



UNIVERSIDADE FEDERAL DO MARANHÃO

Curso de Ciência da Computação

Giovanna Pavani Esteve

**Analisando Redes Neurais HRNet para  
Segmentação de Rins e Tumores em Imagens de  
Tomografia Computadorizada**

São Luís - MA

2023

Giovanna Pavani Esteve

**Analisando Redes Neurais HRNet para Segmentação de  
Rins e Tumores em Imagens de Tomografia  
Computadorizada**

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Geraldo Braz Júnior

São Luís - MA

2023

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).  
Diretoria Integrada de Bibliotecas/UFMA

Pavani Esteve, Giovanna.

Analisando redes neurais HRNet para segmentação de rins e tumores em imagens de tomografia computadorizada / Giovanna Pavani Esteve. - 2023.

33 f.

Orientador(a): Geraldo Braz Júnior.

Monografia (Graduação) - Curso de Ciência da Computação, Universidade Federal do Maranhão, São Luís, 2023.

1. Câncer de Rim. 2. Deep learning. 3. HRNet. 4. Segmentação. I. Braz Júnior, Geraldo. II. Título.

Giovanna Pavani Esteve

# **Analisando Redes Neurais HRNet para Segmentação de Rins e Tumores em Imagens de Tomografia Computadorizada**

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação.

Trabalho aprovado em São Luís - MA, 13 de julho de 2023:

---

**Prof. Dr. Geraldo Braz Júnior**  
Orientador

---

**Profa. Dra. Simara Vieira da Rocha**  
Examinadora

---

**Profa. Dra. Vandecia Rejane Monteiro  
Fernandes**  
Examinadora

São Luís - MA  
2023

# Agradecimentos

Aos meus pais, pela educação que me proporcionaram e pelo apoio e incentivo, sem os quais não estaria graduando.

Aos professores da UFMA não só pelo conhecimento passado, mas pela dedicação em estar sempre motivando os alunos. Em especial ao Prof. Geraldo Braz Júnior pela confiança e pelas oportunidades que me proporcionou ao longo da graduação, sua orientação foi muito importante para meu desenvolvimento durante o curso.

A meus amigos da UFMA, pelas experiências que compartilhamos, desde estudar para provas e produzir trabalhos juntos até as experiências compartilhadas fora da universidade. Mayrla Jansen, José Ribamar Durand, Alana Araujo e Antônio Moreira.

*O pensamento não passa de um relâmpago entre duas longas noites, mas esse relâmpago é tudo.*

(Henri Poincaré)

# Resumo

O câncer de rim está entre os 10 cânceres mais comuns em homens e mulheres e, segundo a GLOBOCAN (*Global Cancer Observatory*), apresenta uma taxa de mortalidade de 54% no Brasil. Um diagnóstico precoce é primordial para um melhor prognóstico dessa doença e os exames de imagem são fundamentais no diagnóstico deste tipo de tumor. A interpretação dessas imagens é um processo manual que depende na observação e experiência do profissional, portanto, técnicas computacionais de automatização desse processo se apresentam como uma alternativa para contribuir para diagnóstico precoce do câncer de rim. Nos últimos anos a rede neural convolucional HRNet vem recebendo atenção crescente na segmentação semântica devido ao seu alto desempenho. Portanto, o objetivo deste trabalho consiste em avaliar o desempenho de redes neurais profundas HRNet no contexto de segmentação de imagens de tomografia computadorizada de pacientes com tumores renais. Serão utilizadas técnicas de otimização para avaliar modelos de redes neurais convolucionais e o dataset consiste no conjunto de imagens do repositório KiTS19. O modelo proposto conseguiu alcançar Dice score de 91,4% e 70,3% para a segmentação do rim e tumor, respectivamente.

**Palavras-chave:** Câncer de Rim, Aprendizagem Profunda, Tomografia Computadorizada, Segmentação, HRNet.

# Abstract

Kidney cancer is among the 10 most common cancers in men and women, and according to with GLOBOCAN (Global Cancer Observatory), presents a mortality rate of 54% in Brazil. An early diagnosis is essential for a better prognosis of this disease and imaging tests are fundamental in the diagnosis of this type of tumor. The interpretation of these images is a manual process that depends on observation and experience of the professional, therefore, computational techniques for automating this process are presented as an alternative to contribute to early diagnosis of kidney cancer. In recent years, the convolutional neural network HRNet has been receiving growing attention on semantic segmentation due to its high performance. Therefore, The objective of this work is to evaluate the performance of deep neural networks HRNet in the context of segmentation of CT images of patients with renal tumors. Optimization techniques will be used to evaluate models of convolutional neural networks and the dataset consists of the set of images from the KiTS19 repository. The proposed model was able to achieve Dice score of 91.4% and 70.3% for kidney and tumor segmentation, respectively.

**Keywords:** Kidney Cancer, Deep Learning, Computed Tomography, Segmentation, HRNet.



# Lista de ilustrações

Figura 1 – Exemplo de imagens de pulmão criadas a partir de exames de (a) Raio-X e (b) TC . . . . .	16
Figura 2 – Modelo de um neurônio . . . . .	18
Figura 3 – Modelo de rede neural artificial. . . . .	18
Figura 4 – Modelo de rede neural convolucional. . . . .	20
Figura 5 – Exemplo de imagem de uma corrida de motocicleta (a) e sua segmentação (b). . . . .	20
Figura 6 – HRNetV1 com 4 canais de resolução (representa n etapas, sendo $n = 4$ )	21
Figura 7 – HRNetV2 - Concatenando as representações (upsampled) das resoluções	21
Figura 8 – Metodologia . . . . .	22
Figura 9 – Exemplo de imagem da base KiTS19 com presença de tumor renal e sua máscara . . . . .	23
Figura 10 – Arquitetura geral da rede HRNet . . . . .	23
Figura 11 – Arquitetura HRNet se dividindo em 4 resoluções. . . . .	24
Figura 12 – Exemplos de segmentação dos tumores renais de 3 pacientes com o modelo HRNet construído. Da esquerda para a direita, fatia de entrada, mascara de segmentação e segmentação efetuada pela rede. . . . .	29

# Lista de tabelas

Tabela 1 – Hiperparâmetros de busca . . . . .	26
Tabela 2 – Resultados da HRNet com dataset completo . . . . .	26
Tabela 3 – Resultados do modelo com melhores parâmetros treinado em 5-folds da base de dados. . . . .	27
Tabela 4 – Resultados da HRNet com dataset de tumores em comparação com a U-Net . . . . .	28

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>1.1</b>	<b>Trabalhos Relacionados</b>	<b>12</b>
<b>1.2</b>	<b>Objetivos</b>	<b>13</b>
1.2.1	Objetivos Específicos	13
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>15</b>
<b>2.1</b>	<b>Câncer Renal</b>	<b>15</b>
<b>2.2</b>	<b>Tomografia computadorizada (TC)</b>	<b>15</b>
<b>2.3</b>	<b>Deep Learning (DL)</b>	<b>16</b>
2.3.1	Redes Neurais Convolucionais	17
2.3.2	HRNET	19
<b>2.4</b>	<b>Considerações Finais</b>	<b>21</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>22</b>
<b>3.1</b>	<b>Aquisição da Base</b>	<b>22</b>
<b>3.2</b>	<b>Construção da Rede Neural</b>	<b>22</b>
<b>3.3</b>	<b>Treinamento</b>	<b>24</b>
<b>3.4</b>	<b>Avaliação</b>	<b>25</b>
<b>4</b>	<b>RESULTADOS</b>	<b>26</b>
<b>5</b>	<b>CONCLUSÃO</b>	<b>30</b>
	<b>REFERÊNCIAS</b>	<b>32</b>

# 1 Introdução

O avanço de tecnologias de software e de hardware como, unidades de processamento centrais (CPUs) e unidades de processamento gráfico (GPUs), além do crescimento da quantidade de dados disponível ao público, foram fundamentais para o avanço do desenvolvimento de algoritmos de aprendizagem de alta desempenho em diversas aplicações (SHEN; WU; SUK., 2017). Destacando experimentos na área de análise de imagens médicas, como segmentação de imagens de órgãos e tumores.

Câncer é um termo que abrange mais de 100 diferentes tipos de doenças malignas que têm em comum o crescimento desordenado de células, podendo invadir tecidos adjacentes ou órgãos a distância (INCA, 2021a). Assim, existem muitos tipos de tumores que podem se desenvolver em qualquer parte do corpo, estes são nomeados conforme o local onde se originaram. Diferentes tipos de tumores podem se comportar de formas distintas, por esse motivo, dois tipos de câncer distintos são doenças muito diferentes, o local de origem determina seu comportamento como sintomas, consequências e a velocidade de desenvolvimento da doença. O câncer renal é um tipo de tumor que se inicia no rim.

De acordo com o American Cancer Society (2022), o câncer de rim está entre os 10 cânceres mais comuns em homens e mulheres. Em geral, o risco de se desenvolver câncer de rim ao longo da vida é de 1 em 46 em homens, enquanto o risco para mulheres é de 1 em 82. No Brasil, essa doença afeta principalmente idosos, com maior incidência na faixa de idade compreendida entre 40 e 60 anos (BARBOSA; LIMA; GUIMARÃES, 2012).

A taxa de novos casos de câncer de rim aumentam desde a década de 1990, embora pareça ter se estabilizado nos últimos anos. Parte desse aumento se deu, provavelmente, devido ao uso de exames de imagem mais recentes para o diagnóstico, como a tomografia computadorizada (TC), que detecta alguns tipos de câncer que antes não seriam diagnosticados (AMERICAN CANCER SOCIETY, 2022).

Dentre os tipos de tumores renais, o carcinoma de células renais (CCR) maligno é responsável por mais de 90% de casos. Como o câncer de rim não apresenta sintomas nos estágios iniciais, muitas vezes é diagnosticado quando se encontra em uma fase mais avançada (SHON HO SUN, 2020). No Brasil, a incidência estimada é de 7 a 10 casos para cada 100 mil habitantes. O prognóstico desse tipo de câncer depende, dentre outros fatores, da idade do paciente e da rapidez em se diagnosticar e tratar a doença (INCA, 2019).

Dessa forma, um diagnóstico precoce é primordial para um melhor prognóstico. Uma anamnese (registro de dados obtidos em uma conversa inicial com o paciente) e exame físico detalhados são essenciais o diagnóstico inicial. Os exames de imagem são fundamentais no diagnóstico deste tipo de tumor. Para uma melhor avaliação do tumor,

é recomendado uma tomografia computadorizada de abdômen e pelve e/ou ressonância magnética de abdômen. Nesses exames é evidenciada a massa de origem renal, sendo de extrema importância a avaliação de ambos os rins. Os exames laboratoriais são úteis para avaliar a condição clínica do paciente, presença de anemia e alteração da função renal (INCA, 2021b).

O uso de imagens médicas como tomografia computadorizada (TC), ressonância magnética (RM), tomografia por emissão de pósitrons (PET), ultrassom, raio-x, entre outros, é fundamental para a detecção, diagnóstico e tratamento de doenças (BRODY, 2013).

A interpretação dessas imagens é um processo manual que depende na observação e experiência do profissional. No entanto, devido a grandes variações na patologia e a possível fadiga de especialistas, pesquisadores e médicos começaram a se beneficiar da automatização desse processo (SHEN; WU; SUK., 2017).

A análise computadorizada para uma melhor interpretação de imagens tem sido um problema altamente estudado na área de imagens médicas. No que diz respeito à compreensão de imagens, os avanços recentes em aprendizado de máquina, especialmente na área de Deep Learning (DL), estão progredindo para identificar, classificar e quantificar padrões em imagens médicas.

Deep Learning (DL) é uma área de aprendizado de máquina onde modelos computacionais compostos de múltiplas camadas de processamento aprendem representações de dados com múltiplos níveis de abstração. Esses métodos têm melhorado dramaticamente os resultados em reconhecimento de fala, reconhecimento de objetos visuais, detecção de objetos e muitos outros domínios (LECUN; BENGIO; HINTON, 2015).

Assim, considerando as taxas de novos casos de câncer renal, assim como a importância do diagnóstico precoce para um melhor prognóstico desta doença e o avanço da tecnologia de aprendizagem de máquina e aprendizagem profunda, foram surgindo trabalhos que visam melhorar a capacidade da assistência computadorizada em ajudar no processo de detecção do câncer de rim.

## 1.1 Trabalhos Relacionados

Esta seção apresenta trabalhos na literatura que objetivam a segmentação de rins em imagens de TC com estratégias que utilizam Deep Learning, assim como trabalhos que utilização de redes HRNets em imagens médicas.

Em (HELLER NICHOLAS, 2021) é apresentado trabalhos de segmentação semântica de rins e tumores de rins em imagem de TC do dataset KITS19 (HELLER

NICHOLAS, 2019) utilizando métodos baseados em Deep Learning e processamento de imagens. Entre eles, a abordagem que apresentou melhores resultados fez uso extenso de data augmentation com ajustes que incluem, escalação, rotações, brilho, contraste, uma transformação gama e a introdução de ruído gaussiano. A arquitetura proposta formado de três redes U-Net 3D foi testada com utilizando k-fold obtendo Dice de 0,912.

Yu Qian (2019) propõem um modelo de CNN de segmentação treinável em cascata denominado Crossbar-Net. Essa arquitetura foi utilizada para a segmentação de imagens de TC de rins com tumores, encontrando um coeficiente de similaridade de Dice (DSC) de 0,913.

Cruz Luana Batista (2020) utilizam uma abordagem para segmentação de imagens de TC de rins e tumores renais onde, primeiro, realiza a redução de escopo usando AlexNet seguido de uma segmentação inicial utilizando U-Net 2D e, por fim, a redução de falso positivo utilizando processamento de imagem para manter os maiores elementos da imagem (os rins). Essa técnica apresentou como resultado um coeficiente de Dice de 93,03%.

Quanto a utilização de redes HRNet em imagens médicas, em (AHMED, 2021) é proposto a detecção de imagens de raio-x do pulmão para a identificação de COVID-19. Nesse trabalho é feita uma segmentação inicial do pulmão com a rede U-Net (RONNEBERGER OLAF; BROX., 2015), que passa então pela HRNet (WANG JINGDONG, 2020) que consegue evitar características pequenas nas imagens. Antes de entrar na HRNet, são utilizadas técnicas de Data Augmentation nas imagens extraídas da U-Net. A classificação final obteve acurácia e sensibilidade de 99,26% e 98,82% respectivamente.

Esses são exemplos de trabalhos desenvolvidos para a segmentação de rins e tumores renais em imagens de TC, assim como o uso da rede neural HRNet no problema de diagnóstico de imagens médicas. Sendo assim, a metodologia proposta busca analisar a aplicabilidade dessa arquitetura no problema de segmentação de imagens de câncer renal.

## 1.2 Objetivos

Este trabalho tem como objetivo analisar o desempenho da rede neural convolucional High Resolution Network (HRNet) na segmentação semântica de imagens de TC para delimitar regiões dos rins e tumor na imagem.

### 1.2.1 Objetivos Específicos

Destaca-se como objetivos específicos deste trabalho:

- Adaptar e aplicar a rede neural convolucional HRNet para a segmentação de imagens de rins de TC, delimitando as regiões dos rins e do tumor na imagem;
- Análise do desempenho da rede neural convolucional High Resolution Network (HRNet) na segmentação semântica de imagens de TC para delimitar regiões dos rins e tumor na imagem.

## 2 Fundamentação Teórica

Este capítulo apresenta os conceitos explorados para o desenvolvimento do estudo bem como as técnicas computacionais utilizadas.

### 2.1 Câncer Renal

As células cancerosas se diferem das normais de algumas formas. Elas perdem o controle sobre a divisão celular, multiplicando-se de maneira indefinida e rapidamente, além de não morrerem como as células normais e, ao invés disso, acumulam-se formando tumor (PRADO, 2014).

Os rins são um par de órgãos com o formato de feijão, cada um do tamanho de um punho. Eles estão presos à parede posterior superior do abdômen e protegidos pela caixa torácica inferior (AMERICAN CANCER SOCIETY, 2020a). A função dos rins é, entre outras, filtrar o sangue, eliminando substâncias que são nocivas ao organismo, como a amônia, a ureia e o ácido úrico. Eles também atuam secretando substâncias importantes para nossa saúde (SBN, 2022).

Quando se origina nos rins, um tumor por ser localizado em diferentes regiões. Uma região comum de ser encontrado é no tubo responsável por filtrar as impurezas do sangue, nesses casos o câncer é chamado de Carcinoma de Células Renais (CCR) e representa a maioria dos casos de câncer renal. Quando é localizado nas vias excretoras do rim, o tumor é classificado como um Carcinoma de Células Transicionais (CCT) (AMERICAN CANCER SOCIETY, 2020b). De cada 100 casos de câncer de rim, em torno de 50% são de CCT (INCA, 2019).

### 2.2 Tomografia computadorizada (TC)

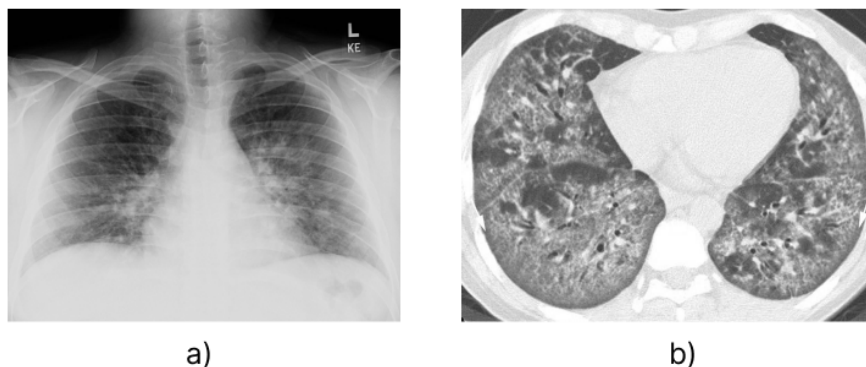
Alguns dos exames de imagem utilizados para identificar um tumor renal são, o Raio-X, a ultrassonografia, a ressonância magnética ou a tomografia computadorizada. Ao contrário da maioria dos outros tipos de câncer, os médicos geralmente podem diagnosticar o câncer de rim com base em exames de imagem sem fazer uma biópsia (AMERICAN CANCER SOCIETY, 2020a).

A tomografia computadorizada é um exame de imagem não invasivo, ele utiliza da tecnologia de raio-x com computadores programados para produzir imagens de alta qualidade dos órgãos internos. Esse exame combina a imagem de múltiplos raio-x e, por



isso, é mais detalhado do que uma radiografia comum. É utilizado para o estudo de diversas partes do corpo e pode diagnosticar diversos tumores (JÚNIOR; YAMASHITA., 2001).

Figura 1 – Exemplo de imagens de pulmão criadas a partir de exames de (a) Raio-X e (b) TC



Fonte: (COHEN et al., 2020) e (YANG et al., 2020)

A tomografia computadorizada usa raios-x para fazer imagens transversais detalhadas dos órgãos internos. Ele pode fornecer informações precisas sobre o tamanho, forma e localização de um tumor. Também serve para verificar se o câncer se espalhou para os linfonodos próximos ou para órgãos e tecidos fora do rim (AMERICAN CANCER SOCIETY, 2020a).

Há 20 anos, os casos de carcinoma de células renais não representavam mais de 5% do total de casos de tumores renais encontrados. Atualmente, graças ao uso rotineiro de técnicas radiológicas não invasivas, como a ultrassonografia e a tomografia computadorizada (TC), houve um aumento acima de 30% no diagnóstico precoce (DALL’OGLIO MARCOS, 2004).

A TC realiza movimentos circulares ao redor do paciente enquanto a fonte de raio-x é acionada. Há também uma série de detectores responsáveis por transformar a radiação em um sinal elétrico, então convertido em uma imagem digital (JÚNIOR; YAMASHITA., 2001).

## 2.3 Deep Learning (DL)

Modelos usados em Aprendizagem Profunda (*Deep Learning*) são chamados de Redes Neurais Artificiais (RNAs). Essas redes são sistemas computacionais baseadas em uma coleção de unidades conectadas chamadas de neurônios, cada conexão entre neurônios transmite sinais de um para o outro. Normalmente neurônios são organizados em camadas, diferentes camadas podem realizar diferentes transformações na informação transmitida da primeira camada, chamada de *input layer*, à última camada, chamada de *output layer* (RAUBER, 2016).

Aprendizagem profunda é um subcampo da aprendizagem de máquina usado para interpretar níveis de maior complexidade. Seus algoritmos dependem de unidades de processamento não lineares estruturados em múltiplas camadas em cascata, geralmente compostas por redes neurais artificiais cujos parâmetros são estimados através do treinamento e utilizados para extração de característica. Ou seja, enquanto a aprendizagem de máquina normalmente trabalha de forma sequencial, a aprendizagem profunda trabalha em camadas encadeadas de forma hierárquica, possibilitando análises ainda mais complexas e profundas (ALZUBAIDI LAITH, 2021).

A estrutura mais simples de uma rede neural artificial é o perceptron (*The perceptron: a probabilistic model for information storage and organization in the brain*), formado de apenas um neurônio. Ele consiste em um modelo matemático que recebe uma entrada, uma função de ativação e produz uma única saída. Um perceptron com várias camadas é chamado de Perceptron de Camadas Múltiplas (MLP). Os perceptrons são limitados uma vez que só conseguem resolver problemas linearmente separáveis (SHEN DINGGANG; SUK, 2017).

Cada nó individual de uma MLP é modelo de regressão linear, composto por dados de entrada, pesos, um bias e uma saída, representado pelas equações 2.1 e 2.2.

$$\sum wixi + b = w1x1 + w2x2 + w3x3 + bias \quad (2.1)$$

$$saída = f(x) = 1if \sum w1x1 + b \geq 0; 0if \sum w1x1 + b < 0 \quad (2.2)$$

Para determinar como a máquina deve alterar seus parâmetros internos é utilizado o algoritmo de *backpropagation* usado para calcular a representação em cada camada a partir da representação da camada anterior (LECUN; BENGIO; HINTON, 2015).

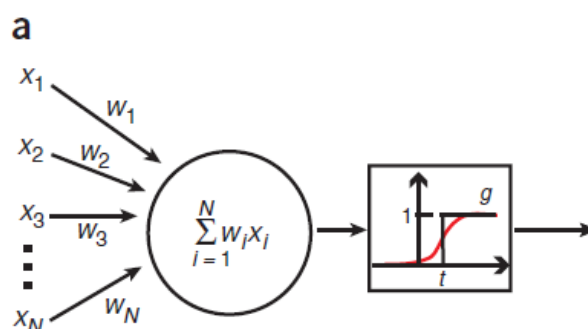
Enquanto aprendizagem de máquina constitui na prática de utilizar algoritmos que analisam dados, aprendem a partir deles e fazem uma predição ou determinação sobre novos dados, o *Deep Learning* utiliza de algoritmos inspirados pela estrutura e função dos neurônios do cérebro para realizar essa tarefa. Quando um modelo de *Deep Learning* aprende e efetua inferências sobre dados que foram previamente rotulados, ele representa um modelo de aprendizagem supervisionado, nesse contexto, aprendizagem não supervisionada acontece quando o modelo efetua inferências sobre dados não rotulados.

### 2.3.1 Redes Neurais Convolucionais

O aprendizado profundo é baseado no conceito de redes neurais artificiais. Essas redes são inspiradas pelos primeiros modelos de processamento do cérebro e podem ser criadas aplicando algoritmo que imitam os processos de neurônios reais. Um neurônio recebe entradas de várias outras unidades ou fontes externas, pesa cada entrada e as soma.

Se a entrada total estiver acima de um limite, a saída da unidade é um; caso contrário, é zero. Portanto, a saída muda de 0 para 1 quando a total ponderada das entradas é igual ao limite. Os pontos no espaço de entrada que satisfazem esta condição definem o chamado hiperplano. Isso significa que um problema de classificação pode ser resolvido por uma unidade limite se as duas classes puderem ser separadas por um hiperplano. Tais problemas são considerados linearmente separáveis. (KROGH, 2008) Sua função é ilustrada na Figura 2.

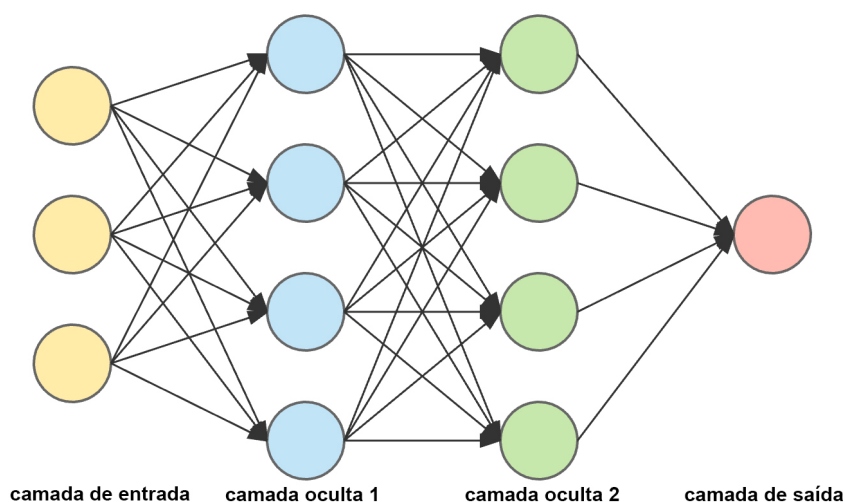
Figura 2 – Modelo de um neurônio



Fonte: Krogh (2008)

Muitos problemas de classificação não são linearmente separáveis. Podemos separar as classes em tais problemas não lineares introduzindo mais hiperplanos; ou seja, introduzindo mais de uma unidade de limiar. Isso geralmente é feito adicionando uma camada extra (oculta) de unidades de limite, cada uma das quais faz uma classificação parcial da entrada e envia sua saída para uma camada final, que reúne as classificações parciais na classificação final.

Figura 3 – Modelo de rede neural artificial.



Fonte: redes-neurais (2020).

A Rede Neural Convolutacional (*Convolutional Neural Network* — CNN) é uma das redes neurais de aprendizagem profunda mais utilizadas. São RNAs normalmente utilizadas para analisar imagens, elas se diferenciam pelas camadas convolucionais e camadas pooling. O termo convolutacional se refere ao processo de filtragem que ocorre nesse tipo de rede (O'SHEA; NASH, 2015).

As redes neurais normais não são bem dimensionadas para receber imagens como dado de entrada. Em redes neurais comuns, cada nó de uma camada é conectado a todos os nós da camada anterior ou seguinte. Assim, ao receber esse tipo de dado, a imagem tridimensional teria suas dimensões reduzidas para um vetor de valores. Já a CNN organiza os neurônios em um dado tridimensional com as dimensões de profundidade, altura e largura, isso permite a codificação de recursos específicos da imagem na arquitetura, tornando a rede mais adequada para tarefas focadas na imagem enquanto reduz ainda mais os parâmetros necessário para configurar o modelo (ALBAWI SAAD; AL-ZAWI, 2017).

As CNNs são compostas por três tipos de camadas. São elas camadas convolucionais, camadas de *pool* e camadas totalmente conectadas. Cada camada convolutacional possui uma quantidade de filtros (*kernels*) de tamanho definido, o processo de filtragem acontece ao se colocar um filtro sobre uma matriz de pixels da imagem criando um mapa de características. Cada filtro produz um mapa de ativação que reconhece certas formas e padrões. Todos esses mapas de ativação são empilhados ao longo da dimensão de profundidade e resultam na matriz de saída (ALBAWI SAAD; AL-ZAWI, 2017).

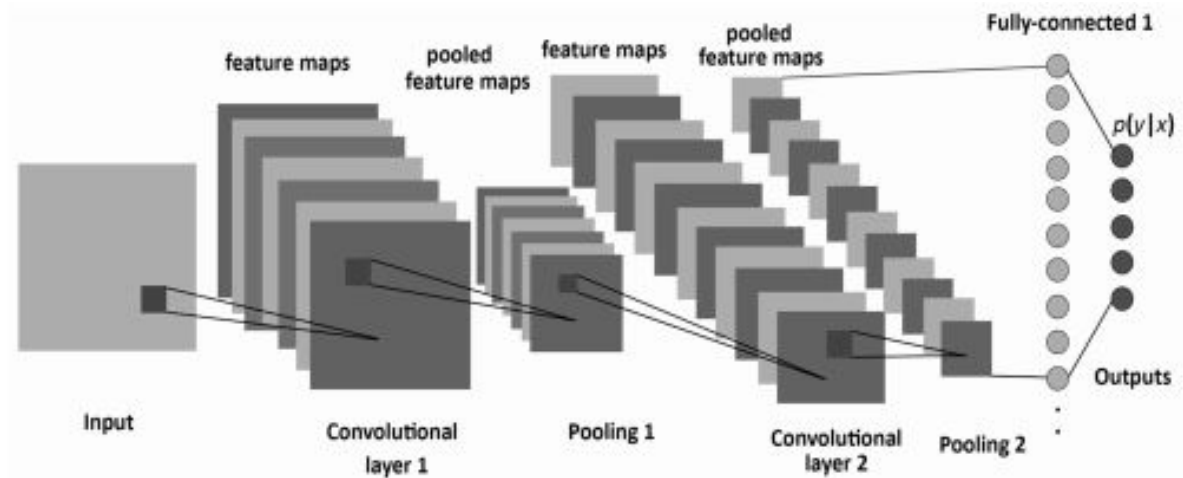
Uma camada de *pooling* é normalmente usada entre camadas convolucionais. É usado para a redução dos parâmetros e do tamanho espacial. Isso ajuda novamente a diminuir o poder computacional exigido. A saída dessa camada pode retornar a maior entrada de uma característica (*max pooling*) ou sua média (*average pooling*). Uma camada totalmente conectada possui conexões com todas as ativações da camada anterior.

### 2.3.2 HRNET

*High-Resolution Network* (HRNet) se refere à alta resolução das imagens que estão sendo processadas. É um algoritmo recente no campo de segmentação semântica, detecção de pontos faciais (WU; JI, 2019) e previsão de pose humana (AMRUTHA K.; PAULOSE, 2021).

Esta rede tem duas características principais. Primeiro, as séries de convolução de resoluções altas a baixas são conectadas em paralelo. Em segundo lugar, as informações entre as resoluções são trocadas repetidamente. A vantagem alcançada inclui a obtenção de uma representação mais precisa no domínio semântico (ALZUBAIDI LAITH, 2021).

Figura 4 – Modelo de rede neural convolucional.



Fonte: Silva (2019).

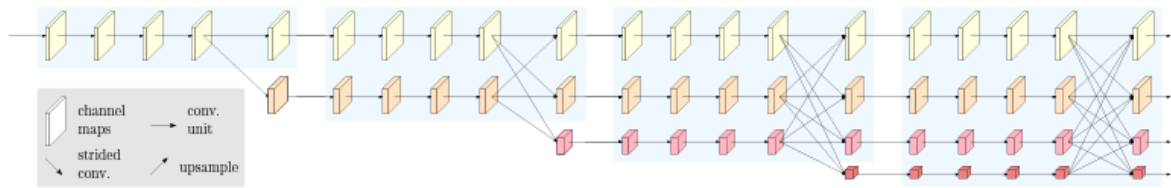
A segmentação semântica é usada para categorizar as estruturas de uma imagem em determinadas classes. Isso é feito rotulando cada pixel da imagem (JORDAN, 2018). No exemplo abaixo, todos os pixels que representam as pessoas são rotulados como pertencentes a classe “pessoa” e todos os pixels que representam a bicicleta estão rotulados como classe “bicicleta” (Figura 5). O objetivo da segmentação de imagens é permitir que o algoritmo segmente a imagem em classes e, portanto, em certas estruturas.

Figura 5 – Exemplo de imagem de uma corrida de motocicleta (a) e sua segmentação (b).



Fonte: (GUO YANMING, 2018)

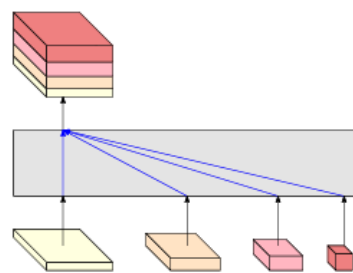
A arquitetura da HRNet é chamada HRNetV1 (Figura 6) e ela visa manter a representação de resolução alta da imagem conectando as diferentes resoluções convolucionais em paralelo (WANG JINGDONG, 2020). Ela começa com um fluxo de convolução de alta resolução, gradualmente adicionando fluxos de convolução de alta a baixa resolução, um por um, e conectando então os fluxos de multi-resolução em paralelo (WANG et al., 2020). A rede resultante consiste em um número  $n$  de etapas com a  $n$ -ésima etapa contendo os fluxos correspondentes de cada resolução. Ela é capaz de manter representações de alta resolução durante todo o processo.

Figura 6 – HRNetV1 com 4 canais de resolução (representa n etapas, sendo  $n = 4$ )

Fonte: (SUN et al., 2019)

Em outras arquiteturas de CNN, a imagem de entrada seria codificada em uma representação de baixa resolução ao passar por uma sub-rede construída como uma série conectada de convoluções de alta a baixa resolução. A representação de baixa resolução seria então recuperada para se tornar uma de alta resolução. Já na HRNet, a abordagem paralela permite que as representações de alta resolução sejam mantidas em toda a rede neural e, portanto, a representação é mais precisa (ALZUBAIDI LAITH, 2021).

Figura 7 – HRNetV2 - Concatenando as representações (upsampled) das resoluções



Fonte: (SUN et al., 2019)

Na arquitetura HRNetV1, apenas as representações de alta resolução entram na camada de saída e, conseqüentemente, características dos subsets das representações de baixa resolução são perdidas. Para que essas informações sejam completamente incluídas na saída é realizado um upsampling das resoluções mais baixas para a resolução mais alta para que a saída de todos os subsets sejam concatenados e incluídos na última camada (Figura 7). Esse modelo é chamado de HRNetV2 (WANG JINGDONG, 2020).

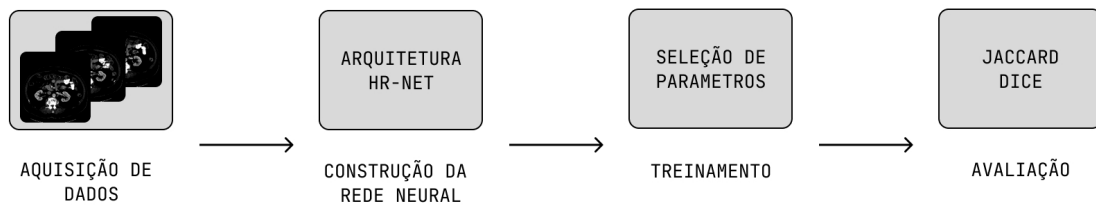
## 2.4 Considerações Finais

Este capítulo apresentou a fundamentação teórica usada no desenvolvimento deste trabalho. Para isso, foram apresentados conceitos sobre a anatomia do câncer renal e o uso da tomografia computadorizada, redes neurais convolucionais, bem como a arquitetura HRNet. No próximo capítulo 3, as etapas da metodologia proposta por este trabalho serão apresentados e discutidos.

## 3 Metodologia

Este capítulo apresenta a metodologia proposta para a utilização da HRNet na segmentação de rins e tumores renais em imagens de TC. O diagrama de fluxo do conjunto de etapas aplicado neste estudo está sumarizado na Figura 8.

Figura 8 – Metodologia



Fonte: acervo do autor

A metodologia proposta para este trabalho envolve 4 etapas: aquisição de dados, construção da rede neural, treinamento e avaliação dos resultados.

### 3.1 Aquisição da Base

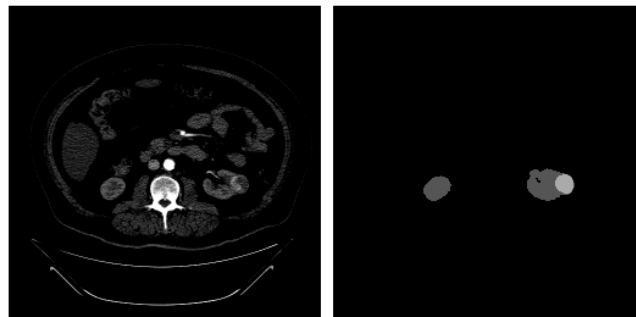
A base utilizada na metodologia foi coletada do repositório do KiTS19 (2019 *Kidney and Kidney Tumor Segmentation Challenge*), uma competição que visou estimular o progresso no problema de segmentação automática. O repositório, que contém um conjunto de treinamento de 210 imagens de TC transversais com tumores renais, foi divulgado publicamente com as correspondentes máscaras de segmentação semântica. Cada caso representa uma imagem 3d segmentada (HELLER NICHOLAS, 2021).

A base para a segmentação foi construída inicialmente separando as camadas da imagem 3d em conjuntos de imagens 2d conforme as classificações dos pixels presentes nas máscaras. Os 210 casos representam 45.242 imagens, dentre elas, 15.831 possuem máscaras com pixels classificados como rim, 5.712 apresentam tumores renais e 29.593 são camadas que não incluem o rim de acordo com suas máscaras.

### 3.2 Construção da Rede Neural

A maioria das técnicas de segmentação semântica eram desenvolvidas inicialmente a partir da classificação de imagens, estendendo a arquitetura para realizar outras tarefas de visão computacional. A HRNet, por outro lado, foi projetada a partir do zero, conectando

Figura 9 – Exemplo de imagem da base KiTS19 com presença de tumor renal e sua máscara

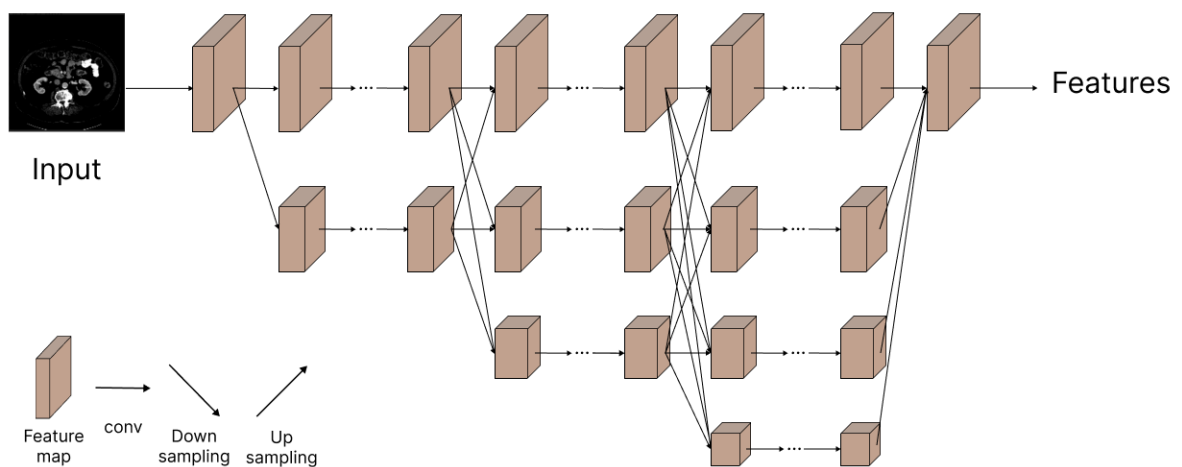


Fonte: acervo do autor

as convoluções em série de alta resolução a baixa resolução em paralelo para manter as representações de alta resolução durante todo o processo.

A arquitetura utilizada neste trabalho recebe como input uma imagem de resolução 256 x 256 pixels. Essa

Figura 10 – Arquitetura geral da rede HRNet



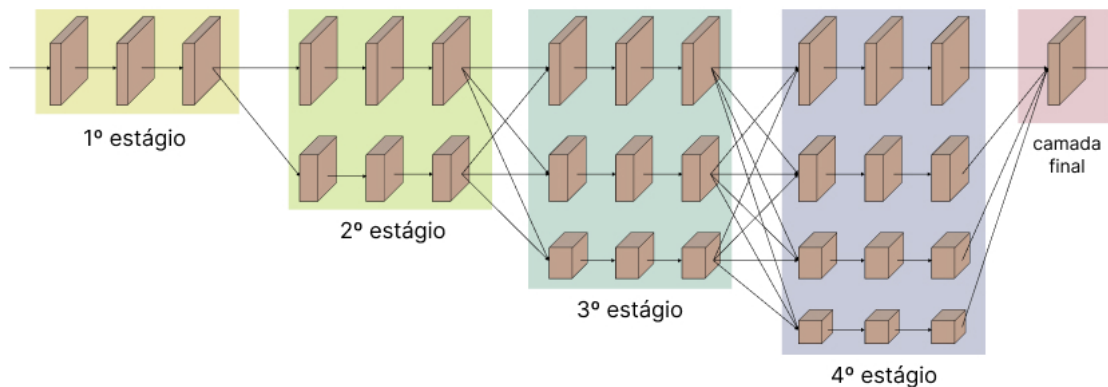
Fonte: acervo do autor

A arquitetura possui quatro estágios. O 1º estágio consiste em convoluções de alta resolução. O 2º (3º, 4º) estágio repete blocos de duas resoluções (três resoluções, quatro resoluções) várias vezes (ou seja, 1, 4, 3). A camada final da rede realiza então o upsampling das representações de baixa resolução para a resolução mais alta e concatena os 4 subconjuntos resultantes (Figura 11).

A segmentação é feita em cada fatia do exame, sendo, portanto, uma segmentação 2D. Após a aplicação da arquitetura obtém-se o melhor modelo gerado para segmentação dos rins.



Figura 11 – Arquitetura HRNet se dividindo em 4 resoluções.



Fonte: acervo do autor

Os seguintes hiperparâmetros foram otimizados via GridSearch durante o treinamento:

- O número de amostras em cada iteração: 32, 16, 8, 2
- A função de otimização: Adam, SGD
- A taxa de aprendizado: 0.1, 0.01, 0.001, 0.0001
- Função de erro: DiceLoss, categorical cross-entropy
- Número de épocas: 30, 50

### 3.3 Treinamento

A base constituiu de imagens com a presença de pixels classificados como tumor, rim e fundo, de acordo com suas máscaras. Ela foi dividida entre (70%), teste (20%) e validação (10%) conforme o número de pacientes, resultando em uma base de treino de 147 casos (31707 imagens) e validação de 21 casos (4250 imagens), deixando 42 casos (9467 imagens) para teste.

Foi adotado um método de regularização para reduzir efeitos de *overfit*. O *early stopping* encerra o treinamento quando há sinal de *overfitting*, como quando o desempenho do dataset de validação passa a diminuir em comparação com o dataset de treino. Como o desempenho de uma rede neural pode variar, subindo e descendo em alguns momentos, ao invés de parar o treinamento no primeiro sinal de *overfitting*, é comum definir um número de vezes que esta variação pode ocorrer antes de encerrar o treinamento.

Para definir o parâmetro foram gerados treinamentos com 30 e 50 épocas assim como *early stopping* baseado nas métricas de IoU e função de perda da validação.

## 3.4 Avaliação

Para avaliação dos modelos, foram adotadas as métricas de IoU (Jaccard) e Dice.

O Dice, também conhecido como *F-Score*, é considerado a métrica de avaliação mais utilizada para a segmentação de imagens médicas. Esta métrica pode ser calculada através da razão entre o dobro da interseção entre as duas regiões segmentadas e a soma das áreas das duas regiões. Matematicamente, pode ser expressa pela Equação 3.1,

$$DSC = \frac{2VP}{2VP + FP + FN}. \quad (3.1)$$

O índice de Jaccard, também conhecido como coeficiente de Jaccard ou *Intersection over Union* (IoU), assim como o coeficiente Dice, é uma métrica aplicada para verificar a similaridade entre dois conjuntos. Pode ser calculado pela razão entre o tamanho da interseção dos conjuntos e o tamanho da união dos conjuntos. Pode ser denotada pela Equação 3.2,

$$JCC = \frac{VP}{VP + FP + FN}. \quad (3.2)$$

## 4 Resultados

Inicialmente, foram executados experimentos para ajuste dos hiperparâmetros. Esses parâmetros foram escolhidos observando o desempenho e avaliando o modelo progressivamente. A Tabela 1 apresenta os hiperparâmetros utilizados.

Tabela 1 – Hiperparâmetros de busca

Hiper-parâmetros	Escopo de treino	Valores selecionados
<b>Batch size</b>	32, 16, 8, 2	32
<b>Optimizer</b>	Adam, SGD	Adam
<b>Taxa de aprendizado</b>	0,1-0,0001	0.01
<b>Função loss</b>	dice, categorical-crossentropy	categorical-crossentrop
<b>Épocas</b>	30, 50	30

A HRNet foi treinada com a base completa, incluindo as imagens que representam as camadas com presença de rim sem tumor e camadas sem presença de rim. O dataset de treino e validação foi consistiu da mesma separação de casos da etapa anterior, desta vez incluindo todas as camadas de imagens 2d para cada caso. Assim, esta base consistiu em 31.707 imagens de treino (147 casos) e 4.250 imagens de validação (21 casos), representando 70% e 10% da base, respectivamente. Os 20% separados para teste, representou 9.467 imagens (42 casos)

O modelo mostrou melhores resultados quando treinado por 30 épocas com 32 amostras por iteração, otimizador Adam com taxa de aprendizado de 0,01 e CategoricalCrossEntropy como função da perda.

A Tabela 2 apresenta os resultados dos testes realizados na rede descrita na segunda etapa da implementação

Tabela 2 – Resultados da HRNet com dataset completo

Classes	Jaccard	Dice
<b>Fundo</b>	0,998	0,999
<b>Rim</b>	0,844	0,915
<b>Tumor</b>	0,580	0,734
<b>Mean</b>	0,807	0,883
<b>Acc</b>	99%	

Segmentações geradas pelo modelo descrito acima estão representadas na Figura 12.

Todos os treinos e testes foram realizados com a mesma divisão do dataset. O modelo com melhores resultados foi então treinado utilizando o método de validação

cruzada k-fold. Esse processo divide a base de dados em diferentes grupos, permitindo que o modelo seja exposto a diferentes amostras de treino e teste.

Neste trabalho foi determinado  $k = 5$ , dividindo o dataset em 5 grupos de treino e teste. Uma vez definido o parâmetro  $k$ , o dataset é dividido em 5 grupos de dados, cada grupo passa pelo treinamento sendo obtido uma estimativa de desempenho do modelo em dados não vistos (Tabela 3).

Tabela 3 – Resultados do modelo com melhores parâmetros treinado em 5-folds da base de dados.

	<b>Rim</b>		<b>Tumor</b>	
	<b>Jaccard</b>	<b>Dice</b>	<b>Jaccard</b>	<b>Dice</b>
<b>Fold 1</b>	0,844	0,915	0,580	0,734
<b>Fold 2</b>	0,870	0,930	0,481	0,649
<b>Fold 3</b>	0,836	0,910	0,620	0,765
<b>Fold 4</b>	0,851	0,919	0,479	0,648
<b>Fold 5</b>	0,814	0,897	0,564	0,721
<b>Mean</b>	<b>0,843</b>	<b>0,914</b>	<b>0,544</b>	<b>0,703</b>

Por se tratar de uma base de dados onde a presença de fatias com pixels de tumor é muito reduzida, escolhemos fazer um segundo experimento complementar somente com fatias que possuam o tumor. O dataset de treino e validação foi separado conforme os casos, e consistiu em 3.935 imagens de treino (147 casos) e 531 imagens de validação (21 casos), representando 70% e 10% da base, respectivamente. Foram separadas 20% para teste, representado por 1.246 imagens (42 casos).

A rede foi treinada por 50 épocas com 2 amostras por iteração, otimizador Adam com taxa de aprendizagem de 0,01 e CategoricalCrossEntropy como função da perda. Este conjunto de parâmetros apresentou os melhores resultados após vários experimentos.

Para comparação, uma rede U-Net foi treinada com o mesmo escopo de hiperparâmetros selecionados por tentativa e erro. Os parâmetros que apresentaram melhores resultados para a U-Net treinou 16 amostras por iteração e utilizou como função de perda o DiceLoss.

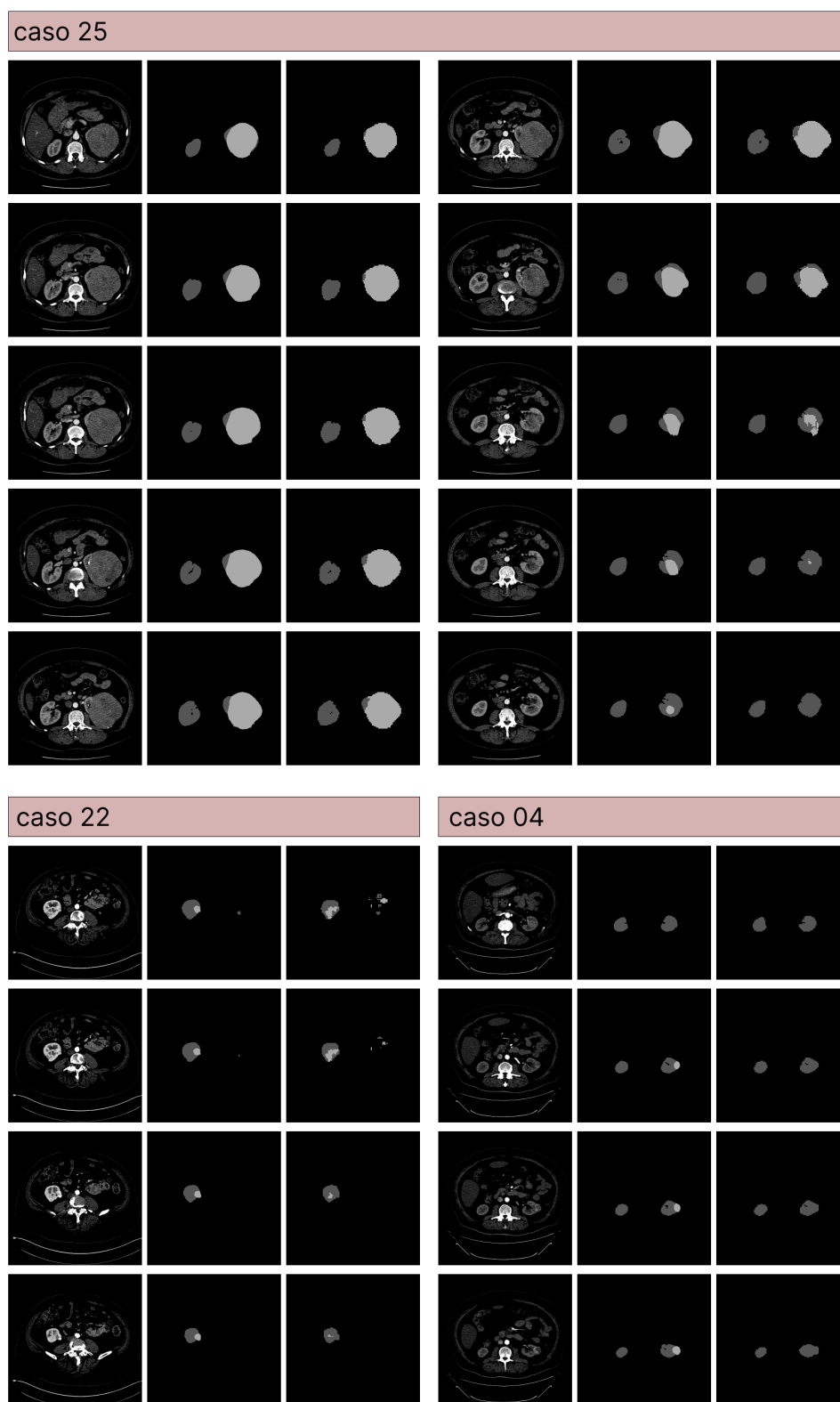
A Tabela 4 apresenta os resultados dos testes realizados nas redes descritas na primeira etapa da implementação.

Os resultados dos experimentos com a arquitetura HRNet apresentaram leve melhora em comparação com a U-Net.

Tabela 4 – Resultados da HRNet com dataset de tumores em comparação com a U-Net

Classes	HRNet		U-Net	
	Jaccard	Dice	Jaccard	Dice
Fundo	0,992	0,996	0,994	0,997
Rim	<b>0,711</b>	<b>0,831</b>	0,701	0,824
Tumor	<b>0,589</b>	<b>0,741</b>	0,536	0,698
Mean	<b>0,764</b>	<b>0,856</b>	0,744	0,840
Acc	99%		99%	

Figura 12 – Exemplos de segmentação dos tumores renais de 3 pacientes com o modelo HRNet construído. Da esquerda para a direita, fatia de entrada, máscara de segmentação e segmentação efetuada pela rede.



Fonte: acervo do autor

## 5 Conclusão

A HRNet é uma rede que vem de tarefas gerais de visão e vence em muitas frentes da visão computacional, incluindo segmentação semântica, estimativa de pose humana e detecção de objetos. Ela vem recebendo atenção crescente na segmentação semântica devido ao seu alto desempenho. A segmentação semântica é um problema que consiste em atribuir um rótulo de classe a cada píxel para classificar as partes de uma imagem. A HRNet aplicada à segmentação semântica utiliza a variante da arquitetura denominada de HRNetV2.

Recentemente, métodos baseados em aprendizado profundo mostraram excelentes resultados na segmentação 3D automática, mas requerem grandes conjuntos de dados para treinamento e ainda há pouco consenso sobre quais métodos têm melhor desempenho. O 2019 Kidney and Kidney Tumor Segmentation challenge (KiTS19) foi uma competição realizada em conjunto com a 2019 International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) que abordou estas questões e estimular o progresso neste problema de segmentação automática.

Este trabalho teve como objetivo avaliar o desempenho da rede neural convolucional High Resolution Network (HRNet) na segmentação semântica de imagens de TC para delimitar regiões dos rins e tumor na imagem. As imagens coletadas do repositório Kits19 não foram alteradas por processamento de imagem.

O modelo proposto alcançou um valor Dice de 0,914 e 0,703 para tumor e rim, respectivamente, resultando em uma pontuação composta de Dice de 0,872 para as 3 classes de segmentação (rim, tumor e fundo).

Ademais, a comparação entre os resultados das otimizações das redes HRNet e U-Net realizadas do dataset reduzido de tumores apresentou melhores resultados no Dice de segmentação de rim e de tumor da HRNet.

Devido a grande quantidade de imagens da base todas as fatias foram previamente redimensionadas de 512x512 para 256x256.

Dessa forma, os resultados dessa pesquisa apresentam Dice e Jaccard menor em comparação com outros trabalhos relacionados. O uso de técnicas de data augmentation como forma de diminuir a disparidade entre fatias com e sem tumor assim como a utilização das imagens em sua resolução original são fatores que poderiam contribuir para a melhora desses resultados.

Como trabalhos futuros, podemos avaliar a metodologia proposta no conjunto de dados atualizado em 2021 (Kits2021) adicionando mais parâmetros em nosso espaço de busca, além de utilizar técnicas de data augmentation.



## Referências

- AHMED, S. e. a. Automated covid-19 detection from chest x-ray images: a high-resolution network (hrnet) approach. *SN computer science*, Springer, v. 2, n. 4, p. 1–17, 2021. Citado na página 13.
- ALBAWI SAAD, T. A. M.; AL-ZAWI, S. Understanding of a convolutional neural network. *international conference on engineering and technology (ICET)*, Ieee, 2017. Citado na página 19.
- ALZUBAIDI LAITH, e. a. Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions. *Journal of big Data*, Springer, v. 8, n. 1, p. 1–74, 2021. Citado 3 vezes nas páginas 17, 19 e 21.
- AMERICAN CANCER SOCIETY. *Tests for Kidney Cancer*. 2020. Disponível em: <<https://www.cancer.org/cancer/kidney-cancer/detection-diagnosis-staging/how-diagnosed.html>>. Citado 2 vezes nas páginas 15 e 16.
- AMERICAN CANCER SOCIETY. *What Is Kidney Cancer?* 2020. Disponível em: <<https://www.cancer.org/cancer/kidney-cancer/about/what-is-kidney-cancer.html>>. Citado na página 15.
- AMERICAN CANCER SOCIETY. *Key Statistics About Kidney Cancer*. 2022. Disponível em: <<https://www.cancer.org/cancer/kidney-cancer/about/key-statistics.html>>. Citado na página 11.
- AMRUTHA K., P. P.; PAULOSE, J. Human body pose estimation and applications. *Innovations in Power and Advanced Computing Technologies (i-PACT)*, IEEE, p. 1–6, 2021. Citado na página 19.
- BARBOSA, G. d. S.; LIMA, R. de; GUIMARÃES, R. M. Série histórica de mortalidade por neoplasias renais no brasil (1996-2010). *Cadernos Saúde Coletiva*, SciELO Brasil, v. 20, p. 537–540, 2012. Citado na página 11.
- BRODY, H. Medical imaging. *Nature*, v. 502, n. 7473, p. S81–S81, 2013. Citado na página 12.
- COHEN, J. P.; MORRISON, P.; DAO, L.; ROTH, K.; DUONG, T. Q.; GHASSEMI, M. *COVID-19 Image Data Collection: Prospective Predictions Are the Future*. 2020. Citado na página 16.
- CRUZ LUANA BATISTA, e. a. da. Kidney segmentation from computed tomography images using deep neural network. *Computers in Biology and Medicine*, Elsevier, v. 123, p. 103906, 2020. Citado na página 13.
- DALL’OGLIO MARCOS, e. a. Carcinoma de células renais incidentais e sintomáticos: fatores patológicos e sobrevida. *Revista da associação médica brasileira*, scielo, v. 50, p. 27–31, 2004. Citado na página 16.

GUO YANMING, e. a. A review of semantic segmentation using deep neural networks. *International journal of multimedia information retrieval*, v. 7, p. 87–93, 2018. Citado na página 20.

HELLER NICHOLAS, e. a. The kits19 challenge data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes. arxiv, 2019. Citado na página 13.

HELLER NICHOLAS, e. a. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced ct imaging: Results of the kits19 challenge. *Medical image analysis*, Elsevier, v. 67, p. 101821, 2021. Citado 2 vezes nas páginas 12 e 22.

INCA. *Ministério da Saúde amplia tratamento para câncer renal*. 2019. Disponível em: <<https://www.inca.gov.br/noticias/ministerio-da-saude-amplia-tratamento-para-cancer-renal>>. Citado 2 vezes nas páginas 11 e 15.

INCA. *Como surge o câncer?* 2021. Disponível em: <<https://www.inca.gov.br/como-surge-o-cancer>>. Citado na página 11.

INCA. *Tumor de Wilms - versão para Profissionais de Saúde*. 2021. Disponível em: <<https://www.inca.gov.br/es/node/2858>>. Citado na página 12.

JORDAN, J. *An overview of semantic image segmentation*. 2018. Disponível em: <<https://www.jeremyjordan.me/semantic-segmentation/>>. Citado na página 20.

JÚNIOR, E. A.; YAMASHITA., H. Aspectos básicos de tomografia computadorizada e ressonância magnética. *Brazilian Journal of Psychiatry*, SciELO Brasil, v. 23, n. 1, p. 2–3, 2001. Citado na página 16.

KROGH, A. What are artificial neural networks? *Nature biotechnology*, 2008. Citado na página 18.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group, v. 521, n. 7553, p. 436–444, 2015. Citado 2 vezes nas páginas 12 e 17.

O'SHEA, K.; NASH, R. *An introduction to convolutional neural networks*. 2015. Citado na página 19.

PRADO, B. B. F. d. Influência dos hábitos de vida no desenvolvimento do câncer. *Ciência e Cultura*, SciELO, v. 66, n. 1, p. 21–24, 2014. Citado na página 15.

RAUBER, T. W. Redes neurais artificiais: Princípios básicos. *Revista Eletrônica Científica Inovação e Tecnologia*, v. 1, n. 13, p. 47–57, 2016. Citado na página 16.

REDES-NEURAI. 2020. Disponível em: <<https://github.com/Sam-Marx/redes-neurais>>. Citado na página 18.

RONNEBERGER OLAF, P. F.; BROX., T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention.*, Springer, p. 234–241, 2015. Citado na página 13.

SBN. *Sociedade Brasileira de Nefrologia*. 2022. Disponível em: <<https://www.sbn.org.br/o-que-e-nefrologia/compreendendo-os-rins/>>. Citado na página 15.

SHEN, D.; WU, G.; SUK., H.-I. Deep learning in medical image analysis. *Annual review of biomedical engineering*, annual reviews, v. 19, p. 221–248, 2017. Citado 2 vezes nas páginas 11 e 12.

SHEN DINGGANG, G. W.; SUK, H.-I. Deep learning in medical image analysis. *Annual review of biomedical engineering*, v. 19, p. 221–248, 2017. Citado na página 17.

SHON HO SUN, e. a. Classification of kidney cancer data using cost-sensitive hybrid deep learning approach. *Symmetry*, mdpi, v. 12, n. 1, p. 154, 2020. Citado na página 11.

SILVA, C. Q. *Redes neurais aplicadas no reconhecimento de símbolos matemáticos manuscritos online*. Tese (Doutorado) — Universidade de São Paulo, 2019. Citado na página 20.

SUN, K.; ZHAO, Y.; JIANG, B.; CHENG, T.; XIAO, B.; LIU, D.; MU, Y.; WANG, X.; LIU, W.; WANG, J. *High-Resolution Representations for Labeling Pixels and Regions*. 2019. Citado na página 21.

WANG, J.; SUN, K.; CHENG, T.; JIANG, B.; DENG, C.; ZHAO, Y.; LIU, D.; MU, Y.; TAN, M.; WANG, X.; LIU, W.; XIAO, B. *Deep High-Resolution Representation Learning for Visual Recognition*. 2020. Citado na página 20.

WANG JINGDONG, e. a. Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 43, n. 10, p. 3349–3364, 2020. Citado 3 vezes nas páginas 13, 20 e 21.

WU, Y.; JI, Q. Facial landmark detection: A literature survey. *International Journal of Computer Vision*, arxiv, v. 127, n. 2, p. 115–142, 2019. Citado na página 19.

YANG, X.; HE, X.; ZHAO, J.; ZHANG, Y.; ZHANG, S.; XIE, P. *COVID-CT-Dataset: A CT Scan Dataset about COVID-19*. 2020. Citado na página 16.

YU QIAN, e. a. Crossbar-net: A novel convolutional neural network for kidney tumor segmentation in ct images. *IEEE transactions on image processing*, IEEE, v. 28, n. 8, p. 4060–4074, 2019. Citado na página 13.