swin) e CNNs (VGG16, ResNet50, Inception v3). O Swin Transformer alcançou a maior acurácia de 98,50% para cálculos. Seguindo por outra direção Caglayan *et al.* (2022) utilizaram a arquitetura xResNet50 em 2.959 imagens de TC de 455 pacientes (405 com cálculos). As imagens foram agrupadas por tamanho da pedra (0–1 cm, 1–2 cm, > 2 cm) e avaliadas nos planos axial, coronal e sagital. A acurácia variou de 63,00% (coronal) a 93,00% (sagital).

Outros estudos exploraram técnicas híbridas e de combinação de conjuntos para aprimorar modelos baseados em CNN. Asif, Zheng e Zhu (2024) projetaram modelos de conjunto combinando InceptionV3, InceptionResNetV2, MobileNet e Xception, alcançando 98,79% de acurácia. Chaki e Uçar (2024) propuseram o FINDWELL, um modelo híbrido que integra DarkNet19, InceptionV3 e ResNet101, alcançando 99,80% de acurácia em conjuntos de dados de alta qualidade. Liu e Ghadimi (2024) introduziram uma CNN otimizada pelo algoritmo Flexible Dwarf Mongoose Optimization (FDMO), alcançando 96,68% de acurácia. Huan Ding Qirui Huang (2025) combinaram AlexNet com Extreme Learning Machine (ELM), otimizado usando Firebug Swarm Optimization, alcançando 95,37% de acurácia.

Apesar de apresentarem resultados promissores para a detecção automática de cálculos urinários em TC, as técnicas de aprendizado de máquina profundo e redes neurais convoluci-

onais implementadas possuem uma alta complexidade, exigindo um custo computacional elevado. É notável que na literatura a quantidade de artigos que apresentam soluções por CNNs é significativamente maior que as diversas técnicas de aprendizado de máquina.

Além disso, as bases de dados utilizadas possuem pouca variabilidade, geralmente possuindo exames de pacientes coletados pelos mesmos equipamentos e que possuem cálculos renais com características semelhantes, dificultando a generalização dos modelos treinados para detectar cálculos em exames diferentes das amostras utilizadas no treinamento. Essa limitação indica enviesamento no treinamento dos modelos dos artigos citados, visto que alguns dos dados utilizados apresentavam pouca variabilidade.

Por esses motivos, a exploração de outras arquiteturas de aprendizado de máquina profundo de baixa complexidade que utilizam uma maior variabilidade de imagens de TC pode contribuir no desenvolvimento de sistemas de diagnóstico de cálculos urinários automáticos para acelerar o tratamento de pacientes que possuem essa patologia.

3.3 MATERIAIS E MÉTODOS

Este estudo apresenta um novo método para detecção automática de cálculos urinários que combina um par de bases de dados para treinar modelos CNN. A Figura 19 apresenta uma visão geral da metodologia proposta. A entrada consiste em dois conjuntos de dados de imagens de TC combinados, cada um contendo amostras rotuladas como cálculos renais e normais. Um pipeline de pré-processamento é aplicado a todas as imagens, incluindo redimensionamento, divisão de dados, normalização e aumento de dados. Posteriormente, dois modelos baseados em CNN são treinados independentemente em cada conjunto de dados. Em seguida, seus extratores de características são combinados para construir o modelo proposto chamado DE-MLP (*Dataset Ensemble Multilayer Perceptron*). Após o treinamento, o DE-MLP gera resultados de classificação e mapas de ativação, fornecendo interpretabilidade visual das previsões do modelo. Cada etapa do método proposto é detalhada nesta seção.

3.3.1 Base de Dados Combinada

Foram utilizadas duas bases de dados coletadas de diferentes fontes, referidos como Base de Dados A e Base de Dados B. Foi utilizada apenas a perspectiva coronal no treinamento da CNN porque ela oferece uma visualização anatômica eficaz dos rins e do trato urinário. Essa abordagem também está alinhada com as situações clínicas realistas, facilitando a interpretação médica e a anotação supervisionada dos dados (Andrabi *et al.*, 2015). Estudos anteriores e resultados preliminares demonstraram que a perspectiva do exame coronal é suficiente para a detecção precisa de cálculos renais (Yildirim *et al.*, 2021a; Baygin *et al.*, 2022; Manudhas; Manoranjitham, 2024). Portanto, as perspectivas axial e sagital não foram aplicadas para evitar o aumento do custo computacional. A base de dados combinada consiste originalmente em 3383

Figura 19 – Fluxograma da metodologia proposta neste trabalho.

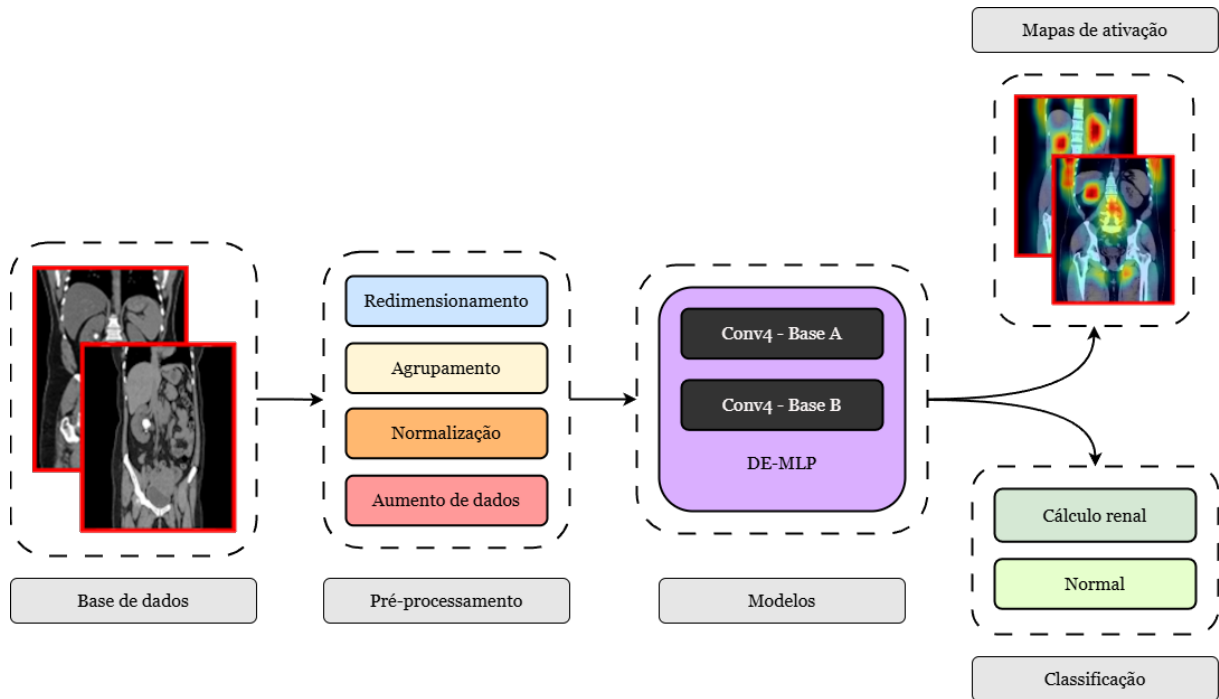


Figura 20 – Fonte: Autor

imagens de TC coronais (1304 com cálculos renais e 2079 sem cálculos renais) e é detalhada a seguir:

a) **Base de Dados A:**

Este conjunto de dados foi apresentado por Yildirim *et al.* (2021a) e está disponível publicamente no *GitHub*². As imagens foram obtidas de pacientes entre 18 e 80 anos no Hospital Municipal Elazığ Fethi Sekin, na Turquia (protocolo de TC 120 kV; corrente automática do tubo 100 mA–200 mA; detector de 128 mm; mAs 203; espessura do corte 5 mm). Cada exame foi analisado por um radiologista e um urologista, que classificaram as imagens como contendo cálculos renais ou normais. O conjunto de dados A consiste em 278 pacientes com cálculos renais e 165 indivíduos normais. Um total de 1799 imagens de TC foram coletadas, das quais 790 correspondem a amostras de cálculos renais e 1009 a indivíduos normais. A Figura 21 apresenta amostras para ambas as classes da base de dados A.

b) **Base de Dados B:**

Este conjunto de dados foi apresentado por Islam *et al.* (2022) e foi coletado do PACS (*Picture Archiving and Communication System*) de hospitais em Dhaka, Bangladesh. O conjunto de dados original inclui tomografias coronais, sagitais e axiais de estudos com e sem contraste, seguindo protocolos padronizados para imagens de abdômen inteiro e urogramas. Todas as informações de identificação do paciente e metadados foram removidos antes de converter as imagens DICOM

²https://github.com/yildirimoza/Kidney_stone_detection

Figura 21 – Exemplos de amostras da Base de Dados A.

Cálculo Renal

Normal



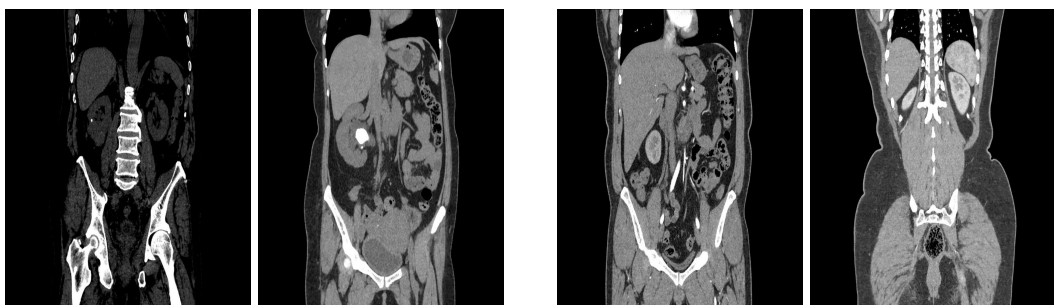
Fonte: Yildirim *et al.* (2021a)

para um formato JPEG sem perdas. O conjunto de dados original contém 12.446 imagens de tomografias computadorizadas axiais, coronais e sagitais, distribuídas da seguinte forma: 3.709 casos de cistos, 5.077 casos normais, 1.377 casos de cálculos renais e 2.283 casos de tumores. Está disponível publicamente no *Kaggle*³. Como este estudo se concentra apenas em cálculos renais, excluimos casos envolvendo tumores ou cistos e selecionamos apenas imagens de tomografia computadorizada coronal rotuladas como casos normais ou de cálculos renais. Após a exclusão, a Base de Dados B consistiu em 1.584 imagens, incluindo 514 amostras de cálculos renais e 1.070 amostras normais. A Figura 22 apresenta amostras de cálculos renais e normais do Conjunto de Dados B.

Figura 22 – Exemplos de amostras da Base de Dados B.

Cálculo Renal

Normal



Fonte: Baygin *et al.* (2022)

3.3.2 Pré-processamento

Antes de serem utilizadas como entrada para os modelos CNN, as imagens dos conjuntos de dados passam por um pipeline de pré-processamento. Todas as imagens são redimensiona-

³www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone

das para 224×224 pixels para normalizar as dimensões de entrada e garantir um treinamento eficiente do modelo.

Foi aplicada uma estratégia de agrupamento baseada na similaridade das imagens para reduzir o viés no treinamento, particularmente o risco de o modelo aprender padrões anatômicos específicos do paciente em vez de se concentrar nas características dos cálculos renais. As imagens foram agrupadas de acordo com o Índice de Similaridade Estrutural (SSIM) (Wang, Z. *et al.*, 2004), definido da seguinte forma:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (26)$$

onde μ_x e μ_y são os valores médios dos pixels das imagens x e y , respectivamente; σ_x^2 e σ_y^2 são as variâncias dos pixels das imagens x e y , respectivamente; σ_{xy} é a covariância entre os pixels das imagens x e y ; C_1 e C_2 são constantes para evitar instabilidades numéricas.

A métrica SSIM é utilizada como uma medida de similaridade de imagens para garantir que duas imagens com $SSIM \geq 0,75$ sejam atribuídas ao mesmo grupo. Esse processo garante que as imagens do mesmo paciente não apareçam em ambos conjuntos de treinamento e validação.

Em seguida, o conjunto de dados foi dividido com base nos grupos de pacientes: 80% dos grupos para treinamento e 20% para validação. Essa estratégia de divisão baseada no paciente garante que o modelo seja avaliado em dados anatomicamente distintos dos dados de treinamento, permitindo uma melhor generalização.

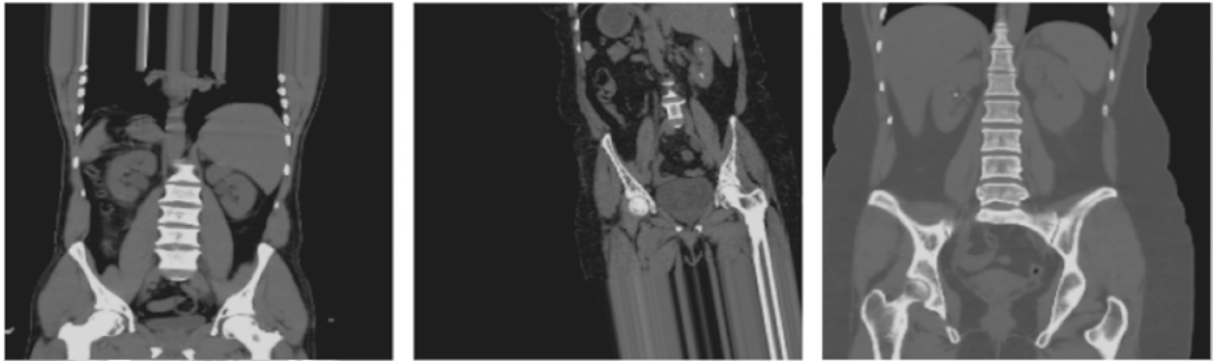
Para melhorar o desempenho da CNN e otimizar o processo de treinamento, a normalização da imagem é aplicada. Essa etapa transforma os valores de intensidade dos pixels da faixa original (0 a 255) em uma faixa normalizada de valores reais (0 a 1).

Para resolver o desequilíbrio de classes e expandir o conjunto de dados, são aplicadas técnicas de aumento de dados com base em imagens reutilizadas, gerando imagens aumentadas e aumentando o tamanho do conjunto de treinamento. Esta etapa é crucial para melhorar o desempenho de classificação e a capacidade de generalização do modelo. As seguintes transformações são usadas para cada lote de treinamento:

- a) Rotação: até 10° ;
- b) Deslocamento da largura: até 2%;
- c) Deslocamento da altura: até 2%;
- d) Faixa de ruído: até 2%;
- e) Zoom: até 2%;
- f) Inversão horizontal: ativada.

A Figura 23 apresenta exemplos de amostras após o aumento de dados.

Figura 23 – Exemplo de amostras após o aumento de dados.



Fonte: Autor

3.3.3 Modelo único (Conv4)

Neste estudo, foi utilizada uma arquitetura baseada em CNN denominada Conv4, anteriormente utilizada por Murillo Bouzon *et al.* (2023) para a detecção automática de cálculos renais em TCs. A arquitetura do Conv4 é ilustrada na Figura 24. O modelo Conv4 é uma arquitetura CNN que compreende quatro camadas convolucionais seguidas por três camadas totalmente conectadas. A estrutura do Conv4 é definida a seguir:

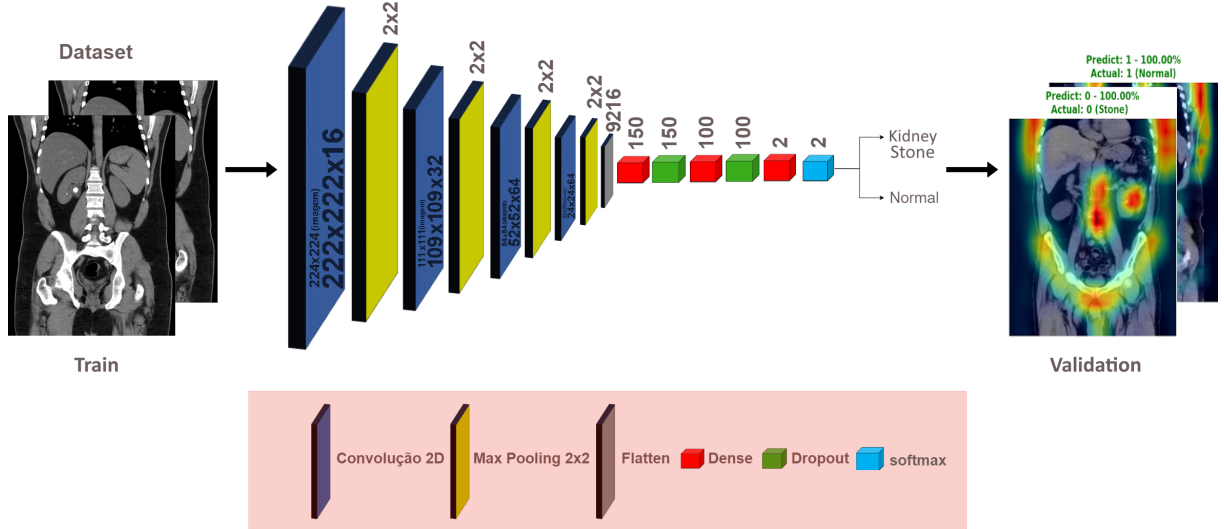
- a) Camada de entrada: A imagem de entrada é uma fatia coronal de 224×224 pixels de uma tomografia computadorizada (TC).
- b) Camadas convolucionais: O modelo Conv4 contém quatro camadas convolucionais com kernels 3×3 , seguidas por ativação ReLU e normalização em lote.
- c) Camada de Pooling (*Max pooling*): O Conv4 tem uma camada de *max pooling* 2×2 após cada camada convolucional.
- d) *Flatten*: Após as camadas convolucionais e de pooling, os mapas de características são transformados em um vetor unidimensional usando uma camada *flatten*. Esta etapa prepara as características extraídas para classificação por uma camada totalmente conectada.
- e) Camadas totalmente conectadas: A camada totalmente conectada do Conv4 consiste em três camadas MLP com 150, 100 e 2 neurônios, respectivamente.
- f) *Dropout*: A regularização do *dropout* é aplicada entre cada camada densa para evitar o sobreajuste, desativando aleatoriamente uma fração dos neurônios durante o treinamento.
- g) *Softmax*: A camada final usa uma função de ativação *softmax*, que converte a saída da rede em valores de probabilidade para cada classe (Cálculo renal vs. Normal), realizando a classificação final.

Neste trabalho, definimos os seguintes modelos Conv4:

- a) *Conv4A*: modelo Conv4 treinado com o conjunto de dados A.
- b) *Conv4B*: modelo Conv4 treinado com o conjunto de dados B.

- c) $Conv4(A + B)$: modelo Conv4 treinado com o conjunto de dados A e o conjunto de dados B combinados.

Figura 24 – Arquitetura Conv4.



Fonte: Autor

3.3.4 Modelo proposto DE-MLP

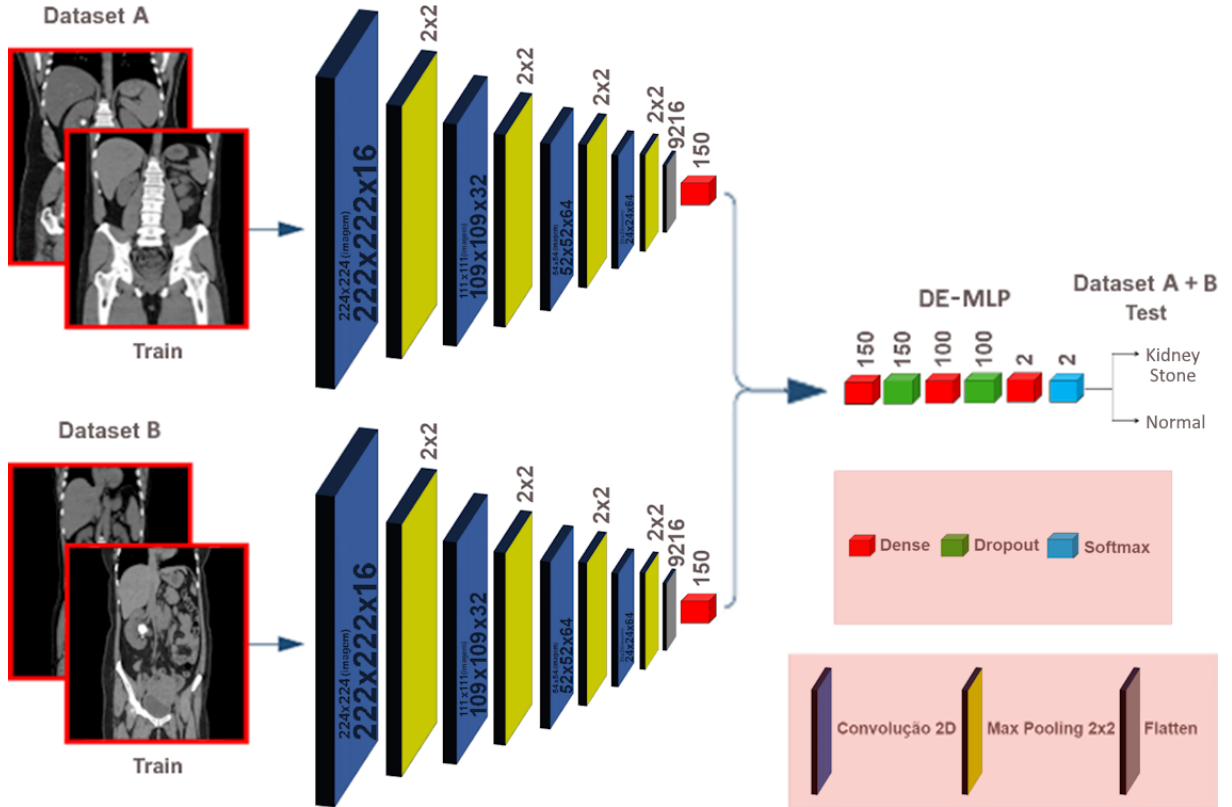
O DE-MLP proposto é uma arquitetura híbrida de aprendizado profundo projetada para aprimorar a classificação de cálculos renais por meio da combinação de vários conjuntos de dados de fontes distintas. O modelo emprega uma estratégia de fusão tardia, que extrai características de duas CNNs independentes baseadas em Conv4, cada uma treinada em conjuntos de dados distintos, e são concatenadas em um espaço de características de alto nível. As características combinadas são então processadas por meio de uma MLP totalmente conectada, que aprende padrões de ambas as fontes para melhorar o desempenho da classificação. A Figura 25 ilustra a arquitetura do DE-MLP proposto.

Seja $\mathbf{x}_A$ e $\mathbf{x}_B$ as imagens de entrada de dois conjuntos de dados distintos (Base de Dados A e Base de Dados B), contendo imagens coronais de TC. Cada conjunto de dados é processado independentemente por meio de um modelo Conv4, denominado $Conv4A(\mathbf{x}_A; \theta_A)$ e $Conv4B(\mathbf{x}_B; \theta_B)$, parametrizado por θ_A e θ_B , respectivamente. As camadas finais de cada Conv4 são removidas, com cada extrator de características $Conv4A$ e $Conv4B$ consistindo em quatro camadas convolucionais seguidas por *max pooling*, *flatten* e uma camada densa com 150 neurônios, transformando a representação espacial em um espaço vetorial F de tamanho $d = 150$. O processo de extração de características é definido como:

$$F_A = Conv4A(\mathbf{x}_A; \theta_A), \quad F_B = Conv4B(\mathbf{x}_B; \theta_B), \quad (27)$$

em que $F_A, F_B \in \mathbb{R}^d$ são os vetores de características extraídos de dimensionalidade d .

Figura 25 – Fluxograma da metodologia da arquitetura DE-MLP proposta neste trabalho.



Fonte: Autor

Os vetores de características extraídos de cada modelo Conv4 são concatenados para formar uma representação unificada:

$$F_{\text{concat}} = \text{concatenate}(F_A, F_B) \in \mathbb{R}^{2d} \quad (28)$$

O vetor de características concatenado F_{concat} é então passado por um classificador MLP, definido como:

$$M^{(i)} = \sigma \left(\mathbf{W}^{(i)} M^{(i-1)} + \mathbf{b}^{(i)} \right), \quad i \in \{1, 2, 3\} \quad (29)$$

sendo:

- a) $M^{(0)} = F_{\text{concat}}$ é a camada de entrada,
- b) $\mathbf{W}^{(i)} \in \mathbb{R}^{n_i \times n_{i-1}}$ e $\mathbf{b}^{(i)} \in \mathbb{R}^{n_i}$ são as matrizes de pesos e *bias* da camada i , respectivamente,
- c) $\sigma(\cdot)$ é a função de ativação ReLU para camadas ocultas,
- d) n_i é o número de neurônios na camada i , seguindo a estrutura:

$$n_1 = 150, \quad n_2 = 100, \quad n_3 = 2 \quad (30)$$

Para mitigar *overfitting*, camadas de *dropout* com probabilidade p são aplicadas no treinamento após cada camada densa:

$$M^{(i)} = \text{Dropout}(p, M^{(i)}) \quad (31)$$

A camada de saída é implementada por uma função de ativação *softmax*:

$$y = \text{Softmax}(\mathbf{W}^{(4)} M^{(3)} + b^{(4)} M^{(4)}), \quad (32)$$

sendo $y \in \mathbb{R}^2$ representa a distribuição de probabilidade nas classes Cálculo Renal e Normal.

O modelo é treinado minimizando a função de perda de entropia cruzada categórica, definida como:

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(\hat{y}_{i,c}), \quad (33)$$

em que N é o número de imagens no lote; C é o número de classes; $y_{i,c}$ é o rótulo padrão-ouro para a imagem i e a classe c (codificado como *one-hot*); $\hat{y}_{i,c}$ é a probabilidade prevista para a imagem i e a classe c .

Durante a primeira etapa, os modelos *Conv4A* e *Conv4B* são treinados independentemente usando seus respectivos conjuntos de dados, Base de Dados A e Base de Dados B. Cada modelo passa por um processo de aprendizado supervisionado, otimizando seus parâmetros para extrair características relevantes das imagens de TC para a detecção de cálculos renais. Após o treinamento e a avaliação, os modelos que alcançaram a maior precisão em seus conjuntos de validação são selecionados. Na segunda fase, esses modelos pré-treinados são usados como extratores de características para o DE-MLP. Em vez de treinar novamente toda a rede, apenas as camadas densas do DE-MLP são treinadas usando as representações de características obtidas dos modelos *Conv4_A* e *Conv4_B* com melhor desempenho. Essa abordagem de aprendizado por transferência melhora o desempenho da classificação, extraindo o conhecimento adquirido na fase inicial de treinamento, ao mesmo tempo em que evita o sobreajuste e reduz os custos computacionais. As principais vantagens do DE-MLP são as seguintes:

- a) Diversidade de recursos: a abordagem híbrida aprende com vários conjuntos de dados para melhorar a generalização.
- b) Redução de *overfitting*: várias fontes de dados aumentam a variabilidade dos dados e melhoram a confiabilidade.
- c) Melhor desempenho de classificação: o classificador MLP refina as representações de recursos extraídas de várias CNNs, alcançando resultados superiores.
- d) Aprendizado por transferência: reutilização de redes neurais já treinadas por outros pesquisadores em outros bancos de dados.

3.3.5 Configuração do treinamento e hiperparâmetros

Os experimentos foram realizados no Jupyter Lab utilizando uma NVIDIA GeForce RTX 3060, um AMD Ryzen 5 5500 e 16 GB de RAM. O conjunto de dados combinado foi dividido

em 80% para treinamento e 20% para validação. O método de validação cruzada de Monte Carlo foi aplicado para avaliar o desempenho do modelo em dados não vistos e generalizar os resultados em 10 execuções diferentes.

A validação cruzada de Monte Carlo foi escolhida ao invés vez de um esquema k-fold fixo porque permite uma proporção de divisão de validação arbitrária, o que é particularmente útil em casos como o conjunto de dados B, em que várias imagens são originárias do mesmo paciente. Ao atribuir aleatoriamente conjuntos completos de imagens de pacientes para treinamento ou validação em cada iteração, evita-se o risco de que imagens de um único paciente apareçam em ambas as divisões, um resultado que poderia influenciar o modelo a aprender características específicas do paciente em vez de características de cálculos renais. No geral, a validação cruzada de Monte Carlo oferece uma estrutura flexível para avaliar classificadores de imagens baseados em CNN.

Os hiperparâmetros foram selecionados com base em resultados empíricos. Os modelos foram treinados por 200 épocas, utilizando uma taxa de aprendizagem de 0,01 junto com o otimizador Adam, conhecido por sua eficácia na otimização de modelos de aprendizado profundo. A função de perda de entropia cruzada na Equação (33) foi aplicada, pois é uma função de perda comumente usada em tarefas de classificação para medir a diferença nas probabilidades entre classes previstas e rotuladas. Foi definido um tamanho de lote de 32 para estabilizar a taxa de convergência do modelo, mantendo a eficiência computacional. Além disso, a função de ativação ReLU foi usada para introduzir não linearidade, e uma taxa de dropout de 0,25 foi implementada para evitar o sobreajuste. A Tabela 6 resume os hiperparâmetros usados para treinar os modelos propostos.

Tabela 1 – Hiperparâmetros utilizados para o treinamento dos modelos de aprendizado profundo.

Hiperparâmetros	Detalhes
Épocas	200
Tamanho do Batch	32
Taxa de aprendizado	0.01
Otimizador	Adam
Função de Ativação	ReLU
Função de Perda	Entropia Cruzada
Dropout	0.25

3.3.6 Métricas de avaliação

O desempenho do DE-MLP foi avaliado usando as métricas Acurácia, Precisão, Recall e F1-Score para verificar sua robustez e eficácia na classificação correta de cálculos renais em exames de TC e para compará-lo com os métodos existentes.

Na análise realizada, a acurácia (Seção 2.15.3) reflete a capacidade do modelo de distinguir imagens de TC de cálculos renais e pacientes normais. A precisão 2.15.4 avalia se o modelo

classifica corretamente as amostras de cálculos renais, quantificando o risco de classificar erroneamente pacientes normais como portadores de cálculos renais.

O *recall* (Seção 2.15.5), representa o número de amostras de cálculos renais do conjunto de dados que o modelo classifica corretamente. Essa métrica é fundamental para quantificar falsos negativos e garantir que o modelo detecte cálculos renais em imagens de TC com a maior assertividade possível. O *F1-Score* (Seção 2.15.6) indica a capacidade do modelo de detectar corretamente cálculos renais nas amostras de dados, considerando um equilíbrio entre todas as amostras previstas como possuidoras de cálculos.

3.4 RESULTADOS E DISCUSSÃO

Os resultados obtidos foram analisados a partir dos experimentos com o modelo proposto. São apresentadas métricas de desempenho para cada modelo treinado com diferentes configurações de conjunto de dados. Para entender melhor os resultados, também foram analisadas a matriz de confusão e as taxas de convergência do modelo DE-MLP.

Além disso, foi avaliado o modelo único Conv4 treinado em três configurações diferentes, conforme definido na Seção 3.3.3: *Conv4A*, *Conv4B* e *Conv4(A+B)*. Esses modelos foram treinados com e sem aumento de dados (DA) para comparar seu desempenho com o DE-MLP.

3.4.1 Métricas de desempenho do modelo proposto

Após 10 execuções da validação cruzada de Monte Carlo, foram calculadas as médias da acurácia, precisão, *recall* e F1-score. As métricas de desempenho para os modelos DE-MLP e Conv4 são apresentadas na Tabela 2. Sem aumento de dados, o DE-MLP alcançou a menor acurácia (85,2%). No entanto, obteve um *recall* maior ao custo de uma precisão menor (74,2%), sugerindo que o modelo prioriza minimizar falsos negativos, mas gera mais falsos positivos. Quando o aumento de dados foi aplicado, a acurácia do DE-MLP aumentou aproximadamente 15%, superando todos os outros modelos. Ele alcançou os maiores valores de acurácia (99,6%), precisão (99,0%), *recall* (99,4%) e *F1-Score* (99,5%). Esse desempenho indica que o DE-MLP se beneficia efetivamente dos dados aumentados no treinamento para aprimorar sua generalização. Os modelos Conv4 também melhoraram com o aumento de dados, particularmente os modelos que usaram o Base de Dados B, que era o conjunto de dados mais desequilibrado.

O desempenho superior do DE-MLP com aumento de dados sugere que a abordagem de combinação de diferentes conjuntos de dados melhora efetivamente a adaptabilidade do modelo para dados de diferentes fontes de coleta. Os altos valores de *recall* e *F1-Score* confirmam ainda mais sua confiabilidade para aplicações clínicas, já que o modelo obteve uma baixa taxa de classificação incorreta de casos de cálculos renais.

Tabela 2 – Comparação de desempenho do DE-MLP com os modelos Conv4.

Modelo	Acurácia (%)	Precisão (%)	Recall (%)	F1-score (%)
<i>Conv4A</i>	96.7 ± 1.19	99.2 ± 1.17	92.9 ± 2.55	95.8 ± 1.25
<i>Conv4B</i>	88.1 ± 4.91	88.3 ± 13.8	82.2 ± 13.7	83.0 ± 4.84
<i>Conv4(A + B)</i>	91.8 ± 4.42	93.4 ± 6.80	85.4 ± 7.10	89.1 ± 5.87
DE-MLP	85.2 ± 0.98	74.2 ± 1.99	95.5 ± 0.67	83.4 ± 1.02
<i>Conv4A</i> + DA	97.2 ± 1.08	97.4 ± 1.85	95.7 ± 1.85	96.6 ± 1.36
<i>Conv4B</i> + DA	94.6 ± 5.02	93.1 ± 9.40	94.0 ± 5.59	93.3 ± 5.31
<i>Conv4(A + B)</i> + DA	95.7 ± 3.35	93.8 ± 7.15	95.0 ± 2.86	94.3 ± 4.17
DE-MLP + DA	99.6 ± 0.49	99.0 ± 0.45	99.4 ± 1.02	99.5 ± 0.67

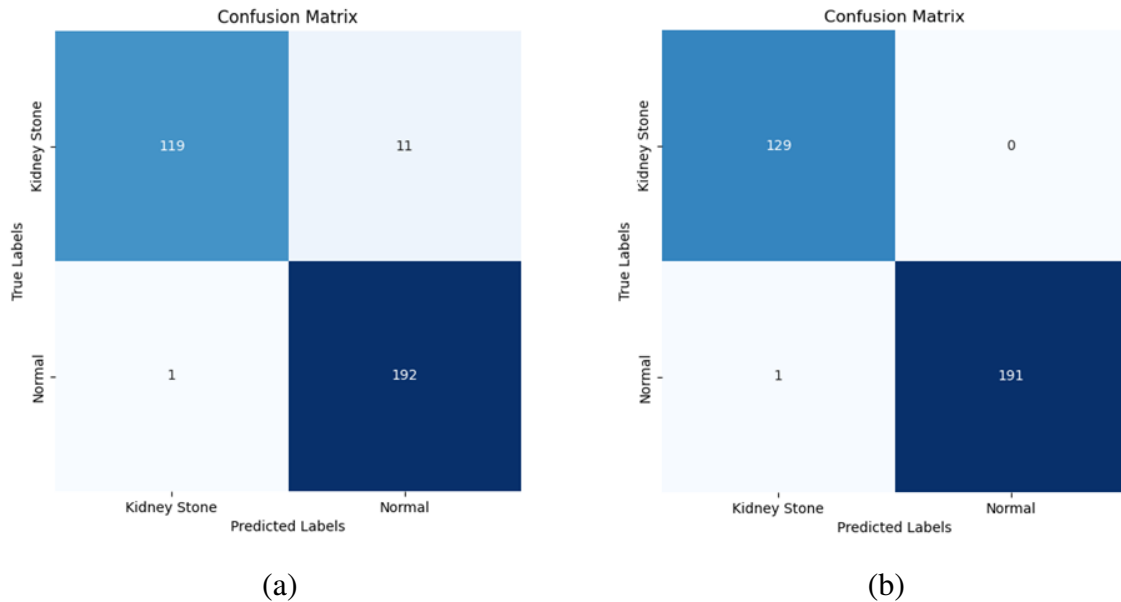
3.4.2 Matriz de confusão: DE-MLP

As matrizes de confusão do modelo DE-MLP foram analisadas para entender melhor o comportamento de classificação do modelo. A Fig. 26 apresenta as matrizes de confusão do melhor entre 10 modelos DE-MLP gerados. Essas matrizes ilustram os resultados da classificação de diferentes imagens de teste e refletem o desempenho de cada modelo treinado com base no número de amostras classificadas incorretamente. Quando treinado sem aumento de dados (Fig. 26a) o DE-MLP classificou incorretamente 11 amostras de cálculos renais como normais e 1 amostra normal como cálculo renal. Essa observação explica o alto *recall* obtido pelo modelo, priorizando a sensibilidade ao custo de menor precisão. O modelo DE-MLP quando treinado com aumento de dados (Fig. 26b) destaca uma notável melhoria do modelo neste quesito. Nesse caso, ele classificou corretamente todos os 129 casos de cálculos renais sem falsos negativos e classificou incorretamente apenas um caso normal como cálculo renal. Esse aprimoramento reflete um aumento significativo na precisão, *recall* e acurácia do modelo.

3.4.3 Taxa de convergência

Foi feita a observação da taxa de convergência para avaliar o impacto do aumento de dados no treinamento dos modelos propostos. A Figura 27 apresenta os gráficos de convergência para os modelos *Conv4A*, *Conv4B* e *Conv4(A + B)*, apresentados na Seção 3.3.3 sem e com aumento de dados. Os gráficos ilustram como a precisão de cada modelo evoluiu ao longo dos períodos de treinamento para os conjuntos de treinamento e validação. O modelo *Conv4A* (27a) obteve uma convergência rápida sem aumento de dados. A precisão da validação atingiu um nível mais baixo, em torno de 90%. No entanto, ela melhorou com o aumento de dados (Figura 27b) e se tornou mais estável ao longo dos períodos, reduzindo a lacuna de generalização. O modelo *Conv4B* (27c) mostra uma lacuna entre a precisão de treinamento e validação e instabilidade na curva de validação devido ao desequilíbrio no conjunto de dados B, o que pode causar dificuldade na generalização. A instabilidade da curva de validação foi suavizada para o modelo *Conv4B* treinado com aumento de dados (27d), e a convergência

Figura 26 – Matrizes de confusão geradas. 26a apresenta a matriz do modelo DE-MLP treinado com modelos que não passaram por aumento de dados. 26b apresenta a matriz do modelo DE-MLP treinado com modelos que passaram por aumento de dados.



Fonte: Autor

melhorou, tornando-o mais estável em comparação com a curva de precisão de validação não aumentada. O modelo $Conv4(A + B)$ (Figura 27e) combina amostras dos conjuntos de dados A e B. Embora essa distribuição de dados mais ampla tenha fornecido uma base de aprendizado diversificada, o comportamento de convergência foi menos favorável. O modelo apresentou flutuações na precisão de validação e não atingiu o mesmo nível de desempenho que o $Conv4A$, mesmo com aumento de dados. Na configuração $Conv4(A + B) + DA$ (Figura 27f), embora o aumento de dados tenha estabilizado a curva de validação, a precisão final permaneceu abaixo da do $Conv4A$, sugerindo que o treinamento de um único modelo diretamente em um conjunto de dados combinado pode introduzir conflitos na representação de características, limitando o desempenho. Como o modelo DE-MLP fundiu os modelos $Conv4A$ e $Conv4B$ pré-treinados em sua arquitetura e apenas suas camadas totalmente conectadas foram treinadas, uma análise direta da curva de convergência não é aplicável. Portanto, os gráficos de convergência para o modelo DE-MLP não são apresentados. No entanto, com base nas métricas de avaliação apresentadas na Seção 3.4.1, o DE-MLP, especialmente quando ambos os modelos $Conv4$ são treinados com aumento de dados, superou todas as outras configurações, incluindo $Conv4(A + B) + DA$, demonstrando a vantagem de combinar modelos específicos treinados em conjuntos de dados distintos, ao invés de unificar os conjuntos de dados em um único processo de treinamento.

3.4.4 Mapas de ativação

A técnica Grad-CAM (Gradient-weighted Class Activation Mapping) (Selvaraju *et al.*, 2016) foi aplicada para visualizar as regiões consideradas pelo modelo para detectar cálculos renais em imagens de TC.

O Grad-CAM utiliza a camada convolucional final do modelo treinado para produzir um mapa de ativação que destaca as regiões mais significativas que afetam a decisão de classificação. O mapa de ativação gerado ajuda a compreender as características que o modelo considera e é útil na área médica para visualizar se o modelo está considerando os mesmos padrões que um especialista clínico para prever a doença corretamente.

As Figuras 28 e 29 apresentam os mapas de ativação para cálculos renais e classificações normais, respectivamente. Cada par de imagens consiste na tomografia computadorizada original (à esquerda) e no mapa de ativação Grad-CAM correspondente (à direita), ilustrando as regiões mais significativas na decisão de classificação.

Ao detectar amostras de cálculos renais (Figura 28), o mapa de ativação do DE-MLP mostra que o modelo ativa as regiões renais onde os cálculos estão localizados. As áreas destacadas em vermelho no mapa sugerem que a decisão de classificação é baseada principalmente na localização real do cálculo renal na maioria dos casos. No entanto, algumas amostras no Conjunto de dados A destacam outras regiões da coluna vertebral (Figura 28d) e no Conjunto de dados B tiveram ativações mais fracas nas regiões renais (Figuras 28e e 28f). Apesar disso, o modelo ainda detecta corretamente os cálculos renais nesses casos. Isso indica que, apesar da intensidade de ativação ser menor, o modelo ainda considera a região renal como uma característica relevante para a classificação.

Para amostras normais classificadas corretamente (Figura 29), o modelo não ativa as regiões renais, conforme observado nos mapas de ativação. Outra observação é que ativações fortes aparecem nas regiões dos ossos, costelas e coluna vertebral. Apesar dessas respostas de alta intensidade em áreas não relacionadas à doença renal, o modelo ainda classifica corretamente a amostra como normal, sugerindo que essas ativações não influenciam negativamente o processo de tomada de decisão. Essas descobertas sugerem que a ausência de ativações de alta intensidade na área renal é um forte indicador de um caso normal.

Os resultados do Grad-CAM indicam que o modelo também se baseia nas regiões renais para a detecção de cálculos renais, reforçando sua interpretabilidade clínica. No entanto, a variabilidade na força de ativação nas imagens do Conjunto de Dados B sugere que alguns casos podem ser classificados com base em outros padrões, o que poderia ser investigado mais a fundo. Para amostras normais, a ausência de ativação renal está de acordo com as expectativas médicas, mas respostas fortes nas estruturas ósseas requerem análise adicional de interpretabilidade. A solução dos problemas apresentados poderia ser resolvida com mais variabilidade de dados disponíveis para o treinamento do modelo e com a segmentação das regiões de interesse para serem comparadas com as regiões ativadas pelo modelo.

3.4.5 Análise do caso de amostra classificada incorretamente

Apesar da alta precisão de classificação alcançada pelo DE-MLP, uma amostra no conjunto de validação/teste foi classificada incorretamente. Para entender melhor a classificação incorreta, a tomografia computadorizada original, juntamente com os mapas de ativação Grad-CAM para ambos os extratores de características, foram analisadas. Na Figura 30, a tomografia computadorizada original (imagem à esquerda) corresponde a um caso normal, mas o modelo a classificou incorretamente como um caso de cálculo renal. Os mapas de ativação para ambos os extratores de características (imagens do centro e da direita) indicam as regiões mais influentes no processo de tomada de decisão. O extrator de características DE-MLP *Conv4A* exibiu ativação generalizada em toda a região abdominal, em vez de se concentrar em estruturas específicas do rim. Esse padrão de ativação provavelmente diminuiu a confiança nesse extrator de características, já que não destacou as regiões mais relevantes para a detecção de cálculos renais. Consequentemente, essa falta de localização precisa pode ter contribuído para a classificação incorreta. Por outro lado, o extrator de características treinado no conjunto de dados B demonstrou ativações mais localizadas, mas ainda enfatizou áreas não diagnósticas. Investigamos mais a fundo os valores de confiança de cada extrator e a confiança do *Conv4B* foi maior do que a do *Conv4A* (0,99 e 0,85, respectivamente). Ao atribuir uma confiança maior à saída do extrator de características *Conv4B*, o modelo classificou erroneamente a amostra como um caso de cálculo renal, embora o rótulo da verdade fundamental fosse normal. Este caso destaca a importância da confiabilidade do extrator de características no desempenho da classificação. Em trabalhos futuros, investigações adicionais poderiam ser conduzidas para refinar o processo de extração de características, garantindo que o modelo priorize informações específicas do rim, minimizando o impacto de regiões não diagnósticas. Tais melhorias poderiam levar a um sistema de diagnóstico mais robusto e menos propenso a erros.

3.4.6 Contextualização com o estado-da-arte

Para avaliar melhor a eficácia do nosso modelo DE-MLP proposto, é apresentado o seu desempenho ao lado dos métodos de detecção de cálculos renais mais recentes, baseados em aprendizado de máquina, que utilizaram a Base de Dados A ou a Base de Dados B. A Tabela 3 resume a comparação dos métodos com a Base de Dados A, mostrando a acurácia obtida e o número de parâmetros treináveis. A literatura relacionada utilizou métodos baseados em redes neurais de aprendizado profundo, desde arquiteturas únicas, como XResNet-50 e ResNet-50 com aproximadamente 21 milhões de parâmetros treináveis, até abordagens mais complexas baseadas em conjuntos, que exigiram mais de aproximadamente 150 milhões de parâmetros treináveis. Embora a maior acurácia (99,71%) tenha sido alcançada pelo método de Baygin *et al.* (2022), ela está associada a um custo computacional mais alto. Além disso, o tamanho limitado do conjunto de dados utilizado levanta preocupações quanto à generalização do modelo para diversos

cenários clínicos em todo o mundo. A maior acurácia entre esses métodos foi de 98,79%, inferior ao nosso modelo DE-MLP + DA proposto, que alcançou uma acurácia significativamente maior de 99,60%, mantendo um baixo custo computacional de aproximadamente 3 milhões de parâmetros treináveis. Isso indica que nosso modelo equilibra efetivamente acurácia e eficiência usando o Base de Dados A, superando arquiteturas maiores e mais complexas na detecção automática de cálculos renais. Para o Base de Dados B (Tabela 4), estudos anteriores exploraram outros métodos de aprendizado de máquina baseados em árvores de decisão e distância do kernel (Kolli *et al.*, 2022; Hama; Majeed; Nariman, 2025). A classificação baseada em aprendizado profundo também foi estudada, com arquiteturas baseadas em conjunto alcançando 99,80% de acurácia, mas ao custo de aumentar significativamente o tamanho do modelo ($\approx 157M$ parâmetros). O XResNet152 (aproximadamente 44,5 milhões de parâmetros treináveis) (Manudhas; Manoranjitham, 2024) também alcançou 99,80% de acurácia; no entanto, sua validação experimental foi limitada, pois os autores realizaram apenas uma validação hold-out (teste único sem validação cruzada). Essa falta de validação cruzada levanta preocupações sobre a generalização do modelo. O modelo DE-MLP proposto alcançou uma acurácia média de 99,60% com menos parâmetros treináveis (aproximadamente 3 milhões), tornando-o uma alternativa leve com um alto desempenho de classificação.

O DE-MLP também apresenta uma vantagem significativa em termos de generalização. Ao contrário dos modelos anteriores treinados exclusivamente na Base de Dados A ou na Base de Dados B, o modelo proposto inclui ambos os conjuntos de dados durante o treinamento. Essa abordagem permite que o DE-MLP aprenda características de uma gama mais ampla de dados demográficos de pacientes, equipamentos de imagem e condições de aquisição, levando a um modelo de diagnóstico mais robusto e generalizado. Como resultado, ele melhora a aplicabilidade no mundo real e é mais confiável em diferentes contextos clínicos. Essas comparações destacam as vantagens do DE-MLP, que atinge alta acurácia, mantendo baixo custo computacional e alta capacidade de generalização. É uma escolha adequada para aplicações do mundo real, onde eficiência, precisão e adaptabilidade são fundamentais.

Tabela 3 – Tabela comparativa do método proposto com métodos anteriores que utilizaram a Base de Dados A para detecção de cálculos renais

Referência	Acurácia (%)	Método	Técnica de validação	Parâmetros treináveis (Milhões)
Asif et al., 2024 (Asif; Zheng; Zhu, 2024)	98.79%	PSOWeightedAvgNet and StackedEnsembleNet	Hold-out validation	$\approx 162M$
Yildirim et al., 2021 (Yildirim <i>et al.</i> , 2021a)	96.82%	XResNet-50	Hold-out validation	$\approx 25M$
Baygin et al., 2022 (Baygin <i>et al.</i> , 2022)	99.71%	ExDark19	10-fold cross-validation	$\approx 21M$
Asif et al., 2023 (Asif <i>et al.</i> , 2023)	97.98%	StoneNet	Hold-out validation	$\approx 3.3M$
Trabalho proposto	99.60%	DE-MLP + DA	Monte Carlo cross-validation (10 rodadas)	$\approx 3M$
Patro et al., 2023 (Patro <i>et al.</i> , 2023)	98.50%	Deep Kronecker Neural Network	10-fold cross-validation	$\approx 2.7M$
Trabalho proposto	97.20%	Conv4A + DA	Monte Carlo cross-validation (10 rodadas)	$\approx 1.45M$

Tabela 4 – Tabela comparativa do método proposto com métodos anteriores que utilizaram a Base de Dados B para detecção de cálculos renais

Referência	Acurácia (%)	Método	Técnica de validação	Parâmetros treináveis (Milhões)
Kolli et al., 2022 (Kolli et al., 2022)	97.00%	Random Forest and StackedEnsembleNet	Hold-out validation	—
Hama et al., 2025 (Hama; Majeed; Nariman, 2025)	96.31%	LBP and SVM	Hold-out validation	—
Chaki et al., 2024 (Chaki; Uçar, 2024)	99.80%	DarkNet19, InceptionV3 and ResNet101	10-fold cross-validation	$\approx 157M$
Ding et al., 2025 (Huan Ding Qirui Huang, 2025)	95.37%	Firebug Swarm and Alex/ELM Network	Hold-out validation	$\approx 62M$
Manudhas et al., 2024 (Manudhas; Manoranjitham, 2024)	99.80%	XResNet152	Hold-out validation	$\approx 44.5M$
Liu et al., 2024 (Liu; Ghadimi, 2024)	96.68%	Flexible Dwarf Mongoose Optimization	Hold-out validation	$\approx 25M$
Trabalho proposto	99.60%	DE-MLP + DA	Monte Carlo cross-validation (10 rodadas)	$\approx 3M$
Trabalho proposto	94.60%	<i>Conv4B</i> + DA	Monte Carlo cross-validation (10 rodadas)	$\approx 1.45M$

3.4.7 Aplicações clínicas do modelo DE-MLP proposto

O DE-MLP proposto obteve alto desempenho e eficiência, demonstrando seu potencial para aplicações clínicas reais na detecção de cálculos renais. As potenciais contribuições nos cenários clínicos são apresentadas a seguir:

3.4.7.1 Assistente de diagnóstico de cálculos renais para radiologistas

O modelo DE-MLP pode ser adaptado a um sistema de diagnóstico assistido por computador para ajudar os radiologistas a detectar cálculos renais com precisão em exames de tomografia computadorizada. O alto desempenho do nosso modelo e o mapa de ativação com regiões importantes destacadas auxiliam a clínica médica na redução do tempo de diagnóstico e da variabilidade entre observadores. Isso pode ser particularmente benéfico em ambientes clínicos com alta carga de trabalho, onde diagnósticos rápidos e confiáveis são cruciais.

3.4.7.2 Ferramenta de diagnóstico inicial

Em hospitais ou clínicas onde o acesso a radiologistas especializados é limitado, o DE-MLP poderia pré-analisar tomografias computadorizadas e identificar casos suspeitos para investigação adicional. Isso poderia melhorar as taxas de detecção precoce, garantindo que os pacientes recebam atendimento médico oportuno, mesmo em regiões com recursos de saúde limitados.

3.4.7.3 Integração automatizada de radiologia

O DE-MLP poderia ser integrado a um sistema automatizado de radiologia, como o PACS, sem exigir recursos computacionais extensos, dado seu baixo custo computacional de $\approx$

3M parâmetros treináveis. Ele pode funcionar em segundo plano para analisar grandes volumes de tomografias computadorizadas sem interromper o fluxo de trabalho do radiologista.

3.4.8 Limitações do modelo

Apesar dos resultados promissores alcançados pelo modelo DE-MLP, as seguintes limitações são reconhecidas: O treinamento e a avaliação dependeram exclusivamente de dois conjuntos de dados disponíveis publicamente: a Base de Dados A da Turquia e a Base de Dados B de Bangladesh. Embora esses conjuntos de dados sejam originários de fontes distintas, o número limitado de locais de coleta restringe a variabilidade dos protocolos de imagem, configurações do scanner e dados demográficos dos pacientes. Além disso, a diversidade demográfica das amostras de pacientes não está disponível, limitando a confiança na robustez do modelo em diferentes populações e ambientes clínicos.

Uma limitação da escolha do projeto foi o uso exclusivo de imagens coronais dos conjuntos de dados, descartando as perspectivas axial e sagital. Embora essa decisão tenha sido motivada pela redução da complexidade computacional, ela pode restringir a abrangência da capacidade diagnóstica do modelo e limitar sua adaptabilidade aos fluxos de trabalho radiológicos padrão.

Outra limitação é que ambos os modelos no conjunto compartilham a mesma arquitetura, o que restringe a exploração de recursos representacionais complementares que poderiam surgir da diversidade de arquiteturas.

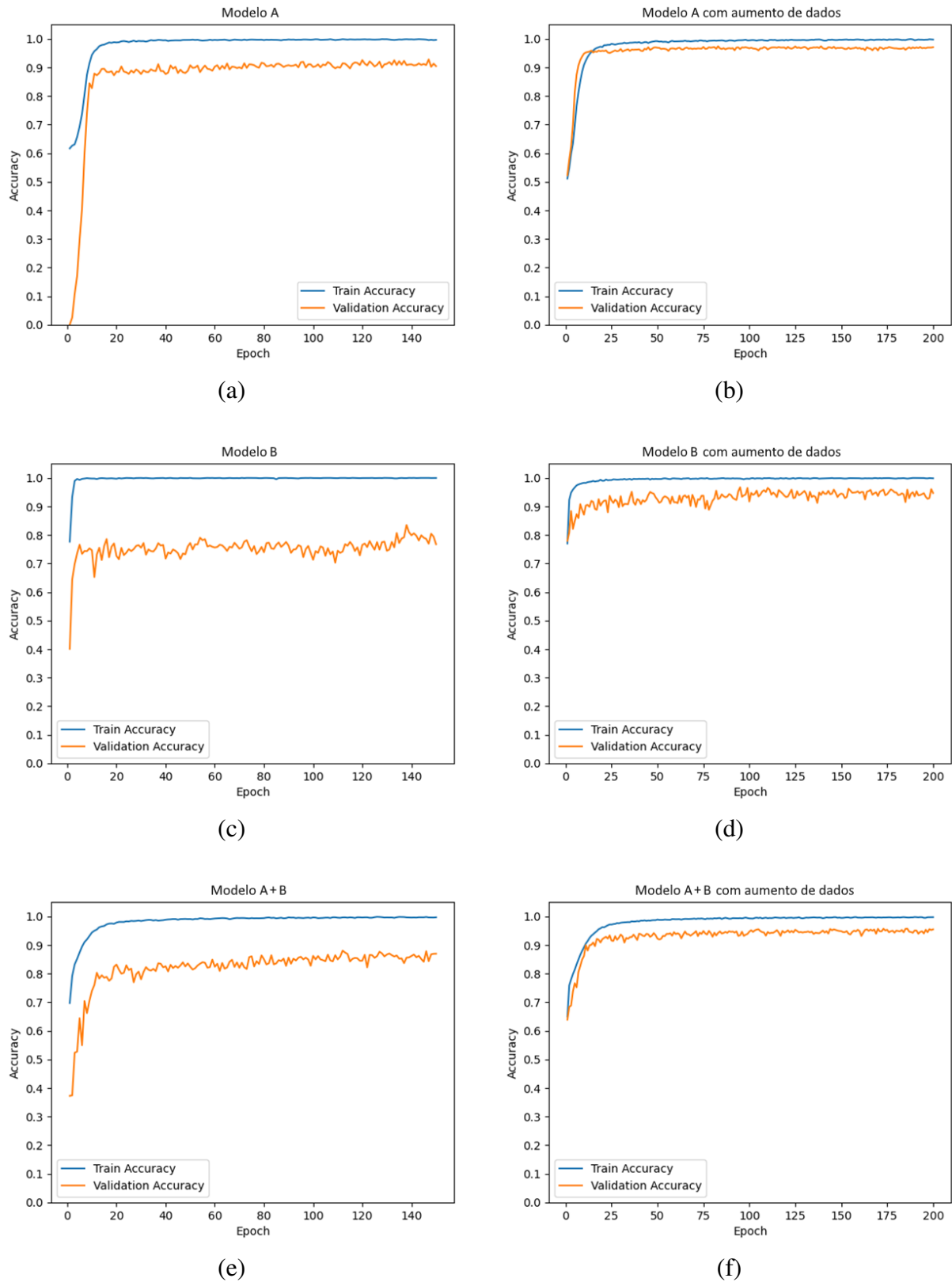
Além disso, o modelo apresenta um aumento significativo na complexidade e no número de parâmetros à medida que novas fontes de imagem distintas são adicionadas ao treinamento, o que pode impactar o tempo de processamento e a escalabilidade do sistema. Por fim, para realizar a predição em um cenário de múltiplas bases, a imagem de entrada precisa ser processada individualmente por cada modelo treinado em cada base distinta, o que eleva o custo computacional de inferência e pode dificultar a aplicação em ambientes clínicos em tempo real.

3.4.9 Direções futuras

As seguintes direções futuras são propostas para superar as limitações apontadas na Seção 3.4.8 e melhorar a aplicabilidade prática da estrutura DE-MLP em ambientes clínicos reais. Primeiro, expandir o conjunto de treinamento para incluir conjuntos de dados públicos ou privados mais diversificados e em grande escala melhorariam a generalização entre diferentes populações e sistemas de imagem por TC. Segundo, a introdução de diversidade de modelos no conjunto pode capturar representações complementares e melhorar o desempenho. Além disso, a integração de múltiplas perspectivas de imagens de TC, como vistas axiais e sagitais, deve ser considerada para refletir a prática clínica. Essa abordagem poderia aumentar a precisão do diagnóstico e fornecer uma representação mais abrangente das características dos cálculos renais.

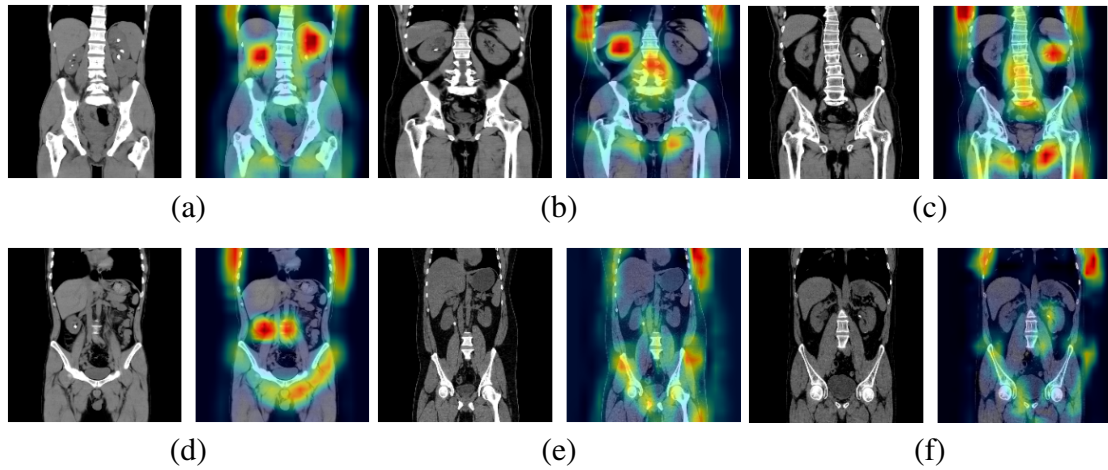
Embora isso possa aumentar a complexidade computacional, a otimização de arquiteturas ou estratégias eficientes de aprendizagem multivista poderiam ser empregadas para reduzir os custos computacionais. Planejamos investigar uma estratégia de Conjunto de Arquitetura de Conjunto de Dados: uma abordagem de fusão tardia que considera a fonte de dados e a arquitetura do modelo como parâmetros do conjunto. Outra direção promissora envolve a análise do efeito dos níveis de dose de radiação da TC no desempenho do modelo para identificar a dose mínima necessária para previsões confiáveis, reduzindo assim a exposição do paciente sem comprometer a precisão do diagnóstico. Ao buscar essas melhorias, esperamos aprimorar o potencial da estrutura DE-MLP para evoluir para uma ferramenta mais robusta e clinicamente viável para a detecção de cálculos renais, com possíveis extensões para outras aplicações de imagens médicas.

Figura 27 – Gráficos de convergência gerados. (a) e (b) apresentam gráficos de modelos da base A. (c) e (d) apresentam gráficos de modelos da base B. (e) e (f) apresentam gráficos de modelos da base conjunta A e B.



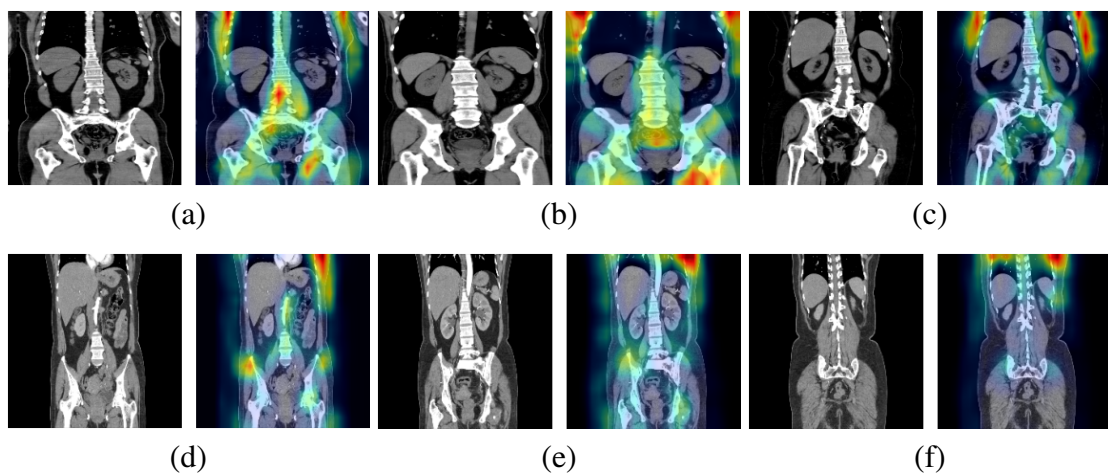
Fonte: Autor

Figura 28 – Mapas de calor Grad-CAM das regiões mais ativadas pelo modelo proposto nas classificações de amostras que possuem cálculos renais.



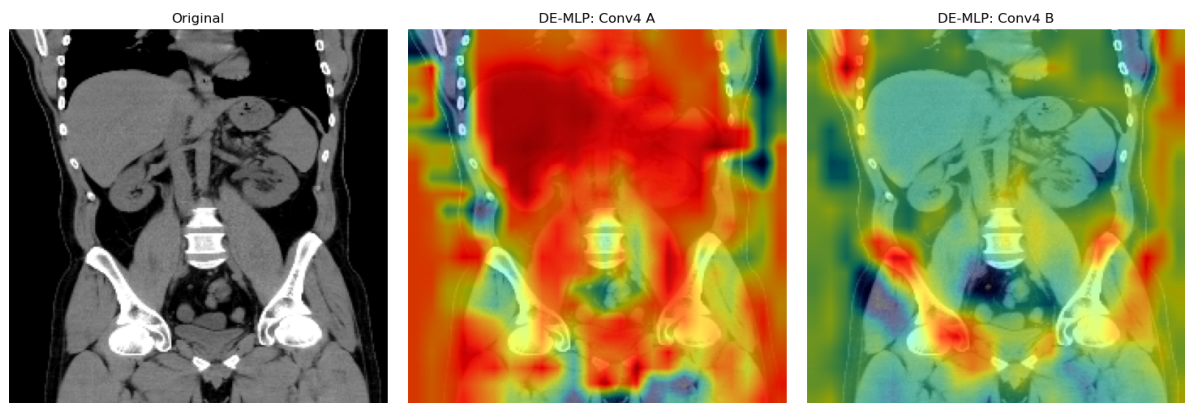
Fonte: Autor

Figura 29 – Mapas de calor Grad-CAM das regiões mais ativadas pelo modelo proposto nas classificações de amostras que não possuem cálculos renais.



Fonte: Autor

Figura 30 – Comparação da amostra que o DE-MLP classificou incorretamente com os mapas de ativação do extrator de características dos modelos *Conv4A* e *Conv4B*.



Fonte: Autor

4 SEGMENTAÇÃO AUTOMÁTICA DE CÁLCULOS URINÁRIOS EM TCS

Neste capítulo é apresentada uma nova base de dados chamada KSSD2025 (Kidney Stone Segmentation Dataset 2025), composta por 838 imagens axiais de TC manualmente anotadas para segmentação de cálculos renais. Foi realizada uma avaliação de quatro modelos de segmentação baseados em arquiteturas U-Net modificadas: U-Net, U-Net++, U-Net3+ e TransU-Net. O desempenho dos modelos foi avaliado utilizando 5-fold *cross-validation*, além de testes adicionais e contextualização com outros trabalhos similares. O método e os resultados apresentados neste capítulo foram publicados por Murillo F. Bouzon *et al.* (2025).

4.1 CONTEXTUALIZAÇÃO

A urolitíase, caracterizada pela formação de cálculos renais, é uma condição de alta prevalência associada ao aumento do risco de evolução para doença renal crônica, configurando-se como um problema crescente de saúde pública global (Jha *et al.*, 2013; Alelign; Petros, 2018). Diante do aumento contínuo da incidência dessa condição, tornam-se necessárias estratégias diagnósticas mais eficientes, capazes de reduzir complicações clínicas e aliviar os encargos sobre os sistemas de saúde (Reicherz *et al.*, 2022).

A TC sem contraste é considerada o padrão-ouro para o diagnóstico de cálculos renais, devido à sua elevada sensibilidade e especificidade (Saw *et al.*, 2000; Kambadakone *et al.*, 2010). Além da confirmação da presença dos cálculos, a TC fornece informações detalhadas sobre tamanho, volume, localização e composição provável, parâmetros essenciais para orientar condutas terapêuticas adequadas (Stewart *et al.*, 2015; Lidén, 2018; Wilhelm *et al.*, 2018). Entretanto, a análise dos exames de TC ainda depende, em grande parte, da inspeção manual realizada por radiologistas, que precisam examinar inúmeras fatias de imagens para delimitar os cálculos. Esse processo é demorado, suscetível a variabilidade inter e intraobservador e dependente da disponibilidade de especialistas experientes, especialmente em contextos com poucos recursos.

Nesse cenário, a segmentação automática de cálculos renais surge como uma alternativa promissora, permitindo acelerar o fluxo de trabalho diagnóstico, reduzir vieses humanos e assegurar maior consistência na análise clínica. Diferentemente da simples detecção, que apenas indica a presença de cálculos, a segmentação fornece a localização precisa das estruturas, informação fundamental para o planejamento de intervenções cirúrgicas minimamente invasivas.

Nos últimos anos, métodos de aprendizado profundo (*deep learning*) têm impulsionado avanços expressivos em imagens médicas, incluindo diagnóstico automatizado (Chan; Hadjiiski; Samala, 2020) e segmentação (Hesamian *et al.*, 2019). Estudos recentes aplicaram essas técnicas à detecção de cálculos renais (Hameed *et al.*, 2021), à previsão de desfechos clínicos (Suarez-Ibarrola *et al.*, 2020) e à estimativa da composição dos cálculos (Kim, U. S. *et al.*, 2022). Contudo, um desafio central permanece: a escassez de bases de dados anotadas manualmente por especialistas. A criação de máscaras de segmentação é trabalhosa, exige conhecimento es-

pecializado e demanda tempo significativo, limitando a disponibilidade de bases de dados para treinar e validar modelos de segmentação. Essa lacuna representa hoje um dos maiores obstáculos para a consolidação de sistemas de diagnóstico inteligente baseados em segmentação automática de cálculos renais.

4.2 TRABALHOS RELACIONADOS

Abordagens anteriores baseavam-se em técnicas clássicas de processamento de imagens. Thein *et al.* (2019) propuseram um método automatizado para detectar cálculos em exames de TC, utilizando limiarização de Otsu, operações morfológicas e análise volumétrica. A abordagem atingiu sensibilidades de 47%, 64% e 70% em diferentes variações de limiarização. Jain *et al.* (2023) introduziram uma abordagem para detecção e segmentação, empregando segmentação por limiar, o algoritmo de watershed e técnicas de redução de ruído. Testada em mais de 150 imagens do conjunto de dados Kaggle Kidney Stone Images, essa abordagem alcançou um SSIM de 0,863 e um PSNR (*Peak Signal-to-Noise Ratio*) de 30,637.

Com os avanços em aprendizado profundo, as CNNs (LeCun *et al.*, 1998) tornaram-se uma solução popular para a análise de imagens médicas, incluindo o diagnóstico de cálculos renais. Yildirim *et al.* (2021b) treinaram uma rede XResNet-50 em 1.799 imagens 2D de TC, obtendo 96,82% de acurácia. Neha (2023) combinou técnicas de aprendizado de máquina para localização renal e segmentação de cálculos, atingindo um Intersecção sobre União (IoU) de 76,79%.

Entre as técnicas de aprendizado profundo, a U-Net (Ronneberger; Fischer; Brox, 2015) tem sido amplamente utilizada em tarefas de segmentação, demonstrando sua eficácia em diversas aplicações médicas (Song *et al.*, 2022; Peng; Sonka; Chen, 2023; Wang, H. *et al.*, 2023; Azad *et al.*, 2022), e vários estudos a aplicaram na segmentação de cálculos renais em imagens de TC. Cui *et al.* (2020) desenvolveram um modelo baseado em U-Net para segmentação renal e de cálculos, alcançando sensibilidade de 95,9% e Valor Preditivo Positivo (PPV) de 98,7%. Dan Li *et al.* (2022) exploraram cinco variações de redes profundas de segmentação, com o modelo Res U-Net atingindo um coeficiente Dice de 96,54%. Young In Kim *et al.* (2023) utilizaram U-Net++ para segmentação de cálculos urinários em exames de TC sem contraste, relatando um coeficiente Dice de 87,56%. Preedanan *et al.* (2023) desenvolveram um pipeline em cascata baseado em U-Net para segmentação em radiografias abdominais, alcançando F2-score de 71,28%. Rabby *et al.* (2023) aplicaram U-Net para segmentação renal e YOLOv5/YOLOv7 para detecção de cálculos em imagens de TC, em que YOLOv7 obteve acurácia de 99,5% e IoU de 91,4%. Jang *et al.* (2024) propuseram uma U-Net modificada para segmentação do sistema urinário, alcançando coeficiente Dice de 0,8673. Além de abordagens baseadas em CNNs, pesquisadores também exploraram modelos híbridos. Elton *et al.* (2022) combinaram uma U-Net 3D para segmentação renal com um classificador CNN de 13 camadas, alcançando sensibilidade de 0,86 e Área sob a Curva ROC (AUC) de 0,95. Zih-Hao Huang *et al.* (2023) combinaram Inception-

ResNetV2 para classificação com U-Net (ResNet34) para segmentação, atingindo 99,7% de acurácia e 93,7% de IoU. Huang *et al.* (2024) introduziram uma V-Net aprimorada com ponderação de exemplos difíceis para segmentação em TC de dupla energia, alcançando AUC de 0,96. Kumar *et al.* (2024) usaram segmentação baseada em level set e classificação Kernel-ELM em imagens de ultrassom, obtendo 90,73% de acurácia.

Embora arquiteturas baseadas em U-Net tenham dominado a segmentação médica, estudos recentes exploraram modelos baseados em Transformers para capturar dependências de longo alcance e melhorar a acurácia em diferentes doenças clínicas. Chen (2024) introduziram o TransU-Net, um modelo híbrido CNN-Transformer que combina a capacidade de localização da U-Net com o modelamento de contexto global dos Transformers. A abordagem obteve um índice Dice de 88,39% para segmentação multi-órgão e 67,67% para segmentação tumoral. De forma semelhante, Guangju Li *et al.* (2024) propuseram o MultiIB-TransU-Net, integrando blocos de gargalo de informação a uma única camada Transformer, o que reduziu o overfitting e manteve um alto desempenho (Dice: 81,83% no conjunto Synapse). Lv *et al.* (2025) introduziram o ECM-TransU-Net, uma U-Net Transformer aprimorada por atenção multi-escala e realce de bordas. Avaliado em grandes conjuntos de TC e RM, o ECM-TransUNet atingiu um DSC de 84,68%, com um custo computacional de 50,68G FLOPs e 66,47M parâmetros. Ji *et al.* (2024) propuseram o BGRD-TransUNet para segmentação de lesões mamárias, introduzindo módulos residuais multi-escala e mecanismos de atenção em bordas, o que melhorou significativamente o IoU e o DSC em relação ao TransUNet.

Outras abordagens híbridas e não convolucionais também foram investigadas. Modelos baseados em LSTM, por exemplo, têm sido propostos para melhorar previsões baseadas em sequências clínicas ou de imagem. Zhu *et al.* (2024) desenvolveram um modelo LSTM adaptativo com pesos de atributos para prever precocemente a presença de cálculos a partir de exames de sangue e urina, alcançando 89,5% de acurácia e um AUC de 0,95. Em outro trabalho, Jadhav *et al.* (2025) propuseram um modelo híbrido CNN-LSTM treinado em imagens de ultrassom para detecção de cálculos, relatando uma acurácia de 97% e uma forte capacidade de generalização em cenários clínicos. Além disso, Choi *et al.* (2025) introduziram o PFSH-Net, uma rede híbrida paralela espaço-frequência para segmentação de cálculos em imagens de TC veterinária, alcançando um índice Dice de 0,7667.

Apesar do sucesso dos modelos baseados em aprendizado profundo, especialmente variantes de U-Net, um grande desafio persiste: a escassez de conjuntos de dados anotados publicamente disponíveis. Essa limitação afeta a capacidade de generalização dos modelos e sua aplicabilidade clínica. Embora técnicas como aumento de dados com redes neurais generativas (Preedanani *et al.*, 2022) e uso de radiômica (Kaviani *et al.*, 2023) tenham sido exploradas, elas não suprimem totalmente a necessidade de bases de dados segmentadas de alta qualidade.

Para suprir essa lacuna, foi construído o KSSD2025 (Kidney Stone Segmentation Dataset 2025), um novo conjunto de dados com 838 imagens de TC e máscaras de segmentação de cálculos renais anotadas. A sua aplicabilidade foi avaliada em um estudo comparativo que

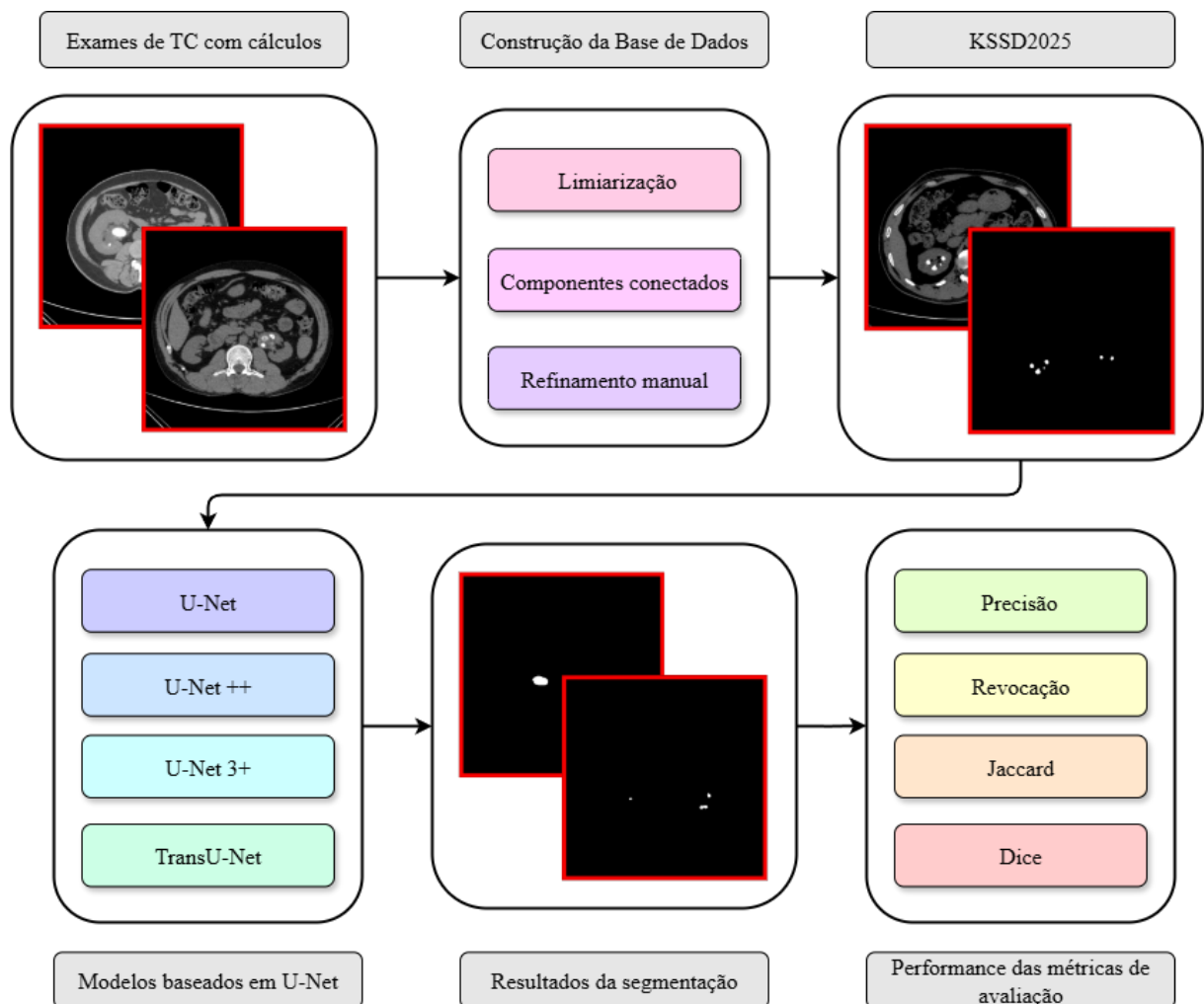
envolve quatro arquiteturas baseadas em U-Net: U-Net, U-Net++, U-Net3+ e TransU-Net. Os experimentos demonstram que esse conjunto de dados representa um recurso valioso para o avanço da pesquisa em segmentação automática de cálculos renais.

4.3 MATERIAIS E MÉTODOS

A metodologia proposta é apresentada na Fig. 31. O fluxograma inicia com imagens axiais de TC contendo cálculos renais, seguido pelo pré-processamento das imagens e pela anotação manual das máscaras de segmentação de referência para a construção do dataset. Em seguida, foram treinados quatro modelos baseados em U-Net utilizando o dataset construído.

As seções seguintes apresentam uma descrição detalhada de cada etapa da metodologia proposta.

Figura 31 – Diagrama da construção do KSSD2025 e da validação dos modelos baseados em U-Net.



Fonte: Autor.

4.3.1 Imagens de TC de Cálculos Renais

As imagens utilizadas neste estudo foram introduzidas por Islam *et al.* (2022). Os exames foram coletados no Dhaka Central International Medical College and Hospital (DCIMCH), Bangladesh, com aprovação ética das autoridades hospitalares. Pacientes diagnosticados com cálculos renais, cistos e tumores participaram voluntariamente do estudo, fornecendo consentimento para a coleta dos dados.

A base de dados contém cortes coronais, sagitais e axiais de TC, adquiridos seguindo protocolos padronizados de imagem para exames na região abdominal. Para garantir a privacidade dos pacientes, todas as informações de identificação e metadados foram removidos antes da conversão das imagens DICOM para o formato JPEG sem perdas. Ao todo, foram extraídas 12446 imagens, compreendendo 3709 amostras de cistos, 2283 de tumores, 1377 de cálculos renais e 5077 de casos normais. Esta base de dados não possui nenhuma informação de máscara de segmentação anotada.

Os dados originais estão disponíveis publicamente no Kaggle¹.

4.3.2 Anotação dos dados

O treinamento de modelos de aprendizado profundo para segmentação médica depende de bases de dados de alta qualidade com anotações de referência. No entanto, obter tais datasets é particularmente desafiador em imagens médicas devido à complexidade das anotações manuais e à limitada disponibilidade de dados rotulados por especialistas. A segmentação de cálculos renais em exames de TC é especialmente difícil devido às variações no tamanho, formato e níveis de contraste dos cálculos. Para superar essa dificuldade e permitir o treinamento de modelos robustos de segmentação, foi desenvolvido o KSSD2025, em que foram adicionadas máscaras de segmentação anotadas manualmente a partir das regiões de cálculos renais com base nas imagens previamente apresentadas por Islam *et al.* (2022).

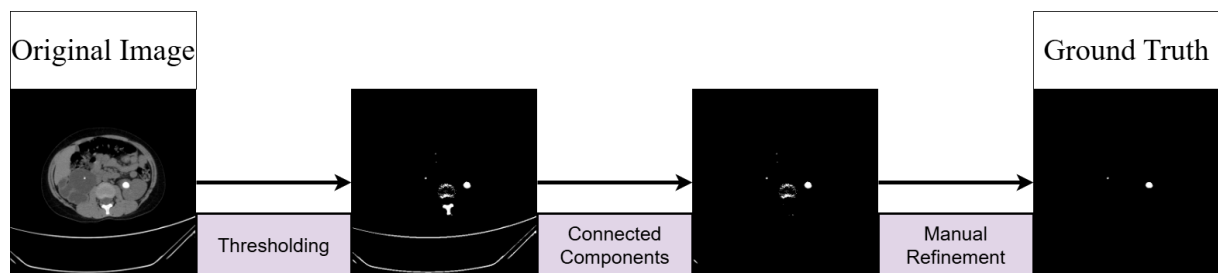
Das 12446 imagens disponíveis no dataset original, selecionamos 838 imagens que continham especificamente cálculos renais na visão axial. A perspectiva axial foi escolhida porque forneceu o maior número de amostras e apresentou uma referência anatômica mais adequada para localizar e estimar os tamanhos dos cálculos durante a anotação manual. Para preparar a base de dados para tarefas de segmentação, foi aplicada uma metodologia de pré-processamento para gerar as máscaras de referência (padrão-ouro):

- a) **Limiarização:** Foi definido empiricamente um limiar de intensidade de pixel em 150, para manter apenas as regiões com tons mais claros da imagem, que correspondem a estruturas de alta densidade.

¹<https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone>

- b) **Componentes Conectados:** O algoritmo de componentes conectados foi aplicado para segmentar diferentes regiões da imagem e calcular seus tamanhos. Regiões com mais de 300 pixels foram descartadas, pois nenhum cálculo renal nos exames ultrapassava esse tamanho. Essa etapa removeu efetivamente grandes estruturas ósseas e a região da mesa de exame das imagens.
- c) **Refinamento Manual:** Um refinamento manual final foi realizado para eliminar ruídos remanescentes e regiões que não correspondiam a cálculos, garantindo segmentações precisas. Esse processo foi conduzido em colaboração com especialistas em urologia e radiologia.

Figura 32 – Ilustração do pipeline de anotação do dataset.



Fonte: Autor.

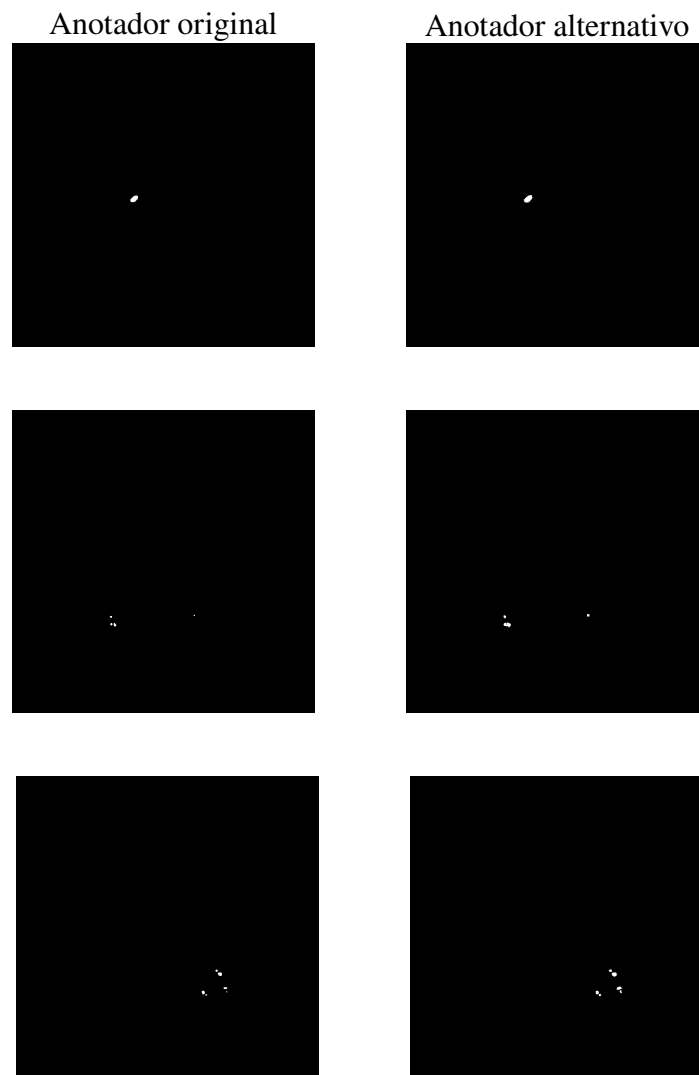
As máscaras de segmentação de referência (padrão-ouro) foram sistematicamente pareadas com suas imagens originais correspondentes para formar uma base de dados estruturada após as etapas de pré-processamento e anotação. Cada máscara anotada delimita precisamente as regiões de cálculos renais, garantindo um conjunto de dados adequado para treinamento, validação e testes em aplicações de aprendizado profundo. O pipeline de anotação do dataset é ilustrado na Fig. 32. Para avaliar a consistência entre anotadores, um segundo anotador humano rotulou independentemente um subconjunto de 100 cortes de TC. Essas anotações foram então comparadas com as anotações originais criadas com o suporte de especialistas. O coeficiente de similaridade Dice médio entre os dois conjuntos de máscaras foi de 74,53%. Exemplos representativos dessa comparação são mostrados na Fig. 33. Quando comparado à anotação totalmente manual, um anotador rotulou aproximadamente 100 imagens por hora utilizando o pipeline de anotação, contra apenas 10 imagens por hora sem ele, resultando em um ganho de produtividade de 10 vezes. Esse ganho de eficiência destaca a importância das etapas de pré-processamento para acelerar o processo de rotulagem, tornando viável a anotação em larga escala sem comprometer a qualidade.

Dada a escassez de datasets de segmentação de cálculos renais com referência anotada por especialistas, o desenvolvimento do KSSD2025 representa uma contribuição significativa para a área, fornecendo um dataset de alta qualidade para modelos de aprendizado profundo.

O KSSD2025 difere de outros datasets de segmentação de cálculos renais nos seguintes aspectos:

- a) É composto exclusivamente por cortes axiais de TC, seguindo os padrões PACS utilizados em ambientes clínicos.
- b) O processo de anotação combinou técnicas clássicas de visão computacional (limiarização e componentes conectados) como pré-filtro para auxiliar os anotadores humanos.
- c) O refinamento manual foi realizado com suporte de radiologistas e urologistas experientes.

Figura 33 – Exemplos de comparação de máscaras anotadas entre anotadores.



Fonte: Autor.

O KSSD2025 está disponível no Kaggle².

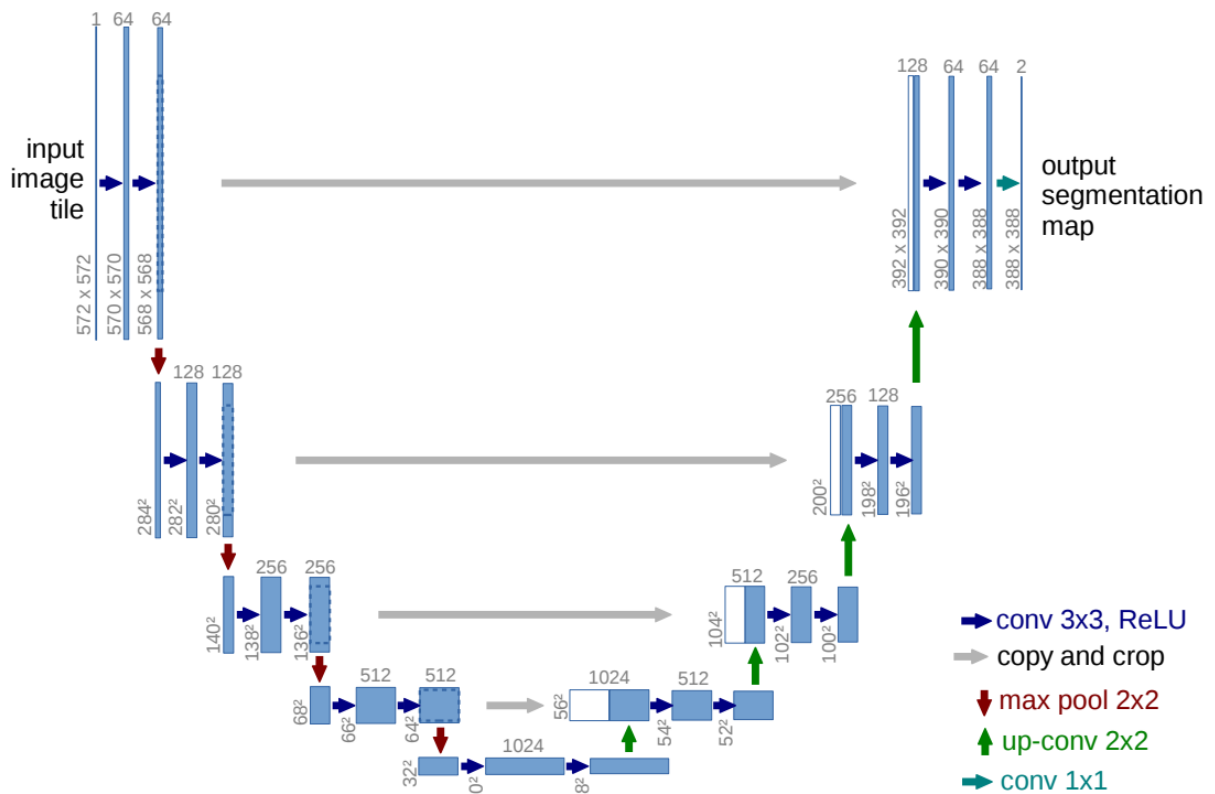
²<https://www.kaggle.com/datasets/murillobouzon/kssd2025-kidney-stone-segmentation-dataset>

4.3.3 Modelos U-Net Modificados

Os modelos treinados para segmentação de cálculos renais neste estudo são baseados na arquitetura U-Net (Ronneberger; Fischer; Brox, 2015), que tem demonstrado desempenho notável em segmentação de imagens médicas. Originalmente proposta para processamento de imagens médicas, a U-Net é amplamente adotada devido à sua capacidade de aprender eficientemente características espaciais e segmentar com precisão estruturas pequenas em imagens médicas. Embora arquiteturas alternativas, como transformers, modelos recorrentes e abordagens híbridas no domínio da frequência, tenham ganhado atenção, foi optado avaliar o KSSD2025 apenas com modelos baseados em U-Net devido ao seu alto desempenho na literatura relacionada e ampla adoção clínica.

A U-Net original segue uma estrutura encoder-decoder com desenho simétrico em forma de “U”, conforme ilustrado na Fig. 34. O caminho do encoder, também conhecido como *contracting path*, extrai progressivamente características hierárquicas usando camadas convolucionais com ativação ReLU, seguidas de camadas de max-pooling que reduzem as dimensões espaciais enquanto aumentam o número de canais de características. O caminho do decoder, ou *expanding path*, reconstrói as regiões segmentadas por meio de upsampling, convoluções transpostas e concatenação com mapas de características correspondentes do encoder via conexões de *skip*.

Figura 34 – Arquitetura original da U-Net.



Fonte: Ronneberger, Fischer e Brox (2015)

Para melhorar os resultados na segmentação de cálculos renais no KSSD2025, propusemos uma modificação na U-Net, ajustando seus três principais componentes para uma arquitetura mais leve, conforme segue:

- Encoder:** extrai características espaciais progressivamente usando convoluções (3×3) com normalização em lote (Batch Normalization) e ativação ReLU. Camadas Dropout (0.1–0.3) foram aplicadas para melhorar a generalização, e max pooling (2×2) foi usado para o downsampling. O encoder segue uma hierarquia estruturada com mapas de características de 16, 32, 64 e 128 filtros.
- Bottleneck:** funciona como ponte entre encoder e decoder, processando as características extraídas com 256 filtros para capturar representações de alto nível.
- Decoder:** utiliza convoluções transpostas (2×2) para realizar o upsampling, reconstruindo progressivamente a máscara de segmentação e incorporando informações espaciais via conexões de *skip*. O decoder reduz as dimensões dos mapas de características de 128 para 64, 32 e 16 filtros, gerando a saída final. A última camada aplica uma convolução 1×1 com ativação sigmoide, produzindo uma máscara binária representando a segmentação dos cálculos renais.

A U-Net modificada utilizada neste trabalho é ilustrada na Fig. 35. Essa versão melhora a eficiência em comparação com o design original, produzindo resultados similares com um menor custo computacional. A incorporação de normalização em lote após convoluções e a aplicação estratégica de Dropout ajudam a mitigar overfitting. A redução do número inicial de filtros de 64 para 16 diminui significativamente o custo computacional e o uso de memória, tornando o modelo mais adequado para implantação em ambientes com recursos limitados. Entretanto, a simplificação pode reduzir a precisão da segmentação em regiões de fronteira. Esses trade-offs posicionam o modelo como uma alternativa leve para tarefas menos complexas. A Tabela 5 mostra a comparação de uso de recursos entre a U-Net modificada e a original.

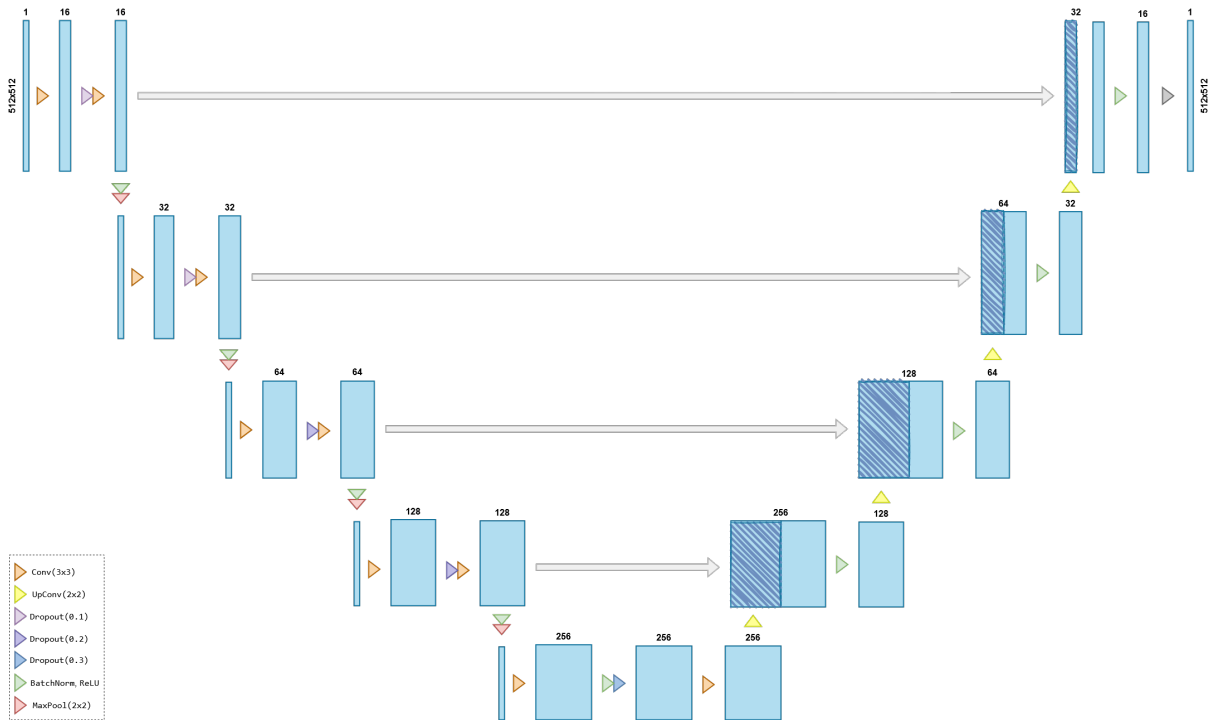
Tabela 5 – Comparação entre a U-Net modificada e a original em termos de custo computacional e uso de recursos.

Modelo	Tamanho do filtro inicial	Parâmetros treináveis	Tamanho
U-Net Modificada	16	1.4M	5.3 MB
U-Net Original	64	22.3M	85.2 MB

Para avaliar o desempenho de diferentes variações baseadas em U-Net, foram treinados e analisados quatro modelos de segmentação, todos incorporando as modificações propostas. As arquiteturas foram:

- U-Net:** O modelo de base, seguindo o design padrão U-Net com uma estrutura codificador-decodificador e *skip connections* para reter informações espaciais.
- U-Net++** (Zhou *et al.*, 2018): Uma versão aprimorada do U-Net que introduz *skip pathways* densos entre o codificador e o decodificador, refinando a fusão de

Figura 35 – Arquitetura U-Net modificada.



Fonte: Autor.

recursos ao incorporar blocos convolucionais aninhados entre pares codificador-decodificador, em vez de *skip connections* diretas.

- c) **U-Net3+** (Huang, H. *et al.*, 2020): Uma extensão que integra *skip connections* em escala completa, permitindo que o decodificador agregue informações multi-escala de todos os níveis do codificador simultaneamente, construindo assim uma representação aprimorada de recursos ao mesclar saídas de múltiplas resoluções.
- d) **TransU-Net** (Chen *et al.*, 2021): Um modelo híbrido que incorpora Vision Transformers (ViTs) no codificador para introduzir mecanismos de autoatenção. Os blocos transformadores substituem parte das camadas convolucionais tradicionais, permitindo que o TransU-Net capture dependências de longo alcance e aprimore a precisão da segmentação em imagens médicas complexas.

Cada um desses modelos foi treinado com o dataset KSSD2025 e avaliado com métricas de segmentação. Nas próximas seções deste capítulo, as U-Nets citadas são consideradas com as modificações apresentadas anteriormente. A seção a seguir fornece detalhes da configuração experimental.

4.3.4 Configuração de Treinamento e Hiperparâmetros

Os experimentos foram realizados utilizando o Jupyter Lab em um computador desktop com GPU NVIDIA GeForce RTX 3060, CPU AMD Ryzen 5 5500 e 16 GB de RAM. Para avaliar a capacidade de generalização dos modelos, empregamos uma estratégia de validação

cruzada de 5 folds para avaliar o desempenho de todos os modelos em dados não vistos. Para melhorar ainda mais o desempenho de classificação e a capacidade de generalização, aplicamos aumento de dados em cada batch de treinamento, utilizando as seguintes transformações:

- a) Rotação: até $2,5^\circ$;
- b) Deslocamento horizontal: até 0,75%;
- c) Deslocamento vertical: até 0,75%;
- d) Intervalo de ruído: até 0,75%;
- e) Zoom: até 0,75%;
- f) Flip horizontal: habilitado.

Os hiperparâmetros foram selecionados empiricamente para otimizar o desempenho da segmentação. Os modelos foram treinados por 150 épocas com batch size de 16. Utilizamos o otimizador Adam, com taxa de aprendizado de 0,001, devido às suas capacidades de aprendizado adaptativo. A função de perda Dice (Equação (24)) foi aplicada para otimizar a sobreposição da segmentação entre previsões e o padrão-ouro, sendo eficaz na segmentação de imagens médicas. Como descrito anteriormente, a função de ativação ReLU foi utilizada em cada camada convolucional. A Tabela 6 resume os hiperparâmetros utilizados no processo de treinamento.

As quatro arquiteturas foram treinadas utilizando os mesmos hiperparâmetros para garantir uma comparação justa. Optamos por não ajustar cada modelo individualmente para evitar viesar os resultados para configurações específicas. Embora escalas alternativas de taxa de aprendizado possam potencialmente beneficiar modelos individuais, nosso objetivo era avaliar diferenças arquitetônicas sob condições de treinamento idênticas.

Tabela 6 – Hiperparâmetros utilizados para o treinamento das arquiteturas de deep learning.

Hiperparâmetros	Detalhes
Épocas	150
Batch size	16
Taxa de aprendizado	0,001
Otimizador	Adam
Função de ativação	ReLU
Função de perda	Dice

4.4 RESULTADOS E DISCUSSÃO

Nesta seção, é apresentada a avaliação experimental do conjunto de dados KSSD2025 para segmentação de pedras nos rins em imagens de TC axiais. A investigação de desempenho dos modelos baseados em U-Net é detalhada, sendo apresentadas as métricas de segmentação obtidas na validação cruzada em 5 dobras, os resultados visuais no conjunto de dados desenvolvido, a análise de casos de erro e o desempenho em amostras de outras fontes de dados. Além

disso, é feita uma análise de complexidade dos modelos propostos. Para contextualizar a eficácia da abordagem proposta, comparamos o desempenho dos modelos treinados com o KSSD2025 com estudos anteriores que empregaram metodologias semelhantes para segmentação de pedras nos rins. Finalmente, é discutida as implicações clínicas de nossos achados e potenciais aplicações para pesquisas futuras.

4.4.1 Análise de Desempenho com Validação Cruzada 5-Fold

Seguindo o procedimento de validação cruzada 5-fold, a precisão média, recall, índice de Jaccard e coeficiente de Dice foram avaliadas para cada modelo baseado em U-Net treinado com o conjunto de dados KSSD2025. Os resultados médios obtidos nos cinco folds estão resumidos na Tabela 7. Entre as arquiteturas avaliadas, a U-Net obteve o melhor desempenho geral, com um coeficiente de Dice de 97,06% e um índice de Jaccard de 94,65%, indicando forte similaridade entre as máscaras de segmentação previstas e as máscaras de referência. U-Net++ e U-Net3+ também alcançaram métricas de desempenho elevadas, com valores de Dice de 96,63% e 96,68%, respectivamente. Em contraste, o TransU-Net apresentou o menor desempenho em segmentação, com coeficiente de Dice de 95,53% e índice de Jaccard de 92,21%. Além disso, seus maiores desvios-padrão em recall e índice de Jaccard indicam previsões menos consistentes. Embora os mecanismos de atenção no TransU-Net sejam projetados para melhorar o contexto global, eles podem ser menos eficazes na detecção de pequenas regiões de alto contraste, como cálculos renais, especialmente quando treinados com um número limitado de amostras anotadas.

Tabela 7 – Resultados da validação cruzada 5-fold para modelos baseados em U-Net treinados com KSSD2025.

Modelo	Precisão (%)	Recall (%)	Jaccard (%)	Dice (%)
U-Net	97.83 ± 0.33	96.61 ± 0.77	94.65 ± 0.62	97.06 ± 0.39
U-Net++	97.16 ± 0.68	96.44 ± 0.67	93.97 ± 0.83	96.63 ± 0.63
U-Net3+	96.88 ± 0.34	96.90 ± 1.02	94.02 ± 0.86	96.68 ± 0.61
TransU-Net	96.26 ± 0.52	95.39 ± 1.24	92.21 ± 1.27	95.53 ± 0.89

Em resumo, a U-Net demonstrou superior assertividade na segmentação de cálculos renais. As demais arquiteturas também produziram resultados satisfatórios, todas com valores de Dice superiores a 95%, o que evidencia a qualidade do conjunto de dados KSSD2025 como recurso de treinamento para modelos de segmentação de cálculos renais.

4.4.2 Análise Visual dos Resultados

Para avaliar ainda mais a qualidade da segmentação, foi realizada uma inspeção visual em amostras representativas do conjunto de validação. A Fig. 36 ilustra os resultados de seg-

mentação obtidos pelos modelos U-Net, U-Net++, U-Net3+ e TransU-Net, incluindo casos com cálculo renal único e múltiplo.

Nos exemplos apresentados, todos os modelos produziram máscaras de segmentação que se assemelham bastante à anotação manual de referência, corroborando os altos valores de Dice relatados na análise quantitativa. Os casos selecionados incluem cenários clínicos diversos, como cálculos localizados no rim esquerdo ou direito, variando em tamanho e quantidade (cálculo único versus múltiplos cálculos).

Os modelos U-Net, U-Net++ e U-Net3+ apresentaram segmentações particularmente precisas tanto em casos simples (cálculo grande único) quanto em casos complexos (múltiplos cálculos pequenos), demonstrando forte capacidade de localização. Em contraste, o TransU-Net ocasionalmente produziu previsões ligeiramente mais difusas ou fragmentadas, especialmente na presença de múltiplos cálculos, o que pode explicar seu desempenho relativamente menor em Dice e maior variância.

De modo geral, os resultados qualitativos apoiam a eficácia de todos os modelos avaliados para segmentação de cálculos renais, destacando, ao mesmo tempo, a consistência superior fornecida pela arquitetura U-Net.

4.4.3 Análise Estatística

Para avaliar a significância estatística das diferenças de desempenho entre os modelos de segmentação avaliados, foi realizada uma análise estatística com base nos valores do coeficiente Dice.

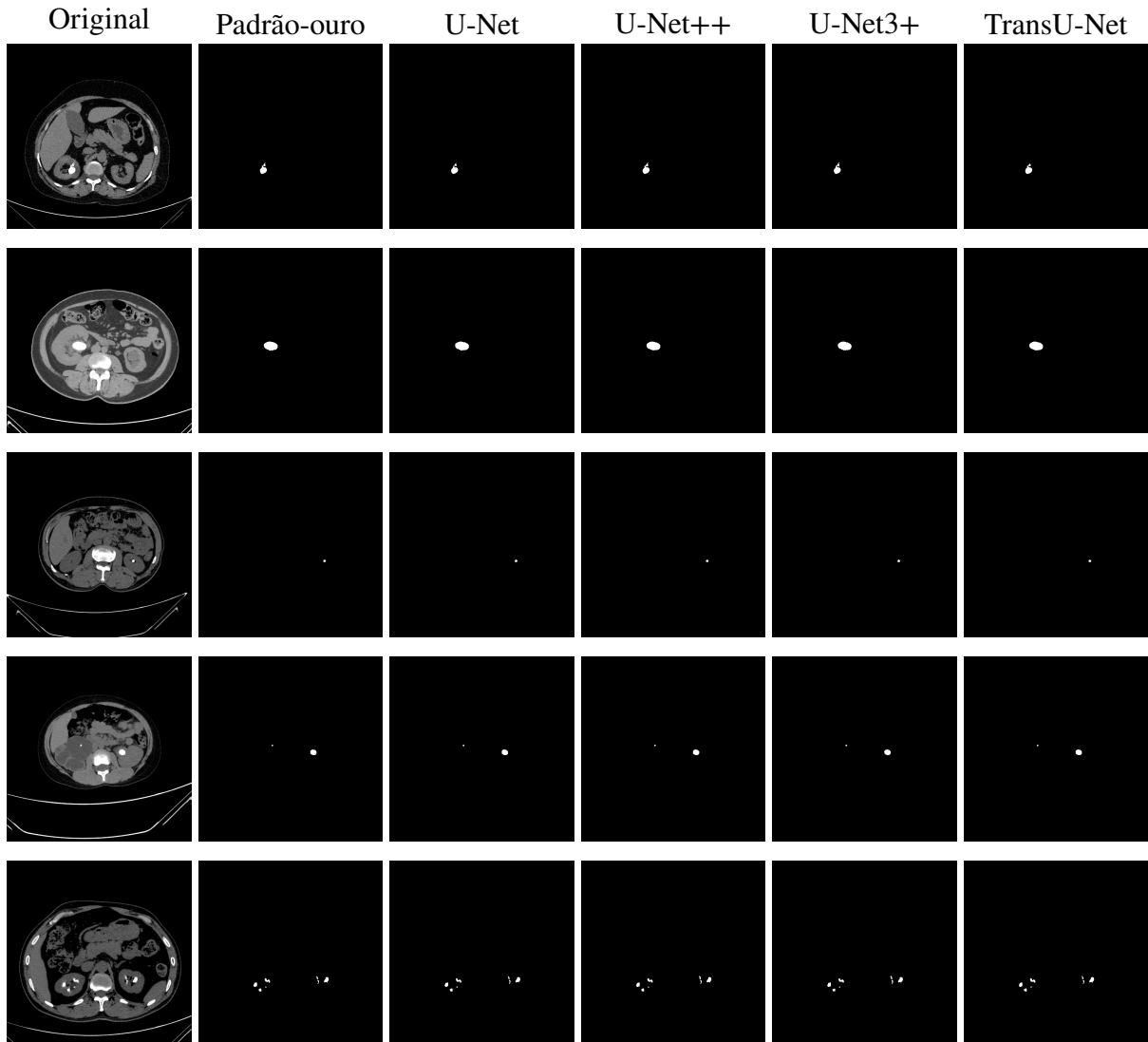
Para verificar se os valores de Dice seguiam uma distribuição aproximadamente normal, foi aplicado o teste de Shapiro–Wilk (Shaphiro; Wilk, 1965). A Tabela 8 apresenta o coeficiente Dice médio, os intervalos de confiança de 95% e os resultados do teste de normalidade de Shapiro-Wilk para cada arquitetura. Em todos os casos, os valores de p foram superiores a 0,87, indicando ausência de evidências para rejeitar a hipótese nula de normalidade ($p > 0,05$). Assim, considerou-se apropriado o uso de testes paramétricos nas comparações.

Tabela 8 – Coeficiente Dice médio, intervalos de confiança e teste de normalidade para cada arquitetura.

Arquitetura	Dice Médio	IC 95%	Shapiro-Wilk (W/p)
U-Net	0.9706	[0.9651, 0.9761]	0.9704 / 0.8775
U-Net++	0.9663	[0.9575, 0.9751]	0.9842 / 0.9556
U-Net3+	0.9669	[0.9598, 0.9740]	0.9791 / 0.9296
TransU-Net	0.9553	[0.9429, 0.9677]	0.9889 / 0.9757

Em seguida, foram realizados testes t pareados unilaterais ($A > B$) entre os modelos, a fim de verificar se uma arquitetura apresentava desempenho médio significativamente superior

Figura 36 – Resultados de segmentação obtidos por cada modelo baseado em U-Net treinado com KSSD2025. A primeira coluna mostra as amostras originais. A segunda coluna mostra as amostras de referência manualmente anotadas. A terceira coluna apresenta a previsão obtida pelo modelo U-Net. A quarta coluna apresenta a previsão obtida pelo modelo U-Net++. A quinta coluna apresenta a previsão obtida pelo modelo U-Net3+. A sexta coluna apresenta a previsão obtida pelo modelo TransU-Net.

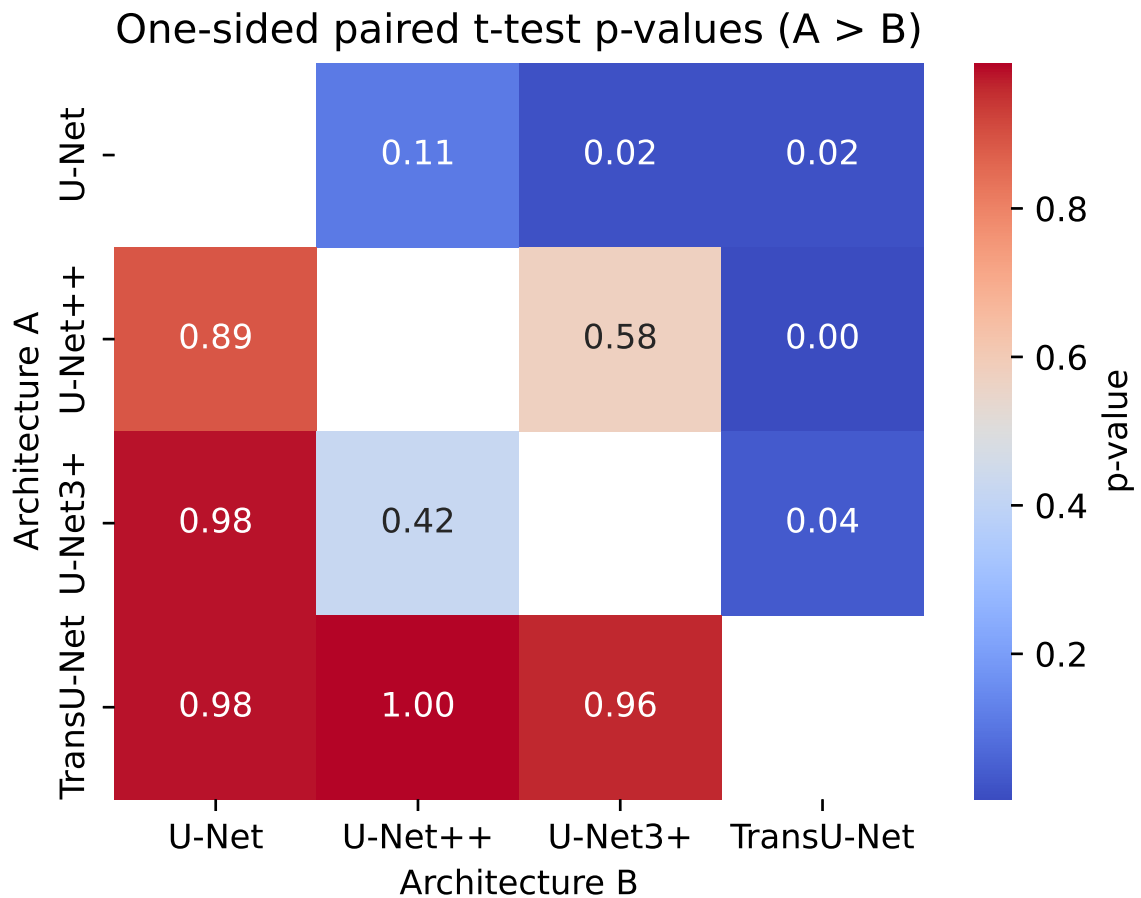


Fonte: Autor.

à outra. Para comparar os modelos de forma mais detalhada, foram realizados testes t pareados unilateral ($A > B$) entre cada par de arquiteturas. Os resultados são mostrados na Figura 37, apresentando valores de p inferiores a 0.05 que indicam rejeição da hipótese nula ($A \leq B$), sugerindo superioridade estatisticamente significativa do modelo A sobre o modelo B..

A análise indica que a U-Net alcançou o maior desempenho médio e superou significativamente tanto a U-Net3+ quanto a TransU-Net ($p = 0,02$ em ambos os casos). A U-Net++ também superou significativamente a TransU-Net ($p < 0,001$), embora seu desempenho não tenha apresentado diferença estatisticamente significativa em relação à U-Net ou à U-Net3+.

Figura 37 – Valores de p do teste t pareado unilateral ($A > B$). Valores de p menores indicam superioridade estatisticamente significativa da Arquitetura A sobre B.



Fonte: Autor.

Estes resultados corroboram a conclusão de que, sob as condições experimentais selecionadas, a arquitetura U-Net apresentou o desempenho de segmentação mais assertivo.

4.4.4 Análise dos Casos de Erro

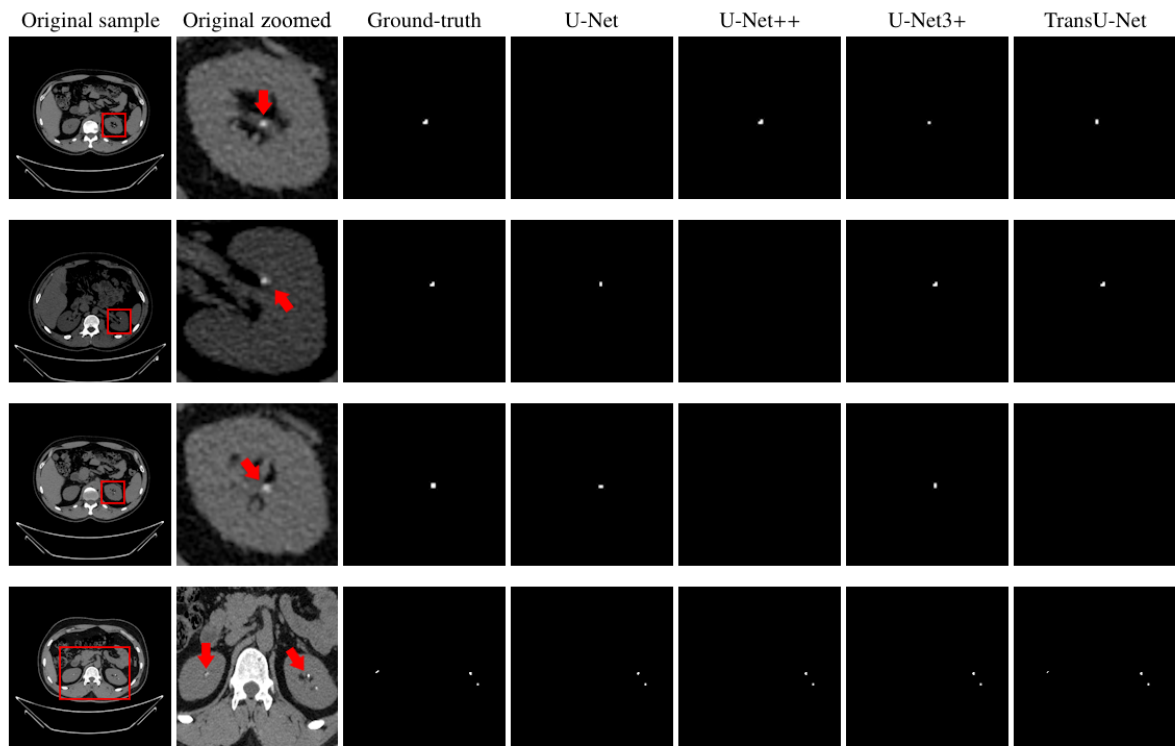
Apesar do alto desempenho observado em todos os modelos, alguns casos desafiadores levaram a erros de segmentação, como ilustrado na Fig. 38. Essas amostras destacam casos de falha em que cálculos renais foram parcial ou completamente não segmentados por um ou mais modelos.

Nos exemplos mostrados, as anotações padrão-ouro revelam a presença de pedras menores com baixo contraste, que podem ser difíceis de distinguir do tecido circundante. Alguns modelos apresentaram dificuldades para segmentar com precisão essas regiões, ocasionalmente não gerando nenhum sinal de predição, como observado nos casos do U-Net, U-Net++ e TransU-Net. Tais erros são particularmente evidentes quando as pedras estão próximas de texturas heterogêneas. Curiosamente, embora o modelo TransU-Net tenha alcançado os menores valores

médios de métricas de desempenho nos resultados da validação cruzada, ele foi o modelo que mais se aproximou da segmentação de referência no exemplo da última linha da Fig. 38. Esse resultado sugere que o TransU-Net pode ter uma maior capacidade de generalização em cenários complexos, apesar de seu desempenho geral inferior na validação cruzada do KSSD2025.

Esses casos de erro enfatizam as limitações dos modelos ao lidar com alvos sutis em condições desafiadoras. Abordar essas questões poderia envolver a incorporação de mecanismos de atenção ajustados à detecção de pequenos objetos ou o aumento do conjunto de dados com casos difíceis mais representativos.

Figura 38 – Os resultados de segmentação revelaram que os modelos baseados em U-Net apresentaram erros. A primeira coluna exibe as amostras originais. A segunda coluna exibe as amostras originais ampliadas. A terceira coluna exibe a amostra padrão-ouro ampliada. A quarta coluna mostra a imagem de predição obtida pelo modelo U-Net. A quinta coluna mostra a imagem de predição obtida pelo modelo U-Net++. A sexta coluna mostra a imagem de predição obtida pelo modelo U-Net3+. A sétima coluna mostra a imagem de predição obtida pelo modelo TransU-Net.



Fonte: Autor.

4.4.5 Avaliação dos Modelos em Dados Externos

Para avaliar ainda mais a capacidade de generalização dos modelos, foi realizado um teste suplementar utilizando amostras de um conjunto de dados externo, coletado por radiologistas do Hospital Universitário da Universidade de São Paulo, não incluído no processo de treinamento

ou validação do KSSD2025. Este conjunto de dados externo continha amostras divididas por partes anatômicas do sistema urinário onde os cálculos estavam localizados. A Tabela 9 resume os valores de Dice obtidos para cada localização anatômica, especificamente o rim direito (5 amostras), rim esquerdo (3 amostras) e ureter (9 amostras), totalizando 17 amostras do conjunto de teste externo.

Tabela 9 – Coeficientes de Dice (%) obtidos por cada modelo quando testados em um conjunto de dados externo, segmentado por região anatômica.

Modelo	Rim Direito (5 amostras)	Rim Esquerdo (3 amostras)	Ureter (9 amostras)	Média (17 amostras)
U-Net	84,05%	81,78%	60,49%	75,44%
U-Net++	84,05%	83,45%	73,94%	80,48%
U-Net3+	84,90%	83,45%	91,99%	86,78%
TransU-Net	84,05%	84,67%	72,57%	80,45%

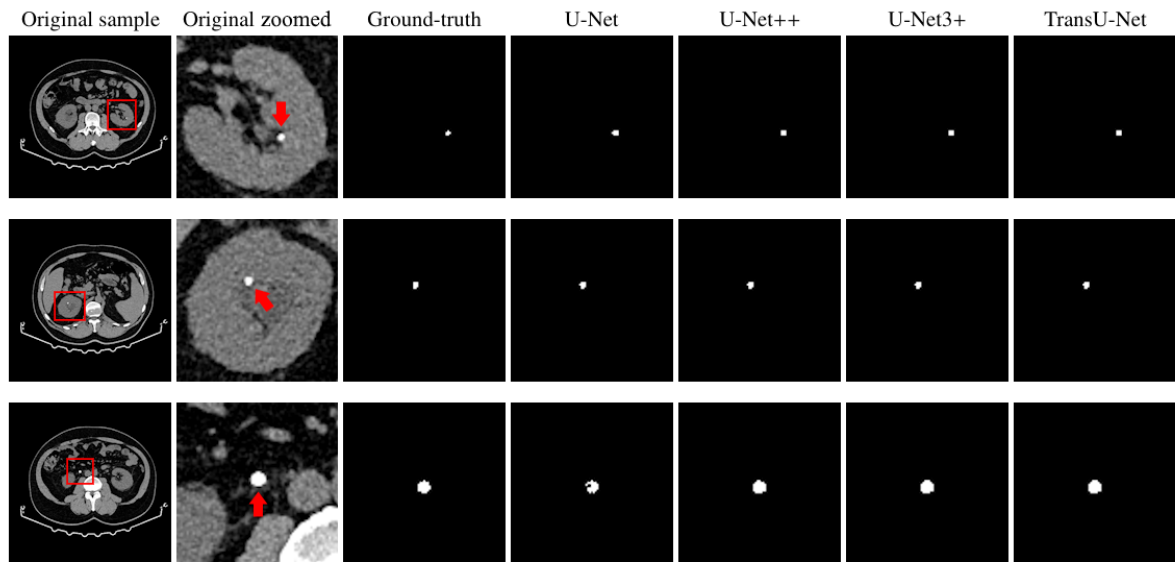
Os resultados revelam uma queda perceptível de desempenho em comparação com a validação cruzada do KSSD2025 (Seção 4.4.1), particularmente na segmentação de cálculos ureterais. Esse declínio era esperado devido a possíveis variações nos parâmetros de aquisição de imagem, níveis de contraste e anatomia dos pacientes entre os conjuntos de dados.

Entre os modelos avaliados, o U-Net3+ alcançou a maior média de Dice (86,78%). U-Net, U-Net++ e TransU-Net apresentaram desempenho ligeiramente inferior, especialmente na detecção de cálculos ureterais, como mostrado na Tabela 9, os quais são mais difíceis de detectar devido ao seu pequeno tamanho e baixo contraste.

A Figura 39 ilustra exemplos representativos dos resultados de segmentação. A análise visual confirma que a maioria das máscaras de segmentação geradas pelos modelos permanece razoavelmente próxima ao padrão-ouro, especialmente para cálculos renais. No entanto, a segmentação de cálculos ureterais continua inconsistente, devido à localização anatômica distinta, ao menor tamanho dos cálculos e à posição mais variável dentro do trato urinário. Diferentemente dos cálculos renais, que tendem a ser maiores e centrados nas regiões dos rins, os cálculos ureterais frequentemente aparecem em áreas estreitas e tortuosas, podendo ser facilmente confundidos com tecidos moles adjacentes. Essas diferenças anatômicas e espaciais desafiam a capacidade de generalização dos modelos, sobretudo quando treinados principalmente em casos de cálculos renais. Essa análise sugere a necessidade de melhorias ou ajuste fino dos modelos ao lidar com diferentes cenários clínicos.

Esses achados destacam a importância de avaliar os modelos com dados externos e demonstram que as arquiteturas treinadas no KSSD2025 alcançaram um nível promissor de generalização. Uma maior robustez poderia ser atingida com a inclusão de amostras de treinamento mais diversas, provenientes de outras fontes.

Figura 39 – Resultados de segmentação em amostras de dados externos. A primeira linha mostra os resultados obtidos em uma amostra de cálculo no rim esquerdo. A segunda linha mostra os resultados obtidos em uma amostra de cálculo no rim direito. A terceira linha mostra os resultados obtidos em uma amostra de cálculo no ureter. Todos os modelos foram treinados exclusivamente no conjunto de dados KSSD2025.

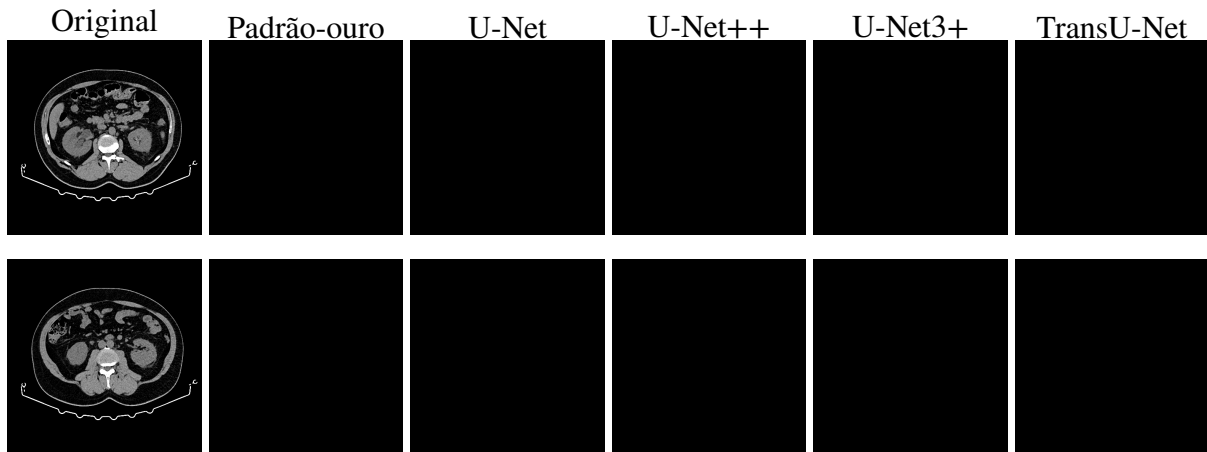


Fonte: Autor.

4.4.6 Avaliação dos Modelos em Amostras Sem Cálculos Renais

Para avaliar e verificar a capacidade dos modelos treinados em evitar falsos positivos, também foi avaliada o desempenho em amostras de tomografia computadorizada sem cálculos renais. Esses casos são críticos em ambientes clínicos, em que a segmentação incorreta de pedras inexistentes poderia levar a intervenções desnecessárias. As imagens foram coletadas por radiologistas do Hospital Universitário da Universidade de São Paulo e não foram incluídas no processo de treinamento ou validação do KSSD2025. A Fig. 40 apresenta resultados qualitativos nessas amostras. Todos os quatro modelos baseados na arquitetura U-Net, U-Net, U-Net++, U-Net3+ e TransU-Net, produziram máscaras de segmentação vazias, identificando corretamente a ausência de cálculos renais. Essas saídas são consistentes com as anotações padrão-ouro, que não contêm regiões rotuladas. Esse comportamento demonstra que os modelos são capazes não apenas de detectar cálculos renais quando presentes, mas também de evitar falsos positivos em rins saudáveis. Nesta avaliação específica, todos os modelos alcançaram 100% de acurácia na segmentação e taxa de falsos positivos de 0%, correspondendo perfeitamente às máscaras de referência e confirmando sua confiabilidade em casos negativos. Embora a correspondência tenha sido perfeita nas amostras testadas neste estudo, experimentos adicionais com um número maior de amostras sem cálculos renais são necessários para validar a eficácia dos modelos propostos na prevenção de falsos positivos.

Figura 40 – Resultados de segmentação em amostras sem cálculos renais. A primeira coluna mostra as fatias originais de TC. A segunda coluna apresenta as máscaras padrão-ouro anotadas manualmente. Da terceira à sexta colunas estão as máscaras de segmentação previstas pelos modelos U-Net, U-Net++, U-Net3+ e TransU-Net, respectivamente. Todos os modelos foram treinados exclusivamente no conjunto de dados KSSD2025.



Fonte: Autor.

4.4.7 Análise de Complexidade

A Tabela 10 resume a complexidade computacional de cada modelo baseado em U-Net. Como esperado, a U-Net é a arquitetura mais eficiente, exigindo menos memória (3,2 GB), possuindo o menor número de parâmetros (1,39M) e apresentando o menor tempo de treinamento. A U-Net++ e a U-Net3+ introduzem *skip connections* e mecanismos de supervisão mais sofisticados, resultando em aumentos moderados de complexidade. Em contraste, a TransU-Net, que integra blocos do tipo transformer, é a mais exigente computacionalmente, com 4,51M de parâmetros e consumo de 18 GB de memória.

Apesar de sua simplicidade, a U-Net alcançou o melhor desempenho durante a validação cruzada, enquanto a U-Net3+ obteve os melhores resultados em dados externos. Isso evidencia que, nesta tarefa de segmentação binária com um conjunto de dados limitado, arquiteturas convolucionais mais simples podem superar modelos mais complexos. A relativa baixa dificuldade da tarefa e o pequeno número de amostras no conjunto de treinamento provavelmente limitaram as vantagens de modelos mais profundos e baseados em transformers, que normalmente requerem bases de dados maiores para alcançar seu pleno potencial.

Tabela 10 – Informações de complexidade de cada modelo baseado em U-Net utilizado neste trabalho.

Modelo	Parâmetros Treináveis (M)	Uso de Memória (GB)	Tempo Médio de Treinamento (s)	Tamanho do Modelo (MB)
U-Net	1.39	3.2	1416	16.5
U-Net++	1.94	8.5	1859	23
U-Net3+	2.31	12	2066	27.3
TransU-Net	4.51	18	3254	53.3

4.4.8 Paralelo com o estado da arte

Os resultados obtidos foram comparados com dois estudos recentes que também abordaram a tarefa de segmentação de cálculos renais (Kim, Y. I. *et al.*, 2023; Choi *et al.*, 2025). A Tabela 11 resume as principais diferenças entre essas abordagens e o estudo proposto, destacando aspectos metodológicos e indicadores de desempenho.

Tabela 11 – Comparação entre o estudo proposto e abordagens do estado da arte.

Artigo	Nosso trabalho	Young In Kim <i>et al.</i> (2023)	Choi <i>et al.</i> (2025)
Tamanho do dataset	838	410	63
Arquiteturas de rede	U-Net, U-Net++, U-Net3+, and TransU-Net	U-Net++	PFSH-Net
Tamanho do batch	16	12	2
Épocas	150	100	150
Early stop	No	10 epochs	No
Função de perda	Dice	Dice + 0.1CE	0.3Dice + 0.7CE
Validação cruzada	5-fold	5-fold	5-fold
Aumento de dados	Yes	Yes	No
Perspectiva da TC	Axial	Axial, Coronal and Sagittal	Axial
Dice score	97.06%	88.44%(additive model) 87.56% (axial only)	65.08% (Kidney stone)
Clinical Analysis	Não	Sim	Não
Subject	Humano	Humano	Cão

O KSSD2025 compreende 838 imagens de TC anotadas, mais do que o dobro das 410 amostras usadas por Young In Kim *et al.* (2023) e significativamente maior que as 63 imagens usadas por Choi *et al.* (2025). Esse conjunto de dados mais amplo provavelmente contribuiu para o melhor desempenho de segmentação observado em nosso trabalho, já que uma maior quantidade de amostras normalmente melhora a capacidade de generalização dos modelos de aprendizado profundo.

Em termos de arquitetura de rede, Young In Kim *et al.* (2023) utilizaram U-Net++, enquanto Choi *et al.* (2025) empregaram a PFSH-Net, uma arquitetura personalizada. O estudo proposto realizou uma comparação mais ampla entre quatro arquiteturas: U-Net, U-Net++, U-Net3+ e TransU-Net. A U-Net, apesar de ser a menos complexa computacionalmente, obteve a melhor pontuação de Dice durante a validação cruzada (97,06%). Em contraste, Young In Kim *et al.* (2023) relataram 87,56% para fatias axiais usando U-Net++ e 88,44% ao combinar vistas axiais, coronais e sagitais. Choi *et al.* (2025) obtiveram uma pontuação inferior de 65,08% com a PFSH-Net. É importante notar que o trabalho de Choi *et al.* (2025) focou na detecção de cálculos renais em cães ao invés de humanos. Essa diferença de espécie pode introduzir variabilidade e complexidade adicionais na tarefa de segmentação devido a diferenças anatômicas e fisiológicas, o que pode explicar o desempenho consideravelmente inferior observado nos resultados deles.

Curiosamente, embora Young In Kim *et al.* (2023) tenham utilizado três perspectivas de TC e uma estratégia de *early stop*, seu modelo não superou os resultados obtidos neste trabalho, que usaram apenas a perspectiva axial e sem *early stopping*. Isso sugere que a vista axial sozinha

foi suficiente para alcançar alto desempenho na segmentação binária de cálculos renais em nosso dataset. Além disso, isso sugere que a acurácia do modelo pode depender mais da qualidade e quantidade dos dados do que da complexidade arquitetural ou do uso de entradas de diferentes perspectivas.

Quanto aos protocolos de treinamento, aqui foi utilizado um tamanho de batch maior (16 contra 12 e 2) e o modelo foi treinado por mais épocas que Young In Kim *et al.* (2023) (100 épocas) e igual a Choi *et al.* (2025) (150 épocas). Também foi aplicado aumento de dados, ao contrário de Choi *et al.* (2025), o que pode explicar parte da diferença de desempenho. Além disso, Young In Kim *et al.* (2023) e Choi *et al.* (2025) usaram funções de perda compostas combinando Dice e termos de Cross-Entropy, enquanto nós utilizamos apenas a função Dice.

Um aspecto notável do trabalho de Young In Kim *et al.* (2023) é a etapa de análise clínica, na qual compararam as saídas do modelo com medições clínicas obtidas de registros de pacientes. Embora o estudo proposto não tenha incluído tal análise, o seu valor é reconhecido como uma direção relevante para trabalhos futuros.

Em resumo, a abordagem proposta alcançou um alto desempenho de segmentação utilizando apenas uma única vista de TC e arquiteturas relativamente simples. Os resultados obtidos podem ser atribuídos ao uso de um conjunto de dados maior e bem anotado, além de pequenas modificações na U-Net proposta, resultando em um modelo leve e de alto desempenho para segmentação.

4.4.9 Aplicações Clínicas

A segmentação automática de cálculos renais utilizando técnicas de deep learning representa um avanço valioso no fluxo de trabalho clínico para o diagnóstico e tratamento de doenças renais. O alto desempenho de segmentação alcançado pelos modelos treinados neste estudo, em especial pelas arquiteturas U-Net e U-Net3+, demonstra seu potencial para integração em sistemas de diagnóstico auxiliado por computador, oferecendo suporte a radiologistas e urologistas na análise de exames de tomografia computadorizada com maior eficiência e precisão.

Um passo natural em direção à adoção clínica é a implementação do nosso modelo em sistemas de Arquivamento e Comunicação de Imagens (PACS), permitindo o processamento em tempo real das imagens de TC e fornecendo aos clínicos sobreposições automáticas das regiões segmentadas de cálculos renais diretamente na interface de visualização. Essa automação pode reduzir significativamente o tempo e a carga cognitiva da anotação manual, aumentar a reprodutibilidade e minimizar a variabilidade entre observadores. Entretanto, a implementação clínica requer validação rigorosa em diferentes ambientes hospitalares para garantir a robustez do modelo. É necessário considerar a variabilidade da anatomia dos pacientes, dos modelos de tomógrafos, dos protocolos de aquisição de imagens e dos níveis de ruído.

Uma das principais preocupações é a capacidade dos modelos de generalizar entre diferentes dispositivos de imagem, protocolos de aquisição, instituições e perfis demográficos de

pacientes. Modelos treinados com dados de um único scanner ou população podem apresentar queda de desempenho quando aplicados em novos ambientes clínicos devido à mudança de domínio. Para lidar com isso, trabalhos futuros podem envolver técnicas de adaptação de domínio, métodos de harmonização de dados ou treinamento com conjuntos de dados multi-institucionais. Além disso, estratégias como aprendizado por transferência e aprendizado federado podem promover robustez ao mesmo tempo em que preservam a privacidade dos dados. Antes da implementação, é necessária uma validação extensiva em cenários clínicos reais para assegurar segurança, confiabilidade e integração nos fluxos de diagnóstico.

Embora o modelo proposto tenha apresentado alto desempenho no conjunto de dados utilizado, testes adicionais em diversas condições clínicas são essenciais para avaliar sua capacidade de generalização e evitar potenciais erros diagnósticos.

4.4.10 Limitações do Estudo

Apesar dos resultados encorajadores, algumas limitações devem ser reconhecidas. Primeiramente, o conjunto de dados utilizado neste estudo foi coletado em condições relativamente controladas, com imagens adquiridas em um número limitado de fontes (apenas um hospital); isso pode afetar a capacidade do modelo de generalizar para imagens de TC obtidas com diferentes dispositivos ou protocolos.

Em segundo lugar, embora estudos anteriores tenham demonstrado a utilidade de incorporar múltiplas perspectivas de TC (axial, coronal e sagital) para melhorar o desempenho da segmentação, o modelo apresentado foi treinado e validado apenas com imagens axiais. Embora os resultados obtidos sugiram que a visão axial sozinha pode fornecer uma segmentação altamente precisa para esta tarefa, a incorporação de informações multivisuais pode aumentar ainda mais a robustez e deve ser investigada em pesquisas futuras.

Em terceiro lugar, o presente trabalho não inclui correlação clínica entre as segmentações e métricas relevantes em nível de paciente, tais como carga litiásica, densidade de Hounsfield ou desfechos clínicos. Essa omissão limita a interpretabilidade clínica direta dos resultados encontrados e dificulta sua comparação com decisões diagnósticas e terapêuticas baseadas em informações de referência.

Além disso, é reconhecida a ausência de um estudo de ablação neste trabalho. A contribuição deste trabalho não envolve a adição de novos módulos arquiteturais, mas sim uma modificação mínima em modelos baseados na U-Net, reduzindo o tamanho dos filtros nas camadas convolucionais com o objetivo de explorar configurações mais leves e de menor custo computacional. Assim, uma análise de ablação focada em isolar o efeito de múltiplos componentes arquiteturais não é diretamente aplicável aqui. No entanto, avaliar o impacto isolado dessa modificação em diferentes arquiteturas é uma linha promissora para estudos futuros.

Por fim, embora arquiteturas baseadas em *transformers*, como a TransUNet, tenham sido incluídas, seu desempenho não superou o dos modelos baseados em U-Net; isso pode ser atri-

buído ao tamanho relativamente pequeno do nosso conjunto de dados, visto que modelos *transformer* normalmente requerem uma quantidade significativamente maior de dados para explorar plenamente sua capacidade de representação.

4.4.11 Direções Futuras

Trabalhos futuros se concentrarão em melhorar a aplicabilidade clínica e a capacidade de generalização dos modelos propostos. Uma prioridade é expandir o conjunto de dados com imagens adicionais de TC provenientes de múltiplos centros médicos, incluindo exames obtidos com diferentes dispositivos e protocolos de imagem, a fim de possibilitar a generalização de domínio e aumentar a robustez dos modelos em cenários clínicos reais.

Pretende-se também investigar a fusão de diferentes perspectivas multiplanares de TC, incluindo axial, coronal e sagital, utilizando estratégias de fusão precoce ou tardia. Essa abordagem pode capturar características espaciais mais abrangentes e potencialmente levar a melhorias adicionais na precisão da segmentação, especialmente para cálculos renais de formato irregular ou de pequeno tamanho.

Outra direção é a realização de um estudo de validação clínica no qual os cálculos renais segmentados serão comparados com medições manuais de tamanho, volume e densidade das pedras, permitindo a extração automática de características clinicamente relevantes que possam auxiliar médicos no diagnóstico, planejamento terapêutico e monitoramento da progressão ou recorrência da doença.

Do ponto de vista técnico, a exploração mais aprofundada de arquiteturas híbridas, como CNN-transformers, pode ser benéfica, particularmente em conjunto com a expansão do conjunto de dados. Além disso, a otimização da eficiência computacional por meio de técnicas de *model pruning* pode viabilizar aplicações em tempo real em ambientes clínicos com recursos limitados.

Por fim, o objetivo é reduzir a lacuna entre desempenho técnico e utilidade clínica. Para isso, pretende-se desenvolver um sistema protótipo integrado ao PACS, exportando a máscara de segmentação como um arquivo JSON contendo as coordenadas dos contornos ou como um objeto DICOM com anotações incorporadas. Esses formatos podem ser automaticamente sobrepostos às imagens originais de TC em estações de trabalho de radiologia, permitindo uma visualização aprimorada e reduzindo a carga manual para os radiologistas. Essa funcionalidade poderá ser avaliada futuramente em colaboração com departamentos de radiologia, contribuindo para a aplicação clínica de ferramentas de segmentação de cálculos renais baseadas em aprendizado profundo.

5 PLANEJAMENTO AUTOMÁTICO DE TRAJETÓRIA 3D DE URETEROSCOPIA

Neste capítulo é apresentada uma prova de conceito para planejamento de trajetórias 3D de cirurgias de ureteroscopia para tratamento de cálculos urinários. O método se baseia na combinação das informações de diagnóstico automático obtidas nas etapas dos Capítulos 3 e 4, e em informações de mapeamento do sistema urinário do paciente obtidas a partir de outro exame rotulado manualmente. A partir dos dados de entrada, utilizou-se métodos de busca baseados em IA para encontrar o caminho cirúrgico ótimo e, então, é feita uma comparação entre filtros de processamento de sinais e geração de curvas paramétricas para a suavização da trajetória. Devido à escassez de bases de dados disponíveis, a metodologia proposta foi avaliada em apenas um caso simulado a partir de exames coletados de um único paciente.

5.1 CONTEXTUALIZAÇÃO

A ureteroscopia flexível é um procedimento endoscópico minimamente invasivo utilizado para diagnóstico e tratamento de cálculos no trato urinário superior. Um dos principais desafios nesse tipo de procedimento é a definição de uma trajetória segura, eficiente e anatômicamente viável para alcançar o cálculo urinário, respeitando a geometria do sistema urinário de cada paciente.

Com os avanços recentes em processamento de imagens médicas e IA, tornou-se possível reconstruir modelos tridimensionais personalizados do trato urinário a partir de imagens de TC com contraste. Essa reconstrução permite a visualização detalhada das estruturas anatômicas, possibilitando também a segmentação precisa do cálculo urinário e o planejamento individualizado da navegação endoscópica.

No entanto, o problema do planejamento automático de trajetória apresenta desafios significativos. A anatomia do trato urinário é altamente variável entre pacientes e frequentemente inclui regiões de curvatura acentuada, estreitamentos e bifurcações, o que dificulta a geração de caminhos viáveis. Além disso, é necessário garantir que a trajetória proposta seja suficientemente suave para evitar sobrecarga mecânica sobre o ureteroscópio e minimizar o risco de colisões com as paredes do órgão.

Para lidar com essas questões, este estudo propõe um método semi-automático, considerando as seguintes etapas: (i) segmentação do cálculo urinário e das estruturas anatômicas relevantes; (ii) busca da trajetória inicial utilizando algoritmos clássicos de busca; (iii) aplicação de técnicas de suavização para otimizar o trajeto, garantindo maior compatibilidade anatômica. A definição do caminho final é guiada por uma função objetivo multiobjetivo que considera métricas geométricas e anatômicas, como curvatura, torção, risco de colisão e erro de alinhamento com o alvo final.

5.2 TRABALHOS RELACIONADOS

O planejamento de trajetória consiste em encontrar um caminho ótimo e livre de colisões, partindo de um ponto inicial até um ponto de destino, considerando restrições físicas e geométricas de um modelo virtual do ambiente em que será feita a navegação. Para isso, é necessário que se tenha um ambiente estruturado para calcular automaticamente os possíveis caminhos e então pode-se encontrar a trajetória otimizada utilizando técnicas de IA para busca e otimização. Esse procedimento tem sido utilizado para auxiliar no planejamento de cirurgias, sendo considerado um modelo 3D de algum órgão de interesse como ambiente navegável (Qin *et al.*, 2023), (Han *et al.*, 2022). Foram considerados apenas os trabalhos em que foi feita a geração automática de trajetórias 3D em cirurgias a partir de TCs.

Um algoritmo de planejamento de trajetória para cirurgia vascular cerebral foi proposto por Guo *et al.* (2017), a partir de imagens de tomografia computadorizada. Para pré-processar as imagens, foi utilizado o programa *Mimic* para segmentação por limiarização. Para a reconstrução 3D, foi utilizado o programa *Maya*¹, com o modelo 3D exportado para o programa *Unity*², em que foi conduzido o problema de planejamento de trajetória. O algoritmo de planejamento de trajetória proposto é baseado em matrizes de projeção, utilizando informações de todo o ambiente para evitar obstáculos e gerar uma rota até o ponto de destino. Para os experimentos, foi utilizado um dispositivo que simula a operação de um cateter e captura a trajetória real realizada pelo médico. Essa trajetória capturada foi comparada com a trajetória gerada pelo algoritmo, sendo obtido um erro médio de 0,3 milímetros.

O trabalho de Ciszkiwicz e Milewski (2018) teve como objetivo desenvolver um procedimento para planejamento de trajetória de cirurgias minimamente invasivas de joelho. A partir de tomografias, foi construída uma malha 3D que representasse o joelho. Para encontrar uma trajetória livre de colisões, foi utilizado o Algoritmo Genético de codificação real. A trajetória encontrada foi otimizada localmente pelo método Nelder-Mead (Singer; Nelder, 2009). Assim, foram realizadas simulações em três cenários diferentes, ocorrendo a colisão da trajetória com o osso da tíbia em todas as tentativas. Apesar disso, os autores afirmaram que a abordagem híbrida proposta conseguiu reduzir o tempo computacional e pode permitir otimização de cirurgias em tempo real.

No contexto de cirurgias de correção interna da coluna vertebral, o trabalho de Qi Zhang *et al.* (2018) propôs um método para planejamento de trajetória baseado em aprendizado por reforço. Primeiro, foi feita a segmentação de imagens de tomografia e a reconstrução 3D dos órgãos internos para construir a estrutura navegável. Em seguida, a trajetória cirúrgica foi obtida utilizando o algoritmo *Q-learning* (Watkins; Dayan, 1992). Como resultado, foi obtida uma trajetória 3D partindo de um ponto inicial até um ponto final, evitando órgãos vitais e tecidos do corpo humano.

¹<https://www.autodesk.com.br/products/maya/overview>

²<https://unity.com/pt>

O objetivo do trabalho de Cai *et al.* (2019) foi realizar um planejamento de trajetória em cirurgias de inserção de parafusos pediculares utilizando técnicas de aprendizado profundo. Foram utilizadas imagens de tomografias com anotações de pontos de controle que representam o ponto de inserção do parafuso e a sua direção. Então, utilizaram redes do tipo *U-Net 3D* (Çiçek *et al.*, 2016) para extrair características da tomografia e segmentar o corpo vertebral da coluna do paciente. Também foi utilizada uma rede neural do tipo perceptron multicamadas (Rumelhart; Hinton; Williams, 1986) para localizar os pontos de controle na segmentação realizada na etapa anterior. Para validar o trabalho, foram utilizados 21 conjuntos de tomografias de pacientes diferentes. Sobre os valores de métricas utilizadas, foi obtido um coeficiente de Dice médio de 96,67 e coeficiente de Jaccard médio de 93,05.

No trabalho de Fan Yang *et al.* (2019) foi proposto um algoritmo de planejamento de trajetória dentro do sistema urinário para um ureteroscópio flexível utilizando imagens de tomografia. Inicialmente, foi feito o pré-processamento da tomografia utilizando um filtro bilateral, sendo um filtro não-linear que considera a distância geométrica e a diferença de intensidade de cor entre os píxeis. (Tomasi; Manduchi, 1998). Em seguida, foi utilizada limiarização e a técnica *Region Growing* (Adams; Bischof, 1994) para segmentar as regiões de lúmen do ureter e da pélvis renal. Para a reconstrução 3D volumétrica e da superfície, foi utilizado o algoritmo *Marching Cubes*. Para planejar a rota, foi proposto um algoritmo de extração de linha central baseado em um algoritmo de distância de fronteira. Os pontos da imagem segmentada são classificados em pontos de fronteira, internos e externos. Em seguida, foi estabelecido o campo de distância de fronteira entre os pontos internos e pontos de fronteira. Posteriormente, foi calculado o campo de distância entre os pontos internos e o ponto inicial. Simultaneamente, foi construída uma árvore de custo máximo utilizada para traçar a linha central partindo do ponto inicial até o ponto final. Na última etapa, foram aplicadas restrições matemáticas para simular as limitações de flexibilidade e distância do ureteroscópio. Como resultados, apenas foi apresentada a trajetória gerada do ureteroscópio por dentro da reconstrução do sistema urinário humano.

A dissertação de Scarzanella (2020) também seguiu no escopo de planejamento de trajetória para ureteroscopia. Esse trabalho teve como objetivo projetar um protótipo de ureteroscópio e implementar uma estratégia de planejamento de trajetória específica. O protótipo foi construído de forma que possa ser controlado remotamente por *joysticks*. Para a estratégia de planejamento de trajetória, primeiro foi feita a modelagem 3D manualmente do sistema urinário. Em seguida, a estratégia foi dividida em dois segmentos: quando a trajetória está na região da uretra, ela foi definida pelo seu ponto médio. Na região da bexiga, a trajetória foi definida pelos arcos de circunferência para cada plano XZ , YZ , e YX , respectivamente. Para a validação do trabalho, o experimento foi configurado de forma que o protótipo construído operasse em um modelo de silicone do sistema urinário feito por uma impressora 3D. O experimento foi realizado com 3 participantes, em 2 tipos de testes, guiado de forma visual e guiado por voz. Para as métricas de avaliação, foi considerada a posição final, o tempo em segundos e o tamanho da trajetória. Sobre os resultados obtidos, o orifício alvo foi alcançado em todos os testes. Nos

testes realizados, o grupo guiado por voz realizou a trajetória de menor tamanho. No entanto, os autores apontaram que os resultados não foram estatisticamente significativos.

Dentro do contexto de cirurgia de biópsia pulmonar, Bao *et al.* (2020) propuseram um método de planejamento inteligente de trajetória. Foi feita a aquisição de imagens de tomografia do tórax, que foram pré-processadas utilizando operações morfológicas para segmentar os pulmões. Em seguida, foram encontradas as áreas de inserções para trajetórias efetivas por meio de restrições definidas. O planejamento de trajetória foi considerado como um problema de otimização multi-objetivo, sendo utilizado o método *Pareto* (Craven; Islam, 2005) para resolvê-lo. A trajetória automática foi comparada com uma trajetória gerada por especialistas em 6 pacientes distintos. Os resultados mostraram que todas as trajetórias calculadas pelo método alcançaram as necessidades cirúrgicas, obtendo um desempenho melhor em 75% comparado com a trajetória ótima escolhida por um especialista.

No trabalho de Yucheng He *et al.* (2020) foi proposto um método de planejamento de trajetória para abordagens cirúrgicas endoscópicas nasais baseado em buscas em mapas de píxel de imagens médicas. Através da análise da trajetória da ponta do endoscópio nasal durante a cirurgia, os requisitos para o planejamento de rota da ponta do endoscópio foram obtidos. Em seguida, considerando a anatomia da cavidade nasal, foi construído um mapa 3D binarizado, baseado na tomografia do paciente, contendo a anatomia espacial da cavidade nasal. Assim, foi obtida uma trajetória segura partindo do ponto de entrada nasal até a área de operação utilizando o algoritmo de busca A* (Hart; Nilsson; Raphael, 1968). Por último, foi desenvolvido um sistema endoscópico nasal virtual para testar o método em um modelo contendo tecido nasal. O método foi validado através de uma navegação virtual endoscópica nasal automática, porém, não foi feita uma avaliação clínica com médicos especialistas.

Outro trabalho que utilizou um algoritmo de busca para planejamento de trajetória foi o de Li, Dankelman e De Momi (2021). Neste caso, foi proposta uma rápida abordagem de planejamento de rota para cateteres flexíveis, considerando a curvatura de um robô cirúrgico e restrições de tempo. Nessa proposta, a estrutura vascular foi extraída, representada por linhas de centro vasculares e os seus correspondentes raios vasculares. Então, foi feita a busca do caminho pela linha vascular central utilizando o algoritmo BFS (Breadth-First Search) (Lee, 1961) e otimizada localmente com o Algoritmo Genético. Os resultados mostraram a eficiência do método em cenários de espaço bi-dimensional e tri-dimensional. Também foi mostrada a habilidade de satisfazer as restrições de curvatura do robô em baixo tempo computacional, comparado com outros métodos, atingindo uma frequência de atualização de aproximadamente 6,4 Hz

Um método para planejamento automático de trajetória cirúrgica de inserção de parafuso pedicular foi proposto por Qi *et al.* (2022). A partir de tomografias computadorizadas, a coluna vertebral do paciente foi segmentada utilizando uma arquitetura de rede neural baseada em 3D *U-Net* (Ronneberger; Fischer; Brox, 2015), sendo feita a classificação do tipo de vértebras com uma rede do tipo perceptron multicamadas. Em seguida, foi estabelecido um sistema de coordenadas local em cada vértebra, de acordo com suas características. Baseado nesse sistema de

coordenadas, foi feita a correção da orientação da vértebra por meio de uma matriz de rotação que permite o cálculo da trajetória dos parafusos através da interseção entre pontos da vértebra. Para validar o método, foram utilizados exames de tomografia de 8 pacientes diferentes, sendo a trajetória calculada avaliada por cirurgiões especialistas. Os cirurgiões consideraram como trajetórias aplicáveis em cirurgias reais 93% dos resultados gerados.

No trabalho de Hu, Jiang e Wang (2022) foi proposto um método de planejamento de trajetória para cirurgia de punção de tumor hepático guiada por tomografia computadorizada. As imagens de tomografia foram segmentadas utilizando o método *nnUNet* (Ma *et al.*, 2021). O número de voxels do conjunto de imagens foi otimizado para diminuir o volume de renderização. Então, foi feita a reconstrução 3D de um modelo do fígado e dos vasos hepáticos. Em seguida, foram selecionados múltiplos pontos de entrada da trajetória cirúrgica utilizando agentes de aprendizado por reforço, baseando-se em restrições da punção realizada na cirurgia. O ponto de entrada ótimo foi selecionado baseado em seu tamanho, sendo a trajetória ótima completa obtida através do treinamento de uma rede do tipo *Q-learning* (Seo *et al.*, 2019) de dupla profundidade. A trajetória gerada foi avaliada por dois cirurgiões especialistas, que deram uma pontuação de 1 a 5 para cada resultado. Apenas uma das trajetórias foi avaliada com 4 pontos, enquanto o restante recebeu 5 pontos pela análise dos médicos, mostrando a relevância clínica do método.

Outro trabalho que propôs um algoritmo para planejamento de trajetória em cirurgias de punção foi o de Ling He *et al.* (2023), porém no contexto de biópsia pulmonar. O algoritmo recebe um conjunto de imagens de tomografia, segmentadas utilizando o método de Otsu (Otsu, 1979) e operações morfológicas para extração de bordas. Em seguida, foram definidas as posições de entrada de punção e as posições alvo para determinar a trajetória da agulha de punção. Baseando-se nos requisitos clínicos de biópsia de pulmão, foram propostas 9 restrições para determinar a trajetória ótima. Para otimizar a trajetória e obter a trajetória ótima, foi utilizado o algoritmo de otimização Pareto (Craven; Islam, 2005). Para os experimentos, foram utilizados exames de tomografia de 33 pacientes. A trajetória gerada pelo algoritmo foi comparada com uma trajetória cirúrgica realizada por um cirurgião, com os 8 melhores casos obtendo uma distância média de 5,77mm do trajeto real.

Acar *et al.* (2024) propuseram um sistema de navegação para cirurgia renal endoscópica baseado no registro entre reconstruções 3D do sistema coletor renal obtidas a partir de vídeos de endoscopia, utilizando Structure from Motion (SfM), e modelos segmentados a partir de TCs pré-operatórias utilizando 3D U-Net. Eles utilizaram algoritmos de extração de características (SuperPoint e SuperGlue) para melhorar a reconstrução em ambientes com baixa textura e debris, e realizaram o registro dos modelos usando pontos de correspondência e ICP. O método demonstrou erros de registro submilimétricos e oferece um mapa 3D que auxilia na redução da carga mental do cirurgião durante a navegação intraoperatória.

Chongan Zhang *et al.* (2024) desenvolveram um algoritmo de planejamento de trajetória multiobjetivo para ureteroscópios flexíveis em cirurgia robótica, considerando critérios de segurança como o custo de impacto contra a parede do lúmen e a área de varredura do escopo.

Eles modelaram cinematicamente o comportamento do ureteroscópio em cenários com colisão (formas em C e S) e utilizaram um algoritmo de evolução diferencial (DE) combinado com uma rede neural bioinspirada (BINN) para otimizar a trajetória no espaço de configuração. Os resultados mostraram reduções significativas no impacto e na área de varredura em comparação com métodos baseados na linha central do ambiente, validados em modelos renais impressos em 3D.

Han, Kwon e Kyung (2025) introduziram um método híbrido de rastreamento de ureteroscópio que combina dados de movimento robótico com identificação de lúmen por visão computacional para navegação em cirurgia intrarenal robótica. O sistema utiliza segmentação de instância (Mask R-CNN) para detectar lúmens em imagens do ureteroscópio e os associa a pontos projetados da linha central do modelo 3D do rim, permitindo a tomada de decisão em bifurcações. Validado em um fantoma impresso em 3D, o método alcançou alta taxa de sucesso na localização (89,2% para cálices maiores) e reduziu significativamente a carga cognitiva e o tempo de procedimento em testes de usabilidade com usuários novatos.

Em relação aos trabalhos de planejamento automático de trajetória 3D, é importante destacar o ponto afirmado por Liang Yang *et al.* (2016) em relação às abordagens híbridas para completar essa tarefa, como foi feito nos trabalhos de Ciszkievicz e Milewski (2018) e Li, Dankelman e De Momi (2021). Os resultados obtidos mostraram ser possível obter um desempenho superior com um tempo computacional reduzido em comparação com os métodos utilizados individualmente. Isso demonstra que a investigação da combinação de outras técnicas ainda não exploradas na literatura, tanto de otimização quanto de busca, é interessante e pode contribuir com novas técnicas para gerar automaticamente trajetórias 3D em cirurgias. Um exemplo recente dessa tendência é o trabalho de Chongan Zhang *et al.* (2024), que combina um algoritmo de evolução diferencial com uma rede neural bioinspirada (BINN) para planejamento de trajetórias multiobjetivo em ureterosopia, considerando critérios de segurança como impacto e área de varredura, e alcançando ganhos significativos de desempenho.

Outro ponto importante é a forma de validação da trajetória gerada nesses trabalhos. Uma das formas é a comparação da distância com uma trajetória real gerada por cirurgiões, como foi feito nos trabalhos de Guo *et al.* (2017), Bao *et al.* (2020) e Ling He *et al.* (2023), ou pela avaliação da trajetória por cirurgiões especialistas, feito nos trabalhos de Wang *et al.* (2022) e Qi *et al.* (2022). Além dessas abordagens, métodos de validação baseados em navegação e rastreamento em tempo real, como o sistema híbrido proposto por Han, Kwon e Kyung (2025) que combina dados robóticos e identificação de lúmen para fornecer feedback visual da posição do escopo, surgem como uma alternativa promissora para validar e guiar a execução de trajetórias planejadas.

Ainda sobre planejamento automático de trajetória 3D, especificamente em cirurgias de ureterosopia, destaca-se o trabalho de Fan Yang *et al.* (2019) e a dissertação de Scarzanella (2020). Apesar da semelhança com este trabalho, existem possibilidades que não foram exploradas por eles, como: a consideração da posição do cálculo renal detectada de forma automática; validação com cirurgiões especialistas; e a integração com sistemas de navegação intraopera-

tória que utilizam o registro entre vídeo e modelos 3D pré-operatórios, como demonstrado por Acar *et al.* (2024), para criar um ciclo fechado entre planejamento, execução e verificação em tempo real.

5.3 MATERIAIS E MÉTODOS

A metodologia proposta para o planejamento automático de trajetórias tridimensionais em ureteroscopia flexível é composta por diversas etapas sequenciais, ilustradas no fluxograma da Figura 41. O método integra segmentação automática e manual, reconstrução tridimensional, planejamento de caminho com algoritmos de busca clássicos e técnicas de suavização, além de uma função de avaliação multiobjetivo para a seleção da trajetória ótima final.

5.3.1 Aquisição e Pré-processamento das Imagens

Foram utilizados exames de TC axial e coronal com contraste do sistema urinário. As imagens foram organizadas em volumes tridimensionais, utilizados como entrada para os processos de segmentação e reconstrução 3D.

5.3.2 Segmentação do Cálculo Ureteral e do Trato Urinário

A segmentação do cálculo ureteral foi realizada automaticamente utilizando um modelo *TransU-Net*, previamente treinado em outra base de dados de imagens médicas, no Capítulo 4. Esse modelo foi selecionado por ter segmentado os cálculos desse caso de teste com maior assertividade. Essa etapa permitiu localizar com precisão o cálculo no volume da TC utilizada.

Já a segmentação do trato urinário (ureter, pelve renal e bexiga) foi realizada manualmente, assegurando maior acurácia anatômica e eliminando possíveis falhas decorrentes de segmentação automática.

5.3.3 Reconstrução Tridimensional

As máscaras segmentadas foram utilizadas para reconstrução tridimensional do sistema urinário, empregando a biblioteca *PyVista*. Esse processo consistiu em projetar os volumes segmentados em superfícies 3D, permitindo a visualização e manipulação espacial dos órgãos e do cálculo ureteral.

5.3.4 Planejamento da Trajetória Inicial

O planejamento do caminho entre o ponto de entrada (bexiga) e o cálculo ureteral foi realizado utilizando o algoritmo A^* . A distância euclidiana foi utilizada como função de custo, assegurando a busca do menor trajeto entre os pontos definidos.

Além disso, foram incorporadas duas estratégias adicionais para garantir a viabilidade anatômica:

- a) **Verificação com algoritmo de Bresenham 3D:** responsável por assegurar que todos os voxels ao longo do caminho gerado permanecessem dentro do volume reconstruído (Seção 2.6.1).
- b) **Campo de repulsão:** implementado a partir de uma convolução gaussiana aplicada às paredes do trato urinário, criando uma penalização crescente para trajetórias próximas às superfícies anatômicas. Esse mecanismo mantém o caminho preferencialmente no centro do lúmen ureteral, reduzindo riscos de colisão (Seção 2.6.2).

No contexto deste trabalho, o Bresenham 3D é utilizado para verificar a viabilidade geométrica do caminho gerado pelo algoritmo A^* . A cada iteração, o algoritmo percorre a sequência de voxels correspondente ao segmento entre dois nós consecutivos e verifica se todos eles estão contidos dentro do volume segmentado do trato urinário. Caso algum voxel esteja fora da região anatômica permitida, o trajeto é penalizado ou descartado.

Essa verificação garante que a trajetória gerada permaneça confinada às estruturas internas do trato urinário, evitando colisões com as paredes ureterais e garantindo que o caminho resultante seja anatomicamente plausível.

A integração do campo de repulsão ao A^* proporciona um comportamento mais realista para o planejamento da trajetória, simulando o comportamento físico de um ureteroscópio flexível que naturalmente evita contato direto com as paredes do ureter. Essa estratégia resulta em trajetórias mais seguras e anatomicamente coerentes, especialmente quando combinada à validação voxel-a-voxel realizada pelo algoritmo de Bresenham 3D.

5.3.5 Suavização da Trajetória

A trajetória inicial obtida pelo algoritmo A^* frequentemente apresentava mudanças abruptas de direção devido à discretização volumétrica. Para tornar o caminho mais similar a movimentação de um ureteroscópio, três técnicas de suavização foram aplicadas e comparadas:

- a) Filtro **Savitzky-Golay** (Seção 2.8);
- b) Interpolação por curvas **B-Spline** (Seção 2.9);
- c) Filtro **Laplaciano** (Seção 2.10).

O filtro Savitzky-Golay é aplicado para suavizar a trajetória tridimensional gerada pelo algoritmo A^* , reduzindo oscilações abruptas entre os pontos e preservando a forma global do

trajeto. Essa característica é particularmente relevante na simulação da navegação do ureteroscópio, uma vez que o método mantém a continuidade das curvas e evita distorções excessivas no caminho.

A B-Spline também é utilizada para refinar a trajetória obtida pelo algoritmo A*, suavizando as transições angulares e reduzindo a curvatura máxima sem alterar o caminho anatômico principal. Essa abordagem permite simular de forma mais realista o comportamento de um ureteroscópio flexível, cuja movimentação exige continuidade de curvatura e limitação de torção. Além disso, a interpolação B-Spline possui a vantagem de ser computacionalmente ajustável, permitindo controlar o grau de suavização por meio da ordem da curva e do espaçamento dos nós.

O filtro Laplaciano atua reduzindo pequenas irregularidades na curva gerada, tornando a transição entre pontos mais contínua e diminuindo a torção. Embora apresente simplicidade e baixo custo computacional, o método tende a suavizar excessivamente em regiões de alta curvatura, podendo distorcer o trajeto anatômico se aplicado com alta intensidade.

5.3.6 Função de Avaliação Multiobjetivo

A escolha da trajetória final foi realizada por meio de uma função de avaliação que combina variáveis geométricas, anatômicas e de risco, resultando em uma medida única de qualidade da trajetória gerada. Os parâmetros considerados foram:

- a) **Extrapolações (E_x)**: número de pontos da curva suavizada que ficaram fora do volume reconstruído;
- b) **Erro quadrático médio (MSE)**: distância média entre a curva suavizada e o caminho original gerado pelo algoritmo A*;
- c) **Curvatura média e máxima (C_{mean}, C_{max})**;
- d) **Torção média e máxima (T_{mean}, T_{max})**;
- e) **Pontos de risco (R_p)**: número de pontos da trajetória a uma distância D da parede urinária mais próxima;
- f) **Erro final (E_f)**: distância entre o último ponto da trajetória e a posição do cálculo ureteral.

A função de avaliação multiobjetivo foi definida como:

$$F = 10E_x + 0.1MSE + C_{mean} + C_{max} + T_{mean} + T_{max} + 0.5R_p + 0.4E_f \quad (34)$$

onde valores menores de F indicam trajetórias mais suaves, seguras e anatomicamente plausíveis.

As métricas foram extraídas diretamente a partir da trajetória gerada e avaliadas em conjunto com um urologista, a fim de definir o grau de importância de cada variável. Com base

nessa análise, os pesos de cada termo da Equação 34 foram ajustados empiricamente, representando o impacto relativo de cada parâmetro na segurança e viabilidade clínica da trajetória. Em etapas futuras, pretende-se realizar uma otimização sistemática desses pesos utilizando métodos baseados em aprendizado de máquina ou análise estatística para aprimorar a calibração da função objetivo proposta.

5.4 EXPERIMENTOS E RESULTADOS

Para avaliar a eficácia do método proposto de planejamento automático de trajetória 3D para ureteroscopia, foram conduzidos experimentos utilizando imagens de TC, com reconstrução tridimensional do sistema urinário. O objetivo principal foi validar a correspondência anatômica e a suavidade das trajetórias geradas, considerando métricas geométricas e anatômicas relevantes.

5.4.1 Resultados Quantitativos

Cada uma das técnicas foi testada com três níveis de resolução: 100%, 75% e 50% dos pontos da trajetória original. A Tabela 12 apresenta os valores da função objetivo (*fitness*) para cada técnica e nível de resolução. Valores mais baixos indicam trajetórias mais suaves e compatíveis com a anatomia.

Tabela 12 – Pontuações da função objetivo para cada técnica em diferentes resoluções.

Técnica	100%	75%	50%
Savitzky-Golay	0.5957	0.6532	0.4684
B-Spline	0.6945	0.8000	0.4400
Laplacian	0.6041	0.6707	0.5564

Observa-se que, para todas as técnicas, o uso de apenas 50% dos pontos da trajetória original produziu os melhores resultados. Entre as três abordagens, o método *B-Spline* obteve o menor valor de *fitness* (0.440), indicando a trajetória mais suave e anatomicamente consistente.

A Tabela 13 resume as métricas geométricas e anatômicas para as melhores trajetórias de cada técnica.

Tabela 13 – Comparação quantitativa entre as melhores trajetórias.

Técnica	Fitness	MSE	Curv. Média	Curv. Máx.	Torção Média	Torção Máx.	Erro Final	Tempo (s)
Savitzky-Golay	0.468	0.158	0.039	0.184	0.021	0.191	0.046	12.46
B-Spline	0.440	0.132	0.037	0.195	0.005	0.145	0.112	12.45
Laplacian	0.556	0.095	0.065	0.168	0.114	0.199	0.000	12.46

Os resultados mostram que o método *B-Spline* apresentou o melhor equilíbrio entre suavidade, baixa torção e precisão no alinhamento final, tornando-se o mais promissor para o planejamento automático de trajetórias em ureteroscopia.

5.4.2 Resultados Visuais

A Figura 42 apresenta a comparação visual das trajetórias suavizadas pelas três técnicas. Apesar de todas evitarem contato com as paredes do ureter, a trajetória obtida com *B-Spline* mostrou-se mais contínua e com menor oscilação geométrica, seguida pelo filtro *Savitzky-Golay*. A trajetória suavizada com Laplacian apresentou maior irregularidade e variações angulares.

Além disso, a trajetória final otimizada com *B-Spline* foi sobreposta ao modelo 3D do trato urinário reconstruído, como ilustrado na Figura 43. Essa análise confirma que o caminho permanece inteiramente contido dentro do sistema urinário, respeitando os limites anatômicos e evitando colisões com as paredes urinárias.

5.5 DISCUSSÃO

Este estudo apresenta uma abordagem computacional para a geração de trajetórias cirúrgicas anatomicamente plausíveis no contexto de procedimentos invasivos de ureteroscopia. Ao integrar a segmentação automática de cálculos renais, a reconstrução tridimensional do trato urinário e a otimização da trajetória por meio de uma função de custo multiobjetivo, o método foi capaz de simular com sucesso um caminho completo de navegação endoscópica, partindo da bexiga até a localização do cálculo urinário alvo.

5.5.1 Análise dos resultados

Dentre as técnicas de suavização avaliadas, o filtro baseado em *B-Spline* apresentou os melhores resultados, combinando baixos valores de curvatura e torção com alta precisão de alinhamento e baixa taxa de erro em relação à trajetória original. Essas observações são confirmadas pelas métricas quantitativas (Tabela 13) e pela análise visual (Figuras 42 e 43), nas quais a trajetória gerada por *B-Spline* manteve a coerência anatômica e apresentou transições suaves de curvatura ao longo do sistema. Além disso, todos os métodos testados alcançaram tempos de execução similares, aproximadamente 12,4 segundos, confirmando a viabilidade computacional da estratégia para aplicações em tempo real ou próximo ao tempo real. A análise visual confirma a capacidade do método em respeitar a geometria dos órgãos e sugere potencial para aplicação clínica no planejamento pré-operatório.

5.5.2 Potenciais Clínicos

A adoção de um sistema automático para planejamento de trajetórias na ureteroscopia tem impacto direto sobre a segurança, a eficiência e a personalização do tratamento. Em termos clínicos, a possibilidade de gerar trajetórias que respeitam as curvaturas fisiológicas do trato urinário reduz o risco de perfuração, trauma ureteral e complicações pós-operatórias. Além

disso, a abordagem proposta pode oferecer suporte pré-operatório a urologistas, permitindo uma visualização detalhada do caminho ótimo antes da execução do procedimento.

Outro aspecto relevante é o potencial para aplicação em plataformas de treinamento médico e simulação. Ao gerar trajetórias personalizadas com base em imagens de TC específicas do paciente, o método possibilita a criação de ambientes realistas para treinamento de residentes e especialistas, reproduzindo cenários de navegação endoscópica de forma segura, sem riscos ao paciente.

A integração dessa metodologia a sistemas de cirurgia assistida por robôs representa outro caminho promissor. Combinando os resultados obtidos com controles cinemáticos de robôs cirúrgicos, seria possível automatizar parcialmente a navegação do ureteroscópio, otimizando o acesso a regiões complexas do trato urinário e reduzindo o tempo total de procedimento.

5.5.3 Limitações

Apesar dos resultados promissores, algumas limitações devem ser reconhecidas. Primeiramente, a análise foi realizada com base em um único caso clínico, o que limita a generalização dos resultados. Em segundo lugar, a segmentação do sistema urinário foi realizada manualmente para garantir a acurácia, mas a integração de ferramentas automáticas e robustas de segmentação será necessária para viabilizar a aplicação clínica em larga escala. Embora o método evite colisões com os limites anatômicos, ele ainda não considera as propriedades biomecânicas dos tecidos, como elasticidade ou resistência local. Por último, a metodologia ainda não foi validada em um ambiente cirúrgico real, o que exige avaliações clínicas de urologistas e testes em múltiplos pacientes.

5.5.4 Trabalhos Futuros

Trabalhos futuros devem buscar validar a abordagem em uma amostra maior de pacientes, incorporando análises clínicas de urologistas. A função de otimização da trajetória também poderá ser refinada com a integração de restrições cinemáticas de ureteroscópios flexíveis, incluindo curvatura máxima permitida, angulação da ponta e orientação do canal de trabalho. Além disso, a adaptação do método para outros procedimentos endoscópicos, como a nefrolitotomia percutânea, pode ampliar significativamente sua aplicabilidade no campo da endourologia.

Em resumo, o método apresentado representa um avanço promissor no planejamento cirúrgico personalizado com o uso de métodos computacionais. Ao combinar otimização geométrica e informações anatômicas, este estudo de prova de conceito abre caminho para uma navegação endoscópica mais segura, eficiente e adaptada às características individuais dos pacientes.

Figura 41 – Fluxograma da metodologia proposta para o planejamento automático da trajetória 3D em ureteroscopia.

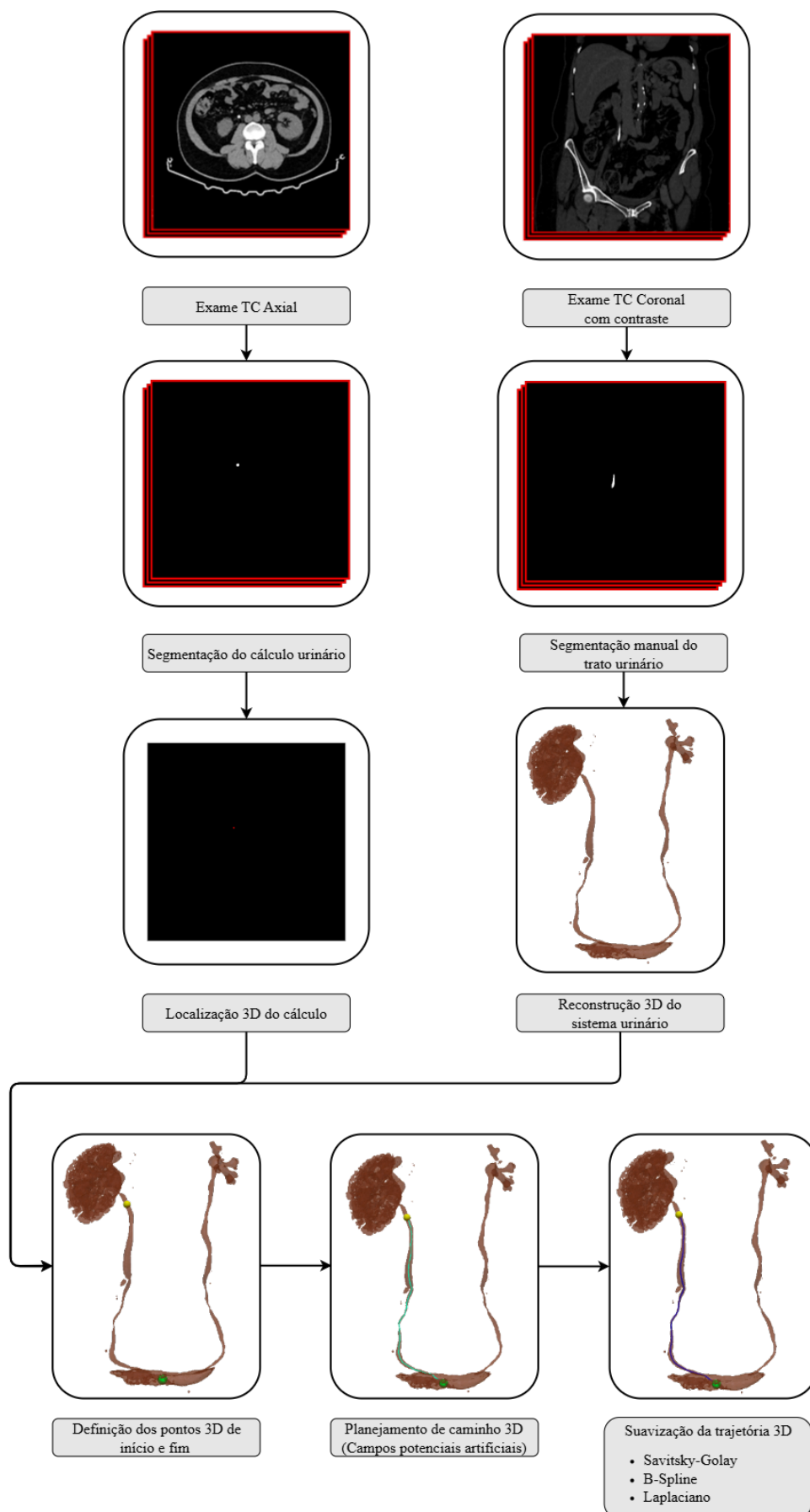
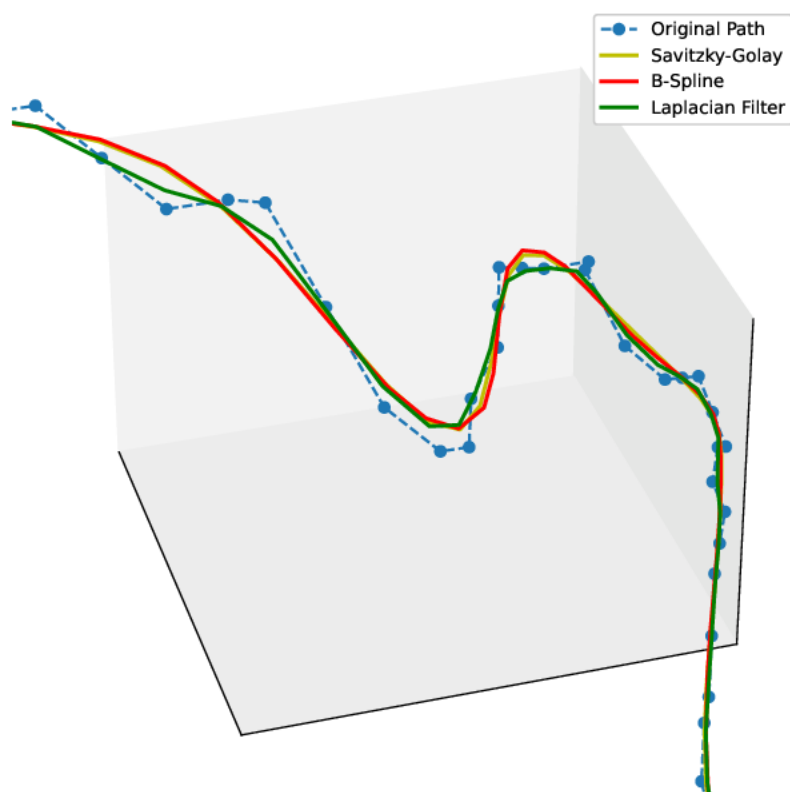


Figura 42 – Visualização 3D das trajetórias suavizadas: comparação entre Savitzky-Golay, B-Spline e Laplacian.



Fonte: Autor.

Figura 43 – Trajetória otimizada com *B-Spline* (azul) sobreposta à reconstrução 3D do trato urinário.



Fonte: Autor.

6 CONCLUSÃO

Esta tese apresentou uma nova abordagem para a automação das etapas de diagnóstico e planejamento pré-operatório em cirurgias de ureteroscopia assistidas por IA. O trabalho é composto por três etapas principais: (i) classificação automática de imagens de TC para detecção de cálculos urinários, (ii) segmentação automática de cálculos e estruturas do trato urinário, e (iii) planejamento semi-automático da trajetória tridimensional do ureteroscópio até o cálculo. Embora os módulos tenham sido implementados e avaliados de forma parcialmente independente, os resultados obtidos demonstram a viabilidade de uma futura integração em um sistema unificado de suporte ao planejamento cirúrgico.

Na primeira etapa, foi proposto o modelo *DE-MLP*, uma estratégia de fusão tardia baseada em *ensemble learning* entre diferentes bases de dados públicas. O modelo atingiu acurácia média superior a 99% na detecção de cálculos urinários em imagens coronais de TC, mostrando desempenho superior a abordagens tradicionais, com baixo custo computacional e boa capacidade de generalização. Essa etapa consolidou o potencial da IA em fornecer diagnósticos automáticos, capazes de identificar casos clínicos a partir de exames radiológicos.

Na segunda etapa, foi introduzido o conjunto de dados *KSSD2025*, composto por 838 imagens axiais de TC com anotações manuais de cálculos urinários, e avaliado o desempenho de quatro arquiteturas de segmentação profunda: U-Net, U-Net++, U-Net3+ e TransU-Net. Todas as arquiteturas apresentaram desempenho elevado (coeficiente Dice médio acima de 95%), com destaque para a U-Net pela sua assertividade. Essa etapa teve como principal contribuição a criação de um recurso público relevante para pesquisas futuras em segmentação e análise de cálculos urinários em tomografias.

Por fim, a terceira etapa propôs um método de planejamento semi-automático de trajetórias endoscópicas 3D entre a bexiga e o cálculo ureteral, a partir de reconstruções volumétricas derivadas de exames de TC. O método combina o algoritmo de busca A* modificado e filtros de suavização geométrica. Entre os métodos avaliados, a interpolação por curvas B-Spline apresentou o melhor desempenho em termos de suavidade, conformidade anatômica e precisão espacial, demonstrando a viabilidade de simular trajetórias anatomicamente plausíveis de ureteroscopia.

Apesar de as três partes terem sido apresentadas de forma parcialmente integradas, a arquitetura modular da tese foi concebida para permitir a futura integração completa entre os módulos de diagnóstico e planejamento. Em um cenário de integração total, os resultados de classificação alimentariam automaticamente o módulo de segmentação, que, por sua vez, serviria como entrada direta para o módulo de planejamento de trajetória, resultando em um sistema completo de planejamento cirúrgico assistido por IA. Essa integração permitiria automatizar todo o fluxo, desde a análise inicial da TC até a geração da trajetória otimizada, possibilitando aplicações em plataformas de suporte à decisão clínica, sistemas de navegação cirúrgica e simuladores de treinamento em realidade virtual.

Como perspectivas futuras, pretende-se aprimorar a generalização dos modelos por meio de bases de dados coletadas de fontes distintas, explorar a fusão de diferentes perspectivas tomográficas (axial, coronal e sagital), expandir o conjunto de dados utilizado e incorporar restrições biomecânicas e cinemáticas dos ureteroscópios reais no planejamento da trajetória. Além disso, a integração completa dos módulos será avaliada em ambientes simulados e clínicos, com o objetivo de validar o desempenho do sistema de forma quantitativa e qualitativa, em colaboração com especialistas em urologia.

REFERÊNCIAS

- ABID, Rayyan; HUSSEIN, Ahmed A; GURU, Khurshid A. Artificial intelligence in urology: current status and future perspectives. **Urologic Clinics**, Elsevier, v. 51, n. 1, p. 117–130, 2024.
- ACAR, Ayberk *et al.* Towards navigation in endoscopic kidney surgery based on preoperative imaging. **Healthcare Technology Letters**, Wiley Online Library, v. 11, n. 2-3, p. 67–75, 2024.
- ADAMS, Rolf; BISCHOF, Leanne. Seeded region growing. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, v. 16, n. 6, p. 641–647, 1994.
- AKSHAYA, M *et al.* Kidney Stone Detection Using Neural Networks. *In: IEEE. 2020 International Conference on System, Computation, Automation and Networking (ICSCAN). [S. l.: s. n.], 2020. p. 1–4.*
- ALELIGN, Tilahun; PETROS, Beyene. Kidney stone disease: an update on current concepts. **Advances in urology**, Wiley Online Library, v. 2018, n. 1, p. 3068365, 2018.
- ANDRABI, Yasir *et al.* Advances in CT imaging for urolithiasis. **Indian journal of urology: IJU: journal of the Urological Society of India**, Wolters Kluwer–Medknow Publications, v. 31, n. 3, p. 185, 2015.
- ANTON, Howard; BIVENS, Irl; DAVIS, Stephen. **Cálculo-Volume II-10**. [S. l.]: Bookman Editora, 2014.
- ANTONY, J Rabin Singh *et al.* Refine Observation of Kidney Stones using Neural Network, 2020.
- ASIF, Sohaib; ZHENG, Xiaolong; ZHU, Yusen. An optimized fusion of deep learning models for kidney stone detection from CT images. **Computer and Information Sciences**, King Saud University, v. 36, p. 102130, 2024.
- ASIF, Sohaib *et al.* StoneNet: An Efficient Lightweight Model Based on Depthwise Separable Convolutions for Kidney Stone Detection from CT Images. **Interdisciplinary Sciences: Computational Life Sciences**, Springer, v. 15, p. 633–652, 2023. Disponível em: <https://doi.org/10.1007/s12539-023-00578-8>.
- AZAD, Reza *et al.* Medical Image Segmentation Review: The Success of U-Net. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 46, p. 10076–10095, 2022. Disponível em: <https://api.semanticscholar.org/CorpusID:254043694>.
- BAO, Nan *et al.* Surgery Path Planning for Lung Biopsy Based on Pareto Optimization. **2020 5th International Conference on Technologies in Manufacturing, Information and Computing (ICTMIC 2020)**, 2020.

BAYGIN, Mehmet *et al.* Exemplar Darknet19 feature generation technique for automated kidney stone detection with coronal CT images. **Artificial Intelligence in Medicine**, Elsevier, v. 127, p. 102274, 2022.

BEVERINI, Martina *et al.* A review on training models in urology. **AME Medical Journal**, v. 7, n. 0, 2021. Disponível em: <https://amj.amegroups.com/article/view/6437>.

BISHOP, Christopher M. **Pattern recognition and machine learning**. [S. l.]: springer, 2006.

BOUZON, Murillo *et al.* Automatic Kidney Stone Detection using Low-cost CNN with Coronal CT Images. In: ANAIS do XVIII Workshop de Visão Computacional. São Bernardo do Campo/SP: SBC, 2023. p. 24–29. Disponível em: <https://sol.sbc.org.br/index.php/wvc/article/view/27527>.

BOUZON, Murillo F. *et al.* KSSD2025: A New Annotated Dataset for Automatic Kidney Stone Segmentation and Evaluation With Modified U-Net Based Deep Learning Models. **IEEE Access**, v. 13, p. 171790–171806, 2025.

BREIMAN, Leo. Random forests. **Machine learning**, Springer, v. 45, p. 5–32, 2001.

BRESENHAM, Jack. Algorithm for Computer Control of a Digital Plotter. **IBM Systems Journal**, v. 4, p. 25–30, fev. 1965.

CAGLAYAN, Alper *et al.* Deep learning model-assisted detection of kidney stones on computed tomography. **International Brazilian Journal of Urology**, 2022.

CAI, Dongyang *et al.* Automatic Path Planning for Navigated Pedicle Screw Surgery Based on Deep Neural Network. In: IEEE. 2019 WRC Symposium on Advanced Robotics and Automation (WRC SARA). [S. l.: s. n.], 2019. p. 62–67.

CAUBERG, Evelyne CC *et al.* A new generation of optical diagnostics for bladder cancer: technology, diagnostic accuracy, and future applications. **European urology**, Elsevier, v. 56, n. 2, p. 287–297, 2009.

CHAKI, Jyotismita; UÇAR, Ayşegül. An Efficient and Robust Approach Using Inductive Transfer-Based Ensemble Deep Neural Networks for Kidney Stone Detection. **IEEE Access**, IEEE, v. 12, p. 32894–32910, 2024.

CHAN, Heang-Ping; HADJIISKI, Lubomir M.; SAMALA, Ravi K. Computer-aided diagnosis in the era of deep learning. **Medical physics**, Wiley Online Library, v. 47, n. 5, e218–e227, 2020.

CHEN, Jieneng *et al.* Transunet: Transformers make strong encoders for medical image segmentation. **arXiv preprint arXiv:2102.04306**, 2021.

CHEN, Ruiying. Segmentation of subcutaneous fat CT images based on the TransUNet model. **Applied and Computational Engineering**, v. 74, p. 239–244, 2024.

CHOI, Jiho *et al.* PFSH-Net: Parallel frequency-spatial hybrid network for segmentation of kidney stones in pre-contrast computed tomography images of dogs. **Computers in Biology and Medicine**, Elsevier, v. 186, p. 109609, 2025.

ÇIÇEK, Özgün *et al.* 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: SPRINGER. MEDICAL Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19. [S. l.: s. n.], 2016. p. 424–432.

CISZKIEWICZ, Adam; MILEWSKI, Grzegorz. Path planning for minimally-invasive knee surgery using a hybrid optimization procedure. **Computer methods in biomechanics and biomedical engineering**, Taylor & Francis, v. 21, n. 1, p. 47–54, 2018.

CORTES, Corinna; VAPNIK, Vladimir. Support-vector networks. **Machine learning**, Springer, v. 20, p. 273–297, 1995.

CRAVEN, Bruce D; ISLAM, Sardar MN. **Optimization in economics and finance: some advances in non-linear, dynamic, multi-criteria and stochastic models**. [S. l.]: Springer Science & Business Media, 2005. v. 7.

CUI, Yingpu *et al.* Automatic Detection and Scoring of Kidney Stones on Noncontrast CT Images Using S.T.O.N.E. Nephrolithometry: Combined Deep Learning and Thresholding Methods. **Molecular Imaging and Biology**, v. 23, p. 436–445, 2020. Disponível em: <https://api.semanticscholar.org/CorpusID:225080761>.

DE BOOR, Carl; DE BOOR, Carl. **A practical guide to splines**. [S. l.]: springer New York, 1978. v. 27.

DICE, Lee R. Measures of the amount of ecologic association between species. **Ecology**, JSTOR, v. 26, n. 3, p. 297–302, 1945.

ELTON, Daniel C. *et al.* A deep learning system for automated kidney stone detection and volumetric segmentation on non-contrast CT scans. **Medical physics**, 2022. Disponível em: <https://api.semanticscholar.org/CorpusID:246813463>.

GALLAGHER, Anthony G; TRAYNOR, Oscar. Simulation in surgery: opportunity or threat? **Irish journal of medical science**, Springer, v. 177, p. 283–287, 2008.

GLAZER, Kelly; BREA, Isabel J; VAITLA, Pradeep. Ureterolithiasis. In: STATPEARLS [Internet]. [S. l.]: StatPearls Publishing, 2022.

GONZALES, Rafael C; WINTZ, Paul. **Digital image processing**. [S. l.]: Addison-Wesley Longman Publishing Co., Inc., 1987.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep learning**. [S. l.]: MIT press, 2016.

GUO, Jian *et al.* A novel path planning algorithm for the vascular interventional surgical robotic doctor training system. *In: IEEE. 2017 IEEE International Conference on Mechatronics and Automation (ICMA)*. [S. l.: s. n.], 2017. p. 45–50.

GUPTA, Soumya *et al.* Mi-unet: Improved segmentation in ureteroscopy. *In: IEEE. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. [S. l.: s. n.], 2020. p. 212–216.

HAMA, Hawkar K.; MAJEED, Hamsa D.; NARIMAN, Goran Saman. Enhanced Kidney Stone Detection and Classification Using SVM and LBP Features. **UHD Journal of Science and Technology**, v. 9, n. 1, p. 10–17, 2025.

HAMEED, BM Zeeshan *et al.* Application of Virtual Reality, Augmented Reality, and Mixed Reality in Endourology and Urolithiasis: An Update by YAU Endourology and Urolithiasis Working Group. **Frontiers in Surgery**, Frontiers Media SA, v. 9, 2022.

HAMEED, BM Zeeshan *et al.* Artificial intelligence and its impact on urological diseases and management: A comprehensive review of the literature. **Journal of Clinical Medicine**, MDPI, v. 10, n. 9, p. 1864, 2021.

HAN, Jinpei *et al.* A systematic review of robotic surgery: From supervised paradigms to fully autonomous robotic approaches. **The International Journal of Medical Robotics and Computer Assisted Surgery**, Wiley Online Library, v. 18, n. 2, e2358, 2022.

HAN, Jung-Min; KWON, Dong-Soo; KYUNG, Ki-Uk. A Novel Hybrid Ureteroscope Tracking for Robotic-assisted Retrograde Intrarenal Surgery via Recognition of Pathway with Lumen Identification. **IEEE Robotics and Automation Letters**, IEEE, 2025.

HART, Peter E; NILSSON, Nils J; RAPHAEL, Bertram. A formal basis for the heuristic determination of minimum cost paths. **IEEE transactions on Systems Science and Cybernetics**, IEEE, v. 4, n. 2, p. 100–107, 1968.

HE, Ling *et al.* Preoperative path planning algorithm for lung puncture biopsy based on path constraint and multidimensional space distance optimization. **Biomedical Signal Processing and Control**, Elsevier, v. 80, p. 104304, 2023.

HE, Yucheng *et al.* Endoscopic path planning in robot-assisted endoscopic nasal surgery. **IEEE Access**, IEEE, v. 8, p. 17039–17048, 2020.

HESAMIAN, Mohammad Hesam *et al.* Deep learning techniques for medical image segmentation: achievements and challenges. **Journal of digital imaging**, Springer, v. 32, p. 582–596, 2019.

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

HOWARD, Andrew G *et al.* Mobilenets: Efficient convolutional neural networks for mobile vision applications. **arXiv preprint arXiv:1704.04861**, 2017.

HU, Wenrui; JIANG, Huiyan; WANG, Meng. Flexible needle puncture path planning for liver tumors based on deep reinforcement learning. **Physics in Medicine & Biology**, IOP Publishing, v. 67, n. 19, p. 195008, 2022.

HUAN DING QIRUI HUANG, Navid Razmjooy. An improved version of firebug swarm optimization algorithm for optimizing Alex/ELM network kidney stone detection. **Biomedical Signal Processing and Control**, v. 99, p. 106898, 2025.

HUANG, Huimin *et al.* UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation. **ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**, p. 1055–1059, 2020. Disponível em: <https://api.semanticscholar.org/CorpusID:215828394>.

HUANG, Jianping *et al.* Automatic kidney Stone Composition Analysis Method based on dual-energy CT. **Current Medical Imaging**, Bentham Science Publishers, v. 20, n. 1, e080923220827, 2024.

HUANG, Zih-Hao *et al.* Design and validation of a deep learning model for renal stone detection and segmentation on kidney–ureter–bladder images. **Bioengineering**, MDPI, v. 10, n. 8, p. 970, 2023.

INGIMARSSON, Johann P; KRAMBECK, Amy E; PAIS, Vernon M. Diagnosis and management of nephrolithiasis. **Surgical Clinics**, Elsevier, v. 96, n. 3, p. 517–532, 2016.

ISLAM, Nazmul *et al.* Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from CT-radiography. **scientific reports**, 2022.

JACCARD, Paul. The distribution of the flora in the alpine zone. 1. **New phytologist**, Wiley Online Library, v. 11, n. 2, p. 37–50, 1912.

JADHAV, Amolkumar N *et al.* AI-Powered Early Detection of Kidney Stones Using a Hybrid CNN-LSTM Model. **Journal of Neonatal Surgery**, v. 14, 9s, 2025.

JAIN, Shraman *et al.* Detection of Kidney Stone and Estimation of its Size using Image Segmentation Techniques. **Journal of Advanced Research in Applied Sciences and Engineering Technology**, 2023. Disponível em: <https://api.semanticscholar.org/CorpusID:266760699>.

JANG, Dong-Hyun *et al.* Kidney, ureter, and urinary bladder segmentation based on non-contrast enhanced computed tomography images using modified U-Net. **Scientific Reports**, Nature Publishing Group UK London, v. 14, n. 1, p. 15325, 2024.

JHA, Vivekanand *et al.* Chronic kidney disease: global dimension and perspectives. **The Lancet**, Elsevier, v. 382, n. 9888, p. 260–272, 2013.

JI, Zhanlin *et al.* BGRD-TransUNet: A Novel TransUNet-Based Model for Ultrasound Breast Lesion Segmentation. **IEEE Access**, v. 12, p. 31182–31196, 2024. Disponível em: <https://api.semanticscholar.org/CorpusID:267785661>.

KAMBADAKONE, Avinash R *et al.* New and evolving concepts in the imaging and management of urolithiasis: urologists' perspective. **Radiographics**, Radiological Society of North America, v. 30, n. 3, p. 603–623, 2010.

KANAVATI, Fahdi *et al.* Automatic L3 slice detection in 3D CT images using fully-convolutional networks. **ArXiv**, abs/1811.09244, 2018.

KANAVATI, Fahdi *et al.* Automatic L3 slice detection in 3D CT images using fully-convolutional networks. **ArXiv**, abs/1811.09244, 2018.

KAVIANI, Parisa *et al.* Performance of threshold-based stone segmentation and radiomics for determining the composition of kidney stones from single-energy CT. **Japanese Journal of Radiology**, Springer, v. 41, n. 2, p. 194–200, 2023.

KHATIB, Oussama. Real-time obstacle avoidance for manipulators and mobile robots. **The international journal of robotics research**, Sage Publications Sage CA: Thousand Oaks, CA, v. 5, n. 1, p. 90–98, 1986.

KIM, Ui Seok *et al.* Prediction of the composition of urinary stones using deep learning. **Investigative and Clinical Urology**, Korean Urological Association, v. 63, n. 4, p. 441, 2022.

KIM, Young In *et al.* Deep-learning segmentation of urinary stones in noncontrast computed tomography. **Journal of Endourology**, Mary Ann Liebert, Inc., publishers 140 Huguenot Street, 3rd Floor New ... , v. 37, n. 5, p. 595–606, 2023.

KOHAVI, Ron *et al.* A study of cross-validation and bootstrap for accuracy estimation and model selection. In: MONTREAL, CANADA, 2. IJCAI. [S. l.: s. n.], 1995. v. 14, p. 1137–1145.

KOLLI, Chandra Sekhar *et al.* Efficient Development of Supervised Learning Algorithm for Kidney Stone Prediction. **International Conference on Inventive Computation Technologies**, IEEE, p. 1373–1379, 2022.

- KOMORI, Masatsugu *et al.* Complications of flexible ureteroscopic treatment for renal and ureteral calculi during the learning curve. **Urologia Internationalis**, Karger Publishers, v. 95, n. 1, p. 26–32, 2015.
- KUMAR, Davinder *et al.* Enhanced Detection of Kidney Stones using Level Set Segmentation and Kernel-ELM. *In: IEEE. 2024 International Conference on Data Science and Network Security (ICDSNS)*. [S. l.: s. n.], 2024. p. 1–6.
- LECUN, Yann *et al.* Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, Ieee, v. 86, n. 11, p. 2278–2324, 1998.
- LEE, Chin Yang. An algorithm for path connections and its applications. **IRE transactions on electronic computers**, IEEE, n. 3, p. 346–365, 1961.
- LI, Dan *et al.* Deep Segmentation Networks for Segmenting Kidneys and Detecting Kidney Stones in Unenhanced Abdominal CT Images. **Diagnostics**, v. 12, 2022. Disponível em: <https://api.semanticscholar.org/CorpusID:251088259>.
- LI, Guangju *et al.* MultiIB-TransUNet: Transformer with multiple information bottleneck blocks for CT and ultrasound image segmentation. **Medical Physics**, Wiley Online Library, v. 51, n. 2, p. 1178–1189, 2024.
- LI, Zhen; DANKELMAN, Jenny; DE MOMI, Elena. Path planning for endovascular catheterization under curvature constraints via two-phase searching approach. **International Journal of Computer Assisted Radiology and Surgery**, Springer, v. 16, n. 4, p. 619–627, 2021.
- LIDÉN, Mats. A new method for predicting uric acid composition in urinary stones using routine single-energy CT. **Urolithiasis**, Springer, v. 46, n. 4, p. 325–332, 2018.
- LIU, Haozhi; GHADIMI, Noradin. Hybrid convolutional neural network and Flexible Dwarf Mongoose Optimization Algorithm for strong kidney stone diagnosis. **Biomedical Signal Processing and Control**, Elsevier, v. 91, p. 106024, 2024.
- LV, Chunjie *et al.* ECM-TransUNet: Edge-enhanced multi-scale attention and convolutional Mamba for medical image segmentation. **Biomedical Signal Processing and Control**, Elsevier, v. 107, p. 107845, 2025.
- MA, Jun *et al.* Abdomenct-1k: Is abdominal organ segmentation a solved problem? **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE, v. 44, n. 10, p. 6695–6714, 2021.
- MA, Jun *et al.* Segment anything in medical images. **Nature Communications**, Nature Publishing Group UK London, v. 15, n. 1, p. 654, 2024.

MAGER, R *et al.* Clinical outcomes and costs of reusable and single-use flexible ureterorenoscopes: a prospective cohort study. **Urolithiasis**, Springer, v. 46, p. 587–593, 2018.

MALLAT, Stephane G. A theory for multiresolution signal decomposition: the wavelet representation. **IEEE transactions on pattern analysis and machine intelligence**, Ieee, v. 11, n. 7, p. 674–693, 1989.

MANDAL, Swarnendu *et al.* Clavien classification of semirigid ureteroscopy complications: a prospective study. **Urology**, Elsevier, v. 80, n. 5, p. 995–1001, 2012.

MANUDHAS, S. M.; MANORANJITHAM, R. Automatic Kidney Stone Detection using XResNet 152 model for CT images. *In: IEEE. 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE)*. [S. l.: s. n.], 2024. p. 1–6.

MAZZUCCHI, Eduardo *et al.* Single-use flexible ureteroscopes: update and perspective in developing countries. A narrative review. **International braz j urol**, SciELO Brasil, v. 48, p. 456–467, 2022.

MCCULLOCH, Warren S; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, n. 4, p. 115–133, 1943.

MINSKY, Marvin; PAPERT, Seymour. An introduction to computational geometry. **Cambridge tiass., HIT**, 1969.

MOHAMED, Ihab S. **Detection and tracking of pallets using a laser rangefinder and machine learning techniques**. 2017. Tese (Doutorado) – European Master on Advanced Robotics+(EMARO+), University of Genova, Italy.

N, Nisha *et al.* Kidney Stone Detection using CNN Algorithm. **International Journal of Scientific Research in Computer Science, Engineering and Information Technology**, v. 10, p. 1814–1823, dez. 2024.

EL-NAHAS, Ahmed R *et al.* Semirigid ureteroscopy for ureteral stones: a multivariate analysis of unfavorable results. **The Journal of urology**, Wolters Kluwer Philadelphia, PA, v. 181, n. 3, p. 1158–1162, 2009.

NAIR, Vinod; HINTON, Geoffrey E. Rectified linear units improve restricted boltzmann machines. *In: ICML*. [S. l.: s. n.], 2010.

NEHA, Fnu. Kidney localization and stone segmentation from a ct scan image. *In: IEEE. 2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*. [S. l.: s. n.], 2023. p. 1–6.

OJALA, Timo; PIETIKÄINEN, Matti; HARWOOD, David. A comparative study of texture measures with classification based on featured distributions. **Pattern recognition**, Elsevier, v. 29, n. 1, p. 51–59, 1996.

ORDON, Michael *et al.* The surgical management of kidney stone disease: a population based time series analysis. **The Journal of urology**, Wolters Kluwer Philadelphia, PA, v. 192, n. 5, p. 1450–1456, 2014.

OTSU, Nobuyuki. A threshold selection method from gray-level histograms. **IEEE transactions on systems, man, and cybernetics**, IEEE, v. 9, n. 1, p. 62–66, 1979.

PAFFEN, Marc LJE *et al.* A comparison of the physical properties of four new generation flexible ureteroscopes:(de) flection, flow properties, torsion stiffness, and optical characteristics. **Journal of endourology**, Mary Ann Liebert, Inc. 2 Madison Avenue Larchmont, NY 10538 USA, v. 22, n. 10, p. 2227–2234, 2008.

PANTHIER, Frédéric *et al.* Can Artificial Intelligence Accurately Detect Urinary Stones? A Systematic Review. **Journal of Endourology**, Mary Ann Liebert, Inc., publishers 140 Huguenot Street, 3rd Floor New ... , v. 38, n. 8, p. 725–740, 2024.

PATRO, Kiran Kumar *et al.* Application of Kronecker convolutions in deep learning technique for automated detection of kidney stones with coronal CT images. **Information Sciences**, Elsevier, v. 640, p. 119005, 2023.

PENG, Yao; SONKA, Milan; CHEN, Danny Ziyi. U-Net v2: Rethinking the Skip Connections of U-Net for Medical Image Segmentation. **ArXiv**, abs/2311.17791, 2023. Disponível em: <https://api.semanticscholar.org/CorpusID:265498640>.

PREEDANAN, Wongsakorn *et al.* Improvement of urinary stone segmentation using GAN-based urinary stones inpainting augmentation. **IEEE Access**, IEEE, v. 10, p. 115131–115142, 2022.

PREEDANAN, Wongsakorn *et al.* Urinary Stones Segmentation in Abdominal X-Ray Images Using Cascaded U-Net Pipeline With Stone-Embedding Augmentation and Lesion-Size Reweighting Approach. **IEEE Access**, v. 11, p. 25702–25712, 2023.

PUNMIA, B.C. **Surveying: V. 2**. [S. l.]: Laxmi Publications, 2005. Disponível em: <https://books.google.com.br/books?id=oT5EeEDrK04C>.

QI, Xiaozhi *et al.* An Automatic Path Planning Method of Pedicle Screw Placement Based on Preoperative CT Images. **IEEE Transactions on Medical Robotics and Bionics**, IEEE, v. 4, n. 2, p. 403–413, 2022.

QIN, Hongwei *et al.* Review of autonomous path planning algorithms for mobile robots. **Drones**, MDPI, v. 7, n. 3, p. 211, 2023.

RABBY, Salman *et al.* An automated approach for the kidney segmentation and detection of kidney stones on computed tomography using YOLO algorithms. **Journal of Ideas in Health**, v. 6, n. 4, p. 963–970, 2023.

RASULOVA, N.; AMINOVA, A.; ISMAILOVA, F. IMPROVEMENT OF EARLY DIAGNOSIS AND PREVENTION MEASURES OF KIDNEY STONE DISEASES AMONG THE POPULATION. **Science and innovation**, ООО «Science e innovation», v. 2, n. D3, p. 61–66, 2023.

REICHERZ, Alina *et al.* An empirical study on hospital-based prevention of recurrent urinary stone disease in Germany. **World Journal of Urology**, Springer, v. 40, n. 1, p. 237–242, 2022.

ROCHA, Rafael L *et al.* Image inspection of railcar structural components: An approach through deep learning and discrete Fourier transform. *In: SBC. ANAIS do VII Symposium on Knowledge Discovery, Mining and Learning. [S. l.: s. n.]*, 2019. p. 33–40.

RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-net: Convolutional networks for biomedical image segmentation. *In: SPRINGER. MEDICAL Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18. *[S. l.: s. n.]*, 2015. p. 234–241.

ROSENBLATT, Frank. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, American Psychological Association, v. 65, n. 6, p. 386, 1958.

RUMELHART, David E; HINTON, Geoffrey E; WILLIAMS, Ronald J. Learning representations by back-propagating errors. **nature**, Nature Publishing Group, v. 323, n. 6088, p. 533–536, 1986.

RUSSELL, Stuart J. **Artificial intelligence a modern approach**. *[S. l.]*: Pearson Education, Inc., 2010.

SAVITZKY, Abraham; GOLAY, Marcel JE. Smoothing and differentiation of data by simplified least squares procedures. **Analytical chemistry**, ACS Publications, v. 36, n. 8, p. 1627–1639, 1964.

SAW, K. Chee *et al.* Helical CT of Urinary Calculi. **American Journal of Roentgenology**, v. 175, n. 2, p. 329–332, 2000.

SCARZANELLA, Linda. Design of a ureteroscope prototype and of its path panning strategy. Italy, 2020.

SCHOENBERG, Isaac J. Contributions to the problem of approximation of equidistant data by analytic functions. Part A. On the problem of smoothing or graduation. A first class of analytic approximation formulae. **Quarterly of applied mathematics**, v. 4, n. 1, p. 45–99, 1946.

SCHUSTER, Timothy G *et al.* Complications of ureteroscopy: analysis of predictive factors. **The Journal of urology**, Wolters Kluwer Philadelphia, PA, v. 166, n. 2, p. 538–540, 2001.

SELVARAJU, Ramprasaath R. *et al.* Grad-CAM: Why did you say that? Visual Explanations from Deep Networks via Gradient-based Localization. **CoRR**, abs/1610.02391, 2016. arXiv: 1610.02391. Disponível em: <http://arxiv.org/abs/1610.02391>.

SEO, Hyunseok *et al.* Modified U-Net (mU-Net) with incorporation of object-dependent high level features for improved liver and liver-tumor segmentation in CT images. **IEEE transactions on medical imaging**, IEEE, v. 39, n. 5, p. 1316–1325, 2019.

SHAPHIRO, S; WILK, MJB. An analysis of variance test for normality. **Biometrika**, v. 52, n. 3, p. 591–611, 1965.

SHOKEIR, Ahmed A. Renal colic: new concepts related to pathophysiology, diagnosis and treatment. **Current opinion in urology**, LWW, v. 12, n. 4, p. 263–269, 2002.

SINGER, Saša; NELDER, John. Nelder-mead algorithm. **Scholarpedia**, v. 4, n. 7, p. 2928, 2009.

SONG, Fangfang *et al.* Research on U-Net medical image segmentation. *In: OTHER Conferences*. [S. l.: s. n.], 2022. Disponível em: <https://api.semanticscholar.org/CorpusID:247565798>.

SONI, Akanksha; RAI, Avinash. Kidney Stone Recognition and Extraction using Directional Emboss & SVM from Computed Tomography Images. *In: IEEE. 2020 Third International Conference on Multimedia Processing, Communication & Information Technology (MPCIT)*. [S. l.: s. n.], 2020. p. 57–62.

SOROKIN, Igor *et al.* Epidemiology of stone disease across the world. **World journal of urology**, Springer, v. 35, p. 1301–1320, 2017.

STEWART, Gregory *et al.* Stone size limits the use of Hounsfield units for prediction of calcium oxalate stone composition. **Urology**, Elsevier, v. 85, n. 2, p. 292–295, 2015.

SUAREZ-IBARROLA, Rodrigo *et al.* Current and future applications of machine and deep learning in urology: a review of the literature on urolithiasis, renal cell carcinoma, and bladder and prostate cancer. **World journal of urology**, Springer, v. 38, p. 2329–2347, 2020.

SULAKSONO, Nanang; ADI, Kusworo *et al.* Deep learning on medical imaging in identifying kidney stones. *In: EDP SCIENCES. E3S Web of Conferences*. [S. l.: s. n.], 2023. v. 448, p. 02019.

THEIN, Nilar *et al.* Automated False Positive Reduction and Feature Extraction of Kidney Stone Object in 3D CT Images. **International Journal of Intelligent Engineering and Systems**, 2019. Disponível em: <https://api.semanticscholar.org/CorpusID:86580755>.

TOMASI, Carlo; MANDUCHI, Roberto. Bilateral filtering for gray and color images. *In: IEEE. SIXTH international conference on computer vision (IEEE Cat. No. 98CH36271). [S. l.: s. n.], 1998. p. 839–846.*

WANG, Hong *et al.* MCA-UNet: multi-scale cross co-attentional U-Net for automatic medical image segmentation. **Health Information Science and Systems**, v. 11, 2023. Disponível em: <https://api.semanticscholar.org/CorpusID:256362049>.

WANG, Lei *et al.* Application of a three-dimensional visualization model in intraoperative guidance of percutaneous nephrolithotomy. **International Journal of Urology**, Wiley Online Library, 2022.

WANG, Zhou *et al.* Image quality assessment: from error visibility to structural similarity. **IEEE transactions on image processing**, IEEE, v. 13, n. 4, p. 600–612, 2004.

WATKINS, Christopher JCH; DAYAN, Peter. Q-learning. **Machine learning**, Springer, v. 8, p. 279–292, 1992.

WILHELM, Konrad *et al.* Validating automated kidney stone volumetry in CT and mathematical correlation with estimated stone volume based on diameter. **Journal of Endourology**, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, v. 32, n. 7, p. 659–664, 2018.

XU, Qing-Song; LIANG, Yi-Zeng. Monte Carlo cross validation. **Chemometrics and Intelligent Laboratory Systems**, Elsevier, v. 56, n. 1, p. 1–11, 2001.

YANG, Fan *et al.* Path planning of flexible ureteroscope based on ct image. *In: IEEE. 2019 Chinese Control Conference (CCC). [S. l.: s. n.], 2019. p. 4667–4672.*

YANG, Liang *et al.* Survey of robot 3D path planning algorithms. **Journal of Control Science and Engineering**, Hindawi, v. 2016, 2016.

YILDIRIM, Kadir *et al.* Deep learning model for automated kidney stone detection using coronal CT images. **Computers in biology and medicine**, Elsevier, v. 135, p. 104569, 2021.

YILDIRIM, Kadir *et al.* Deep learning model for automated kidney stone detection using coronal CT images. **Computers in biology and medicine**, v. 135, p. 104569, 2021. Disponível em: <https://api.semanticscholar.org/CorpusID:235608847>.

YOUNIS, Hussain A. *et al.* A Systematic Review and Meta-Analysis of Artificial Intelligence Tools in Medicine and Healthcare: Applications, Considerations, Limitations, Motivation and Challenges. **Diagnostics**, v. 14, 2024. Disponível em: <https://api.semanticscholar.org/CorpusID:266833220>.

ZHANG, Aston *et al.* Dive into deep learning. **arXiv preprint arXiv:2106.11342**, 2021.

ZHANG, Chongan *et al.* Multi-Objective Safety-Enhanced Path Planning for the Anterior Part of a Flexible Ureteroscope in Robot-Assisted Surgery. **The International Journal of Medical Robotics and Computer Assisted Surgery**, Wiley Online Library, v. 20, n. 6, e70007, 2024.

ZHANG, Qi *et al.* 3D path planning for anterior spinal surgery based on CT images and reinforcement learning. *In: IEEE. 2018 IEEE International Conference on Cyborg and Bionic Systems (CBS)*. [S. l.: s. n.], 2018. p. 317–321.

ZHAO, Jianchang *et al.* Design and performance investigation of a robot-assisted flexible ureteroscopy system. **Applied bionics and biomechanics**, Hindawi, v. 2021, 2021.

ZHOU, Zongwei *et al.* UNet++: A Nested U-Net Architecture for Medical Image Segmentation. **CoRR**, abs/1807.10165, 2018. arXiv: 1807.10165. Disponível em: <http://arxiv.org/abs/1807.10165>.

ZHU, Quanjing *et al.* Automatic kidney stone identification: an adaptive feature-weighted LSTM model based on urine and blood routine analysis. **Urolithiasis**, Springer, v. 52, n. 1, p. 145, 2024.