

**Aprendizado Profundo em Segmentação de Imagens Médicas:  
Revisão da Literatura e o Estado da Arte**

**Luís Sérgio Rozenkwit**

Trabalho de Conclusão de Curso  
MBA em Inteligência Artificial e Big Data

# UNIVERSIDADE DE SÃO PAULO

## Instituto de Ciências Matemáticas e de Computação

---

Aprendizado Profundo em Segmentação de Imagens  
Médicas: Revisão da Literatura e o Estado da Arte

*Luís Sérgio Rozenkwit*

---

**Luís Sérgio Rozenkwit**

## **Aprendizado Profundo em Segmentação de Imagens Médicas : Revisão da Literatura e o Estado da Arte**

Trabalho de conclusão de curso apresentado ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial

Orientadora: Profa. Dra. Agma Juci M. Traina

USP - São Carlos

2024

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassie  
Seção Técnica de Informática, ICMC/USP,  
com os dados inseridos pelo(a) autor(a)

R893a      Rozenkwit, Luis Sérgio  
Aprendizado profundo em segmentação de imagens  
médicas: revisão da literatura e o estado da arte.  
/ Luis Sérgio Rozenkwit; orientador Professora  
Doutora Agma Juci Machado Traina. -- São Carlos,  
2024.

77 p.

Trabalho de conclusão de curso (MBA em  
Inteligência Artificial e Big Data) -- Instituto de  
Ciências Matemáticas e de Computação, Universidade  
de São Paulo, 2024.

1. Inteligência artificial. 2. Visão  
computacional. 3. Segmentação de imagens médicas. 4.  
Radiologia e diagnóstico por imagem. 5. Aprendizado  
profundo. I. Machado Traina, Professora Doutora  
Agma Juci , orient. II. Título.

Bibliotecários responsáveis pela estrutura de catalogação da publicação de acordo com a AACR2:  
Gláucia Maria Saia Cristianini - CRB - 8/4938  
Juliana de Souza Moraes - CRB - 8/6176



## DEDICATÓRIA

Aos meus familiares e aos professores desta longa especialização, pelo apoio e motivação incansáveis;

À minha irmã Dra Déborah Rozenkwit, por não desistir de mim.

À Professora Doutora Solange Rezende, pela paciência e dedicação a seus alunos;

À Professora Agma Juci Machado Traina, exemplo de inteligência e amor à docência.

*Uma singela homenagem ao meu querido, inigualável e genialmente brilhante papai,*

*Dr. Moizes Rozenkwit, de abençoada memória.*



## RESUMO

ROZENKWIT, LUIS SÉRGIO. **Aprendizado Profundo em Segmentação de Imagens Médicas : Revisão da Literatura e o Estado da Arte**. 2024. 59 f. Trabalho de conclusão de curso (MBA em Inteligência Artificial e Big Data) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2024.

A segmentação de imagens médicas é um componente essencial do processamento de imagens biomédicas e tem contribuído significativamente para o avanço da saúde sustentável. Com o advento da aprendizagem profunda, especialmente por meio das redes neurais convolucionais profundas, houve uma revolução no processamento de imagens médicas. Esta monografia tem como objetivo explorar detalhadamente o campo da segmentação de imagens médicas utilizando técnicas de aprendizagem profunda. Inicialmente, apresenta uma visão geral dos conceitos fundamentais e das características da segmentação de imagens médicas baseada em aprendizagem profunda. Em seguida, discute o estado atual da pesquisa nesse domínio, oferecendo uma análise abrangente dos três principais métodos de segmentação de imagens médicas, juntamente com suas limitações inerentes. Além disso, explora as características distintas de vários tecidos e órgãos patológicos, destacando os desafios específicos que eles representam e os algoritmos clássicos de segmentação empregados para resolvê-los. Apesar dos avanços notáveis, o campo continua a enfrentar obstáculos de pesquisa significativos, como a precisão insuficiente da segmentação e a limitação em quantidade e resolução dos conjuntos de dados de imagens médicas. Para abordar esses problemas, o artigo apresenta um exame detalhado dos métodos atuais baseados em aprendizagem profunda para segmentação de imagens médicas, visando ajudar os pesquisadores a superar esses desafios.

**Palavras-chave:** Segmentação de Imagens Médicas; Inteligência Artificial; Aprendizado de Máquina Profundo; Visão Computacional; Redes Neurais Convolucionais; Processamento de Imagens Médicas; Radiologia e Diagnóstico por Imagem.



## ABSTRACT

ROZENKWIT, LUIS S. **Deep Learning in Medical Image Segmentation:** Literature Review and State of the Art. 2024. 59 f. Final course work (MBA in Artificial Intelligence and Big Data) - Institute of Mathematical and Computer Sciences, University of São Paulo, São Carlos, 2024.

Medical image segmentation is an essential component of biomedical image processing and has contributed significantly to the advancement of sustainable healthcare. With the advent of deep learning, especially through deep convolutional neural networks, there has been a revolution in medical image processing. This monography aims to explore in detail the field of medical image segmentation using deep learning techniques. Initially, it presents an overview of the fundamental concepts and characteristics of medical image segmentation based on deep learning. It then discusses the current state of research in this domain, offering a comprehensive analysis of the three main medical image segmentation methods, along with their inherent limitations. In addition, it explores the distinct characteristics of various pathological tissues and organs, highlighting the specific challenges they pose and the classic segmentation algorithms employed to solve them. Despite notable advances, the field faces significant research obstacles, such as insufficient segmentation accuracy and the limited quantity and resolution of medical image datasets. To address these problems, the text presents a detailed examination of current deep learning-based methods for medical image segmentation, aiming to help researchers overcome these challenges.

**Keywords:** Medical Image Segmentation; Deep Learning; Artificial Intelligence ; Computer Vision; Convolutional Neural Networks; Medical Image Processing ; Radiology ; Medical Image Diagnostic.

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>31</b>
<b>2 REVISÃO BIBLIOGRÁFICA.....</b>	<b>34</b>
<b>3 MÉTRICAS E APLICAÇÕES PRÁTICAS.....</b>	<b>45</b>
<b>4 PRINCIPAIS MÉTODOS DE SEGMENTAÇÃO.....</b>	<b>53</b>
<b>5 O ESTADO DA ARTE.....</b>	<b>70</b>
<b>6 CONCLUSÕES.....</b>	<b>75</b>
<b>7 REFERÊNCIAS.....</b>	<b>77</b>

## 1. INTRODUÇÃO

O estudo e desenvolvimento de técnicas de segmentação de imagens têm se mostrado fator crítico e ao mesmo tempo condição *sine qua non* para o processamento de imagens médicas e seu aproveitamento em inteligência artificial. Embora haja gargalos de diversas matizes (ética, técnica, científica, regulatória, entre outras), há rápida e perceptível evolução deste campo da visão computacional em grande parte pela compreensão da importância de sua aplicação na área médica (significativo impacto no aumento da acurácia diagnóstica, aumento em escala na quantidade de exames liberados no menor tempo possível, potencial impacto positivo na viabilidade financeira dos serviços), que têm levado a uma crescente sinergia entre profissionais de diversas áreas (médicos, programadores, cientistas da computação, engenheiros, cientistas de dados, etc.).

Essencialmente, a segmentação de imagens envolve a divisão de uma imagem inteira em regiões distintas que compartilham certas propriedades comuns. Em termos mais simples, visa separar uma região de interesse (ROI) do restante da imagem (fundo ou *background*). Assim, à partir da abertura de novas frentes de estudo, tendo como consequência natural a elaboração de protocolos e algoritmos de programação (leia-se aprendizagem de máquina), meio da integração de novas teorias e tecnologias de ponta, busca-se um padrão-ouro para a aplicação de segmentação que possa ser utilizado de forma eficiente, efetiva e eficaz nos diversos campos da imagiologia médica. (Lateef, F.; Ruichek, Y. Survey on semantic segmentation using deep learning techniques. *Neurocomputing* 2019, 338, 321–348).

As técnicas para obtenção de imagens médicas comumente utilizadas em ambientes clínicos incluem tomografia computadorizada (TC), ressonância magnética (RM), tomografia por emissão de pósitrons (PET), raios-X (RX), ultrassonografia (USG) e densitometria óssea (DO), e suas sub áreas. Além disso, há também outros tipos de imagens e suas aplicações, como microscopia, histopatologia e oftalmologia (imagens vasculares da retina), que se enquadram na categoria de imagens RGB (red, green, blue) comuns.

As imagens geradas pelos equipamentos espelham a condição de saúde do paciente no momento do exame e, não por acaso, tornaram-se a base principal sobre a qual os médicos fazem suas análises, planejamento clínico, terapêutico e prognósticos. Portanto, inevitavelmente a segmentação de imagens radiológicas tornou-se foco de especial atenção dos cientistas e pesquisadores de visão computacional. (Shen, D.; Wu, G.; Suk, H.I. Deep learning in medical image analysis. *Annu. Rev. Biomed. Eng.* 2017, 19, 221–248).

No campo da segmentação de imagens, os avanços na inteligência artificial e particularmente o aprendizado de máquina profundo (*deep learning*), têm evoluído e entregado resultados impressionantes. A utilização de técnicas de aprendizado profundo provou ser vantajosa em termos de precisão e velocidade quando comparada aos métodos convencionais de aprendizado de máquina e visão computacional.

A implementação de técnicas de aprendizado profundo na segmentação de imagens médicas provou ser uma ferramenta valiosa para os médicos avaliarem com precisão aspectos como forma, tamanho e natureza de eventuais lesões para fins de classificação e planejamento terapêutico.

Tal tecnologia pode aliviar significativamente a sobrecarga de trabalho dos médicos radiologistas, com impacto favorável no desfecho clínico do paciente e na saúde financeira da instituição, uma vez que tem pleno potencial para entregar resultados com alta sensibilidade e especificidade numa escala muito além da capacidade humana. (Goodfellow, I.; Bengio, Y.; Courville, A.; Bengio, Y. *Deep Learning*; MIT Press: Cambridge, UK, 2016).

Para reunir informações abrangentes sobre diferentes metodologias, foi realizada pesquisa no Google Scholar e ArXiv usando as palavras-chave, “inteligência artificial”, “aprendizado de máquina”, “visão computacional” “aprendizado profundo”, “*medical image processing*”, “*computer vision*”, e “*deep learning*” para acessar a literatura mais recente. Além disso, foram avaliados artigos e estudos publicados em congressos, revistas e conferências relevantes ao assunto, tais como ISBI (*International Symposium on Biomedical Imaging*), IPMI (*Information Processing in*

*Medical Imaging*) e MICCAI (*Medical Image Computing and Computer Assisted Intervention Society*), para reunir os materiais relevantes.

Esta pesquisa examina os avanços na segmentação de imagens médicas, especificamente da perspectiva do aprendizado profundo. Ao contrário de outras revisões encontradas em diversas publicações, esta análise abrangente se concentra no progresso recente e compara vários métodos no campo.

Os artigos selecionados, que utilizam predominantemente técnicas de aprendizado profundo, passaram por uma verificação rigorosa. A revisão investiga especificamente a aplicação da tecnologia de aprendizado profundo na segmentação de imagens médicas nos últimos três anos, analisando estruturas de rede, metodologias, pontos fortes e fracos. Além disso, identifica desafios e destaca o potencial para avanços futuros nessa área. A revisão abrange os métodos mais recentes e as abordagens clássicas na tecnologia para aplicação de algoritmos de aprendizado profundo em imagens médicas.

A monografia está estruturada da seguinte forma: o capítulo 2 inicia a exaustiva revisão da literatura apresentando conceitos de segmentação de imagens médicas e as primeiras técnicas de redes neurais e aprendizado profundo e suas aplicações neste campo. No capítulo 3, compartilhamos métricas de avaliação e conjuntos de dados provenientes de alguns dos principais desafios de análise de imagens médicas.

O corpo principal da revisão da literatura é abordado nos capítulos 4 e 5. O capítulo 4 fornece uma visão geral de três modernas e importantes estruturas de rede - FCN (rede totalmente convolucional), U-Net e GAN (rede adversária generativa) - que são usadas para segmentação de imagens médicas baseadas em aprendizado profundo semântico exemplos de arquitetura de rede e aplicações.

Finalmente, o capítulo 5 aborda o estado da arte e o capítulo 6 conclui a monografia resumindo-a.

## 2. REVISÃO BIBLIOGRÁFICA

A aplicação da tecnologia de processamento de imagens médicas permite a segmentação, extração, reconstrução tridimensional e exibição tridimensional de órgãos humanos, tecidos moles bem como de estruturas específicas.

Esse processo envolve a análise e o processamento de imagens 2D ou 3D para dividi-las em diferentes regiões com base em sua semelhança ou diferença.

Ao utilizar tal método, os médicos radiologistas são capazes de analisar qualitativamente ou mesmo quantitativamente lesões e outras áreas de interesse, levando a um aumento significativo na precisão e confiabilidade dos diagnósticos médicos, ou seja, a sua acurácia (Lateef e Ruichek, 2019).

O processo de segmentação de imagens médicas pode ser entendido conceitualmente por meio de um modelo baseado na teoria dos conjuntos. Neste modelo, uma imagem médica, denotada como  $I$ , é analisada em conjunto com uma coleção de restrições de similaridade representada como  $C_i$  (onde  $i = 1, 2, \dots$ ). O objetivo da segmentação é particionar a imagem  $I$  em regiões ou segmentos distintos.

A equação

$N$

$$U = \bigcup_{x=1}^N R_x = I, R_x \cap R_y = \emptyset, \forall x \neq y, x, y \in [1, N]$$

$x=1$

afirma que a união de  $N$  e  $x$  é igual a 1, onde  $R_x$  é igual a  $I$ , e a interseção de  $R_x$  e  $R_y$  é vazia para todo  $x$  não igual a  $y$ , onde  $x$  e  $y$  estão ambos dentro do intervalo de 1 a  $N$ . Essa condição garante que  $R_x$  e  $R_y$  satisfaçam a restrição de similaridade de comunicação  $C_i$  para todos os pixels nas áreas da imagem.

As variáveis  $x$  e  $y$  são usadas para distinguir entre diferentes regiões.  $N$  representa um inteiro positivo maior ou igual a 2, indicando o número de regiões após a segmentação (Shen, Wu e Suk, 2017).

O processo de segmentação de imagens médicas envolve várias etapas. Em primeiro lugar, é obtido um conjunto de dados de imagens médicas que normalmente inclui um conjunto de treinamento, um conjunto de validação e um conjunto de teste.

No processamento de imagens baseado em aprendizado de máquina, o conjunto de dados geralmente é dividido nessas três partes (Goodfellow et al., 2016).

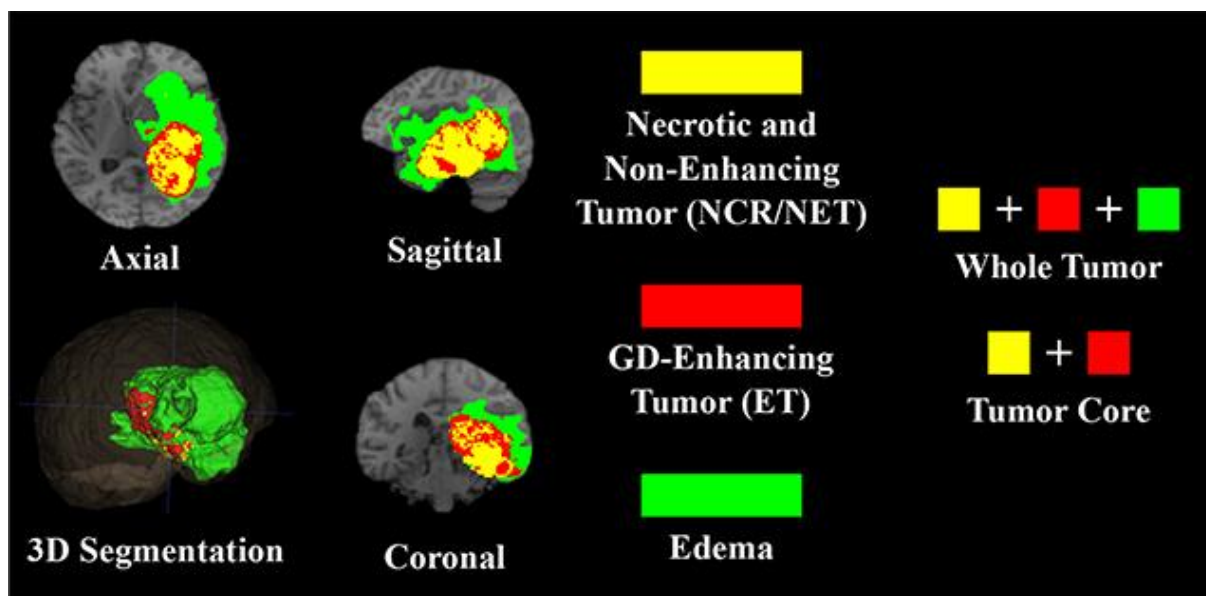


Figura 1. Técnica de segmentação tumoral em ressonância magnética encefálica.

Fonte: Menze et al., 2015.

O conjunto de treinamento é usado para treinar o modelo de rede, enquanto o conjunto de validação é usado para ajustar os hiperparâmetros do modelo.

O impacto final do modelo é confirmado usando o conjunto de teste. A imagem é pré-processada e aumentada, normalmente envolvendo a padronização da imagem de entrada e a aplicação de rotação aleatória e dimensionamento para expandir o conjunto de dados.

Uma técnica adequada de segmentação de imagens médicas é utilizada para segmentar a imagem médica e produzir imagens segmentadas que representem apropriadamente o conteúdo e semântica originais de informação (Almeida e Tavares, 2020).

A avaliação do desempenho da estimativa é realizada para avaliar a eficácia da segmentação de imagens médicas, exigindo o estabelecimento de indicadores de desempenho apropriados para verificação.

A segmentação de imagens é um componente crucial do processo geral. O campo da compreensão de imagens tornou a segmentação de imagens um ponto focal, pois é um paradigma bem estabelecido na pesquisa de visão computacional (Hesamian et al., 2019).

A segmentação de imagens envolve a divisão de uma imagem em áreas distintas com base em vários recursos, como tons de cinza, cor, textura e formas. O objetivo é garantir que esses recursos exibam consistência ou semelhança dentro da mesma área, ao mesmo tempo em que exibem diferenças claras entre as diferentes áreas.

Dependendo do nível de granularidade, a segmentação de imagens pode ser categorizada em segmentação semântica, segmentação de instância e segmentação panorâmica (Altaf et al., 2019). Na área médica, a segmentação de imagens é considerada principalmente uma tarefa de segmentação semântica. Além disso, existem vários ramos de segmentação de imagens, incluindo segmentação de imagens de satélite, segmentação de imagens médicas e direção autônoma.

À medida que o número de estruturas de rede propostas aumenta, a precisão dos métodos de segmentação de imagens continua a melhorar. No entanto, é importante observar que não existe um algoritmo de segmentação universal que possa ser aplicado a todas as imagens. A eficácia dos métodos tradicionais de segmentação de imagens não pode mais ser comparada às aquelas baseadas em aprendizado profundo (Hu et al., 2019).

No entanto, os conceitos subjacentes a esses métodos tradicionais ainda são valiosos para fins de aprendizagem. Exemplos de tais métodos incluem segmentação baseada em limite, segmentação baseada em região e segmentação baseada em detecção de borda.

Esses métodos utilizam princípios de processamento digital de imagens e matemática para segmentar imagens. Embora sejam simples de calcular e rápidos em termos de velocidade de segmentação, eles podem não ter precisão quando se trata de capturar detalhes finos.



Atualmente, os métodos baseados em aprendizado profundo (*deep learning*) alcançaram avanços notáveis na segmentação de imagens. Esta nova abordagem superou as técnicas tradicionais de segmentação em termos de precisão. A introdução de redes totalmente convolucionais marcou um avanço significativo no uso de aprendizado profundo para segmentação semântica de imagens (Geiger, Lenz e Urtasun, 2012).

No domínio das redes de segmentação, existem opções notáveis como U-Net, Mask R-CNN, RefineNet e DeconvNet. Tais redes possuem uma capacidade notável de lidar efetivamente com bordas intrincadas (Ess et al., 2009).

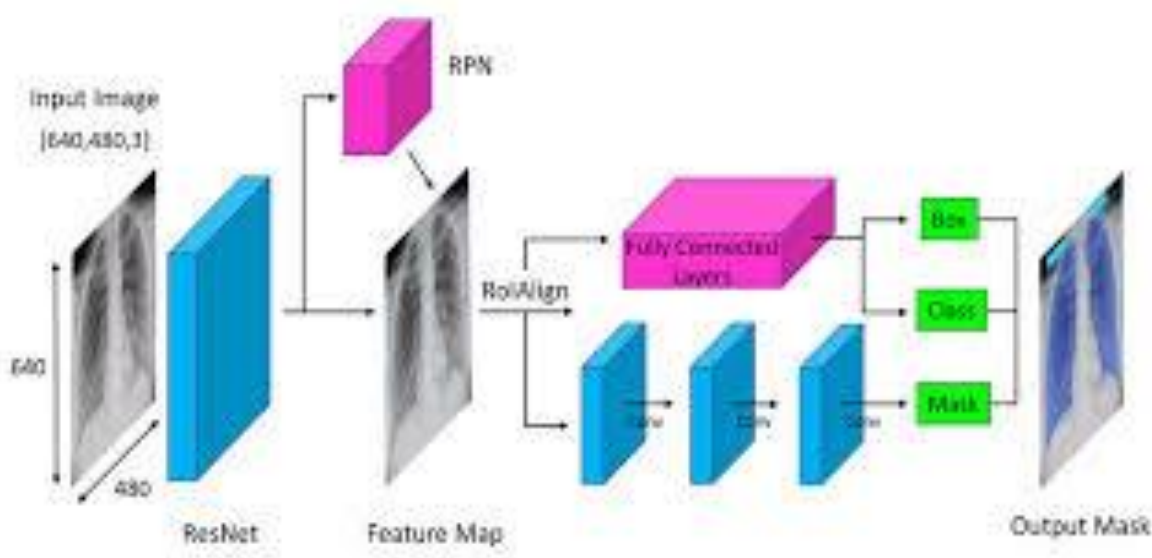


Figura 2. R-CNN na segmentação de imagens em radiografia torácica.

Fonte: Soumyajit et al., 2021.

## **2.1. AS REDES NEURAS CONVOLUCIONAIS DESEMPENHAM UM PAPEL SIGNIFICATIVO NESTE CAMPO DE ESTUDO**

A fusão da tecnologia de aprendizado profundo e processamento de imagem deu origem à rede neural convolucional (CNN).

No campo da análise e processamento de imagens, a CNN alcançou inúmeros avanços solidificando seu status como uma das redes neurais mais proeminentes. Notavelmente, no conjunto de dados ImageNet amplamente utilizado, a CNN tem sido fundamental em várias realizações, como extração de recursos de imagem, classificação e reconhecimento de padrões.

Como um modelo profundo de aprendizado supervisionado, a CNN opera com base no princípio do compartilhamento de peso para mapeamento de recursos em diferentes posições dentro da rede de camadas anterior. Ao alavancar as relações espaciais, essa abordagem reduz efetivamente o número de parâmetros, melhorando assim o desempenho do treinamento (Ma, Tavares e Jorge, 2009).

A jornada da rede neural convolucional desde sua proposta até sua aplicação generalizada pode ser dividida em várias etapas: brotamento teórico, desenvolvimento experimental, aplicação em larga escala e pesquisa aprofundada. Uma teoria crucial que surgiu durante os estágios iniciais foi o conceito de campos receptivos e máquinas neurocognitivas na informação visual humana. Em 1962, Hubel et al. (1962) conduziram pesquisas biológicas revelando que a transmissão de informações visuais no cérebro ocorre por meio de excitação de campo receptivo em vários níveis, introduzindo assim o conceito de campos receptivos. Com base nesse conceito, Fukushima (1982) propôs uma máquina neurocognitiva que é considerada a primeira rede de implementação de redes neurais convolucionais.

O estágio de desenvolvimento experimental começou em 1998, quando LéCun et al. introduziram a LeNet5, uma rede treinada usando um algoritmo de retropropagação baseado em gradiente. O interesse da comunidade acadêmica em redes neurais convolucionais cresceu significativamente com a proposta da rede LeNet5,

particularmente devido à sua aplicação bem-sucedida no reconhecimento de caligrafia.

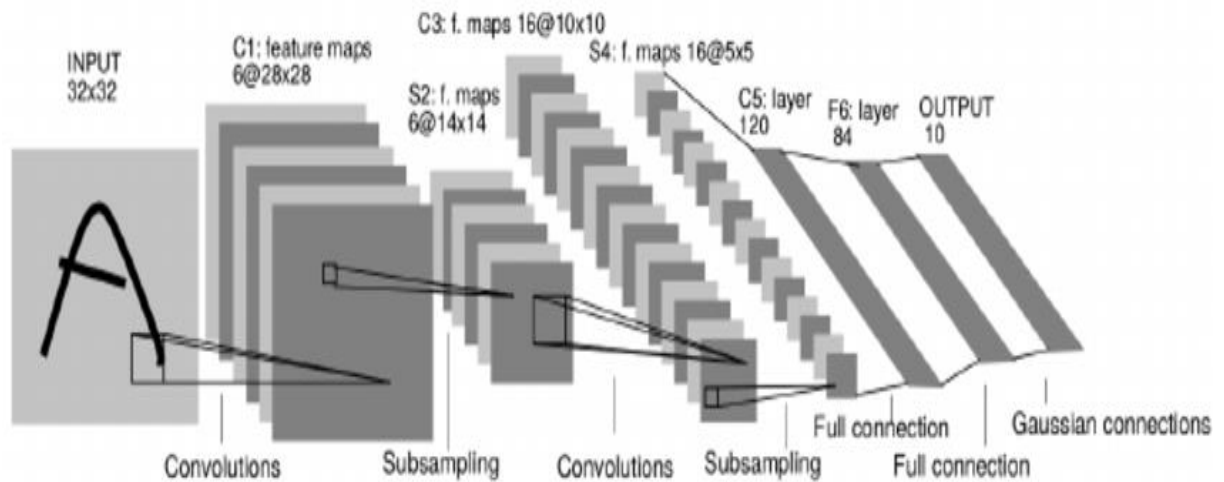


Figura 3. Rede LaNet-5

Fonte: Lécun et.al, 1998.

Após o desenvolvimento inicial da rede LeNet5, a rede neural convolucional entrou em uma fase de avanço experimental.

No entanto, não foi até a introdução da rede AlexNet em 2012 que as redes neurais convolucionais solidificaram sua posição no domínio das aplicações de aprendizado profundo.

A inovadora AlexNet proposta por Krizhevsky et al. (2012) alcançou um sucesso notável na classificação de imagens usando o conjunto de treinamento ImageNet.

Como resultado, as redes neurais convolucionais emergiram como um ponto focal para pesquisas em visão computacional, um campo que continua a se expandir e evoluir.

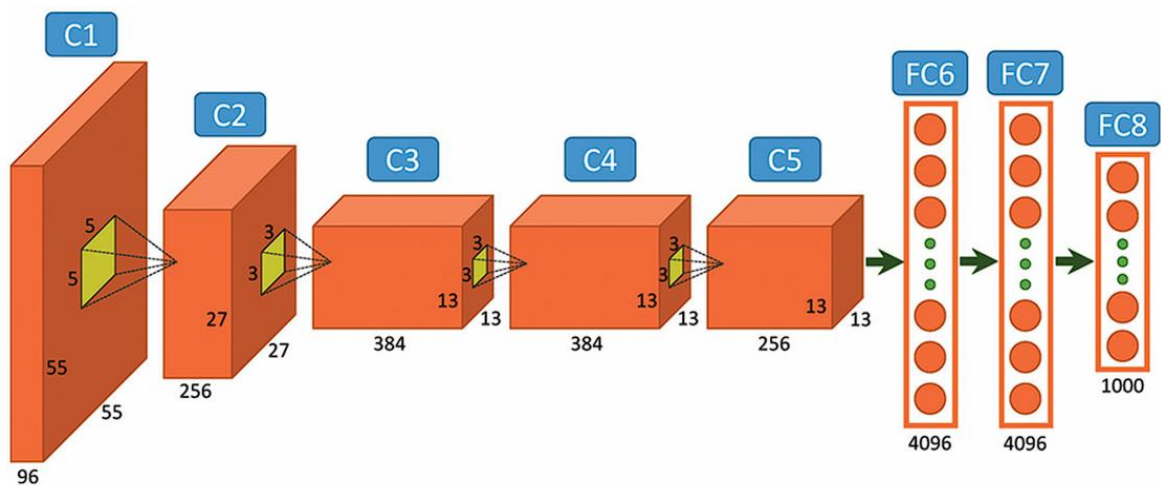


Figura 4. Rede AlexNet.

Fonte: S. Collet, Available Online: 2017. <https://www.saagie.com/blog/object-detection-part1/>

A estrutura típica de uma CNN inclui várias camadas, ou seja, a camada de entrada, a camada de saída e várias camadas ocultas. Dentro das camadas ocultas, cada camada realiza operações distintas como convolução, agrupamento e ativação (Ferreira, Gentil e Tavares, 2014).

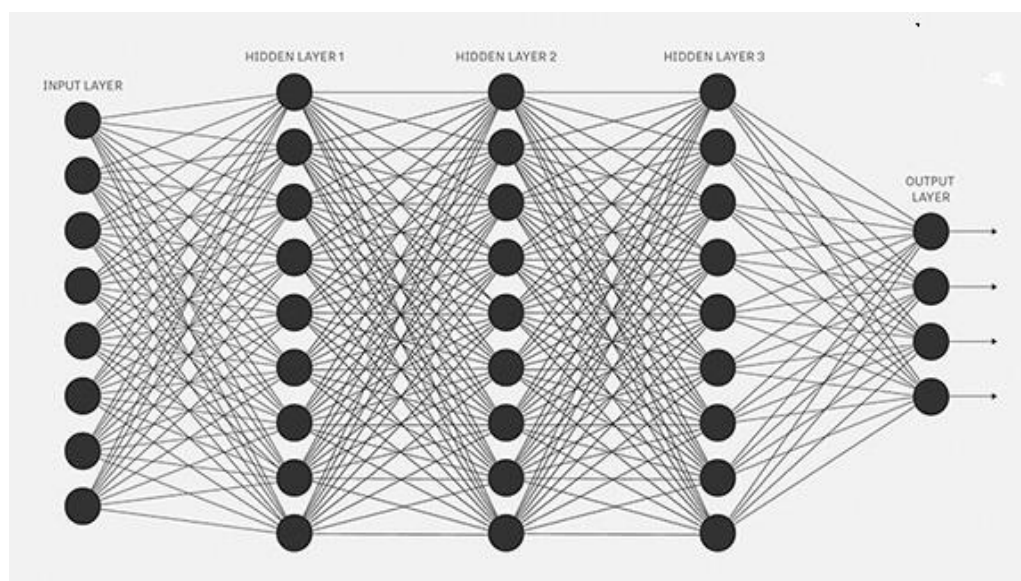


Figura 5. Rede neural convolucional de 3 camadas.

Fonte: Xiaofei, Yang et col.

A imagem de entrada está vinculada à camada de entrada, com o número de neurônios nessa camada correspondendo aos pixels na imagem de entrada.

Por meio de uma operação de convolução, a camada convolucional intermediária realiza a extração de recursos nos dados de entrada, produzindo um mapa de recursos. O resultado da operação de convolução é determinado pela configuração dos parâmetros dentro do kernel de convolução (Ma et al., 2010).

A camada convolucional é seguida pela camada de pooling, que serve para filtrar e escolher mapas de recursos, reduzindo efetivamente a complexidade computacional de toda a rede.

Na camada totalmente conectada, cada neurônio está conectado a todos os neurônios da camada anterior.

A saída resultante é então passada para o classificador, que determina o resultado da classificação.

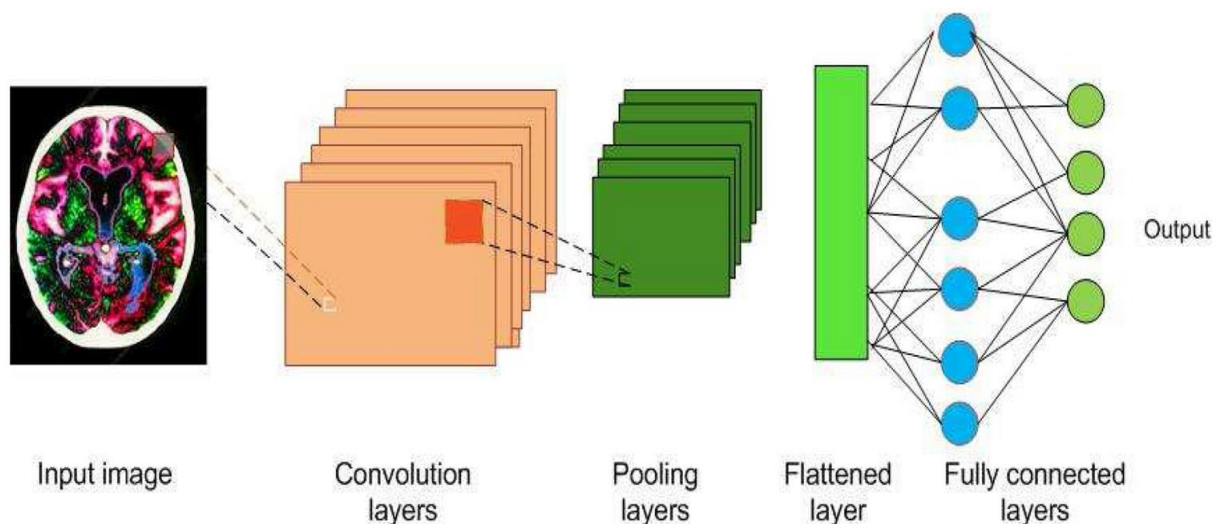


Figura 6. Típica rede neural convolucional e sua aplicação em imagiologia médica.

Fonte: Sarvamangala et al, 2016

## 2.2. REDES NEURAIS CONVOLUCIONAIS BIDIMENSIONAIS (CNN 2D)

A rede neural convolucional típica usada é uma CNN 2D, onde a imagem de entrada é 2D e o kernel de convolução também é 2D, como visto em modelos como ResNet e VGG.

Se considerarmos uma imagem de entrada com dimensões  $H \times W$  e três canais no formato RGB, o kernel de convolução que tem dimensões  $(c, h, w)$  se move pela dimensão espacial da imagem de entrada. Aqui,  $c$ ,  $h$  e  $w$  representam o número de canais, a altura e a largura do kernel de convolução, respectivamente. Para obter um valor, a operação de convolução é realizada em cada canal, inserindo o valor da imagem e o valor de  $(h, w)$ .

A imagem de entrada é dividida em 30 camadas, com cada camada representando um pixel.

A extração de feição é executada nos dados de entrada pela camada convolucional intermediária por meio de uma operação de convolução, resultando em um mapa de feição.

O resultado da operação de convolução é determinado pelos parâmetros definidos no kernel de convolução.

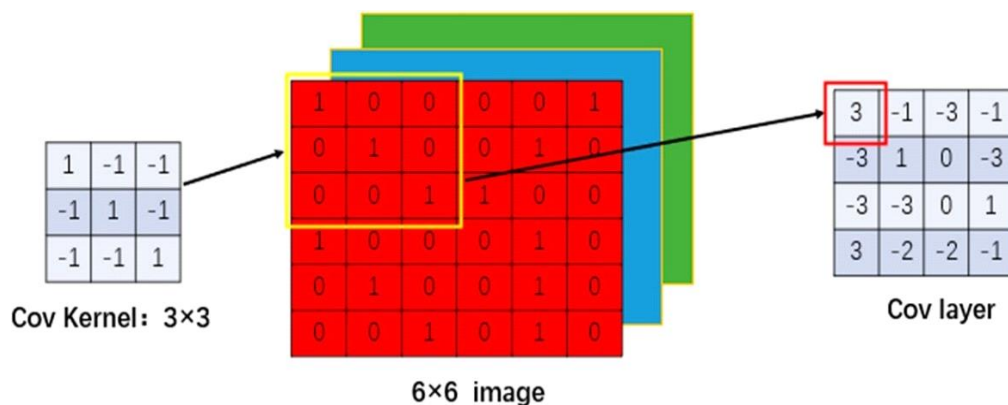


Figura 7. Rede neural convolucional de duas dimensões (2D CNN).

Fonte: Xiaofei, Yang et col.

### 2.3. REDES NEURAIS CONVOLUCIONAIS TRIDIMENSIONAIS (CNN 3D)

A maioria das imagens médicas, como tomografias computadorizadas e ressonâncias magnéticas, são normalmente apresentadas em formato tridimensional.

A imagem de TC convencional, que normalmente é visualizada, é uma representação 2D, fornecendo apenas uma única fatia da imagem geral.

Para segmentar com precisão os tecidos doentes, é necessário um kernel de convolução 3D. Um exemplo disso é a rede 3D U-Net, que utiliza um kernel de convolução 3D projetado especificamente para segmentação de imagens médicas 3D (Xu et al., 2010).

Ao incorporar a natureza tridimensional da imagem, uma CNN 3D pode extrair uma representação de volume mais robusta ao longo dos eixos X, Y e Z, aproveitando efetivamente os benefícios das informações espaciais.

Ao contrário de sua contraparte 2D, o kernel de convolução 3D inclui uma dimensão de profundidade adicional, correspondente ao número de fatias 2D dentro da imagem médica.

No caso de uma imagem 3D denotada como  $C \times N \times H \times W$ , onde C representa o número de canais, N representa o número de camadas de fatia e H e W representam a altura e a largura do kernel de convolução, o kernel de convolução 3D fornece uma abordagem abrangente para a análise de imagens (Ma et al., 2010).

Semelhante à operação de convolução 2D, o processo de obtenção de um valor envolve deslizar a janela pela altura, largura e número de camadas em cada canal em 3D.

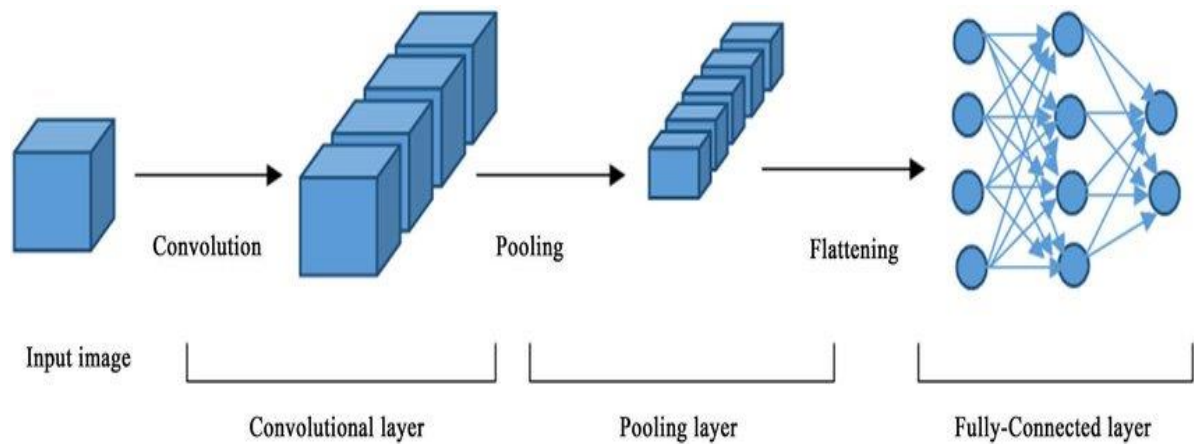


Figura 8. Uma arquitetura CNN 3D básica.

Fonte : Tasnin et al., 2020

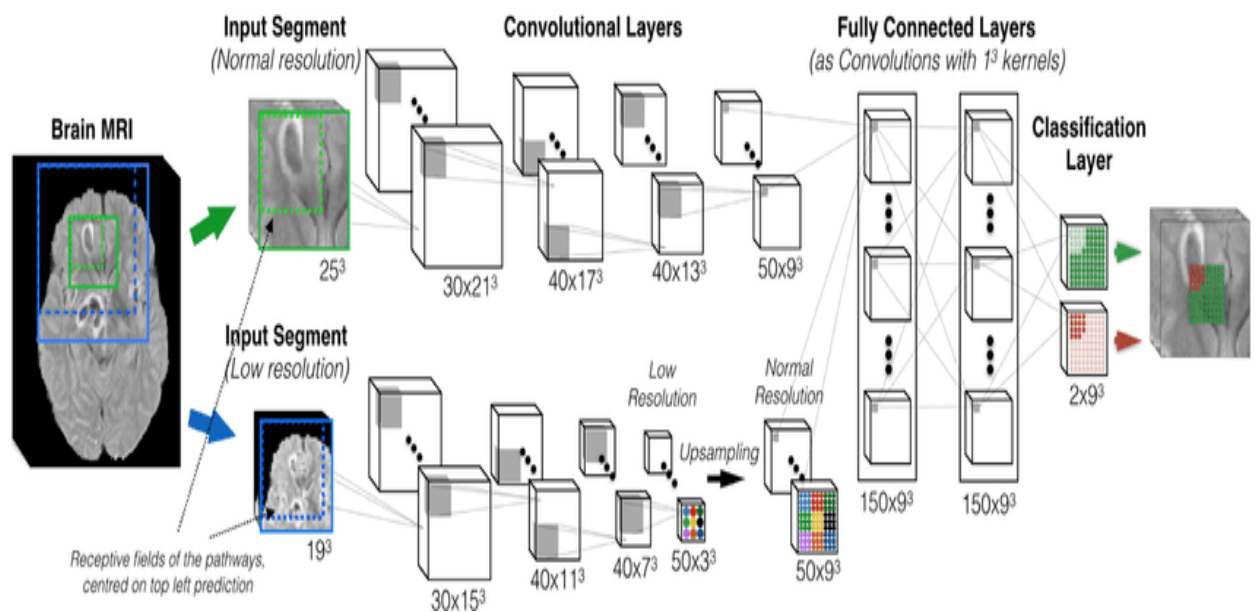


Figura 9. Aplicação de CNN 3D para detecção de lesão cerebral em RM.

Fonte: Kamnitsas et al., 2017



### 3. MÉTRICAS DE AVALIAÇÃO DE SEGMENTAÇÃO E CONJUNTOS DE DADOS

#### 3.1. MÉTRICAS DE AVALIAÇÃO

Avaliar a qualidade de um algoritmo requer um indicador objetivo correto. Em algoritmos de segmentação médica, as anotações desenhadas à mão pelos médicos são geralmente usadas como padrão-ouro (verdade fundamental, GT para abreviar).

Outros resultados da segmentação do algoritmo são os resultados da previsão (Rseg, SEG para abreviar). A avaliação de segmentação de imagens médicas é dividida em métodos baseados em pixels e baseados em sobreposição (He et al., 2017).

**Índice de Dados:** O coeficiente de dados é uma função para avaliar a similaridade. Geralmente é usado para calcular a semelhança ou sobreposição entre duas amostras. É também o mais usado. Seu intervalo de valores é de 0 a 1. Quanto mais próximo o valor estiver de 1, melhor será o efeito de segmentação. Dados dois conjuntos A e B, as métricas são definidas como:

$$\text{Dados}(A, B) = \frac{2|A \cap B|}{|A| + |B|}$$

**Índice de Jaccard:** O índice de Jaccard é semelhante ao coeficiente de dados. Dados dois conjuntos A e B, as métricas são definidas como:

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

**Precisão de Segmentação (SA):** A área de segmentação precisa é responsável pela porcentagem da área real na imagem GT. Entre eles, Rs representa a área de referência da imagem segmentada desenhada manualmente pelo especialista. Ts representa a área real da imagem obtida pela segmentação do algoritmo. |Rs - Ts| indica o número de pixels segmentados incorretamente (Lin et al., 2017).

$$SA = \left(1 - \frac{|Rs - Ts|}{Rs}\right) \times 100\%$$

**Taxa de Supersegmentação:** A proporção de pixels que são divididos na área de referência da imagem GT é calculada da seguinte forma:

$$OU = \frac{Os}{Rs + Os}$$

Os pixels em Os aparecem na imagem segmentada real, mas não aparecem na imagem segmentada teórica Rs. Rs representa a área de referência da imagem segmentada desenhada manualmente pelo especialista (Lin et al., 2017).

**Taxa de Subsegmentação:** A proporção do resultado da segmentação para os pixels ausentes em Imagem GT. Calculado da seguinte forma:

$$UR = \frac{Us}{Rs + Os}$$

Os pixels em Us aparecem na imagem segmentada teórica Rs, mas não aparecem na imagem segmentada real. Rs e Os têm o mesmo significado que acima (Noh, Hong e Han, 2015).

**Distância de Hausdorff:** Descreve uma medida do grau de semelhança entre dois conjuntos de pontos, ou seja, a distância entre os dois limites da verdade fundamental e a entrada do resultado da segmentação para a rede. Sensível ao limite dividido (Zhu, 2005).

$$H = \left( \max_{i \in \text{seg}} \left( \min_{j \in \text{gt}} d(i, j) \right), \max_{j \in \text{gt}} \left( \min_{i \in \text{seg}} d(i, j) \right) \right)$$

onde, i e j são pontos pertencentes a conjuntos diferentes. d representa a distância entre i e j.

### 3.2. CONJUNTOS DE DADOS PARA SEGMENTAÇÃO DE IMAGENS MÉDICAS

Para qualquer segmentação de modelo baseada em aprendizado profundo, é crucial coletar dados suficientes no conjunto de dados. A qualidade do algoritmo de segmentação depende dos dados de imagem de alta qualidade fornecidos pelos especialistas e do conjunto de dados padronizado por rótulos correspondente, o que permite uma comparação justa entre os sistemas. Esta seção apresentará alguns conjuntos de dados públicos frequentemente usados no campo da segmentação de imagens médicas (Zeiler e Fergus, 2014).

**Decatlo de Segmentação Médica (MSD):** Simpson et al. criaram um grande conjunto de dados de imagens médicas anotadas à mão de código aberto de várias partes anatômicas. Esse conjunto de dados pode avaliar objetivamente os métodos gerais de segmentação por meio de benchmarks abrangentes e tornar público o acesso a dados de imagens médicas. O conjunto de dados tem um total de 2633 imagens médicas tridimensionais, envolvendo aplicações clínicas reais de múltiplas estruturas anatômicas, vários modelos e múltiplas fontes (ou instituições).

É dividido em dez categorias ou tarefas (“tasks”) :

- **Task01\_BrainTumour:** Há um total de 750 imagens, e os rótulos são divididos em duas categorias: Glioma (tumor necrótico/ativo) e edema.
- **Task02\_Heart:** Há um total de 30 imagens, e o rótulo é o átrio esquerdo, uma vez que os dados são provenientes do Desafio de Segmentação do Átrio Esquerdo (LASC).
- **Task03\_Liver:** São 201 imagens no total, com rótulos divididos em fígado e tumores. O tipo de imagem é tomografia computadorizada.
- **Task04\_Hippocampus:** Há um total de 394 imagens, e os rótulos são hipocampo normal e hipocampo alterado. O tipo de imagem é ressonância magnética. O conjunto de dados consistiu em ressonância magnética adquirida em 90 adultos saudáveis e 105 adultos com transtorno psicótico não afetivo.

- **Task05\_Prostate:** Há um total de 48 imagens, e os rótulos são glândula central e zona periférica. O tipo de imagem é ressonância magnética. O conjunto de dados da próstata consistiu em 48 estudos de ressonância magnética multiparamétrica fornecidos pela Universidade Radboud (Holanda) relatados em um estudo de segmentação anterior.
- **Task06\_Lung:** Há um total de 96 imagens, e o rótulo é tumor de pulmão. O tipo de imagem é a tomografia computadorizada. O conjunto de dados pulmonares era composto por pacientes com câncer de pulmão de células não pequenas da Universidade de Stanford. A região do tumor foi indicada por um radiologista torácico, em uma seção transversal representativa da tomografia, usando OsiriX.
- **Task07\_Pancreas:** Há um total de 420 imagens, com rótulos divididos em pâncreas e massa pancreática (cisto ou tumor). O tipo de imagem é a tomografia computadorizada. O conjunto de dados do pâncreas consistia em pacientes cujas massas pancreáticas foram removidas.
- **Task08\_HepaticVessel:** Há um total de 443 imagens, e os rótulos são vasos hepáticos. O tipo de imagem é a TC. Este segundo conjunto de dados hepáticos consistia em pacientes com vários tumores hepáticos primários e metastáticos.
- **Task09\_Spleen:** Há um total de 61 imagens, e o rótulo é o baço. O tipo de imagem é a TC. O conjunto de dados do baço é composto por pacientes submetidos a tratamento quimioterápico para metástases hepáticas no Memorial Sloan Kettering Cancer Center.
- **Task10\_Colon:** Há um total de 190 imagens, e o rótulo é câncer de cólon. O tipo de imagem é a TC. (Gu et al., 2018).

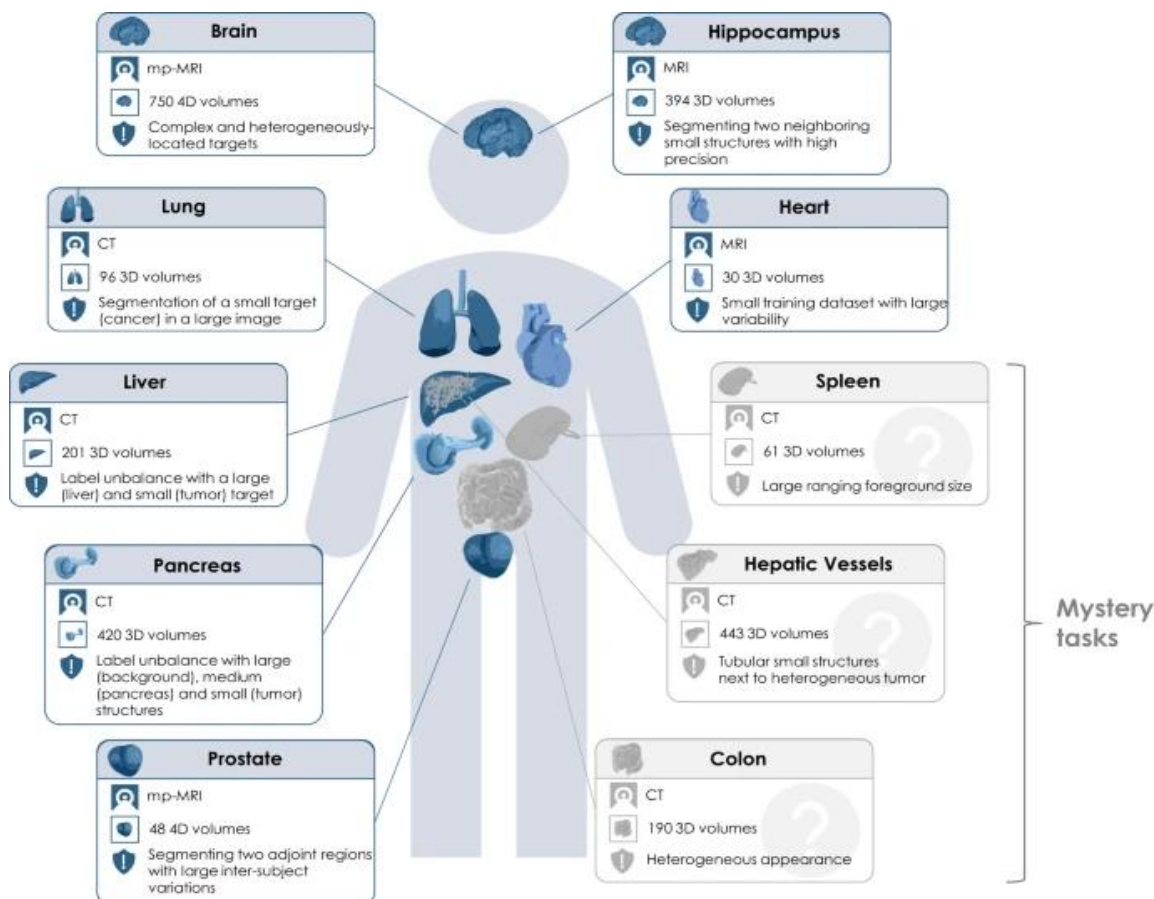


Figura 10. Desafio de dez tarefas do *Medical Segmentation Decathlon*.

Fonte: <https://www.nature.com/articles/s41467-022-30695-9>

**Segmentação em Radiografias de Tórax (SCR):** Todas as radiografias de tórax são retiradas do banco de dados da *Japanese Society of Radiological Technology*. O banco de dados SCR foi criado para simplificar o estudo comparativo da segmentação do campo pulmonar, coração e clavícula em radiografias de tórax posterior padrão. Todos os dados no banco de dados são segmentados manualmente para fornecer padrões de referência. A imagem é digitalizada de filme para 2048 x 2048 pixels, com uma resolução espacial de 0,175 mm/pixel e uma escala de cinza de 12 bits. Cada uma das 154 imagens tem um nódulo pulmonar e as outras 93 imagens não têm nódulos pulmonares (Rundo et al., 2019).

**Segmentação de Tumor Cerebral (BRATS):** É um conjunto de dados de competição para segmentação de tumores cerebrais, que é combinado com a conferência MICCAI (*Medical Image Computing and Computer Assisted Intervention Society*), uma das mais tradicionais instituições da área de computação e medicina. Com o objetivo de avaliar os melhores métodos de segmentação de tumores cerebrais e comparar diferentes métodos, é realizado todos os anos desde 2012. Existem cinco tipos de rótulos: tecido cerebral saudável, área necrótica, área de edema, realce do tumor e área sem realce. Novos conjuntos de treinamento são adicionados a cada ano (Qiu, Yao e Mei, 2017).

**Banco de Dados Digital para Mamografia de Rastreamento (DDSM):** O DDSM é um recurso utilizado pela comunidade de pesquisa em análise de imagens de mamografia e é amplamente utilizado por pesquisadores. O banco de dados contém aproximadamente 2500 estudos. Cada estudo inclui duas imagens de cada mama, bem como algumas informações relevantes do paciente e informações de imagem (Heath, M.; Bowyer, K.; Kopans, D.; Kegelmeyer, P.; Moore, R.; Chang, K.; Munishkumaran, S., 2000).

**Segmentação de Lesão de AVC Isquêmico (ISLES):** Fornece exames de ressonância magnética contendo um grande número de amostras precisas de AVC e parâmetros clínicos relacionados. Este desafio é organizado para avaliar a fisiopatologia do AVC e a previsão de resultados clínicos em imagens precisas de ressonância magnética (Heath, M.; Bowyer, K.; Kopans, D.; Kegelmeyer, P.; Moore, R.; Chang, K.; Munishkumaran, S., 2000).

**Segmentação de Tumor Hepático (LiTS):** Esses dados são fornecidos por diferentes centros clínicos em todo o mundo para a segmentação de tumores hepáticos e hepáticos. O conjunto de dados de treinamento contém 130 tomografias computadorizadas e o conjunto de dados de teste contém 70 tomografias computadorizadas (Bilic, P.; Christ, P.F.; Vorontsov, E.; Chlebus, G.; Chen, H.; Dou, Q.; Fu, C.-W.; Han, X.; Heng, P.-A.; Hesser, J.; et al, 2019).

**Segmentação de Imagem de RM da Próstata (PROMISE12):** Este conjunto de dados é usado para segmentação da próstata e incluem pacientes com doenças benignas (hiperplasia benigna da próstata) e câncer de próstata. Esses casos incluem uma imagem de RM transversal da próstata ponderada em T2 (Bilic, P.; Christ, P.F.; Vorontsov, E.; Chlebus, G.; Chen, H.; Dou, Q.; Fu, C.-W.; Han, X.; Heng, P.-A.; Hesser, J.; et al, 2019).

**Coleção de Imagens do Consórcio do Banco de Dados de Imagens Pulmonares (LIDC-IDRI):** O conjunto de dados é composto por arquivos de imagens médicas de tórax (como tomografia computadorizada e radiografia) e rótulos de lesões de resultados de diagnóstico correspondentes. O objetivo é estudar a detecção precoce do câncer em populações de alto risco. Um total de 1018 exemplos de pesquisa estão incluídos. Para as imagens em cada exemplo, quatro radiologistas torácicos experientes realizaram um diagnóstico e anotação em dois estágios (Armato, S.G., III; McLennan, G.; Bidaut, L.; McNitt-Gray, M.F.; Meyer, C.R.; Reeves, A.P.; Zhao, B.; Aberle, D.A.; Henschke, C.I.; Hoffman, E.A.; et al, 2011)

**Série de Estudos de Imagem de Acesso Aberto (OASIS):** Este é um projeto que visa permitir que a comunidade científica forneça conjuntos de dados de ressonância magnética cerebral gratuitamente. Uma terceira geração foi lançada. O OASIS-3 é uma compilação retrospectiva de dados de mais de 1000 participantes coletados de vários projetos em andamento por meio do WUSTL Knight ADRC nos últimos 30 anos. O OASIS-3 é um conjunto de dados longitudinais de neuroimagem, clínicos, cognitivos e biomarcadores para o envelhecimento normal e a doença de Alzheimer.

Os participantes incluíram 609 adultos cognitivamente normais e 489 pessoas em vários estágios de declínio cognitivo, com idades entre 42 e 95 anos (Marcus, D.S.; Fotenos, A.F.; Csernansky, J.G.; Morris, J.C.; Buckner, R.L., 2010).

**Imagens Digitais da Retina para Extração de Vasos (DRIVE):** Este conjunto de dados é usado para comparar a segmentação dos vasos sanguíneos nas imagens da retina. As fotos no banco de dados DRIVE vieram de um projeto de triagem de retinopatia diabética na Holanda, e 40 fotos foram selecionadas aleatoriamente.

Entre eles, 33 casos não apresentavam sinais de retinopatia diabética e sete casos apresentavam sinais de retinopatia diabética precoce leve. Cada imagem é capturada com  $768 \times 584$  pixels com 8 bits por plano de cor. O campo de visão de cada imagem é circular com um diâmetro de aproximadamente 540 pixels (Staal, J.; Abràmoff, M.D.; Niemeijer, M.; Viergever, M.A.; Van Ginneken, B. Ridge, 2004).

**Sociedade de Análise de Imagens Mamográficas (MIAS):** O MIAS é um banco de dados de imagens mamográficas apresentando neoplasia de mama criado por uma organização de pesquisa britânica em 1995. Cada pixel tem uma escala de cinza de 8 bits. O banco de dados MIAS contém imagens de mama esquerda e direita de 161 pacientes, com um total de 322 imagens, incluindo 208 imagens saudáveis, 63 imagens de câncer de mama benigno e 51 imagens de câncer de mama maligno. O limite da área da lesão também foi calibrado por especialistas (Suckling, J.P., 1994).

**Dados Cardíacos de Sunnybrook (SCD):** Também conhecidos como dados de desafio de segmentação do ventrículo esquerdo de RM cardíaca de 2009 e consistem em 45 imagens de cine-ressonância magnética com uma variedade de pacientes e patologias: saudável, hipertrofia, insuficiência cardíaca com infarto e insuficiência cardíaca sem infarto (Fonseca, C.G.; Backhaus, M.; Bluemke, D.A.; Britten, R.D.; Chung, J.D.; Cowan, B.R.; Dinov, I.D.; Finn, J.P.; Hunter, P.J.; Kadish, A.H.; et al., 2011).



## 4. PRINCIPAIS REDES DE SEGMENTAÇÃO DE IMAGENS MÉDICAS

### 4.1.1. Redes Totalmente Convolucionais (FCN)

FCN's (*Fully Convolutional Networks*) representam uma arquitetura desenvolvida principalmente para segmentação semântica. As camadas usadas são conectadas apenas localmente — convolução, *pooling* e *upsampling*. Ao não usar camadas densas, o número de parâmetros é reduzido (as redes são, portanto, mais rápidas de treinar) e uma FCN pode receber imagens de tamanhos diferentes porque todas as conexões são locais. A arquitetura integra um caminho de *downsampling*, por meio do qual o contexto é extraído e interpretado, e um caminho de *upsampling* para localização.

Nesta arquitetura são utilizadas conexões de salto para o mesmo que recuperar informações espaciais de granulação fina perdidas no caminho de *downsampling*. Essas redes convolucionais são modelos visuais com hierarquias de recursos extremamente poderosas.

Uma vez que uma rede convolucional padrão, treinada de ponta a ponta, em pixels, pode ter um desempenho melhor na segmentação semântica de imagens do que os recursos de engenharia manual de última geração.

Versões totalmente convolucionais das redes de classificação rápida contemporâneas — AlexNet, VGG net e GoogLeNet — aprendem extrações densas de recursos e são diretamente aplicáveis à tarefa de classificação em pixels.

Na arquitetura do FCN, camadas de desconvolução são usadas para que o mapa de características da camada de convolução final possa ser ampliado para corresponder ao tamanho da imagem de entrada — permitindo, portanto, previsões para cada pixel, preservando as informações espaciais dentro da imagem de entrada original.

O mapa de características ampliado é então classificado pixel a pixel na segmentação da imagem final. Ele é baseado no passo de ampliação e é dividido em FCN-32s, FCN-16s e FCN-8. (Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440).

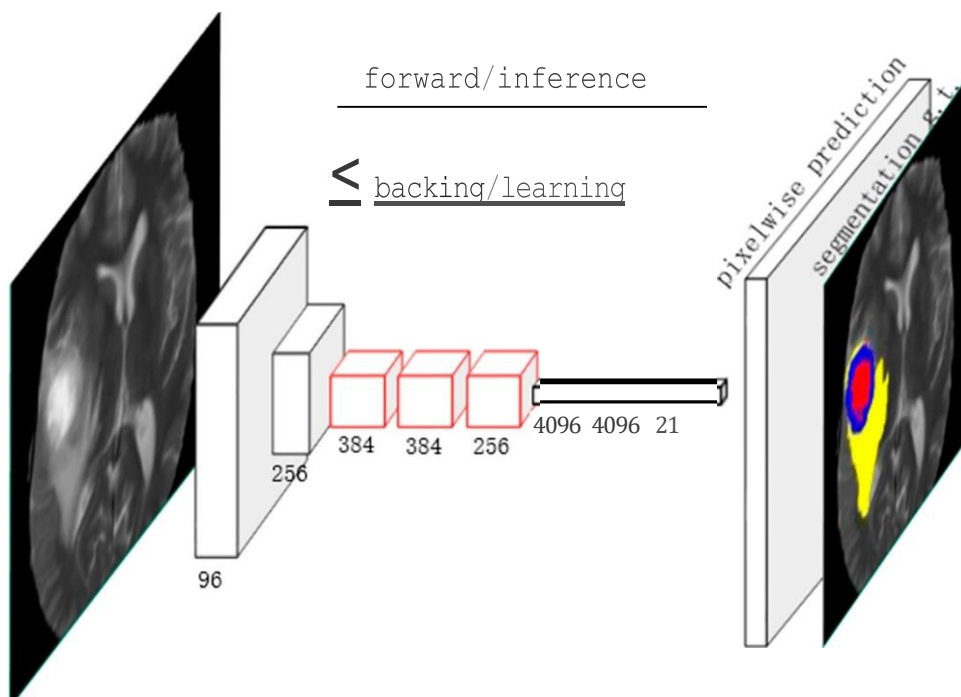


Figura 11. Estrutura de rede totalmente convolucional (FCN)

Fonte: Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.

#### 4.1.2 DEEP LAB v1

Por outro lado, a FCN também apresenta problemas.

Os resultados da segmentação podem ser relativamente confusos e seu *upsampling* pode não ser sensível aos detalhes da imagem podendo resultar em aquisições grosseiras.

Em segundo lugar a ideia da segmentação é essencialmente classificar cada pixel sem consideração completa. A relação entre os pixels carece de consistência espacial.

Para obter um mapa de pontuação mais denso no FCN, pode ser adicionado preenchimento à primeira camada convolucional. O tamanho do preenchimento é igual a 100, o que trará muito ruído.

Chen et al. propuseram o DeepLab v1, que alterou o passo de agrupamento do original 2 para 1 e o tamanho do preenchimento do original 100 para 1. Dessa forma, o tamanho da imagem agrupada não é reduzido e o resultado do mapa de pontuação obtido é mais denso do que o do FCN.

O DeepLab v1 é reescrito com base na rede VGG-16, removendo a última camada totalmente conectada da rede VGG e usando convolução completa em vez disso, porque usar muitas camadas de agrupamento resultará em um tamanho de camada de recurso muito pequeno.

Os recursos contidos são muito esparsos, o que não é propício à segmentação semântica.

Finalmente, o DeepLab v1 usa campo aleatório condicional (CRF) para melhorar a precisão dos limites de segmentação.( Krähenbühl, P.; Koltun, V. Efficient inference in fully connected crfs with gaussian edge potentials. Adv. Neural Inf. Process. Syst. 2011, 24, 109–117).

### 4.1.3 DeepLab v2

DeepLab v2 abordou as dificuldades de segmentação devido a variações de escala do mesmo objeto dentro da mesma imagem.

Quando o objeto semelhante tem tamanhos variantes na mesma imagem ou em imagens diferentes a maneira comum é garantir que as imagens tenham o mesmo tamanho por meio do redimensionamento embora isso possa levar a distorções por perdas de pixels. A inovação no DeepLab v2 vem do uso da convolução atrous (ou dilatada) de uma maneira mais flexível, que os autores propuseram como agrupamento de pirâmide atrous espacial (ASPP).

O ASPP é baseado na ideia de agrupamento de pirâmide espacial (SPP) para derivar uma estrutura para amostragem convolucional paralela de buracos em taxas diferentes de uma maneira equivalente para capturar contexto em imagens em múltiplas escalas.

Em seu trabalho, os autores do DeepLab v2 escolheram uma rede mais complicada e expressiva baseada no ResNet-101. Os recursos de alto nível obtidos após passar pelas camadas profundas do DCNN são espacialmente subamostrados por meio de operações de agrupamento.

Ele extrai recursos de alta resolução porque o DeepLab v2 descarta o subamostramento no conjunto final de camadas de agrupamento máximo e o substitui pela convolução atrous para calcular mapas de recursos com uma densidade de amostragem mais alta.

Eles também retiraram a camada totalmente conectada da rede e colocaram uma camada totalmente convolucional, usando um CRF totalmente conectado para melhorar a precisão do limite para segmentação.

(Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 40, 834–848).

#### 4.1.4 SEGNET

Ele segue a arquitetura de codificador-decodificador simétrico do FCN para segmentação semântica visando atingir segmentação de imagem de ponta a ponta no nível de pixel. A estrutura é dividida principalmente em duas partes: o codificador e o decodificador.

Na segmentação semântica, as redes tomam, como entrada, uma imagem RGB e têm um rótulo que é uma imagem de  $n$  canais. Aqui,  $n$  é o número de rótulos envolvidos. Cada canal corresponde a um rótulo (carros, estradas, etc.) e cada pixel em um determinado canal será 1 ou 0, dependendo se o pixel pertence ao rótulo correspondente a esse canal.

Com base nas informações analisadas no processamento, ele forma uma nova imagem, onde cada pixel é associado a uma cor ou um rótulo para identidade.

A novidade neste caso está em como o decodificador conduz o *upsampling* do mapa de recursos de entrada que ele recebe de um nível mais baixo de resolução, enquanto no FCN utiliza-se a desconvolução para *upsampling*.

Finalmente, um kernel de convolução treinável é aplicado para a operação de convolução para produzir um mapa de recursos denso.

Após a restauração dos mapas de características para sua resolução original, eles são alimentados para o classificador *softmax* para classificação em nível de pixel.

Isso reservou a integridade das informações de alta frequência, caráter aprimorado das bordas e parâmetros reduzidos de treinamento, mas o *depooling* também ignora as informações adjacentes de mapas de características de baixa resolução.

(Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39, 2481–2495)

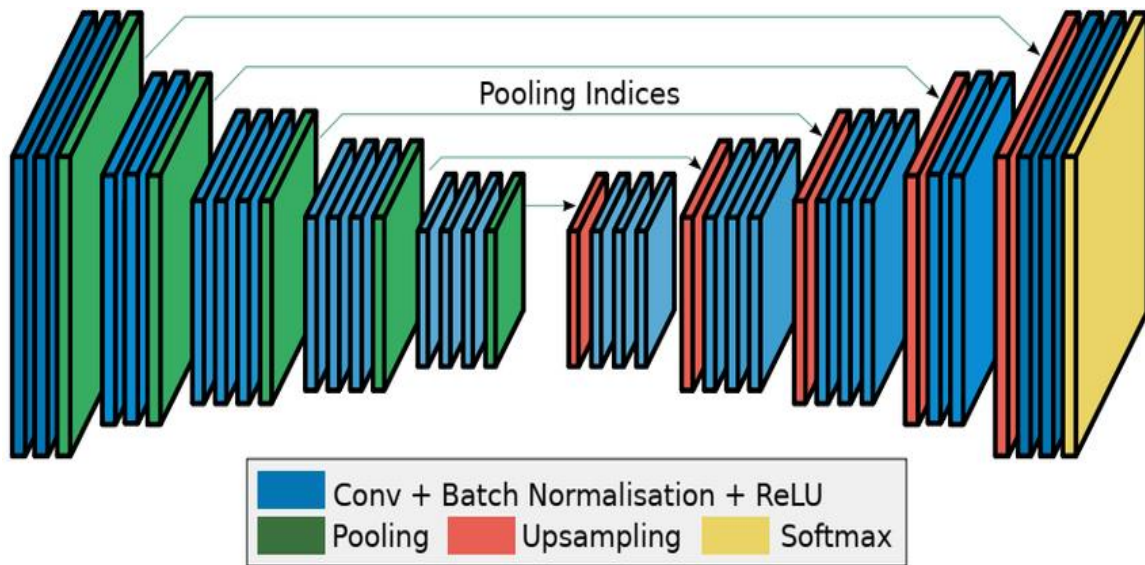


Figura 12. Segnet

Fonte: (Badrinarayanan et al., 2017).

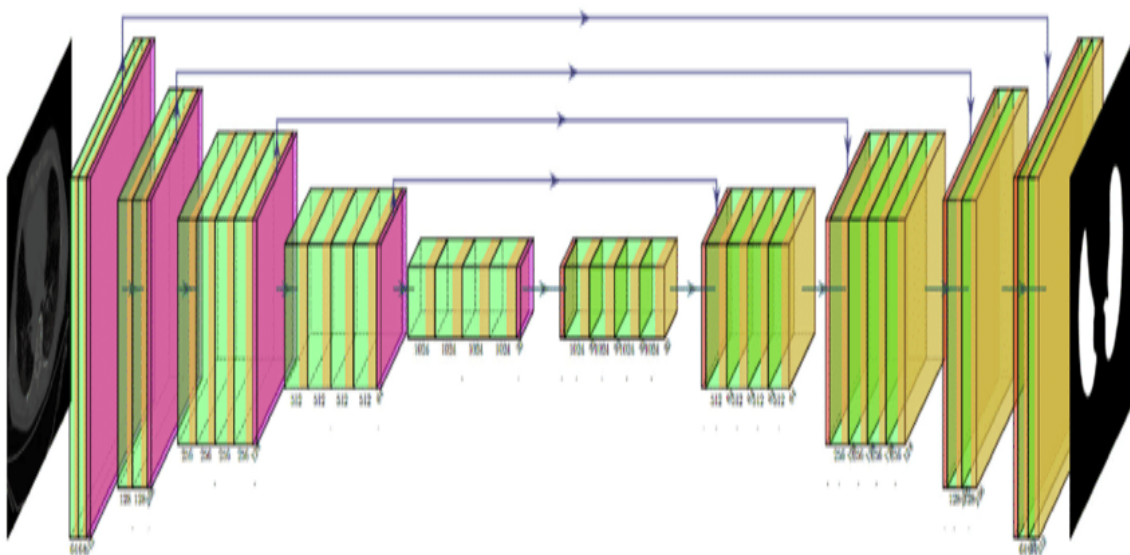


Figura 13. VGGSegnet usada para classificação de pneumopatia intersticial em TC

Fonte: Badrinarayanan et al., 2017.

## 4.2 U-NET

### 4.2.1 U-NET 2D

Baseia-se no paradigma da segmentação semântica, em que há classificação por pixel, ou seja, não se considera a imagem como um todo.

Após a FCN, Ronneberger et al. introduziram ainda mais uma rede específica para imagens médicas: U-Net. Eles nomearam sua proposta após a qual ela encontrou ampla aplicabilidade na tarefa de segmentação de imagens médicas em muitas outras aplicações.

Devido ao seu alto desempenho e à capacidade de capturar detalhes finos em grandes distâncias, entre muitas outras tarefas (e mais), variantes da U-Net têm visto ampla aplicação em diferentes subcampos da visão computacional.

Foi apresentada na conferência MICCAI 2015 e recebida com entusiasmo, com milhares de citações, encorajando o desenvolvimento de uma ampla gama de variantes e aplicações. (Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241).

Embora existam muitas novas maneiras de projetar redes neurais convolucionais, a maioria delas têm se baseado na ideia central da U-Net, que é conhecida por ter canais U e conexões de salto.

O codificador aqui segue a mesma estrutura do SegNet, apresentando quatro submódulos, cada um compreendendo duas camadas convolucionais. Após cada submódulo, o pool máximo é executado para downsampling.

O decodificador também tem quatro submódulos. O *upsampling* aumenta a resolução progressivamente até que as previsões sejam feitas para cada valor de pixel.

Uma entrada de tamanho 572 x 572 é alimentada nela, e uma dimensão de 388 x 388 é obtida como saída.

A saída sendo menor que a entrada é principalmente porque lida com segmentação na área médica, que precisa ser mais precisa.

Como visto na figura 6, esta rede não contém nenhuma camada totalmente conectada; há apenas convoluções e *downsamplings*, com uma conexão de salto para conectar o resultado de upsampling à saída do submódulo com a mesma resolução no codificador que a entrada do próximo submódulo no decodificador que ele usa. Isso significa que esse segmento será mais preciso.

A razão pela qual a U-Net é apropriada para segmentação de imagens médicas é que informações de baixo e alto nível podem ser mescladas ao mesmo tempo por sua estrutura; as informações de baixo nível auxiliam na melhoria da precisão, enquanto as de alto nível auxiliam na extração de características de imagem mais detalhadas.

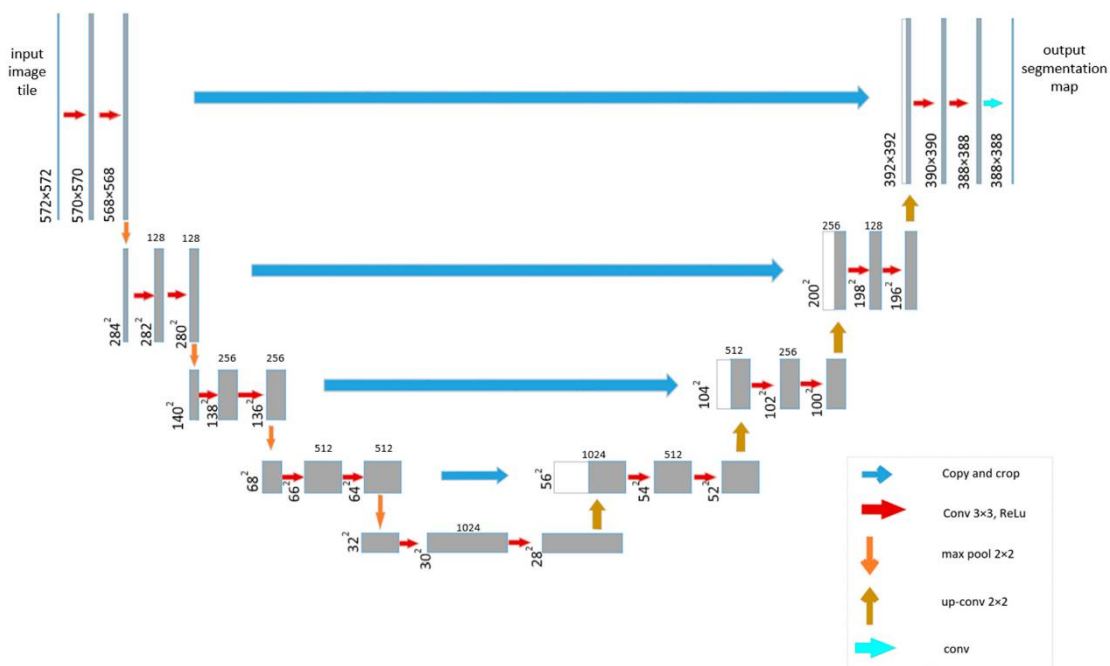


Figura 14. Arquitetura de uma U-Net

Fonte : Ronnenberger et al., 2015.



#### 4.2.2. 3D U-NET

A melhoria propiciada pela U-Net alterou o foco de atenção na pesquisa em segmentação de imagens médicas e muitas variantes foram desenvolvidas, com destaque para Çiçek et al., que propuseram um modelo 3D U-Net.

Este modelo visa fazer com que a estrutura U-Net tenha informações espaciais mais ricas. A estrutura de rede tridimensional é semelhante à U-Net, com um caminho de codificação e um caminho de decodificação.

Cada caminho tem quatro níveis de resolução. Cada camada no caminho de codificação contém duas convoluções  $3 \times 3$ , seguidas por uma camada ReLU.

Utilizando uma camada de pooling máxima para reduzir a dimensionalidade.

No caminho de decodificação, cada camada contém uma camada de desconvolução  $2 \times 2 \times 2$  com um passo de 2, seguida por duas camadas de convolução  $3 \times 3 \times 3$ . Cada convolução é seguida por uma camada ReLU.

Por meio de um atalho, a camada com a mesma resolução no caminho de codificação é passada para o caminho de decodificação, fornecendo a ela recursos originais de alta resolução.

Esta rede não só pode treinar em um conjunto de dados escassamente rotulados e prever outros lugares não rotulados neste conjunto de dados, mas também treinar em vários conjuntos de dados escassamente rotulados e então prever novos dados. (Çiçek, Ö.; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 17–21 October 2016; pp. 424–432).

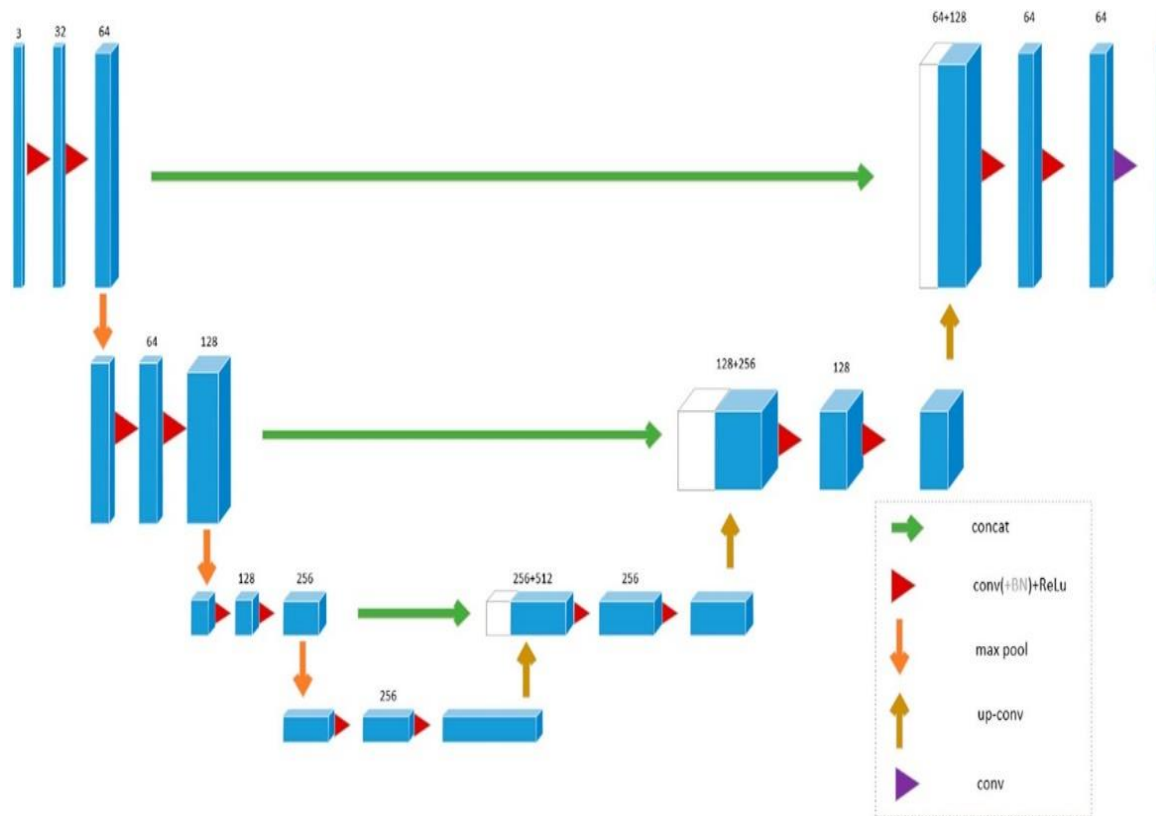


Figura 15. Arquitetura de uma rede 3D U-Net

Fonte: Çiçek et al., 2016.

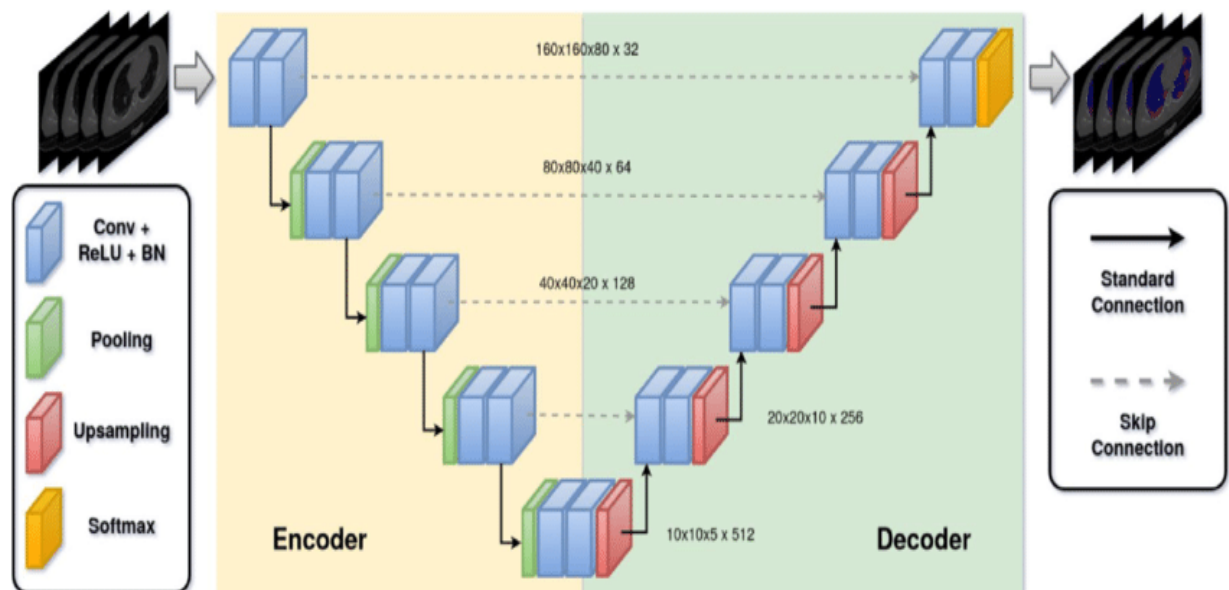


Figura 16: 3D UNet para segmentação de pneumopatia intersticial em imagem de TC

Fonte: Muller et al., 2018.

### 4.2.3. V-NET

Milletari et al. propuseram uma estrutura de deformação 3D V-Net da estrutura de rede U-Net. A arquitetura do V-Net utiliza a função de perda de coeficiente Dice, diferentemente da função de perda de entropia cruzada clássica.

Ele faz o kernel convolucional 3D da imagem e reduz a dimensão do canal por meio de um kernel de convolução  $1 \times 1 \times 1$  no lado esquerdo do caminho gradualmente comprimido da rede, que é dividido em muitos estágios, com cada estágio contendo de uma a três camadas convolucionais para fazer com que cada estágio aprenda uma função de parâmetro, as entradas e saídas de cada estágio são adicionadas para atingir o aprendizado da função residual.

A operação de convolução é usada para extrair características dos dados, enquanto ao mesmo tempo no final de cada "estágio" através do tamanho de passo apropriado a resolução dos dados é reduzida.

No lado direito, a rede é um suporte espacial descompactado gradualmente de recurso de resolução mais baixa para segmentação de volume de saída.

O tamanho de saída final da rede é consistente com o tamanho de entrada original. Um kernel é usado para convolver a imagem e reduz a dimensão do canal através de um kernel de convolução  $1 \times 1 \times 1$ .

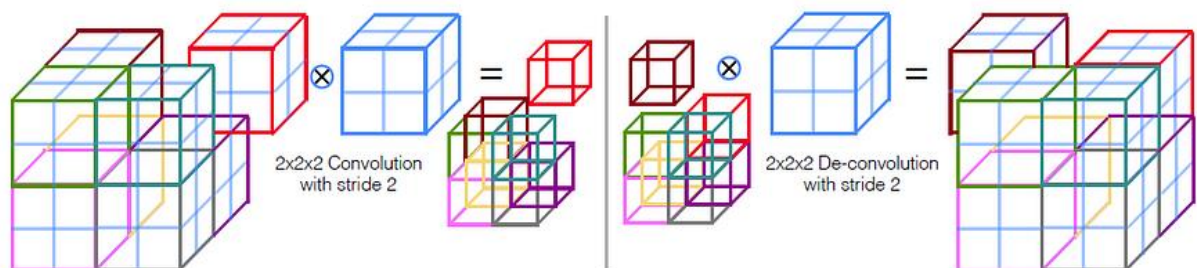
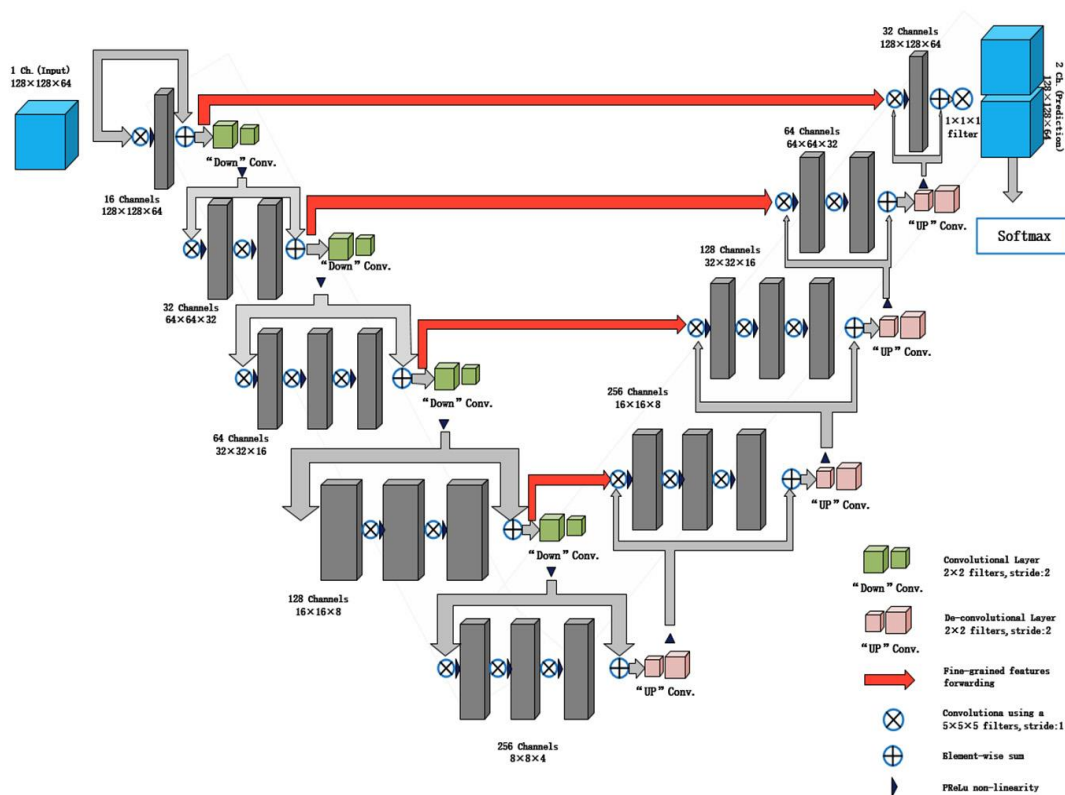
No lado esquerdo da rede há um caminho gradualmente compactado, dividido em muitos estágios, cada um contendo de uma a três camadas convolucionais.

Para fazer cada estágio aprender uma função de parâmetro, adicionamos entrada e saída de cada estágio para obter aprendizado da função residual.

Os dados são passados pela operação de convolução que extrai características dos dados e reduz a resolução no final de cada "estágio" nos dados pelo tamanho de

passo apropriado. Um caminho gradualmente descompactado está na extremidade direita da rede.

Deve-se notar que o tamanho final da saída da rede deve ser o mesmo que o tamanho da entrada original.( Milletari, F.; Navab, N.; Ahmadi, S.-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp.565–571).



Figuras 17 e 18: Rede V-Net. *Downsampling* (esquerda) e *Upsampling* (direita)

Fonte : Milletari et al., 2016.

#### 4.2.4. OUTRAS ESTRUTURAS U-NET

Res-UNet (Weighted Res-UNet) e H-DenseUNet são baseadas em conexões residuais e conexões densas, substituindo cada submódulo do U-Net por uma conexão residual e uma conexão densa, respectivamente. Aplicado à segmentação de imagens para vasos sanguíneos da retina está o Res-UNet.

Apresentando um desafio adequado para Xiao et al., que propuseram um Res-UNet ponderado projetado especificamente com base em um modelo U-Net original com mecanismos de atenção ponderada adicionais ajudando a distinguir características mais aprendidas sobre a manutenção de estruturas de árvore separadas de pixels de vasos. Problemas comumente observados na segmentação de vasos são a ausência de pequenos vasos sanguíneos e a segmentação ruim do disco óptico.

Isso resulta de uma tentativa de usar técnicas padrão na segmentação de uma estrutura que se parece com uma estrutura de bifurcação em forma de árvore, onde os vasos sanguíneos são muito finos para serem detectados.

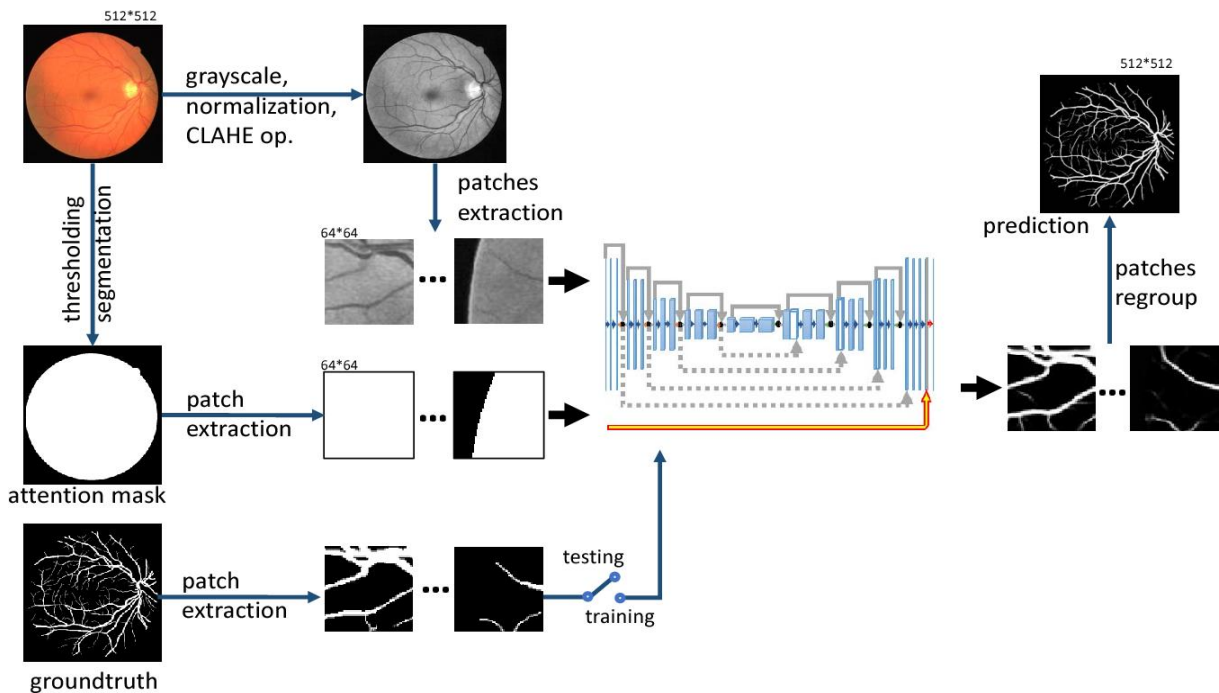


Figura 19: ResUNet e a detecção de finos vasos retinianos.

Fonte: Xiaoli Xiang et al., 2018

O H-DenseUNet representa uma nova abordagem para explorar eficientemente recursos representativos de alto nível *intra-slice* e *inter-slice*, seguido pela otimização dos recursos por meio da camada de fusão de recursos híbridos. A arquitetura lidou elegantemente com os problemas de que as convoluções 2D ignoram contextos volumétricos e as convoluções 3D têm uma enorme carga computacional. Experimentos extensivos no conjunto de dados LiTS e 3DIRCADb de 2017 provaram a eficácia do modelo H-DenseUNet (Xiaomeng Li et al, 2017).

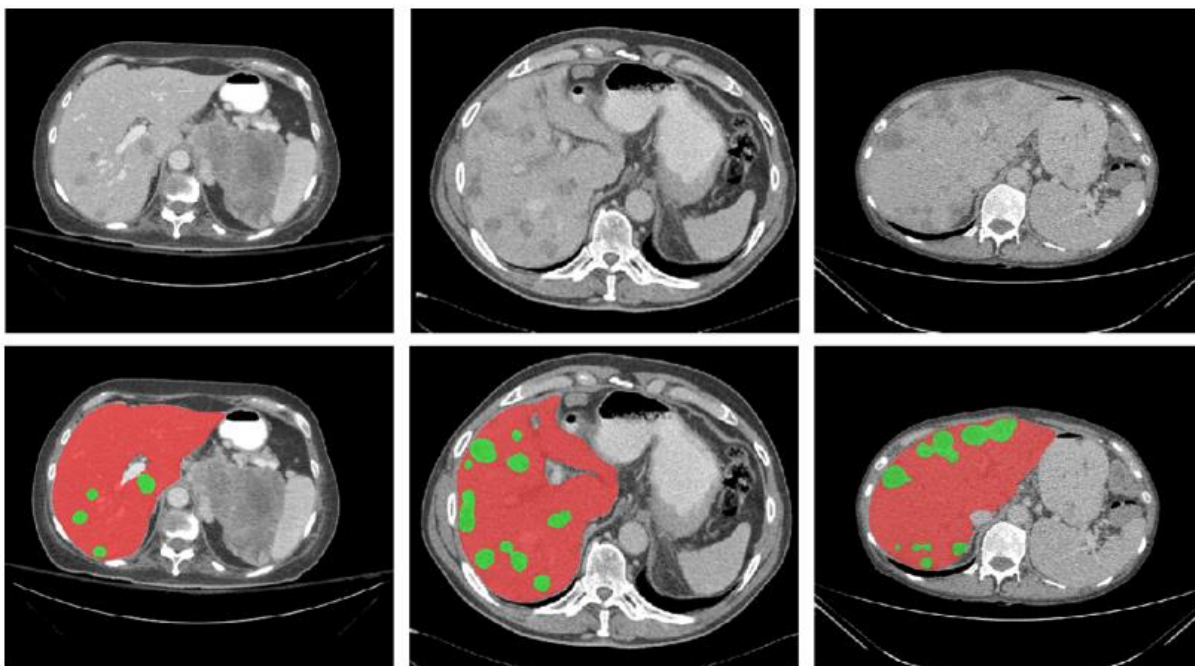


Figura 20: H-DenseUNet aplicada na segmentação de imagens tumorais em tomografias computadorizadas. A cor vermelha representa hepatócitos saudáveis, enquanto as verdes mostram células tumorais.

Fonte: M. Moghbel et al., 2017.

### 4.3 REDES GENERATIVAS ADVERSARIAS (GAN'S)

Em 2014 Goodfellow et al. propuseram uma abordagem adversarial para aprender um modelo generativo profundo — Generative Adversarial Network (GAN). Sua estrutura é composta de duas partes; para a primeira parte — esta é uma rede de geração onde algum ruído aleatório  $z$  (um número aleatório) serve como entrada e a partir deste ruído de entrada uma imagem é gerada. A segunda parte é a rede adversarial: usada para julgar se uma imagem é "real" ou não. Seu parâmetro de entrada é  $x$  (uma imagem), e a saída  $D(x)$  representa a probabilidade de que  $x$  seja uma imagem real. Em palavras simples, é por meio do treinamento que duas redes competem entre si. A rede de geração produz dados falsos; por outro lado, o adversário usa um discriminador para verificar a autenticidade.

Eventualmente, espera-se que o valor gerado pelo gerador possa ser falso.

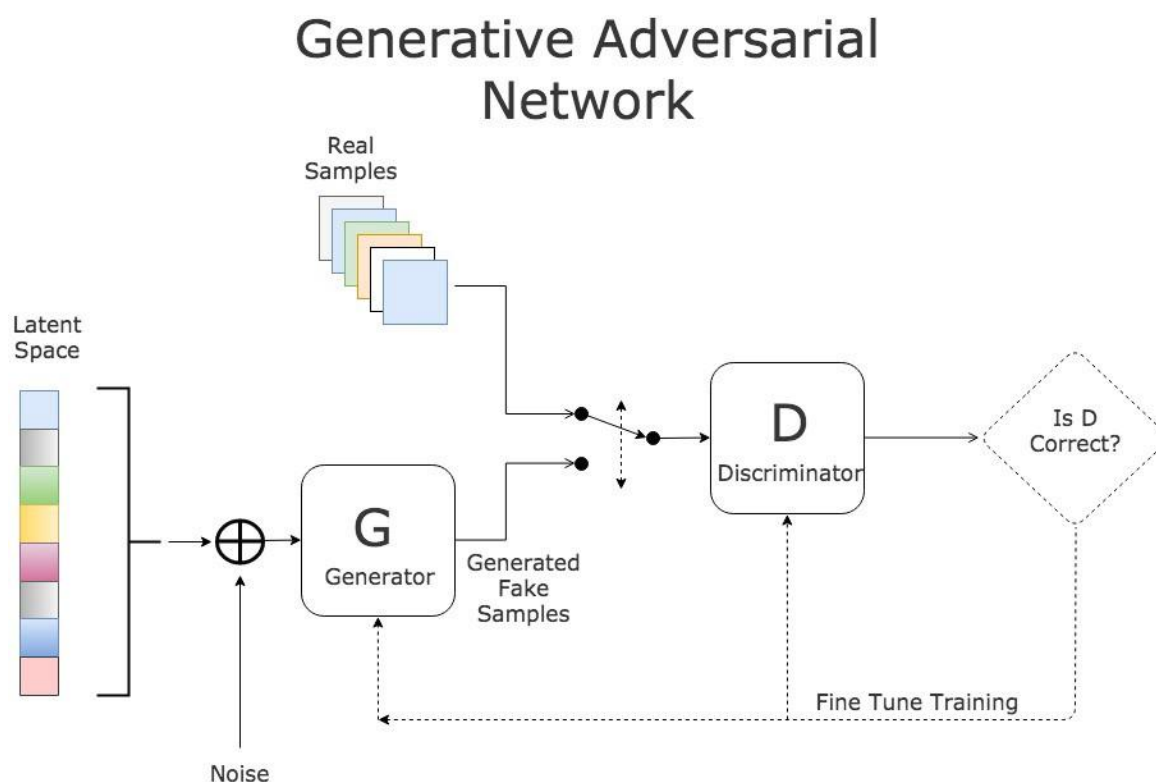


Figura 21. Rede Generativa Adversária.

Fonte: Goodfellow et al., 2014



#### 4.3.1 REDE DE SEGMENTAÇÃO ADVERSÁRIA (SegAN)

Tomando emprestada a motivação das Redes Adversariais Generativas clássicas, Xue et al. introduziram notavelmente uma Rede Neural Adversarial de ponta a ponta para realizar segmentação de imagens médicas: SegAN.

A obtenção de segmentações pixel a pixel infere que a saída de um único valor escalar para real/falso pelo discriminador GAN clássico pode não ser eficaz na produção de dados estáveis ou feedback de gradiente suficiente para as redes.

O segmentador é uma rede neural totalmente convolucional que gera os mapas de rótulos de segmentação.

Propuseram então uma nova Rede Adversarial Crítica com força de perda L1 multiescala, encorajando tanto o crítico quanto o segmentador a aprender características globais e locais capturando relacionamentos espaciais de longo e curto alcance entre pixels, ao aprender a atualizar a partir do feedback adversário do crítico.

Na estrutura SegAN, as redes de segmentação e crítica são treinadas alternadamente em um jogo min-max; o crítico recebe como entrada um par de imagens e é treinado pela maximização de uma função de perda multiescala, onde (imagem original \* mapa de rótulo previsto, imagem original \* mapa de rótulo da verdade básica) e o segmentador é treinado apenas com gradientes realizados pelo crítico, visando minimizar a função de perda multiescala.

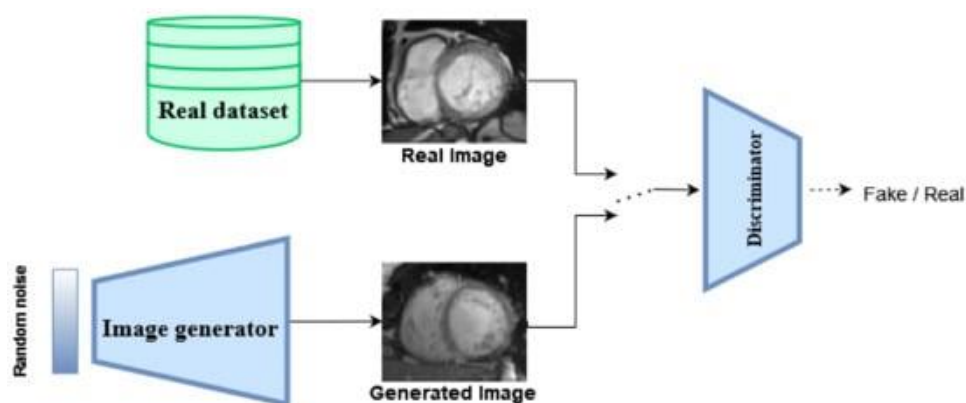
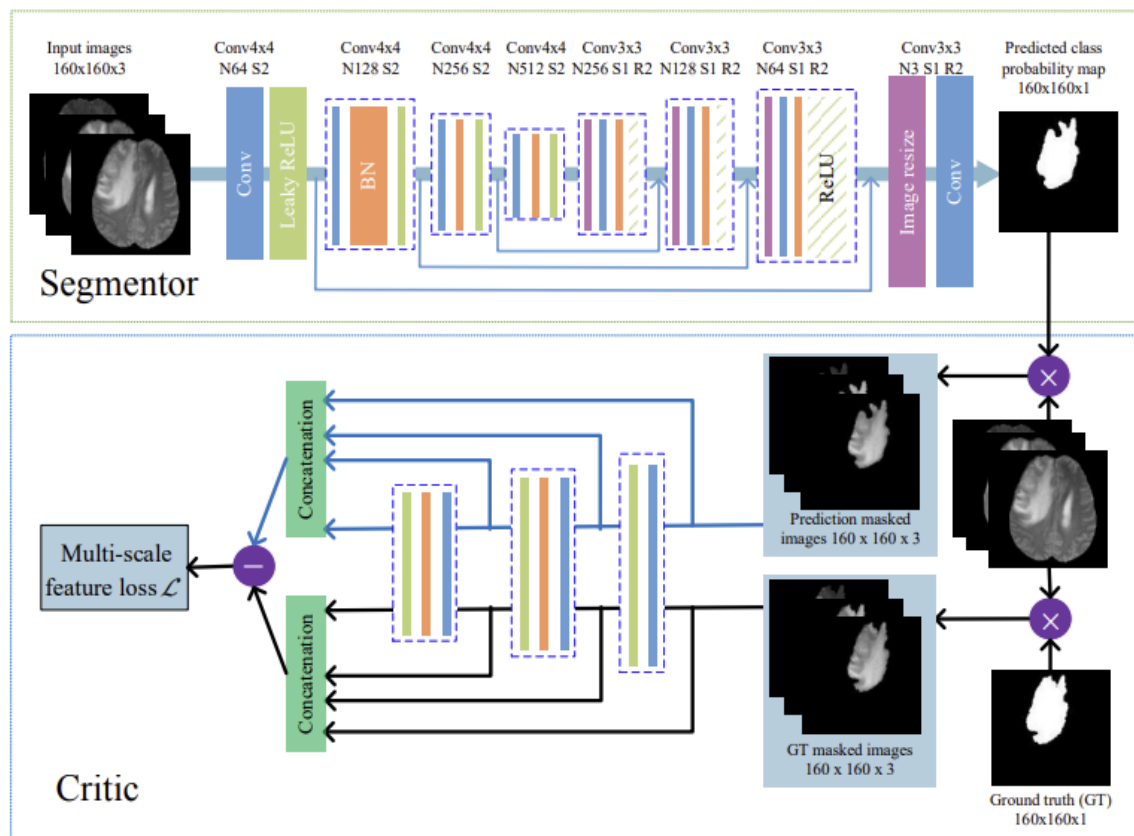


Figura 22. Exemplo de GAN em imagem cardíaca (Skandarani et col.,2022).





1 The architecture of the proposed SegAN with segmentor and critic networks.  $4 \times 4$  convolutional layers with stride 2 (S2) and the corresponding number of feature maps (e.g., N64) are used for encoding, while image resize layers with a factor of 2 (R2) and  $3 \times 3$  convolutional layers with stride 1 are used for decoding. Masked images are calculated by pixel-wise multiplication of a label map and (the multiple channels) input image. Note that, although only one label map (for whole tumor segmentation) is illustrated here, multiple label maps (e.g. also for core and Gd-enhanced tumor core) can be generated by the segmentor in one path.

Figura 23. Rede SegAN.

Fonte: (Xue, Y et col.; 2018).

Os autores afirmaram a estrutura SegAN é mais robusta e eficiente para a tarefa de segmentação e fornece desempenho mais fino do que o método de segmentação de última geração U-net (Xue, Y.; Xu, T.; Zhang, H.; Long, L.R.; Huang, X. SegAN: Adversarial Network with Multi-scale L1 Loss for Medical Image Segmentation. Neuroinformatics 2018, 16, 383–392.).

## 5. O ESTADO DA ARTE

Segmentação precisa de múltiplos órgãos de imagens médicas é uma etapa essencial no diagnóstico auxiliado por computador, assim como na navegação cirúrgica, e radioterapia.

Nos últimos anos, com uma extração de recursos baseada em dados, abordagem e treinamento ponta a ponta, multi-órgãos automáticos baseados em aprendizagem profunda os métodos de segmentação superaram em muito os métodos tradicionais e se tornaram um novo tema de pesquisa.

Esta técnica abrange dois principais fundamentos, que são conjunto de dados (datasets públicos, repositórios de universidades e hospitais universitários, etc.) e a metodologia em que os dados são abordados (totalmente supervisionada, fracamente supervisionadas e semi-supervisionadas). (van Ginneken B, Schaefer-Prokop CM, Prokop M. Computer-aided diagnosis: how to move from the laboratory to the clinic. *Radiology*. 2011;261:719–32.)

O processo inicia-se com a rotulagem manual pelos médicos, o que é muito trabalhoso e demorado.

Além disso, diferentes médicos ou hospitais fornecem resultados diferentes da rotulagem (Hurkmans CW, Borger JH, Pieters BR, Russell NS, Jansen EPM, Mijnheer BJ. Variability in target volume delineation on CT scans of the breast. *Int J Radn Oncol Biol Phys*. 2001;50:1366–72).

Portanto, há uma necessidade urgente de métodos precisos e automatizados de segmentação de múltiplos órgãos na prática clínica.

Alguns dos benchmarks comuns para a segmentação de imagens usam recursos de imagem extraídos manualmente, que incluem métodos de limiarização (Saranathan AM, Parente M. Threshold based segmentation method for hyperspectral images. 2013 5th workshop on hyperspectral image and signal processing: evolution in remote sensing (WHISPERS). Gainesville, FL, USA: IEEE;2013. p. 1–4. <http://ieeexplore.ieee.org/document/8080656/>. Accessed 8 Oct 2022).

Devido ao grande número de recursos de imagem extraídos manualmente e limiares ou sementes não robustos que devem ser fornecidos, seus resultados na

segmentação são geralmente instáveis, geralmente fornecendo apenas um resultado de segmentação aproximado ou trabalhando apenas para alguns órgãos específicos. Métodos baseados em conhecimento fazem isso, onde informações anatômicas detalhadas de vários órgãos são obtidas automaticamente, com base em conjuntos de dados rotulados, em vez de extrair manualmente os recursos da imagem como em outros métodos. Isso provou aumentar a precisão e a robustez dos métodos de segmentação multiorgânica baseados em multiatlas, como fusão de rótulos multiatlas e modelos de forma estatística (Ecabert O, Peters J, Schramm H, Lorenz C, von Berg J, Walker MJ, et al. Automatic model-based segmentation of the heart in CT images. *IEEE Trans Med Imaging*. 2008;27:1189–201.)

O método baseado em multiatlas continua usando o alinhamento de imagem para alinhar contornos estruturais predefinidos à imagem que precisa ser segmentada. No entanto, essa abordagem normalmente é multietapa; portanto, seu desempenho pode depender de vários fatores relevantes que são influenciados em cada etapa. Além disso, atlas fixos tornam muito difícil controlar variações anatômicas de órgãos em diferentes pacientes, levando a uma grande variabilidade entre usuários, tornando a adaptação clínica quase impossível (Qazi AA, Pekar V, Kim J, Xie J, Breen SL, Jaffray DA. Auto-segmentation of normal and target structures in head and neck CT images: a feature-driven model-based approach: feature-driven model-based segmentation. *Med Phys*. 2011;38:6160–70).

Também é altamente computacional e leva tempo para realizar uma tarefa de alinhamento. O modelo atua de forma estatística, usando as inter-relações em posição entre diferentes órgãos e a forma de cada órgão no espaço estatístico como restrições que são aplicadas aos resultados da segmentação para regularizá-los. Sua precisão depende em grande parte da confiabilidade e extensibilidade do modelo de forma; na verdade, um modelo baseado em estruturas anatômicas regulares tem muito pouco efeito em problemas de segmentação com estruturas irregulares (Zhou X, Takayama R, Wang S, Hara T, Fujita H. Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method. *Med Phys*. 2017;44:5221–33).

Assim, os métodos totalmente supervisionados requerem todos os órgãos anotados na segmentação de múltiplos órgãos para realizar a tarefa. Os métodos existentes podem ser divididos em quatro partes que podem estar relacionadas às suas análises, ou seja: arquitetura de rede, dimensão de rede, módulos de segmentação de imagem e função de perda de rede.

Módulos de segmentação de imagem referem-se a módulos que são frequentemente usados na segmentação de múltiplos órgãos para precisão e rapidez.

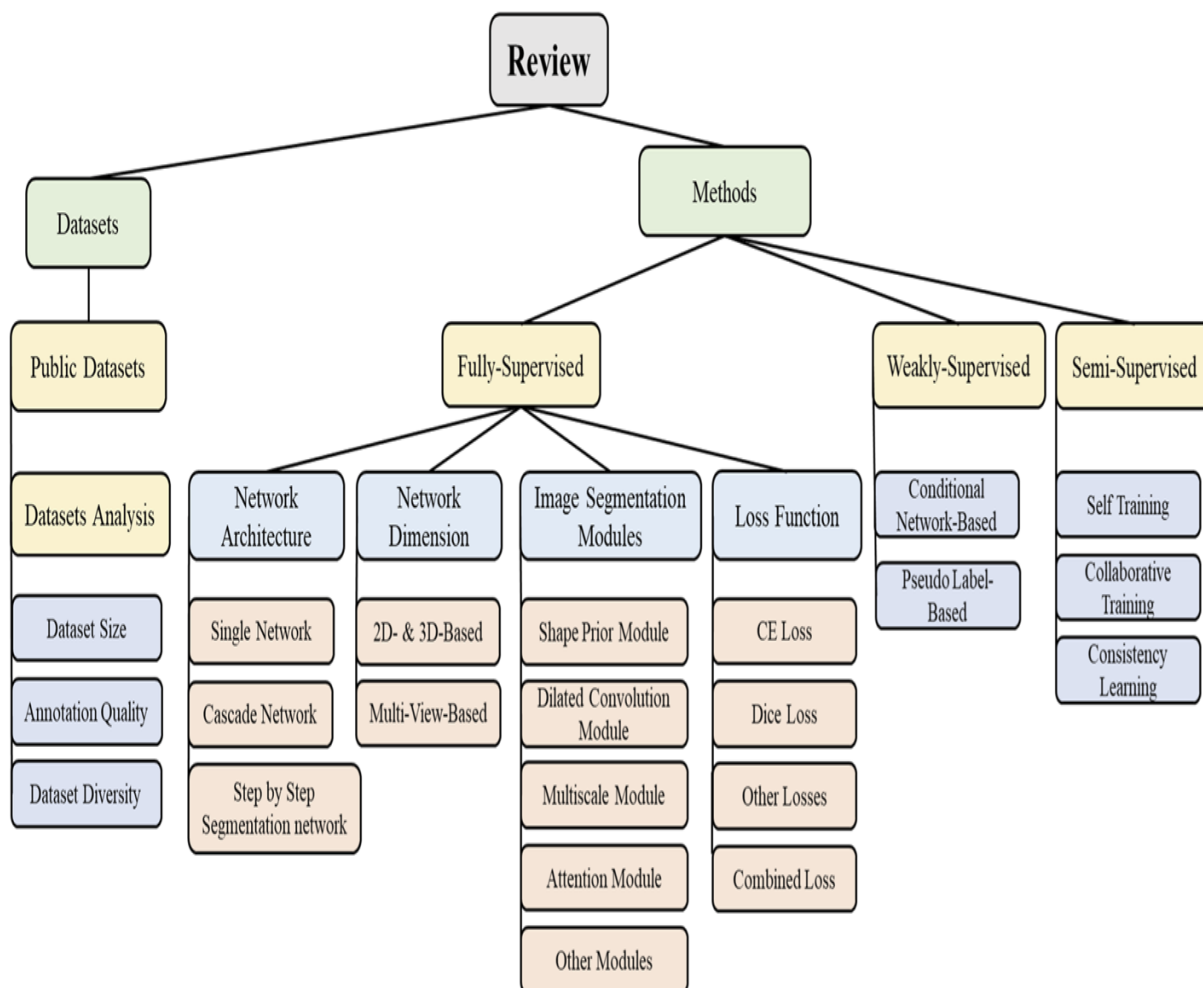


Figura 24: Resumo esquemático de segmentação multiórgãos

Fonte: Xiaoliu, L. et col., 2024.

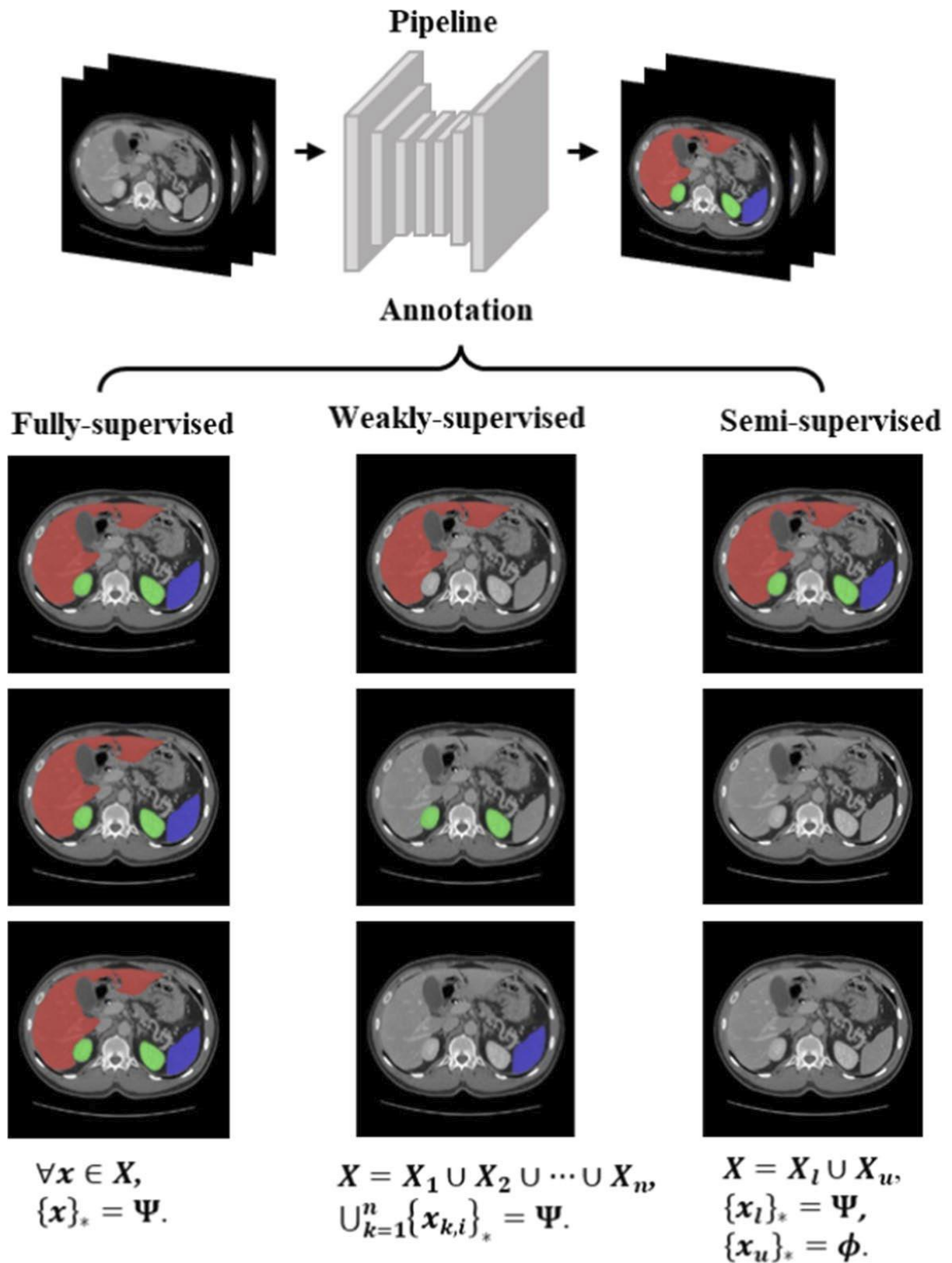


Figura 44. Visão geral dos métodos de segmentação multiórgãos.

Fonte: Segmentation Outside the Cranial Vault Challenge. 2015; <https://prod.sagebase.org/repo/v1/doi/locate?id=syn3193805&type=ENTITY>. Accessed 3 Oct 2022

Nos métodos baseados em múltiplas visualizações, a segmentação precisa de imagens médicas requer o uso efetivo de informações espaciais entre fatias de imagem. Embora a entrada de imagens 3D diretamente na rede neural ajude a evitar alto uso de memória, isso resulta em perda de informações espaciais entre fatias. Uma solução tem sido os métodos baseados em múltiplas visualizações, usando redes neurais 2,5D com múltiplas fatias 2D ou convoluções 2D e 3D.

Isso pode reduzir o uso de memória enquanto mantém as informações espaciais intactas entre fatias. Esse tipo de informação ajudará a melhorar a precisão na segmentação de imagens médicas.

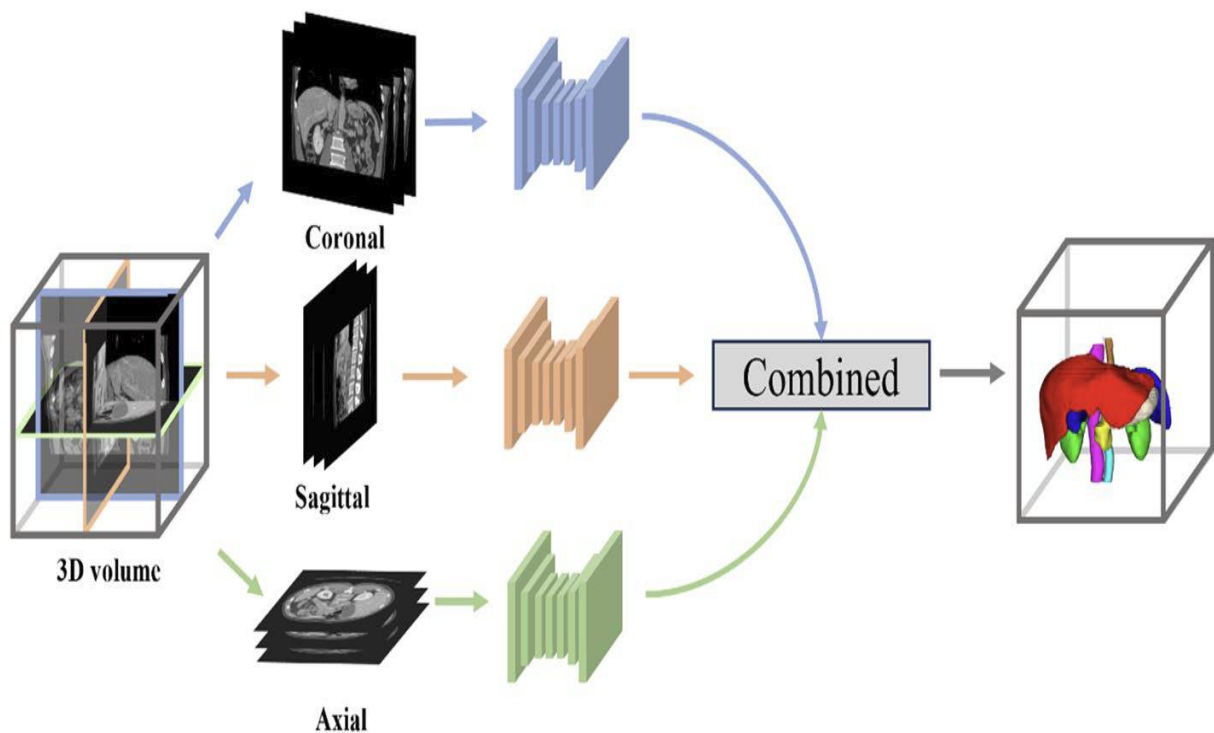


Figura 45. Estrutura de métodos baseados em múltiplas visualizações

Fonte: Zhou et col.; 2017

## 6. CONCLUSÃO

Embora a pesquisa em segmentação de imagens médicas venha evoluindo, o efeito da segmentação automática ainda não pode atender a todas as necessidades de aplicações práticas.

A principal razão é que a atual pesquisa de segmentação de imagens médicas ainda apresenta dificuldades e desafios a serem superados, especialmente em uma área sensível, como a radiologia e diagnóstico por imagem.

É um campo multidisciplinar - A fisiopatologia e a clínica são ciências complexas, muito além do arcabouço geral de conhecimentos de um cientista da inteligência artificial (habitualmente oriundo das ciências exatas como a ciência da computação, engenharia de software, etc), que acaba perdendo nuances preciosas de cada necessidade clínica específica, que demanda longo treinamento e extensa experiência médica. Por outro lado, de modo geral, médicos não compreendem a tecnologia específica da inteligência artificial, seus subcampos e ferramentas específicas para cada necessidade técnica que se apresenta. Como resultado, a inteligência artificial não pode atender completamente às necessidades clínicas específicas desse profissional de saúde, porém a IA traz a possibilidade de desenvolvimento de técnicas e ferramentas que são valiosas para apoiar o trabalho do especialista médico.

Assim, afim de promover a aplicação da inteligência artificial e o aprendizado de máquina na área médica, deve ser estimulada uma irrestrita cooperação entre médicos e programadores (cientistas da computação, engenheiros, etc). Esta sinergia mitigará a questão do livre acesso a dados e imagens médicas, ao passo que potencializará o desenvolvimento de novos algoritmos de aprendizado de máquina alinhados às necessidades específicas de cada serviço médico, tendo como óbvio resultado um significativo impacto positivo na acurácia (e portanto, na confiabilidade) de cada exame.

Há diferenças substanciais entre as diferentes modalidades de imagens médicas. Essas diferenças impactam o desenvolvimento e mesmo a adaptabilidade do modelo de aprendizado profundo durante a segmentação. O ruído e os artefatos das imagens médicas também são um grande problema no pré-processamento de dados.

Os conjuntos de dados de imagens médicas existentes são pequenos em escala. O treinamento de algoritmos de aprendizado profundo requer uma grande quantidade de suporte ao conjunto de dados, o que leva ao problema de sobreajuste (*overfitting*) no processo de treinamento de modelos de aprendizado profundo.

Uma maneira de resolver a quantidade insuficiente de dados de treinamento é o aprimoramento desses dados, como transformação geométrica e o aprimoramento do espaço de cores.

As GAN's usam dados originais para sintetizar novos dados. Outra forma de abordar o problema baseia-se em um modelo de meta-aprendizagem para estudar a segmentação de imagens médicas em condições de amostra pequena.

O modelo de aprendizado profundo tem suas próprias falhas. Ele se concentra principalmente em três aspectos: design da estrutura de rede, design do modelo de segmentação de dados 3D e design da função de perda.

Durante a pesquisa realizada neste trabalho de conclusão de curso percebeu-se que há espaço para se aprimorar as técnicas de segmentação de imagens ao se explorar o design da estrutura da rede.

O desenvolvimento de novas técnicas e algoritmos como resultado do esforço conjunto profissional multidisciplinar tem levado a resultados mais do que satisfatórios e descortinam um futuro bem próximo onde a excelência e confiabilidade dos métodos de imagem por inteligência artificial apoiará o tratamento médico sustentável.

A segmentação multiórgãos baseada em aprendizado profundo é a mais recente fronteira neste campo de estudo, abordando duas frentes principais (conjuntos de dados e métodos para segmentação de múltiplas regiões do corpo), e sua acurácia tem superado em muito os métodos tradicionais, mostrando resultados promissores.



## 7. REFERÊNCIAS

- ALMEIDA, G.; TAVARES, J.M.R.S. Deep learning in radiation oncology treatment planning for prostate cancer: A systematic review. *Journal of Medical Systems*, v. 44, p. 1–15, 2020. [CrossRef]
- ALTAF, F.; ISLAM, S.M.S.; AKHTAR, N.; NANJUA, N.K. Going deep in medical image analysis: Concepts, methods, challenges, and future directions. *IEEE Access*, v. 7, p. 99540–99572, 2019. [CrossRef]
- ANTHIMOPOULOS, M.M.; CHRISTODOULIDIS, S.; EBNER, L.; GEISER, T.; CHRISTE, A.; MOUGIAKAKOU, S. Semantic Segmentation of Pathological Lung Tissue with Dilated Fully Convolutional Networks. *IEEE Journal of Biomedical and Health Informatics*, v. 23, p. 714–722, 2019. [CrossRef]
- ARMATO, S.G., III; MCLENNAN, G.; BIDAUT, L.; MCNITT-GRAY, M.F.; MEYER, C.R.; REEVES, A.P.; ZHAO, B.; ABERLE, D.A.; HENSCHKE, C.I.; HOFFMAN, E.A.; et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Medical Physics*, v. 38, p. 915–931, 2011. [CrossRef] [PubMed]
- ARSHAD, H.; KHAN, M.A.; SHARIF, M.I.; YASMIN, M.; TAVARES, J.M.R.; ZHANG, Y.D.; SATAPATHY, S.C. A multilevel paradigm for deep convolutional neural network features selection with an application to human gait recognition. *Expert Systems*, v. e12541, 2020. [CrossRef]
- BADRINARAYANAN, V.; KENDALL, A.; CIPOLLA, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 39, p. 2481–2495, 2017. [CrossRef] [PubMed]
- BAYRAMOGLU, N.; KAAKINEN, M.; EKLUND, L.; HEIKKILA, J. Towards virtual h&e staining of hyperspectral lung histology images using conditional generative adversarial networks. In: *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 22–29 October 2017, p. 64–71.
- BHANDARY, A.; PRABHU, G.A.; RAJINIKANTH, V.; THANARAJ, K.P.; SATAPATHY, S.C.; ROBBINS, D.E.; SHASKY, C.; ZHANG, Y.-D.; TAVARES, J.M.R.S.; RAJA, N.S.M. Deep-learning framework to detect lung abnormality—A study with chest X-Ray and lung CT scan images. *Pattern Recognition Letters*, v. 129, p. 271–278, 2020. [CrossRef]
- BILIC, P.; CHRIST, P.F.; VORONTSOV, E.; CHLEBUS, G.; CHEN, H.; DOU, Q.; FU, C.-W.; HAN, X.; HENG, P.-A.; HESSER, J.; et al. The liver tumor segmentation benchmark (lits). *arXiv*, arXiv:1901.04056, 2019.
- BORNE, L.; RIVIÈRE, D.; MANGIN, J.F. Combining 3D U-Net and bottom-up geometric constraints for automatic cortical sulci recognition. In: *Proceedings of the International Conference on Medical Imaging with Deep Learning*, London, UK, 8–10 July 2019.
- CASAMITJANA, A.; CATÀ, M.; SÀNCHEZ, I.; COMBALIA, M.; VILAPLANA, V.

Cascaded V-Net using ROI masks for brain tumor segmentation. In: Proceedings of the International MICCAI Brainlesion Workshop, Quebec City, QC, Canada, 14 September 2017, p. 381–391.

CHANG, Q.; QU, H.; ZHANG, Y.; SABUNCU, M.; CHEN, C.; ZHANG, T.; METAXAS, D.N. Synthetic Learning: Learn from Distributed Asynchronized Discriminator GAN Without Sharing Medical Image Data. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Virtual, 14–19 June 2020, p. 13856–13866.

CHEN, C.; QIN, C.; QIU, H.; TARRONI, G.; DUAN, J.; BAI, W.; RUECKERT, D. Deep Learning for Cardiac Image Segmentation: A Review. *Frontiers in Cardiovascular Medicine*, v. 7, p. 25, 2020. [CrossRef]

CHEN, L.C.; PAPANDREOU, G.; KOKKINOS, I.; MURPHY, K.; YUILLE, A.L. Semantic image segmentation with deep convolutional nets and fully connected crfs. *arXiv*, arXiv:1412.7062, 2014.

CHEN, L.C.; PAPANDREOU, G.; KOKKINOS, I.; MURPHY, K.; YUILLE, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 40, p. 834–848, 2017. [CrossRef]

CHEN, L.C.; PAPANDREOU, G.; SCHROFF, F.; ADAM, H. Rethinking atrous convolution for semantic image segmentation. *arXiv*, arXiv:1706.05587, 2017.

CHEN, L.C.; ZHU, Y.; PAPANDREOU, G.; SCHROFF, F.; ADAM, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018, p. 801–818.

CHRIST, P.F.; ELSHAER, M.E.A.; ETTLINGER, F.; TATAVARTY, S.; BICKEL, M.; BILIC, P.; REMPFLE, M.; ARMBRUSTER, M.; HOFFMAN, F.; D'ANASTASI, M.; et al. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 17–21 October 2016, p. 415–423.

CHRIST, P.F.; ETTLINGER, F.; GRÜN, F.; ELSHAER, M.E.A.; LIPKOVA, J.; SCHLECHT, S.; AHMADDY, F.; TATAVARTY, S.; BICKEL, M.; BILIC, P.; et al. Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks. *arXiv*, arXiv:1702.05970, 2017.

ÇİÇEK, Ö.; ABDULKADIR, A.; LIENKAMP, S.S.; BROX, T.; RONNEBERGER, O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 17–21 October 2016, p. 424–432.

CIGLA, C.; ALATAN, A.A. Region-based image segmentation via graph cuts. In: Proceedings of the 2008 15th IEEE International Conference on Image Processing, San Diego, CA, USA, 12–15 October 2008, p. 2272–2275. *Sustainability*, v. 13, p. 1224, 2021.

DAI, W.; DONG, N.; WANG, Z.; LIANG, X.; ZHANG, H.; XING, E.P. Scan: Structure correcting adversarial network for organ segmentation in chest x-rays. In: *Mining Data for Financial Applications*. Cham: Springer, 2018, p. 263–273.

DAR, S.U.H.; YURT, M.; KARACAN, L.; ERDEM, A.; ERDEM, E.; ÇUKUR, T. Image synthesis in multi-contrast MRI with conditional generative adversarial networks. *IEEE Transactions on Medical Imaging*, v. 38, p. 2375–2388, 2019. [CrossRef]

DONG, S.; LUO, G.; WANG, K.; CAO, S.; MERCADO, A.; SHMUILOVICH, O.; ZHANG, H.; LI, S. VoxelAtlasGAN: 3D left ventricle segmentation on echocardiography with atlas guided generation and voxel-to-voxel discrimination. In: *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Granada, Spain, 16–20 September 2018, p. 622–629.

EDUPUGANTI, V.G.; CHAWLA, A.; AMIT, K. Automatic optic disk and cup segmentation of fundus images using deep learning. In: *Proceedings of the 25th IEEE International Conference on Image Processing (ICIP)*, Athens, Greece, 7–10 October 2018, p. 2227–2231.

ESS, A.; MÜLLER, T.; GRABNER, H.; VAN GOOL, L. Segmentation-Based Urban Traffic Scene Understanding. *BMVC*, v. 1, p. 2, 2009.

FERREIRA, A.; GENTIL, F.; TAVARES, J.M.R.S. Segmentation algorithms for ear image data towards biomechanical studies. *Computer Methods in Biomechanics and Biomedical Engineering*, v. 17, p. 888–904, 2014. [CrossRef]

FONSECA, C.G.; BACKHAUS, M.; BLUEMKE, D.A.; BRITTEN, R.D.; CHUNG, J.D.; COWAN, B.R.; DINOV, I.D.; FINN, J.P.; HUNTER, P.J.; KADISH, A.H.; et al. The Cardiac Atlas Project—An imaging database for computational modeling and statistical atlases of the heart. *Bioinformatics*, v. 27, p. 2288–2295, 2011. [CrossRef] [PubMed]

FUKUSHIMA, K.; MIYAKE, S. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In: *Competition and Cooperation in Neural Nets*. Berlin: Springer, 1982, p. 267–285.

GEIGER, A.; LENZ, P.; URTASUN, R. Are we ready for autonomous driving? The kitti vision benchmark suite. In: *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, USA, 16–12 June 2012, p. 3354–3361.

GIACOMELLO, E.; LOIACONO, D.; MAINARDI, L. Brain MRI Tumor Segmentation with Adversarial Networks. *arXiv*, arXiv:1910.02717, 2019.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A.; BENGIO, Y. *Deep Learning*. Cambridge: MIT Press, 2016.

GOODFELLOW, I.; POUGET-ABADIE, J.; MIRZA, M.; XU, B.; WARDE-FARLEY, D.; OZAIR, S.; COURVILLE, A.; BENGIO, Y. Generative adversarial nets. *Advances in Neural Information Processing Systems*, v. 27, p. 2672–2680, 2014.

GU, J.; WANG, Z.; KUEN, J.; MA, L.; SHSHROUDY, A.; SHUAI, B.; LIU, I.; WANG, X.; WANG, G.; CAI, J.; et al. Recent advances in convolutional neural networks. *Pattern Recognition*, v. 77, p. 354–377, 2018. [CrossRef]

HAN, X. Automatic liver lesion segmentation using a deep convolutional neural network method. *arXiv*, arXiv:1704.07239, 2017.

HAN, Z.; WEI, B.; MERCADO, A.; LEUNG, S.; LI, S. Spine-GAN: Semantic segmentation of multiple spinal structures. *Medical Image Analysis*, v. 50, p. 23–35, 2018. [CrossRef]

HE, K.; GKIOXARI, G.; DOLLÁR, P.; GIRSHICK, R. Mask r-cnn. In: *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 22–29 October 2017, p. 2961–2969.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 16 June–1 July 2016, p. 770–778.

HEATH, M.; BOWYER, K.; KOPANS, D.; KEGELMEYER, P.; MOORE, R.; CHANG, K.; MUNISHKUMARAN, S. The digital database for screening mammography. In: *Proceedings of the 5th International Workshop on Digital Mammography*, Toronto, ON, Canada, 11–14 June 2000, p. 212–218.

HESAMIAN, M.H.; JIA, W.; HE, X.; KENNEDY, P. Deep learning techniques for medical image segmentation: Achievements and challenges. *Journal of Digital Imaging*, v. 32, p. 582–596, 2019. [CrossRef] [PubMed]

HU, P.; CAO, Y.; WANG, W.; WEI, B. Computer Assisted Three-Dimensional Reconstruction for Laparoscopic Resection in Adult Teratoma. *Journal of Medical Imaging and Health Informatics*, v. 9, p. 956–961, 2019. [CrossRef]

HUBEL, D.H.; WIESEL, T.N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, v. 160, p. 106, 1962. [CrossRef] [PubMed]

HUO, Y.; XU, Z.; BAO, S.; BERMUDEZ, C.; PLASSARD, A.J.; YAO, Y.; LIU, J.; ASSAD, A.; ABRAMSON, R.G.; LANDMAN, B.A. Splenomegaly segmentation using global convolutional kernels and conditional generative adversarial networks. *Medical Imaging*, v. 10574, p. 1057409, 2018.

HURKMANS CW, BORGER JH, PIETERS BR, RUSSEL NS, et col. Variability in target volume delineation on CT scans of the breast. *Int J Radn Oncol Biol Phys*. 2001;50:1366–72

IBTEHAZ, N.; RAHMAN, M.S. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Networks*, v. 121, p. 74–87, 2020. [CrossRef] [PubMed]

IZADI, S.; MIRIKHARAJI, Z.; KAWAHARA, J.; HAMARNEH, G. Generative adversarial networks to segment skin lesions. In: *Proceedings of the IEEE 15th International*

Symposium on Biomedical Imaging, Washington, DC, USA, 4–7 April 2018, p. 881–884.

J.; JASON, H.; NEELAM, T.; ANDREAS, R.; SEAN, B.L.; JOSEPH, D.O.; HARINI, V. Integrating cross-modality hallucinated MRI with CT to aid mediastinal lung tumor segmentation. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–18 October 2019, p. 221–229.

JUE, al. Craniomaxillofacial bony structures segmentation from MRI with deep-supervision adversarial learning. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, 16–20 September 2018, p. 720–727.

KHOSRAVAN, N.; MORTAZI, A.; WALLACE, M.; BAGCI, U. Pan: Projective adversarial network for medical image segmentation. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–18 October 2019, p. 68–76.

KOHL, S.; BONEKAMP, D.; SCHLEMMER, H.P.; YAQUBI, K.; HOHENFELLNER, M.; HADASCHIK, B.; RADTKE, J.P.; MAIER-HEIN, K. Adversarial networks for the detection of aggressive prostate cancer. arXiv, arXiv:1702.08014, 2017.

KRÄHENBÜHL, P.; KOLTUN, V. Efficient inference in fully connected crfs with gaussian edge potentials. Advances in Neural Information Processing Systems, v. 24, p. 109–117, 2011.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G.E. Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, v. 60, p. 1097–1105, 2012. [CrossRef]

LATEEF, F.; RUICHEK, Y. Survey on semantic segmentation using deep learning techniques. Neurocomputing, v. 338, p. 321–348, 2019. [CrossRef]

LÉCUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. IEEE, v. 86, p. 2278–2324, 1998. [CrossRef]

LEOPOLD, H.A.; ORCHARD, J.; ZELEK, J.S.; LAKSHMINARAYANAN, V. Pixelbnn: Augmenting the pixelcnn with batch normalization and the presentation of a fast architecture for retinal vessel segmentation. Journal of Imaging, v. 5, p. 26, 2019. [CrossRef]

LI, X.; CHEN, H.; QI, X.; DOU, Q.; FU, C.-W.; HENG, P.A. H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Transactions on Medical Imaging, v. 37, p. 2663–2674, 2018. [CrossRef] [PubMed]

LIN, G.; MILAN, A.; SHEN, C.; REID, I. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017, p. 1925–1934.2014. [CrossRef]

LIU, F.; ZHOU, Z.; JANG, H.; SAMSONOV, A.; ZHAO, G.; KIJOWSKI, R. Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging. *Magnetic Resonance in Medicine*, v. 79, p. 2379–2391, 2018. [CrossRef] [PubMed]

LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, 7–12 June 2015, p. 3431–3440.

LUC, P.; COUPRIE, C.; CHINTALA, S.; VERBEEK, J. Semantic segmentation using adversarial networks. *arXiv*, arXiv:1611.08408, 2016.

MA, Z.; TAVARES, J.M.R.S.; JORGE, R.M.N. A review on the current segmentation algorithms for medical images. In: *Proceedings of the 1st International Conference on Imaging Theory and Applications*, Lisbon, Portugal, 5–8 February 2009.

MA, Z.; TAVARES, J.M.R.S.; JORGE, R.N.; MASCARENHAS, T. A review of algorithms for medical image segmentation and their applications to the female pelvic cavity. *Computer Methods in Biomechanics and Biomedical Engineering*, v. 13, p. 235–246, 2010. [CrossRef]

MARCUS, D.S.; FOTENOS, A.F.; CSERNANSKY, J.G.; MORRIS, J.C.; BUCKNER, R.L. Open access series of imaging studies: Longitudinal MRI data in nondemented and demented older adults. *Journal of Cognitive Neuroscience*, v. 22, p. 2677–2684, 2010. [CrossRef]

MENZE, B.H.; JAKAB, A.; BAUER, S.; KALPATHY-CRAMER, J.; FARAHANI, K.; KIRBY, J.; BURRER, Y.; PORZ, N.; SLOTBOOM, J.; WIEST, R.; et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging*, v. 34, p. 1993–2024,

MILLETARI, F.; NAVAB, N.; AHMADI, S.-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: *Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV)*, Stanford, CA, USA, 25–28 October 2016, p. 565–571.

MIRIKHARAJI, Z.; HAMARNEH, G. Star shape prior in fully convolutional networks for skin lesion segmentation. In: *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Granada, Spain, 16–20 September 2018, p. 737–745.

MIRZA, M.; OSINDERO, S. Conditional generative adversarial nets. *arXiv*, arXiv:1411.1784, 2014.

MOESKOPS, P.; VETA, M.; LAFARGE, M.W.; EPPENHOF, K.A.J.; PLUIM, J.P.W. Adversarial training and dilated convolutions for brain MRI segmentation. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham: Springer, 2017, p. 56–64.

MONDAL, A.K.; DOLZ, J.; DESROSIERS, C. Few-shot 3d multi-modal medical image segmentation using generative adversarial learning. *arXiv*, arXiv:1810.12241, 2018.

MYRONENKO, A. 3D MRI brain tumor segmentation using autoencoder regularization. In: Proceedings of the International MICCAI Brainlesion Workshop, Shenzhen, China, 17 October 2018, p. 311–320.

NIE, D.; WANG, L.; ADELI, E.; LAO, C.; LIN, W.; SHEN, D. 3-D fully convolutional networks for multimodal isointense infant brain image segmentation. *IEEE Transactions on Cybernetics*, v. 49, p. 1123–1136, 2019. [CrossRef] [PubMed]

NOH, H.; HONG, S.; HAN, B. Learning deconvolution network for semantic segmentation. In: Proceedings of the IEEE International Conference on Computer Vision, Las Condes, Chile, 11–18 December 2015, p. 1520–1528.

NOVIKOV, A.A.; LENIS, D.; MAJOR, D.; HLADŮVKA, J.; WIMMER, M.; BÜHLER, K. Fully convolutional architectures for multiclass segmentation in chest radiographs. *IEEE Transactions on Medical Imaging*, v. 37, p. 1865–1876, 2018. [CrossRef]

OKTAY, O.; SCHLEMPER, J.; FOLGOC, L.L.; LEE, M.; HEINRICH, M.; MISAWA, K.; MORI, K.; MCDONAGH, S.; HAMMERLA, N.Y.; KAINZ, B.; et al. Attention u-net: Learning where to look for the pancreas. *arXiv*, arXiv:1804.03999, 2018.

QIU, Z.; YAO, T.; MEI, T. Learning spatio-temporal representation with pseudo-3d residual networks. In: Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017, p. 5533–5541.

REN, S.; HE, K.; GIRSHICK, R.; SUN, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, v. 39, p. 91–99, 2015. [CrossRef]

REZAEI, M.; HARMUTH, K.; GIERKE, W.; KELLERMEIER, T.; FISCHER, M.; YANG, H.; MEINEL, C. A conditional adversarial network for semantic segmentation of brain tumor. In: Proceedings of the International MICCAI Brainlesion Workshop, Quebec City, QC, Canada, 14 September 2017, p. 241–252.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015, p. 234–241.

RUNDO, L.; HAN, C.; NAGANO, Y.; ZHANG, J.; HATAYA, R.; MILITELLO, C.; TANGHERLONI, A.; NOBILE, M.S.; FERRETTI, C.; BESOZZI, D.; et al. USE-Net, Incorporating Squeeze-and-Excitation blocks into U-Net for prostate zonal segmentation of multi-institutional MRI datasets. *Neurocomputing*, v. 365, p. 31–43, 2019. [CrossRef]

SARANATHANAM, PARENTEM. Threshold based segmentation method for hyperspectral images. 2013 5th workshop on hyperspectral image and signal processing: evolution in remote sensing (WHISPERS). Gainesville, FL, USA: IEEE;2013. p. 1–4. <http://ieeexplore.ieee.org/document/8080656/>. Accessed 8 Oct 2022

SEKUBOYINA, A.; REMPFLE, M.; KUKAČKA, J.; TETTEH, G.; VALENTINITSCH, A.; KIRSCHKE, J.; MENZE, B.H. Btrfly net: Vertebrae labelling with energy-based adversarial learning of local spine prior. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, 16–20 September 2018, p. 649–657.

SHANKARANARAYANA, S.M.; RAM, K.; MITRA, K.; SIVAPRAKASAM, M. Joint optic disc and cup segmentation using fully convolutional and adversarial networks. In: Fetal, Infant and Ophthalmic Medical Image Analysis. Cham: Springer, 2017, p. 168–176.

SHEN, D.; WU, G.; SUK, H.I. Deep learning in medical image analysis. Annual Review of Biomedical Engineering, v. 19, p. 221–248, 2017. [CrossRef]

SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. arXiv, arXiv:1409.1556, 2014.

SIMPSON, A.L.; ANTONELLI, M.; BAKAS, S.; BILELLO, M.; FARAHANI, K.; VAN GINNEKEN, B.; KOPP-SCHNEIDER, A.; LANDMAN, B.A.; LITJENS, G.; MENZE, B.; et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. arXiv, arXiv:1902.09063, 2019.

SON, J.; PARK, S.J.; JUNG, K.H. Retinal vessel segmentation in fundoscopic images with generative adversarial networks. arXiv, arXiv:1706.09318, 2017.

STAAL, J.; ABRÀMOFF, M.D.; NIEMEIJER, M.; VIERGEVER, M.A.; VAN GINNEKEN, B. Ridge-based vessel segmentation in color images of the retina. IEEE Transactions on Medical Imaging, v. 23, p. 501–512, 2004. [CrossRef]

SUCKLING, J.P. The mammographic image analysis society digital mammogram database. Digital Mammo, v. 17, p. 375–386, 1994.

SZEDEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCHE, V.; RABINOVICH, A. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015, p. 1–9.

TAHA, A.; LO, P.; LI, J.; ZHAO, T. Kid-net: Convolution networks for kidney vessels segmentation from ct-volumes. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, 16–20 September 2018, p. 463–471.

TRAN, P.V. A fully convolutional neural network for cardiac segmentation in short-axis MRI. arXiv, arXiv:1604.00494, 2016.

TRAN, T.; KWON, O.H.; KWON, K.R.; LEE, S.H.; KANG, K.W. Blood cell images segmentation using deep learning semantic segmentation. In: Proceedings of the IEEE International Conference on Electronics and Communication Engineering, Essex, UK, 16–17 August 2018, p. 13–16.

TUAN, T.A.; PHAM, T.B.; KIM, J.Y.; TAVARES, J.M.R. Alzheimer's diagnosis using deep learning in segmenting and classifying 3D brain MR images. International Journal



of Neuroscience, p. 1–10, 2020. [CrossRef] [PubMed]

VAN GINNEKEN, B.; STEGMANN, M.B.; LOOG, M. Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database. *Medical Image Analysis*, v. 10, p. 19–40, 2006. [CrossRef] [PubMed]

VAN GINNEKEN, B.; SCHAEFFER-PROKOP, CM.; Prokop M. Computer-aided diagnosis: how to move from the laboratory to the clinic. *Radiology*. 2011;261:719–32.

WANG, D.; GU, C.; WU, K.; GUAN, X. Adversarial neural networks for basal membrane segmentation of microinvasive cervix carcinoma in histopathology images. In: *Proceedings of the 2017 International Conference on Machine Learning and Cybernetics*, Ningbo, China, 9–12 July 2017.

WANG, S.; YI, L.; CHEN, Q.; MENG, Z.; DONG, H.; HE, Z. Edge-aware Fully Convolutional Network with CRF-RNN Layer for Hippocampus Segmentation. In: *Proceedings of the 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, Chongqing, China, 24–26 May 2019, p. 803–806.

WANG, Y.; CHEN, Y.; YANG, N.; ZHENG, L.; DEY, N.; ASHOUR, A.S.; RAJINIKANTH, V.; TAVARES, J.M.R.S.; SHI, F. Classification of mice hepatic granuloma microscopic images based on a deep convolutional neural network. *Applied Soft Computing*, v. 74, p. 40–50, 2019. [CrossRef]

WANG, Z.; ZOU, N.; SHEN, D.; JI, S. Non-Local U-Nets for Biomedical Image Segmentation. In: *Proceedings of the AAAI*, New York, NY, USA, 7–12 February 2020, p. 6315–6322.

WOLTERINK, J.M.; DINKLA, A.M.; SAVENIJE, M.H.F.; SEEVINCK, P.R.; VAN DEN BERG, C.A.; ISGUM, I. Deep MR to CT synthesis using unpaired data. In: *Proceedings of the International Workshop on Simulation and Synthesis in Medical Imaging*, Quebec City, QC, Canada, 10 September 2017, p. 14–23.

XIA, Q.; YAO, Y.; HU, Z.; HAO, A. Automatic 3D atrial segmentation from GE-MRIs using volumetric fully convolutional networks. In: *Proceedings of the International Workshop on Statistical Atlases and Computational Models of the Heart*, Granada, Spain, 16 September 2018, p. 211–220.

XIAO, X.; LIAN, S.; LUO, Z.; LI, S. Weighted Res-UNet for high-quality retina vessel segmentation. In: *Proceedings of the 9th International Conference on Information Technology in Medicine and Education (ITME)*, Hangzhou, China, 19–21 October 2018, p. 327–331.

XU, A.; WANG, L.; FENG, S.; QU, Y. Threshold-based level set method of image segmentation. In: *Proceedings of the Third International Conference on Intelligent Networks and Intelligent Systems*, Shenyang, China, 1–3 November 2010, p. 703–706.

XU, Z.; WU, Z.; FENG, J. CFUN: Combining faster R-CNN and U-net network for efficient whole heart segmentation. *arXiv*, arXiv:1812.04914, 2018.

XUE, Y.; XU, T.; ZHANG, H.; LONG, L.R.; HUANG, X. SegAN: Adversarial Network with Multi-scale L1 Loss for Medical Image Segmentation. *Neuroinformatics*, v. 16, p. 383–392, 2018. [CrossRef]

YANG, D.; XU, D.; ZHOU, S.K.; GEORGESCU, B.; CHEN, M.; GRBIC, S.; METAXAS, D.; COMANICIU, D. Automatic liver segmentation using an adversarial image-to-image network. In: *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Quebec City, QC, Canada, 10–14 September 2017, p. 507–515.

YE, C.; WANG, W.; ZHANG, S.; WANG, K. Multi-depth fusion network for whole-heart CT image segmentation. *IEEE Access*, v. 7, p. 23421–23429, 2019. [CrossRef]

YU-QIAN, Z.; WEI-HUA, G.; ZHEN-CHENG, C.; TANG, J.-T.; LI, L.-Y. Medical images edge detection based on mathematical morphology. In: *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, Shanghai, China, 17–18 January 2006, p. 6492–6495.

ZEILER, M.D.; FERGUS, R. Visualizing and understanding convolutional networks. In: *Proceedings of the European Conference on Computer Vision*, Zurich, Switzerland, 6–12 September 2014, p. 818–833.

ZHANG, J.; DU, J.; LIU, H.; HOU, X.; ZHAO, Y.; DING, M. LU-NET: An Improved U-Net for Ventricular Segmentation. *IEEE Access*, v. 7, p. 92539–92546, 2019. [CrossRef]

ZHANG, Y.; CHUNG, A.C.S. Deep supervision with additional labels for retinal vessel segmentation task. In: *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Granada, Spain, 16–20 September 2018, p. 83–91.

ZHANG, Y.; YANG, L.; CHEN, J.; FREDERICKSEN, M.; HUGHES, D.P.; CHEN, D.Z. Deep adversarial networks for biomedical image segmentation utilizing unannotated images. In: *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Quebec City, QC, Canada, 10–14 September 2017, p. 408–416.

ZHAO, H.; SHI, J.; QI, X.; WANG, X.; JIA, J. Pyramid scene parsing network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21–26 July 2017, p. 2881–2890.

ZHAO, M.; WANG, L.; CHEN, J.; NIE, D.; CONG, Y.; AHMAD, S.; HO, A.; YUAN, P.; FUNG, S.H.; DENG, H.H.; et

ZHOU, X.; TAKAYAMA, R.; WANG, S.; HARA, T.; FUJITA, H. Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method. *Medical Physics*, v. 44, p. 5221–5233, 2017. [CrossRef] [PubMed]

ZHOU, X.Y.; SHEN, M.; RIGA, C.; YANG, G.-Z.; LEE, S.-L. Focal fcn: Towards small

object segmentation with limited training data. arXiv, arXiv:1711.01506, 2017.

ZHU, J.Y.; PARK, T.; ISOLA, P.; EFROS, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017, p. 2223–2232.

ZHU, X.J. Semi-Supervised Learning Literature Survey. Madison: University of Winsconsin, 2005.

