v7.n10.2026

REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218

RENAI: SISTEMA BASEADO EM APRENDIZADO PROFUNDO PARA DETECÇÃO E SEGMENTAÇÃO DE TUMORES RENAIS EM TOMOGRAFIA COMPUTADORIZADA

RENAI: A DEEP LEARNING-BASED SYSTEM FOR DETECTION AND SEGMENTATION OF RENAL TUMORS IN COMPUTED TOMOGRAPHY

RENAI: UN SISTEMA BASADO EN APRENDIZAJE PROFUNDO PARA LA DETECCIÓN Y SEGMENTACIÓN DE TUMORES RENALES EN TOMOGRAFÍA COMPUTARIZADA

Yuri Nascimento dos Santos[1], Leonardo de Sousa Ramos[1], Cleyson Pereira da Silva[1], Marta de Oliveira Barreiros[2]

e7109106
https://doi.org/10.47820/recima21.v7i10.9106

PUBLICADO: 10/2026

RESUMO

O câncer renal representa aproximadamente 2% dos cânceres diagnosticados no mundo, com incidência mais que dobrada nas últimas cinco décadas em países desenvolvidos. Mais da metade dos casos é descoberta de forma incidental, em exames realizados por outros motivos clínicos, e uma proporção relevante dos pacientes ainda chega ao diagnóstico com doença metastática. A dependência da interpretação humana em exames de tomografia computadorizada introduz variabilidade diagnóstica relevante, agravada pela desigualdade no acesso a serviços especializados nas regiões Norte e Nordeste do país. Este trabalho tem como objetivo geral desenvolver e avaliar o sistema RenAI, baseado em aprendizado profundo, para detecção e segmentação automática de tumores renais em imagens de tomografia computadorizada. Como objetivos específicos, buscou-se: comparar o desempenho de três arquiteturas da família YOLO (YOLOv8x-seg, YOLO11x-seg e YOLO26x-seg) em tarefas de segmentação multiclasse; propor e validar um fluxo de processamento e balanceamento do conjunto de dados KiTS23-2D; e desenvolver uma aplicação *web* funcional para apoio à predição em ambiente simulado. Os experimentos foram conduzidos no Google Colaboratory com GPU NVIDIA A100, utilizando 9.000 imagens balanceadas de 489 pacientes. O modelo YOLO11x-seg obteve o melhor desempenho, com mAP50-95 de 0,439 e F1 médio de 0,714, sendo selecionado para implantação na aplicação. O sistema alcançou precisão de 77% e *recall* de 70% na tarefa de detecção de tumor. O desbalanceamento severo entre classes, especialmente para cistos, permanece como principal limitante. Os resultados demonstram a viabilidade computacional da abordagem, mediante validação clínica futura, em contextos de saúde acessíveis.

PALAVRAS-CHAVE: Câncer renal. Aprendizado profundo. YOLO. Segmentação de imagens médicas. Tomografia computadorizada.

ABSTRACT

Renal cancer accounts for approximately 2% of all cancers diagnosed worldwide, with incidence more than doubling over the past five decades in developed countries. More than half of the cases are discovered incidentally, in examinations performed for other clinical reasons, and a relevant proportion of patients still present with metastatic disease at diagnosis. The reliance on human interpretation of computed tomography scans introduces relevant diagnostic variability, further compounded by inequalities in access to specialized services in the North and Northeast regions of the country. This study aims to develop and evaluate the RenAI system, based on deep learning, for the automatic detection and segmentation of renal tumors in computed tomography images.

[1]Graduando em Bacharelado em Engenharia de Software pela UEPA.
[2] Doutora em Engenharia Elétrica. Professora do curso de Engenharia de Software na UEPA.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Specific objectives included: comparing the performance of three architectures from the YOLO family (YOLOv8x-seg, YOLO11x-seg, and YOLO26x-seg) in multiclass segmentation tasks; proposing and validating a preprocessing and balancing workflow for the KiTS23-2D dataset; and developing a functional web application for prediction support in a simulated environment. Experiments were conducted on Google Colaboratory with an NVIDIA A100 GPU, using 9,000 balanced images from 489 patients. The YOLO11x-seg model achieved the best performance, with a mask mAP50-95 of 0.439 and mean F1 of 0.714, and was selected for deployment. The system achieved 77% precision and 70% recall in tumor detection. Severe class imbalance, particularly for cysts, remains the primary performance limitation. The results demonstrate the computational feasibility of the approach, pending future clinical validation, in accessible healthcare settings.

KEYWORDS: *Renal cancer. Deep learning. YOLO. Medical image segmentation. Computed tomography.*

RESUMEN

El cáncer renal representa aproximadamente el 2% de los cánceres diagnosticados en el mundo, con incidencia más que duplicada en las últimas cinco décadas en países desarrollados. Más de la mitad de los casos se descubre de forma incidental, en exámenes realizados por otros motivos clínicos, y una proporción relevante de los pacientes llega al diagnóstico con enfermedad metastásica. La dependencia de la interpretación humana en tomografías computarizadas introduce variabilidad diagnóstica relevante, agravada por las desigualdades en el acceso a servicios especializados en las regiones Norte y Nordeste del país. Este trabajo tiene como objetivo desarrollar y evaluar el sistema RenAI, basado en aprendizaje profundo, para la detección y segmentación automática de tumores renales en tomografía computarizada. Los objetivos específicos incluyeron comparar el rendimiento de tres arquitecturas YOLO (YOLOv8x-seg, YOLO11x-seg y YOLO26x-seg) en segmentación multiclase, proponer un flujo de preprocesamiento y balanceo del conjunto KiTS23-2D, y desarrollar una aplicación web funcional para apoyo a la predicción en un entorno simulado. Los experimentos se realizaron en Google Colaboratory con GPU NVIDIA A100, utilizando 9.000 imágenes balanceadas de 489 pacientes. El modelo YOLO11x-seg obtuvo el mejor desempeño, con mAP50-95 de 0,439 y F1 medio de 0,714, siendo seleccionado para su implantación. El sistema alcanzó una precisión del 77% y un recall del 70% en la detección de tumor. El severo desequilibrio entre clases, especialmente para quistes, sigue siendo la principal limitación del sistema. Los resultados demuestran la viabilidad computacional del enfoque, sujeta a validación clínica futura, en entornos de salud accesibles.

PALABRAS CLAVE: *Cáncer renal. Aprendizaje profundo. YOLO. Segmentación de imágenes médicas. Tomografía computarizada.*

1. INTRODUÇÃO

O câncer renal representa aproximadamente 2% de todos os cânceres diagnosticados no mundo, com incidência mais que dobrada nas últimas cinco décadas em países desenvolvidos (PADALA et al., 2020). Com a ampla utilização de exames de imagem, mais da metade dos casos passou a ser descoberta de forma incidental, durante exames realizados por outros motivos clínicos (VALLECILLA VALENCIA; GARCIA-PERDOMO, 2026; PADALA et al., 2020). Ainda assim, uma parcela importante dos pacientes chega ao diagnóstico com doença

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

metastática, mais de 25%, segundo revisão da literatura (ROSSI et al., 2018). Quando identificado em estágios avançados, o tratamento torna-se mais complexo, oneroso e com menores chances de sucesso terapêutico.

A detecção precoce por meio de exames de tomografia computadorizada é o método mais utilizado na identificação de lesões renais, porém ainda depende fortemente da interpretação humana, o que introduz variabilidade diagnóstica relevante (XIONG *et al.*, 2025). Esse cenário é agravado pela desigualdade no acesso a serviços especializados no Brasil, especialmente nas regiões Norte e Nordeste, onde as taxas de internação por câncer renal são significativamente menores em comparação com o Sul e o Sudeste (SILVA *et al.*, 2025), sugerindo dificuldades no diagnóstico e não necessariamente menor incidência da doença.

Nesse contexto, técnicas de inteligência artificial baseadas em aprendizado profundo (*deep learning*), em particular redes neurais convolucionais, têm demonstrado resultados promissores na análise automatizada de imagens médicas. Modelos da família YOLO (*You Only Look Once*), originalmente desenvolvidos para detecção de objetos em tempo real, vêm sendo adaptados para tarefas de segmentação em imagens clínicas, permitindo identificar não apenas a presença de uma lesão, mas também delimitar com precisão a região afetada (ZHAO *et al.*, 2024). Essa capacidade é especialmente relevante no contexto do câncer renal, no qual a localização e o tamanho do tumor influenciam diretamente a conduta terapêutica.

Apesar dos avanços observados na literatura, ainda são escassos sistemas que integrem modelos de segmentação de última geração com fluxos de processamento de imagens médicas voltados especificamente para tumores renais. Diante disso, o presente trabalho é orientado pelo seguinte problema de pesquisa: qual arquitetura da família YOLO apresenta melhor desempenho quantitativo na segmentação de tumores renais, cistos e parênquima renal em imagens de tomografia computadorizada, mesmo diante do severo desbalanceamento entre classes?

A justificativa para esta pesquisa reside na necessidade de soluções computacionais acessíveis que reduzam a variabilidade diagnóstica e ampliem o acesso a tecnologias de análise de imagens médicas, especialmente em regiões com menor disponibilidade de especialistas, como o Norte e o Nordeste do Brasil. O uso de dados públicos e de infraestrutura de baixo custo reforça a aplicabilidade prática da proposta.

Este trabalho tem como objetivo geral desenvolver e avaliar o sistema RenAI, baseado em aprendizado profundo, para detecção e segmentação automática de tumores renais em imagens de tomografia computadorizada. Como objetivos específicos, busca-se: (i) comparar o desempenho das arquiteturas YOLOv8x-seg, YOLO11x-seg e YOLO26x-seg em tarefas de

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

segmentação multiclasse; (ii) propor e validar um fluxo de processamento e balanceamento do conjunto de dados KiTS23-2D; e (iii) desenvolver uma aplicação *web* funcional para apoio à predição em ambiente simulado, com suporte a múltiplos formatos de imagem médica.

Os resultados obtidos indicam a viabilidade computacional e abrem caminho para o desenvolvimento de ferramentas de suporte diagnóstico acessíveis em diferentes contextos de saúde, mediante validação clínica futura.

2. REFERENCIAL TEÓRICO

2.1. Inteligência artificial aplicada ao diagnóstico por imagem

A inteligência artificial (IA) consiste em um conjunto de técnicas computacionais desenvolvidas para permitir que sistemas realizem tarefas que, tradicionalmente, dependem da inteligência humana, como reconhecimento de padrões, tomada de decisões e resolução de problemas. Conforme apontado por Yu et al. (2018), a IA busca desenvolver sistemas capazes de processar dados complexos de saúde, reconhecer padrões clínicos e subsidiar tomadas de decisão de forma precisa e eficiente. Nas últimas décadas, o avanço da capacidade computacional, aliado ao crescimento da disponibilidade de dados, impulsionou significativamente o desenvolvimento dessa tecnologia, tornando-a uma das principais ferramentas empregadas em diferentes áreas do conhecimento, especialmente na saúde.

Na medicina, a inteligência artificial tem sido aplicada para auxiliar profissionais na interpretação de exames, no apoio ao diagnóstico e na identificação de alterações clínicas que poderiam passar despercebidas durante uma avaliação convencional. Sistemas inteligentes são capazes de analisar grandes volumes de informações em um curto intervalo de tempo, contribuindo para aumentar a precisão diagnóstica e reduzir o tempo necessário para a emissão de laudos.

Entre as diversas aplicações da IA na área médica, destaca-se a análise de imagens obtidas por exames como tomografia computadorizada (TC), ressonância magnética (RM), radiografia e ultrassonografia. Esses exames produzem grande quantidade de informações visuais, cuja interpretação pode ser complexa e depende diretamente da experiência do especialista. Nesse contexto, algoritmos inteligentes atuam como ferramentas de apoio à decisão clínica, auxiliando na identificação de lesões, tumores e outras anormalidades com maior rapidez e consistência. Conforme observado por Chen *et al*. (2022), os avanços obtidos com técnicas de aprendizado profundo têm ampliado significativamente a capacidade dos sistemas

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

computacionais de analisar imagens médicas, contribuindo para diagnósticos mais precisos e eficientes.

No caso específico do câncer renal, a tomografia computadorizada representa um dos principais métodos de imagem utilizados tanto para a identificação quanto para o planejamento terapêutico da doença. Entretanto, a análise manual das imagens pode demandar tempo considerável e estar sujeita à variabilidade entre diferentes profissionais. Dessa forma, a aplicação da inteligência artificial possibilita automatizar parte desse processo, contribuindo para a identificação precoce de regiões suspeitas e oferecendo suporte ao especialista durante a avaliação do exame.

É importante destacar que a utilização da inteligência artificial não tem como objetivo substituir a atuação médica, mas sim fornecer ferramentas capazes de ampliar a eficiência e a confiabilidade do diagnóstico. A combinação entre o conhecimento clínico do profissional e os recursos computacionais favorece diagnósticos mais precisos e contribui para o desenvolvimento de soluções voltadas à medicina de precisão (JIANG *et al.,* 2017).

Nesse cenário, o projeto RenAI insere-se como uma aplicação da inteligência artificial voltada ao auxílio na detecção precoce de tumores renais em imagens de tomografia computadorizada. Para isso, emprega modelos baseados em aprendizado profundo capazes de identificar e segmentar automaticamente estruturas de interesse, fornecendo informações que podem auxiliar profissionais da saúde durante o processo de diagnóstico.

2.2. Aprendizado profundo e redes neurais convolucionais

O aprendizado profundo (*deep learning*) é um subcampo do aprendizado de máquina (*machine learning)* que utiliza redes neurais artificiais com múltiplas camadas para aprender padrões complexos a partir de grandes volumes de dados. Segundo Alzubaidi *et al*. (2021), diferentemente das abordagens tradicionais de aprendizado de máquina, os modelos de *deep learning* são capazes de aprender automaticamente representações hierárquicas dos dados, reduzindo a necessidade de extração manual de características. Essa capacidade torna essa tecnologia especialmente eficiente em tarefas relacionadas ao reconhecimento de imagens, processamento de linguagem natural e análise de sinais.

Na área da saúde, o aprendizado profundo tem proporcionado avanços significativos na interpretação de exames médicos, permitindo a identificação automática de alterações em imagens de tomografia computadorizada, ressonância magnética, radiografias e outros exames de imagem. Essa capacidade torna a tecnologia uma importante ferramenta de apoio ao

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

diagnóstico, contribuindo para aumentar a precisão das análises e reduzir o tempo necessário para a avaliação dos exames.

Entre as arquiteturas mais utilizadas em tarefas de visão computacional destacam-se as redes neurais convolucionais (*Convolutional Neural Networks* – CNNs). Essas redes foram desenvolvidas para processar dados visuais e são capazes de reconhecer automaticamente padrões, formas, texturas e características presentes nas imagens. Seu funcionamento baseia-se na aplicação de filtros convolucionais que extraem informações em diferentes níveis de complexidade, permitindo que a rede aprenda desde características simples, como bordas e contornos, até estruturas mais complexas, como órgãos e lesões. Conforme demonstrado por Khan *et al*. (2020), as CNNs representam um dos principais avanços do aprendizado profundo para aplicações de reconhecimento de padrões em imagens, consolidando-se como uma das arquiteturas mais empregadas em visão computacional.

Nas aplicações médicas, as CNNs têm apresentado resultados expressivos em atividades de classificação, detecção e segmentação de imagens. Enquanto a classificação determina a presença ou ausência de determinada condição, a detecção localiza regiões de interesse e a segmentação delimita com maior precisão a área ocupada pela estrutura analisada. Essa última característica é particularmente importante em exames de tomografia computadorizada, pois permite identificar com maior exatidão os limites de tumores e outras alterações anatômicas.

Os avanços obtidos pelas redes neurais convolucionais impulsionaram o desenvolvimento de arquiteturas cada vez mais eficientes para tarefas de detecção e segmentação de objetos. Entre essas arquiteturas destaca-se a família YOLO (*You Only Look Once*), que combina alta velocidade de processamento com elevada precisão na identificação de objetos em imagens (TERVEN *et al*., 2023). Ao longo de sua evolução, a família YOLO incorporou mecanismos capazes de realizar não apenas a detecção, mas também a segmentação de instâncias, tornando-se uma alternativa amplamente utilizada em aplicações de visão computacional, inclusive na área médica (ULTRALYTICS, 2024).

Neste trabalho, os conceitos de aprendizado profundo e redes neurais convolucionais constituem a base para o desenvolvimento do sistema RenAI. A pesquisa utiliza modelos da família YOLO voltados à segmentação de imagens médicas, permitindo comparar diferentes arquiteturas quanto ao desempenho na identificação de tumores renais em exames de tomografia computadorizada.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

2.3. Segmentação de imagens médicas

A segmentação de imagens consiste no processo de identificar e delimitar automaticamente regiões específicas de interesse dentro de uma imagem digital. Na área da visão computacional, essa técnica é amplamente utilizada para separar objetos do restante da imagem, permitindo que estruturas anatômicas, órgãos ou lesões sejam analisados individualmente. Segundo Zhou *et al*. (2021), a segmentação automática representa uma das tarefas mais relevantes da visão computacional aplicada à saúde, pois possibilita a identificação precisa de estruturas anatômicas e alterações patológicas em exames médicos. Em exames de imagem, a segmentação desempenha um papel fundamental ao fornecer informações mais precisas sobre a localização, o formato e a extensão de alterações presentes nos tecidos.

Diferentemente das tarefas de classificação e detecção, que indicam respectivamente a presença de determinada condição ou a localização aproximada de um objeto, a segmentação produz uma delimitação detalhada da região analisada. Essa característica torna a técnica especialmente relevante em aplicações médicas, nas quais pequenas diferenças na dimensão ou no contorno de uma lesão podem influenciar diretamente o diagnóstico e o planejamento do tratamento.

Nos exames de tomografia computadorizada, a segmentação permite destacar estruturas anatômicas específicas, facilitando a visualização de órgãos e possíveis anormalidades. No contexto do câncer renal, essa abordagem possibilita identificar com maior precisão os limites do tumor, contribuindo para uma análise mais detalhada das imagens e auxiliando profissionais da saúde na interpretação dos exames (XIONG *et al*., 2025).

Com os avanços do aprendizado profundo, modelos baseados em redes neurais convolucionais passaram a realizar a segmentação de imagens de forma automática e com elevado desempenho. Entre essas soluções, destacam-se os modelos da família YOLO voltados para segmentação de instâncias, capazes de localizar objetos e gerar máscaras que representam seus contornos de maneira precisa. Essa evolução ampliou significativamente as possibilidades de aplicação da visão computacional em exames médicos, permitindo desenvolver sistemas capazes de oferecer suporte ao diagnóstico de diferentes patologias, como observado por Minaee *et al*. (2022) em um amplo levantamento sobre técnicas de segmentação baseadas em aprendizado profundo.

Neste trabalho, a segmentação constitui uma das principais técnicas empregadas para a identificação de tumores renais em imagens de tomografia computadorizada. A utilização de modelos especializados nessa tarefa possibilita comparar diferentes arquiteturas quanto à

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

capacidade de delimitar corretamente as regiões tumorais, contribuindo para avaliar seu potencial como ferramenta de apoio ao diagnóstico precoce.

2.4. Evolução da família YOLO e modelos utilizados na pesquisa

A família YOLO (*You Only Look Once*) representa uma das principais arquiteturas desenvolvidas para tarefas de detecção de objetos em imagens e vídeos. Essa arquitetura introduziu uma abordagem inovadora ao realizar a identificação de objetos em uma única etapa de processamento, diferentemente dos métodos tradicionais que executavam múltiplas fases para localizar e classificar elementos presentes na imagem. Segundo Terven *et al.* (2023), essa estratégia permitiu alcançar elevadas taxas de processamento em tempo real, mantendo níveis competitivos de precisão e tornando a arquitetura uma referência em aplicações de visão computacional.

Desde sua primeira versão, a arquitetura passou por sucessivas evoluções, incorporando melhorias relacionadas à precisão, desempenho computacional e capacidade de generalização. Cada nova geração buscou reduzir limitações identificadas nas versões anteriores, aprimorando mecanismos de extração de características, otimização do treinamento e capacidade de detectar objetos em diferentes escalas. Como consequência, a família YOLO expandiu sua aplicação para áreas que exigem alta precisão, como veículos autônomos, monitoramento inteligente, agricultura de precisão e diagnóstico por imagens médicas (SAPKOTA *et al*., 2025).

Além da evolução voltada à detecção de objetos, versões mais recentes passaram a oferecer suporte nativo à segmentação de instâncias (*instance segmentation*), permitindo que, além da localização dos objetos, fossem produzidas máscaras capazes de delimitar precisamente seus contornos. Essa funcionalidade tornou os modelos particularmente adequados para aplicações médicas, nas quais a identificação exata da região de interesse possui papel fundamental no auxílio ao diagnóstico e ao planejamento terapêutico, conforme descrito na documentação técnica da Ultralytics (2023).

Entre os modelos empregados nesta pesquisa destaca-se o YOLOv8x-seg, desenvolvido pela Ultralytics como uma das primeiras versões da família a integrar, de forma consolidada, recursos de segmentação em uma arquitetura otimizada. O modelo apresenta elevada capacidade de aprendizado, boa precisão na delimitação de objetos e desempenho satisfatório em tarefas que envolvem imagens complexas, características que favoreceram sua ampla adoção em aplicações de pesquisa e desenvolvimento.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Como continuidade da evolução da arquitetura, o YOLO11x-seg introduziu melhorias relacionadas à eficiência computacional e ao refinamento da extração de características, buscando aumentar a precisão da segmentação sem comprometer significativamente o desempenho (ULTRALYTICS, 2024). Essas otimizações ampliaram sua capacidade de representar detalhes presentes nas imagens, aspecto especialmente relevante em exames médicos, nos quais pequenas diferenças entre tecidos saudáveis e regiões lesionadas podem influenciar diretamente o resultado da análise.

Mais recentemente, o YOLO26x-seg introduziu inferência sem *Non-Maximum Suppression* (NMS), remoção da *Distribution Focal Loss* (DFL), um otimizador híbrido voltado à estabilidade de treinamento e, para segmentação, um módulo de protótipos atualizado e mecanismos de atribuição de rótulos sensíveis a alvos pequenos, potencialmente relevantes para classes minoritárias (ULTRALYTICS, 2026).

A escolha dos modelos YOLOv8x-seg, YOLO11x-seg e YOLO26x-seg fundamenta-se justamente na possibilidade de comparar diferentes gerações da mesma arquitetura, avaliando sua evolução quanto à capacidade de segmentar tumores renais em exames de tomografia computadorizada. Por pertencerem à mesma família, esses modelos compartilham princípios arquiteturais semelhantes, permitindo que as diferenças observadas durante os experimentos estejam relacionadas principalmente às melhorias introduzidas em cada versão. Dessa forma, a comparação torna-se mais consistente e contribui para identificar qual arquitetura apresenta melhor desempenho para a aplicação proposta pelo sistema RenAI.

Registra-se que arquiteturas baseadas em U-Net (RONNEBERGER; FISCHER; BROX, 2015), incluindo variantes especializadas para segmentação renal como o nnU-Net, dominam as soluções vencedoras do desafio KiTS23 em sua variante 3D (MYRONENKO et al., 2024), cuja comparação direta com os resultados deste trabalho é limitada pela diferença de dimensionalidade (3D vs. 2D), ainda que o cálculo do coeficiente de *Dice* por classe (Seção 4.2) aproxime a métrica utilizada da literatura de referência.

2.5. Métricas de avaliação de modelos de segmentação

A avaliação de modelos de inteligência artificial aplicados à visão computacional é uma etapa fundamental para verificar a capacidade do algoritmo em identificar corretamente os objetos de interesse. No contexto da segmentação de imagens médicas, a utilização de métricas quantitativas permite comparar diferentes modelos de forma objetiva, possibilitando identificar aquele que apresenta melhor desempenho na delimitação das regiões analisadas. Segundo Tharwat (2021), a escolha de métricas adequadas é indispensável para interpretar corretamente

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

o desempenho de algoritmos de classificação e detecção, permitindo análises mais confiáveis e consistentes.

Entre as métricas mais utilizadas destacam-se a Precisão (*Precision*), o *Recall* (Revocação) e o *F1-Score*, amplamente empregadas na avaliação de modelos de classificação, detecção e segmentação. A Precisão indica a proporção de predições positivas classificadas corretamente pelo modelo, enquanto o Recall representa sua capacidade de identificar todos os casos positivos existentes no conjunto de dados. *O F1-Score*, por sua vez, corresponde à média harmônica entre Precisão e *Recall*, oferecendo uma medida equilibrada do desempenho quando ambas as métricas possuem a mesma importância.

Outra métrica amplamente utilizada em tarefas de detecção e segmentação é a Interseção sobre União (*Intersection over Union* – IoU). Essa métrica mede o grau de sobreposição entre a região prevista pelo modelo e a região real anotada no conjunto de dados. Quanto maior o valor de IoU, maior é a correspondência entre a segmentação produzida pelo modelo e a área efetivamente ocupada pelo objeto de interesse, indicando maior precisão na delimitação das estruturas analisadas.

Além da IoU, destaca-se a Precisão Média (*Mean Average Precision* – mAP), considerada uma das principais métricas para avaliação de modelos da família YOLO. A mAP sintetiza o desempenho do modelo considerando diferentes limiares de confiança e diferentes níveis de sobreposição entre as predições e as anotações reais, permitindo uma avaliação mais abrangente da capacidade de detecção e segmentação. Por esse motivo, essa métrica tornou-se referência na comparação de arquiteturas de visão computacional e é frequentemente empregada em pesquisas envolvendo aprendizado profundo, conforme descrito por Padilla *et al*. (2020) em um estudo comparativo sobre métricas de detecção de objetos.

No contexto desta pesquisa, essas métricas constituem a base para a comparação entre os modelos YOLOv8x-seg, YOLO11x-seg e YOLO26x-seg, permitindo avaliar sua eficiência na segmentação de tumores renais em imagens de tomografia computadorizada. A utilização conjunta desses indicadores possibilita analisar diferentes aspectos do desempenho dos modelos, contribuindo para uma comparação mais confiável e fundamentada quanto à capacidade de identificar corretamente as regiões tumorais.

3. METODOLOGIA

Este capítulo descreve o processo de desenvolvimento do sistema RenAI, desde a seleção e preparação da base de dados até a construção de uma aplicação *web* para apoio à

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

predição em ambiente simulado. O desenvolvimento foi conduzido de forma iterativa, cobrindo etapas de pré-processamento, treinamento, otimização e implantação.

3.1. Classificação da Pesquisa

Esta pesquisa classifica-se como pesquisa aplicada, pois tem como finalidade o desenvolvimento do sistema para apoio à decisão clínica na detecção de tumores renais. Quanto aos objetivos, assume caráter exploratório e descritivo, ao comparar arquiteturas de segmentação de imagens e descrever o processo de construção do fluxo de pré-processamento, treinamento e implantação. A abordagem é quantitativa, apoiada na mensuração e comparação objetiva de métricas de desempenho entre arquiteturas. Quanto aos procedimentos técnicos, articula pesquisa bibliográfica, pesquisa experimental sobre a base pública selecionada e desenvolvimento tecnológico.

3.2. Levantamento Bibliográfico

A revisão bibliográfica foi conduzida no Google Acadêmico, com recorte temporal entre 2018 e 2026, em buscas nos idiomas português e inglês. Foram utilizados termos relacionados a câncer renal, aprendizado profundo, arquiteturas YOLO, segmentação de imagens médicas e tomografia computadorizada. Como critérios de exclusão, adotaram-se: eliminação de duplicatas, exclusão de trabalhos sem relação direta com o tema e desconsideração de estudos sem texto completo disponível.

3.3. Seleção e Análise da Base de Dados

Para o desenvolvimento do presente trabalho, adotou-se a base de dados KiTS23-2D, uma reformatação bidimensional do conjunto do Kidney Tumor Segmentation Challenge *2023* (KITS23 CHALLENGE, 2023), disponibilizada na plataforma Kaggle. O conjunto é composto por exames pré-operatórios com contraste, em fase corticomedular, de pacientes submetidos a nefrectomia parcial ou radical. Na variante 2D, cada volume tridimensional foi seccionado no plano axial (sem exploração da continuidade volumétrica entre fatias adjacentes), resultando em 95.221 cortes de 489 pacientes, com resolução nativa de 512 × 512 pixels. As imagens são fornecidas em formato PNG, e cada corte é acompanhado de uma máscara de segmentação com quatro classes: fundo (0), rim (1), tumor (2) e cisto (3).

A escolha dessa base justifica-se por fatores técnicos e metodológicos. Por ser pública, bem documentada e derivada de um desafio de referência, o conjunto assegura comparabilidade com a literatura e reprodutibilidade dos experimentos. Sua formatação bidimensional, verificada

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

sem erros em relação aos volumes originais, dispensa o fatiamento manual dos arquivos e mostra-se diretamente compatível com arquiteturas de segmentação. A escala do conjunto oferece volume adequado ao treinamento de modelos de aprendizado profundo.

A análise prévia evidenciou três desafios centrais que orientaram as decisões metodológicas subsequentes. O primeiro é o desbalanceamento em nível de fatia: 65,4% dos cortes são compostos exclusivamente por fundo, comportamento anatomicamente esperado, uma vez que os rins ocupam apenas uma faixa do abdômen. O segundo é o desbalanceamento em nível de pixel: o rim corresponde a 0,63% dos pixels totais, o tumor a 0,18% e o cisto a apenas 0,015%, contra 99,17% de fundo. O terceiro é a sobreposição anatômica, pois tumores e cistos quase sempre coexistem na mesma fatia, exigindo a distinção entre estruturas adjacentes e visualmente semelhantes. Tais características reforçam a adequação da base como cenário de pesquisa em segmentação com classes severamente desbalanceadas, justificando as estratégias de balanceamento e de escolha da função de perda adotadas nas etapas seguintes.

3.4. Preparação e Balanceamento da Base de Dados

Os experimentos foram conduzidos no Google Colaboratory (Google Colab), ambiente de desenvolvimento em nuvem que oferece acesso simplificado a recursos de GPU/TPU e integração com o Google Drive, dispensando a instalação local de software (BISONG, 2019). Sua compatibilidade nativa com bibliotecas como PyTorch e Ultralytics viabilizou o treinamento e a avaliação dos modelos YOLOv8x-seg, YOLO11x-seg e YOLO26x-seg com aceleração por GPU, reduzindo o tempo de processamento e a dependência de infraestrutura local de alto desempenho.

O conjunto original KiTS23-2D foi obtido na plataforma Kaggle e, inicialmente, baixado para um computador local e, em seguida, transferido para um repositório no Google Drive, a partir do qual as rotinas de leitura, balanceamento e anotação foram executadas no Colab. O conjunto resultante já balanceado e convertido ao formato YOLO foi gravado diretamente no próprio Google Drive.

Diante do desbalanceamento identificado na seção anterior, desenvolveu-se um fluxo de processamento para a preparação de dados no ambiente do Google Colab, estruturado nas etapas de agrupamento, amostragem, particionamento e normalização. Inicialmente, as fatias foram organizadas em três grupos mutuamente exclusivos, atribuindo-se cada fatia à classe mais rara nela presente, segundo a ordem de prioridade cisto > tumor > rim saudável: (i) fatias contendo cisto; (ii) fatias contendo tumor, mas não cisto; e (iii) fatias contendo apenas parênquima renal saudável, sem tumor ou cisto. Essa estratégia de priorização assegura a

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

retenção das classes minoritárias, que seriam suprimidas por uma amostragem aleatória simples.

De cada grupo foram amostradas, sem reposição e com semente aleatória fixa (*seed* = 42), 3.000 fatias únicas, totalizando 9.000 imagens. A ausência de duplicatas foi confirmada por verificação automática sobre o número de nomes de arquivos distintos. Em seguida, realizou-se o particionamento em proporções aproximadas de 60% para treinamento, 20% para validação e 20% para teste, em nível de paciente: todas as fatias de um mesmo caso permanecem na mesma partição, procedimento que elimina o vazamento de dados entre os conjuntos, condição assegurada por verificações de interseção vazia entre os grupos de pacientes. Verificou-se também a presença de ao menos cinco pacientes com cisto nas partições de validação e teste, de modo a garantir diversidade suficiente da classe mais rara durante a avaliação.

Por se tratar de divisão em nível de paciente, a distribuição de fatias por grupo e partição não é exatamente proporcional, situando-se em torno dos valores de projeto apresentados na Tabela 1.

Tabela 1. Distribuição aproximada de fatias por partição e grupo

Partição	Cisto	Tumor	Rim saudável
Treino (≈60%)	1507	1673	1642
Validação (≈20%)	689	633	617
Teste (≈20%)	804	694	741
Total	3000	3000	3000

Fonte: elaborada pelos autores (2026). *Nota:* valores aproximados; as contagens podem variar ligeiramente em razão do particionamento por paciente.

Por fim, calculou-se a normalização *z-score* utilizando exclusivamente as fatias de treino com presença de rim, evitando a contaminação das estatísticas por dados de validação e teste. As estatísticas globais obtidas — média de 0,1882 e desvio-padrão de 0,2304 — computadas sobre os valores de pixel normalizados ao intervalo [0, 1] foram persistidas e aplicadas a todas as partições.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

3.5. Conversão para o Formato YOLO

As máscaras de segmentação foram convertidas para o formato de segmentação por polígonos do *framework* YOLO. Para cada classe de interesse (rim, tumor e cisto), remapeadas respectivamente para os índices 0, 1 e 2 (com a classe de fundo descartada), gerou-se uma máscara binária e extraíram-se os contornos externos. Contornos com área inferior a 10 pixels foram descartados como ruído de borda. Cada contorno remanescente foi simplificado, adotando-se epsilon equivalente a 0,3% do perímetro, o que reduz o número de vértices e o tamanho dos arquivos de anotação sem perda significativa de fidelidade geométrica; polígonos resultantes com menos de três vértices foram igualmente descartados. As coordenadas dos vértices foram então normalizadas ao intervalo [0, 1] em relação às dimensões da imagem (512 × 512), conforme exige o formato.

As imagens, originalmente em 16 bits (uint16), foram convertidas para 8 bits mediante divisão por 256. Essa operação assume que os valores de pixel do KiTS23-2D correspondem a um janelamento linear da faixa de Unidades Hounsfield [−200, 300] HU — janela habitualmente empregada na avaliação renal com contraste — mapeada para a profundidade de 16 bits original; essa correspondência foi verificada empiricamente pela inspeção do intervalo de valores presentes em uma amostra de fatias do conjunto, tendo o intervalo observado se mostrado compatível com o janelamento descrito. Ainda assim, a redução de 16 para 8 bits reduz a resolução radiométrica disponível, com possível perda de sutilezas de contraste entre tecidos, retomada como limitação na Seção 5. Para cada fatia gerou-se o respectivo arquivo de anotação .txt, mantendo-se arquivos vazios nos casos sem estrutura anotável, conforme requisito do formato. Concluído o processamento, organizou-se a estrutura de diretórios padrão do YOLO (images/ e labels/, cada qual subdividida em treino, validação e teste) e gerou-se o arquivo de configuração dataset.yaml, declarando as três classes (rim, tumor e cisto).

3.6. Treinamento e Otimização

Os experimentos foram conduzidos em uma GPU NVIDIA A100, com treinamento em precisão mista automática (*AMP*) e salvamento periódico de *checkpoints* (a cada 10 épocas), recurso que permite retomar o treinamento após eventuais interrupções de sessão. A estratégia foi desenvolvida de forma iterativa, refinando-se progressivamente os hiperparâmetros, a augmentação e o balanceamento do conjunto, com o objetivo de elevar a sensibilidade às classes minoritárias, sobretudo tumor e cisto, até convergir para uma configuração consolidada, posteriormente aplicada de forma idêntica a todos os modelos comparados.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Definidos a configuração e o conjunto, procedeu-se à etapa central deste trabalho: a comparação controlada entre três arquiteturas de segmentação, todas na escala *x* (*extra-large*): YOLOv8x-seg, YOLO11x-seg e YOLO26x-seg. As três dispõem de pesos pré-treinados no conjunto COCO, de modo que todos os modelos partiram do mesmo regime de aprendizado por transferência. Cada modelo foi treinado com o mesmo *script*, os mesmos hiperparâmetros e a mesma partição de dados, isolando a arquitetura como única variável e assegurando uma comparação justa entre gerações sucessivas do YOLO. O objetivo foi identificar o modelo de melhor desempenho na segmentação de rim, tumor e cisto, a fim de selecioná-lo para a implantação na API.

Os limiares de confiança (0,12) e de IoU (0,45), bem como o limiar de área (10 pixels) e o epsilon de simplificação (0,3%) descritos na Seção 3.5, foram definidos empiricamente no conjunto de validação, maximizando o F1-Score como critério de equilíbrio entre precisão e recall, dada a relevância clínica de ambos os tipos de erro nas classes minoritárias; as métricas reportadas na Seção 4 foram calculadas por fatia (2D), e não por paciente. A avaliação final no conjunto de teste utilizou inferência simples, sem *test-time augmentation* (TTA), de modo consistente com o procedimento empregado no cálculo de *Dice* e IoU por classe (Seção 4.2). O coeficiente de *Dice* e o IoU médio por classe foram calculados sobre os pares predição–anotação corretamente casados, utilizando a implementação oficial do protocolo COCO (pycocotools), com o mesmo limiar de IoU (0,45) adotado nas demais métricas.

Tabela 2. Configuração final de treinamento (comum às três arquiteturas)

Grupo	Parâmetro	Valor
Imagem	imgsz / batch	512 / 64 (A100 40 GB; ajustado pela VRAM)
Treinamento	epochs / patience	300 / 60
Otimização	optimizer	AdamW
Otimização	lr0 / lrf	0,0008 / 0,008
Otimização	escalonamento / *warmup*	cosseno / 15 épocas
Augmentação	mosaic / mixup	1,0 / 0,10

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Augmentação	copy_paste (modo *flip*)	0,7
Augmentação	fliplr / flipud	0,5 / 0,3
Augmentação	degrees / scale / translate	15 / 0,5 / 0,15
Augmentação	hsv_h / hsv_s / hsv_v	0 / 0 / 0,25
Perda e máscara	box / cls	10,0 / 3,0
Perda e máscara	label_smoothing	0,05
Perda e máscara	mask_ratio / overlap_mask	1 / ativado

Fonte: elaborada pelos autores (2026).

3.7. Desenvolvimento da Aplicação Web

Para viabilizar o uso do modelo em um ambiente simulado, desenvolveu-se uma aplicação *web* completa, denominada RenAI.

O *frontend* foi desenvolvido com HTML, CSS e JavaScript puros, sem o uso de *frameworks*. Foi adotada uma paleta com fundo branco e azul-saúde como cor primária, remetendo ao ambiente hospitalar. As cores das classes seguem a codificação visual: verde-escuro para rim, vermelho para tumor e âmbar para cisto, assegurando contraste sobre fundo claro. A interface permite o envio por clique ou por *drag and drop*, a troca de imagem sem recarregamento da página, o ajuste em tempo real dos limiares de confiança e de IoU por meio de *sliders*, e exibe um indicador de status de conexão com o *backend*, com reconexão automática. O *layout* responsivo adapta-se a dispositivos móveis (≤ 600 px), *tablets* (≤ 900 px) e *desktops*, estes com dois painéis lado a lado.

A visualização das detecções adotou marcadores de canto em vez de *bounding boxes* completas, reduzindo a oclusão das estruturas anatômicas subjacentes. Os rótulos foram posicionados com afastamento de 4 pixels acima da linha do canto e submetidos a um algoritmo de detecção de colisão que reposiciona automaticamente rótulos sobrepostos, evitando a sobreposição de texto em fatias com múltiplas detecções simultâneas.

O *backend* foi implementado, expondo dois *endpoints*: *GET* /health, para verificação de disponibilidade do serviço, e *POST* /predict, para inferência. O *endpoint* de predição aceita

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

imagens em múltiplos formatos (PNG de 8 e de 16 bits, JPEG e TIFF), com detecção automática de formato. A cada formato é aplicada a conversão adequada: imagens PNG de 16 bits são convertidas para 8 bits por divisão por 256, preservando o janelamento HU do conjunto de treinamento. A resposta da API inclui a classe detectada, a confiança, as coordenadas da *bounding box*, a área em pixels, a imagem anotada em base64 e os metadados da imagem processada.

A suíte de testes automatizados foi estruturada em três camadas, organizadas por tipo e marcadas com decoradores para execução seletiva. Os testes unitários cobrem as funções de conversão de imagem. Os testes de integração validam os *endpoints* /health e /predict com múltiplos formatos de entrada e com parâmetros nos limites válidos e inválidos, garantindo que entradas corrompidas retornem HTTP 400 ou 422, e nunca HTTP 500. Os testes de ponta a ponta (*E2E*) de desempenho verificam que o *endpoint* /predict responde em menos de 15 segundos em CPU, que chamadas sequenciais não apresentam degradação progressiva superior a 2× e que os resultados são determinísticos para uma mesma entrada. A suíte totaliza 47 casos de teste distribuídos entre as três camadas.

4. RESULTADOS E DISCUSSÃO

4.1. Comparação entre Arquiteturas e Seleção do Modelo

As três arquiteturas treinadas sob configuração idêntica foram avaliadas no conjunto de teste, com limiar de confiança de 0,12 e limiar de IoU de 0,45. A Tabela 3 resume o desempenho global de cada modelo.

Tabela 3. Comparação entre as três arquiteturas (conjunto de teste, métricas de máscara)

Modelo	mAP50-95 (M)	mAP50 (M)	F1 médio	*Recall* médio	Inferência CPU (ms)	Parâmetros (M)
YOLOv8x-seg	0,4287	0,6325	0,7167	0,6564	783,5	71,72
YOLO11x-seg	0,4392	0,6408	0,7138	0,6591	593,4	62,01
YOLO26x-seg	0,4238	0,6201	0,6919	0,6414	644,6	62,73

Fonte: elaborada pelos autores (2026). Nota: (M) indica métrica de máscara.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

O YOLO11x-seg obteve o melhor desempenho na métrica principal de segmentação (mAP50-95 de máscara), superando o YOLOv8x-seg e o YOLO26x-seg. Embora o F1 médio das três arquiteturas seja muito próximo, a vantagem do YOLO11x-seg concentra-se nas classes patológicas, com o maior *recall* de tumor e o menor número de falsos negativos dessa classe, critério de maior relevância clínica, pois um falso negativo corresponde a uma lesão não detectada. Adicionalmente, o YOLO11x-seg é o mais compacto dos três e o mais rápido na inferência em CPU. O YOLO26x-seg, apesar de ser a geração mais recente, registrou o menor desempenho em todas as métricas de qualidade, com tempo de inferência intermediário entre os outros dois modelos. Diante desse conjunto de evidências, selecionou-se o YOLO11x-seg para a implantação na API, por combinar o melhor desempenho de segmentação, o menor número de parâmetros e o menor tempo de inferência entre os modelos avaliados.

4.2. Desempenho por Classe do Modelo Selecionado

A Tabela 4 detalha o desempenho do YOLO11x-seg desagregado por classe.

Tabela 4. Desempenho por classe (YOLO11x-seg, métricas de máscara)

Classe	Precisão	*Recall*	F1	mAP50-95 (M)	IoU médio (TP)	*Dice* (TP)
Rim	0,9480	0,8458	0,8940	0,6766	0,8888	0,9411
Tumor	0,7684	0,6978	0,7314	0,4391	0,8328	0,9088
Cisto	0,6188	0,4337	0,5100	0,2020	0,7518	0,8583

Fonte: elaborada pelos autores (2026). Nota: IoU médio e *Dice* foram calculados sobre as instâncias efetivamente pareadas como verdadeiro positivo.

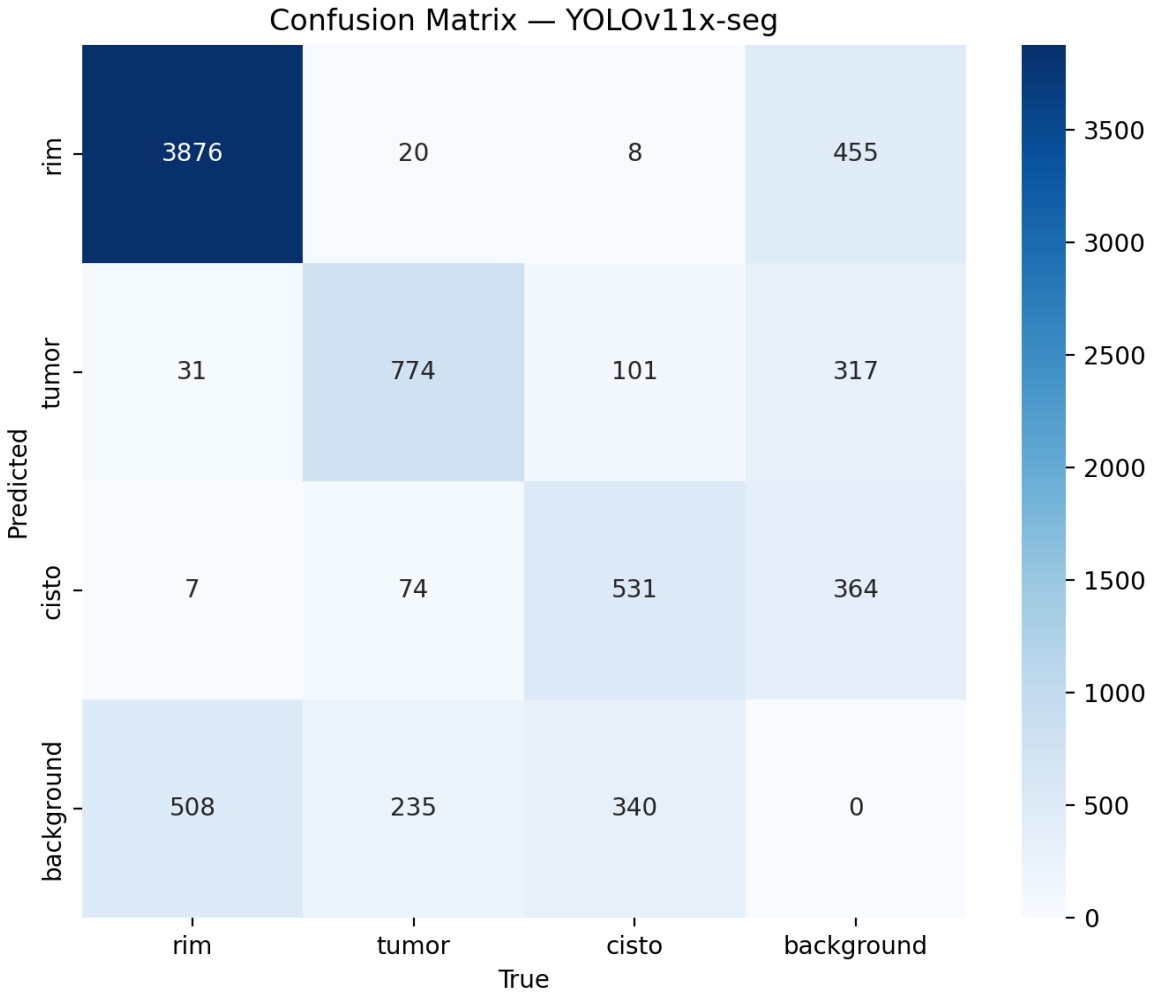
O rim apresentou o melhor desempenho (F1 = 0,894; *Dice* = 0,941), resultado esperado por ser a classe mais frequente e a de maior área média por fatia. O tumor obteve desempenho intermediário (F1 = 0,731; *Dice* = 0,909), com confusão em relação ao cisto explicada pela similaridade de aparência na TC, em que ambas as estruturas surgem como massas hipodensas. O cisto registrou os menores índices (F1 = 0,510), com *recall* de apenas 0,434 — menos da metade dos cistos presentes foi detectada —, refletindo tanto sua raridade no conjunto quanto seu reduzido tamanho médio nas imagens, ainda que, nos casos efetivamente detectados, a qualidade da delimitação permaneça elevada (*Dice* = 0,858), indicando que a principal limitação do sistema está no recall das classes minoritárias, e não na precisão de contorno quando a

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

detecção ocorre. Observa-se um gradiente de dificuldade decrescente com a frequência e o tamanho das estruturas (rim > tumor > cisto), padrão coerente com o desequilíbrio em nível de pixel descrito na metodologia.

4.3. Matriz de Confusão

Figura 1. Matriz de confusão do YOLO11x-seg



Fonte: elaborada pelos autores (2026). Nota: valores por IoU de caixa; tabelas usam IoU de máscara.

A análise da matriz de confusão do YOLO11x-seg evidenciou que o erro dominante é a não detecção de estruturas patológicas, o tipo de erro clinicamente mais grave. O cisto foi a classe mais afetada, com 340 ocorrências preditas como fundo (não detectadas), o que explica

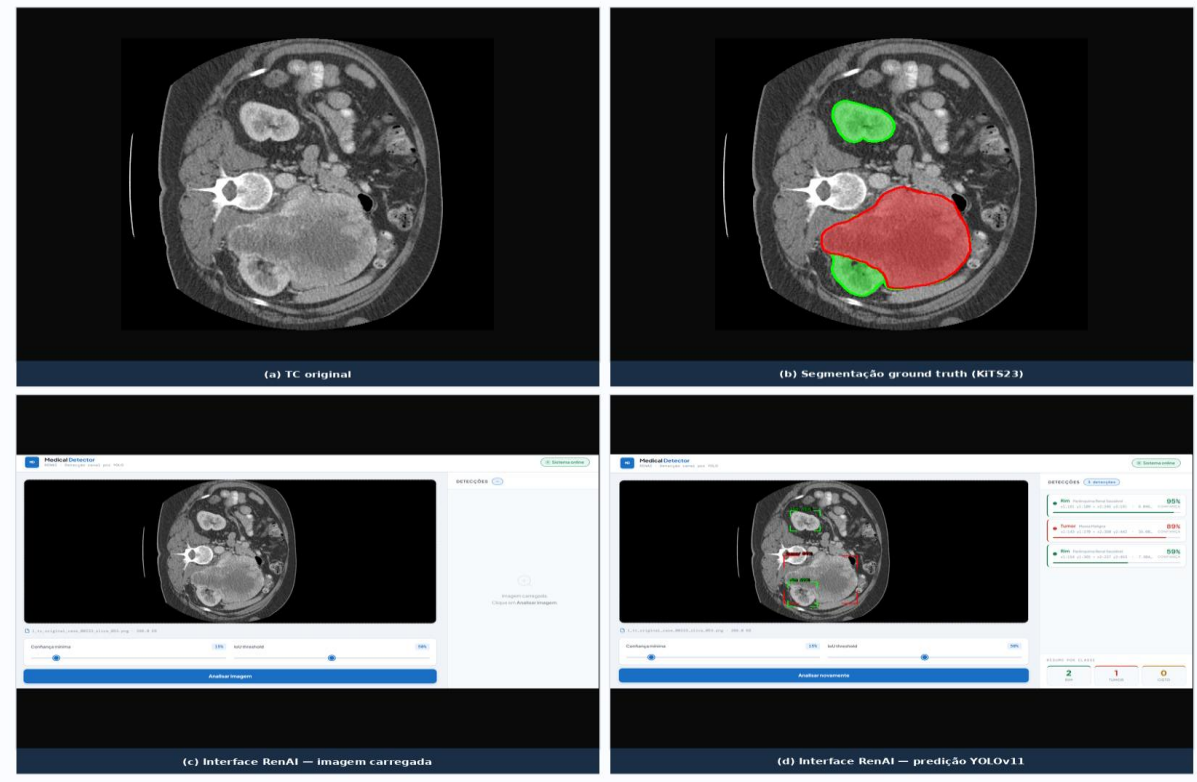
Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

seu baixo *recall*. O tumor apresentou 235 ocorrências preditas como fundo e o rim, 508, número que, no caso do rim, representa uma fração pequena diante de sua frequência total, mantendo o *recall* alto (0,846). Esse padrão de lesões perdidas concentra-se nas classes minoritárias e de menor área, confirmando o diagnóstico da metodologia e justificando as estratégias de mitigação adotadas. Ressalta-se que essa saída corresponde à detecção e segmentação de regiões de imagem, e não a diagnóstico clínico de malignidade, que depende de avaliação histopatológica não incluída neste estudo.

4.4. Aplicação Web

A Figura 2 ilustra o funcionamento da aplicação *web* RenAI, apresentando o fluxo completo desde a imagem de tomografia computadorizada original até a predição do modelo YOLO11x-seg na interface desenvolvida.

Figura 2. Corte axial de TC: (a) imagem original; (b) segmentação de referência; (c) interface RenAI com imagem carregada; (d) predição YOLO11x-seg com detecção de rim (95% e 59% de confiança) e tumor (89% de confiança).



Fonte: elaborada pelos autores (2026).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

A aplicação *web* operou com tempo médio de inferência de 5,8 segundos em CPU, dentro do limite estabelecido de 15 segundos, e o *endpoint* /health respondeu em menos de 50 ms. A execução da suíte de testes resultou na aprovação dos 47 casos distribuídos entre as camadas unitária, de integração e de ponta a ponta (*E2E*), incluindo a verificação de que entradas inválidas retornam HTTP 400 ou 422, nunca HTTP 500. A interface processou corretamente os múltiplos formatos de entrada com detecção automática por *magic bytes*, e a visualização por marcadores de canto, com reposicionamento automático de rótulos, evitou a sobreposição de texto em fatias com múltiplas detecções**.**

5. CONSIDERAÇÕES FINAIS

Este trabalho demonstrou a viabilidade do uso de modelos da família YOLO para detecção e segmentação de estruturas renais em tomografias computadorizadas, com desenvolvimento completo desde a preparação dos dados até a implantação em uma aplicação web funcional. O sistema RenAI alcançou precisão de 77% e *recall* de 70% na tarefa de detecção de tumor, com desempenho superior para rim em comparação com tumor e cisto.

As principais contribuições técnicas incluem: um fluxo de processamento para balanceamento orientado por grupos clínicos com garantia de separação por paciente; uma estratégia de ajuste progressivo de hiperparâmetros orientado pelas métricas de desempenho; e uma aplicação *web* com suporte a múltiplos formatos de imagem médica, testes automatizados em três camadas e interface acessível segundo diretrizes WCAG AA.

O desbalanceamento severo do conjunto de dados permanece como o principal limitante do desempenho, especialmente para a classe cisto, com a menor fração de fatias no conjunto de dados original. A confusão entre tumor e cisto reflete uma limitação inerente ao domínio: ambas as estruturas apresentam aparência similar na TC com contraste, sendo frequentemente distinguidas clinicamente por características que vão além da morfologia 2D, como o padrão de realce pelo contraste ao longo do tempo.

Como limitações do estudo, destacam-se tanto aspectos metodológicos quanto etapas ainda não contempladas no processo de validação clínica. No plano metodológico, o trabalho apoiou-se em uma única base pública, o que, apesar de assegurar comparabilidade com a literatura, restringe a generalização dos resultados a diferentes protocolos de aquisição, equipamentos e populações de pacientes. A representação bidimensional das imagens constitui outra limitação relevante, uma vez que a segmentação de tumores renais possui forte componente volumétrico, não explorado pelos modelos treinados exclusivamente sobre fatias

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

axiais isoladas. Some-se a isso a possível perda de informação decorrente da conversão radiométrica de 16 para 8 bits, operação necessária à compatibilidade com o formato YOLO, mas que pode atenuar sutilezas de contraste relevantes para a distinção entre tecidos. No plano da validação clínica, o sistema ainda carece de etapas indispensáveis antes de qualquer consideração sobre uso assistencial: não foi conduzido estudo prospectivo, tampouco validação multicêntrica capaz de avaliar o desempenho do modelo diante da variabilidade de equipamentos e protocolos entre diferentes serviços de saúde; não houve comparação direta com o desempenho de especialistas humanos; e não foram realizadas avaliações de calibração do modelo nem de seu impacto clínico efetivo. Tais lacunas reforçam que os resultados obtidos demonstram viabilidade computacional do sistema RenAI, e não prontidão para aplicação clínica imediata.

Como perspectivas futuras, destacam-se a utilização de arquiteturas especializadas para segmentação médica que exploram volumes 3D em vez de fatias 2D isoladas; a adoção de técnicas de aprendizado com poucos exemplos para classes raras como cisto; a integração de informações de múltiplas fases de contraste para melhor distinção entre tumor e cisto; e a validação em bases de dados clínicas independentes, incluindo calibração do modelo e comparação com especialistas. A aplicação *web* desenvolvida constitui uma base sólida para futuras expansões, como suporte a volumes DICOM completos e integração com sistemas de laudos clínicos.

REFERÊNCIAS

ALZUBAIDI, L. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. **Journal of Big Data**, v. 8, art. 53, 2021. DOI: 10.1186/s40537-021-00444-8. Disponível em: https://pubmed.ncbi.nlm.nih.gov/33816053/. Acesso em: 20 set. 2026.

BISONG, E. Google Colaboratory. In: BISONG, E. **Building machine learning and deep learning models on Google Cloud Platform**. Berkeley: Apress, 2019. DOI: 10.1007/978-1-4842-4470-8_7. Disponível em: https://link.springer.com/chapter/10.1007/978-1-4842-4470-8_7. Acesso em: 20 set. 2026.

CHEN, X. et al. Recent advances and clinical applications of deep learning in medical image analysis. **Medical Image Analysis**, v. 79, art. 102444, 2022. DOI: 10.1016/j.media.2022.102444. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC9156578/. Acesso em: 20 set. 2026.

JIANG, F. et al. Artificial intelligence in healthcare: past, present and future. **Stroke and Vascular Neurology**, v. 2, n. 4, p. 230-243, 2017. DOI: 10.1136/svn-2017-000101. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC5829945/. Acesso em: 20 set. 2026.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

KHAN, A. et al. A survey of the recent architectures of deep convolutional neural networks. **Artificial Intelligence Review**, v. 53, n. 8, p. 5455-5516, 2020. DOI: 10.1007/s10462-020-09825-6. Disponível em: https://link.springer.com/article/10.1007/s10462-020-09825-6. Acesso em: 20 set. 2026.

KITS23 CHALLENGE. **KiTS23**: the 2023 kidney and kidney tumor segmentation challenge. 2023. Disponível em: https://kits-challenge.org/kits23/. Acesso em: 20 set. 2026.

MINAEE, S. et al. Image segmentation using deep learning: a survey. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 44, n. 7, p. 3523-3542, 2022. DOI: 10.1109/TPAMI.2021.3059968. Disponível em: https://pubmed.ncbi.nlm.nih.gov/33596172/. Acesso em: 20 set. 2026.

MYRONENKO, A. et al. Automated 3D segmentation of kidneys and tumors in MICCAI KiTS 2023 challenge. In: INTERNATIONAL CHALLENGE ON KIDNEY AND KIDNEY TUMOR SEGMENTATION (KiTS 2023), 2023. **Kidney and Kidney Tumor Segmentation**. Cham: Springer, 2024. p. 1-7. (Lecture Notes in Computer Science, v. 14540). DOI: 10.1007/978-3-031-54806-2_1. Disponível em: https://arxiv.org/abs/2310.04110. Acesso em: 20 set. 2026.

PADALA, S. A. et al. Epidemiology of renal cell carcinoma. **World Journal of Oncology**, v. 11, n. 3, p. 79-87, 2020. DOI: 10.14740/wjon1279. Disponível em: https://pubmed.ncbi.nlm.nih.gov/32494314/. Acesso em: 20 set. 2026.

PADILLA, R.; NETTO, S. L.; SILVA, E. A. B. da. A survey on performance metrics for object-detection algorithms. In: INTERNATIONAL CONFERENCE ON SYSTEMS, SIGNALS AND IMAGE PROCESSING (IWSSIP), 2020, Niterói. **Proceedings** [...]. Niterói: IEEE, 2020. p. 237-242. DOI: 10.1109/IWSSIP48289.2020.9145130. Disponível em: https://dblp.dagstuhl.de/rec/conf/iwssip/PadillaNS20.html. Acesso em: 20 set. 2026.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: convolutional networks for biomedical image segmentation. In: MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION (MICCAI 2015), 2015, Munich. **Proceedings** [...]. Cham: Springer, 2015. p. 234-241. (Lecture Notes in Computer Science, v. 9351). DOI: 10.1007/978-3-319-24574-4_28. Disponível em: https://arxiv.org/abs/1505.04597. Acesso em: 20 set. 2026.

ROSSI, S. H. et al. Epidemiology and screening for renal cancer. **World Journal of Urology**, v. 36, n. 9, p. 1341-1353, 2018. DOI: 10.1007/s00345-018-2286-7. Disponível em: https://link.springer.com/article/10.1007/s00345-018-2286-7. Acesso em: 20 set. 2026.

SAPKOTA, R. et al. YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series. **Artificial Intelligence Review**, v. 58, art. 274, 2025. DOI: 10.1007/s10462-025-11253-3. Disponível em: https://link.springer.com/article/10.1007/s10462-025-11253-3. Acesso em: 20 set. 2026.

SILVA, C. E. da et al. Epidemiological profile of kidney cancer in Brazil: a multiregional ecological study. **Brazilian Journal of Nephrology**, v. 47, n. 2, e20240180, 2025. DOI: 10.1590/2175-8239-JBN-2024-0180en. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC11913370/. Acesso em: 20 set. 2026.

TERVEN, J.; CORDOVA-ESPARZA, D.-M.; ROMERO-GONZÁLEZ, J.-A. A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Machine Learning and Knowledge Extraction, v. 5, n. 4, p. 1680-1716, 2023. DOI: 10.3390/make5040083. Disponível em: https://www.mdpi.com/2504-4990/5/4/83. Acesso em: 20 set. 2026.

THARWAT, A. Classification assessment methods. **Applied Computing and Informatics**, v. 17, n. 1, p. 168-192, 2021. DOI: 10.1016/j.aci.2018.08.003. Disponível em: https://www.emerald.com/insight/content/doi/10.1016/j.aci.2018.08.003/full/html. Acesso em: 20 set. 2026.

ULTRALYTICS. **Ultralytics YOLO documentation**. 2023. Disponível em: https://docs.ultralytics.com/. Acesso em: 20 set. 2026.

ULTRALYTICS. **YOLO11 documentation**. 2024. Disponível em: https://docs.ultralytics.com/models/yolo11/. Acesso em: 20 set. 2026.

ULTRALYTICS. **YOLO26 documentation**. 2026. Disponível em: https://docs.ultralytics.com/models/yolo26. Acesso em: 20 set. 2026.

VALLECILLA VALENCIA, K.; GARCIA-PERDOMO, H. A. Overdiagnosis in renal cancer: the hidden consequence of modern imaging. **International Urology and Nephrology**, 2026. DOI: 10.1007/s11255-026-05173-6. Disponível em: https://link.springer.com/article/10.1007/s11255-026-05173-6. Acesso em: 20 set. 2026.

XIONG, Y. et al. Artificial intelligence links CT images to pathologic features and survival outcomes of renal masses. **Nature Communications**, v. 16, art. 1425, 2025. DOI: 10.1038/s41467-025-56784-z. Disponível em: https://www.nature.com/articles/s41467-025-56784-z. Acesso em: 20 set. 2026.

YU, K.-H.; BEAM, A. L.; KOHANE, I. S. Artificial intelligence in healthcare. **Nature Biomedical Engineering**, v. 2, n. 10, p. 719-731, 2018. DOI: 10.1038/s41551-018-0305-z. Disponível em: https://doi.org/10.1038/s41551-018-0305-z. Acesso em: 20 set. 2026.

ZHAO, X. et al. A review of convolutional neural networks in computer vision. **Artificial Intelligence Review**, v. 57, art. 99, 2024. DOI: 10.1007/s10462-024-10721-6. Disponível em: https://link.springer.com/article/10.1007/s10462-024-10721-6. Acesso em: 20 set. 2026.

ZHOU, S. K. et al. A review of deep learning in medical imaging: imaging traits, technology trends, case studies with progress highlights, and future promises. **Proceedings of the IEEE**, v. 109, n. 5, p. 820-838, 2021. DOI: 10.1109/JPROC.2021.3054390. Disponível em: https://ieeexplore.ieee.org/abstract/document/9363915. Acesso em: 20 set. 2026.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.