



**UNIVERSIDADE FEDERAL DO CEARÁ  
CENTRO DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA  
MESTRADO ACADÊMICO EM ENGENHARIA DE TELEINFORMÁTICA**

**ROBERTO FERNANDES IVO**

**REDES VOLUMÉTRICAS TOTALMENTE CONVOLUCIONAIS EM CASCATA  
PARA SEGMENTAÇÃO 3D DE TODO O CORAÇÃO E GRANDE VASO**

**FORTALEZA  
2021**

ROBERTO FERNANDES IVO

**REDES VOLUMÉTRICAS TOTALMENTE CONVOLUCIONAIS EM CASCATA PARA  
SEGMENTAÇÃO 3D DE TODO O CORAÇÃO E GRANDE VASO**

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas.

Orientador: Prof. Dr. Pedro Pedrosa Rebouças Filho

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

I1r Ivo, Roberto Fernandes.

Redes Volumétricas Totalmente Convolucionais em Cascata para Segmentação 3D de todo o Coração e Grande Vaso / Roberto Fernandes Ivo. – 2021.  
48 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2021.

Orientação: Prof. Dr. Pedro Pedrosa Rebouças Filho.

1. Segmentação em cascata. 2. V-Net. 3. Doença Cardíaca Congênita. I. Título.

CDD 621.38

---

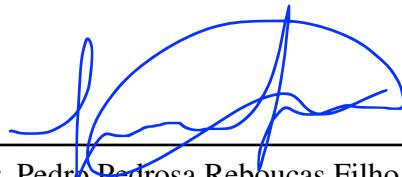
ROBERTO FERNANDES IVO

REDES VOLUMÉTRICAS TOTALMENTE CONVOLUCIONAIS EM CASCATA PARA  
SEGMENTAÇÃO 3D DE CORAÇÃO E GRANDE VASO

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas.

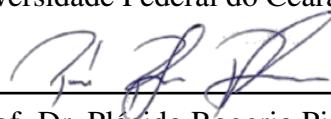
Aprovada em: 07 de Abril de 2021

BANCA EXAMINADORA



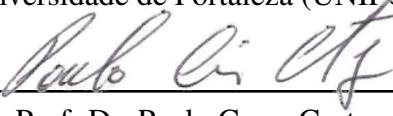
---

Prof. Dr. Pedro Pedrosa Rebouças Filho (Orientador)  
Universidade Federal do Ceará (UFC)



---

Prof. Dr. Plácido Rogerio Pinheiro  
Universidade de Fortaleza (UNIFOR)



---

Prof. Dr. Paulo Cesar Cortez  
Universidade Federal do Ceará (UFC)

*Aos meus pais e toda minha família.*

## **AGRADECIMENTOS**

Primeiramente a Deus, por me conceder saúde e sabedoria para seguir sempre em frente. Obrigado por ser a minha força e o meu guia em todos os momentos. A Ti, Senhor, toda honra e toda a glória.

Aos meus pais, Antônio Ivo e Eliuza, pelo apoio e incentivo em todos os momentos da minha vida. Obrigado por acreditarem em mim, e não medirem esforços para a concretização dos meus sonhos. Sem vocês, nada seria possível.

Ao Prof. Dr. Pedro Pedrosa Rebouças Filho, pela orientação, dedicação, incentivo, paciência e, principalmente, pela amizade construída durante todo o processo. Muito obrigado por todos os seus conselhos e por ter incentivado a conclusão deste ciclo.

A minha namorada, Valéria, por todo o incentivo, carinho e companheirismo durante este período. Obrigado por sempre acreditar na minha capacidade e por todo o amor e dedicação.

Aos meus amigos do Laboratório de Processamento de Imagens, Sinais e Computação Aplicada (LAPISCO), pela convivência e companherismo nesses últimos anos. Obrigado pelo apoio de sempre.

A todos os professores do Departamento de Engenharia de Teleinformática pelo conhecimento não apenas racional, mas a manifestação do caráter e afetividade da educação no processo de formação profissional.

A FUNCAP, pelo financiamento da pesquisa de mestrado.

*“Por mais inteligente que alguém possa ser, se não for humilde, o seu melhor se perde na arrogância. A humildade ainda é a parte mais bela da sabedoria.”*

(Nelson Swel)

## RESUMO

A ausência de câmaras cardíacas, orifícios no coração e/ou conexões anormais vem causando a morte de centenas de pessoas todos os anos. Desta forma, as decisões tomadas pela equipe clínica envolvida no diagnóstico e tratamento para doenças cardíacas congênitas (DCC) devem ser consistentes. Nos casos de DCC, faz-se necessário, portanto, um sistema que seja capaz de segmentar todo o coração e grandes vasos sanguíneos em três dimensões (3D) de maneira eficiente, rápida e precisa. Assim, esta dissertação propõe sanar essa problemática utilizando redes volumétricas totalmente convolucionais em cascata, nomeadamente com o uso de duas Redes Volumétricas de Convolução Total (V-Net) em sequência. A primeira rede visa localizar a região cardíaca, enquanto a segunda segmenta as subestruturas da região cardíaca e dos grandes vasos. Ambas as redes foram treinadas com o conjunto de dados de 2016 do *MICCAI Workshop on Whole-Heart and Great Vessel Segmentation of 3D Cardiovascular MRI in Congenital Heart Disease* (HVSMR). Os resultados experimentais mostraram que o método proposto apresenta um potencial promissor na tomada de decisão em casos de DCC (a partir de imagens de ressonância magnética). A abordagem proposta obteve em média 98,15% para acurácia, 94,89% para precisão, 98,81% para coeficiente de especificidade, 94,27% para coeficiente de sensibilidade, 93,24% para coeficiente de *Matthews*, 80,65% para índice de *Jaccard*, 94,20% para coeficiente de Dice e 1,61mm para a distância de *Hausdorff*. O método proposto possibilitou a visualização e interação do volume segmentado em 3D, que poderá contribuir como recurso de suporte a decisões médicas associados a melhores resultados nas análises estruturais do coração em conjunto com a rede circulatória.

**Palavras-chave:** Segmentação em cascata, V-Net, Doença Cardíaca Congênita.

## ABSTRACT

The absence of cardiac chambers, holes in the heart and abnormal connections cause the death of hundreds of people every year. The clinical team involved in the diagnosis and treatment decisions for congenital heart disease (CHD) must be consistent in their choices. Therefore, in cases of CHD there is a need for a system that is capable of segmenting the whole heart and large blood vessels in 3D efficiently, quickly, and accurately. This article proposes to fill this need by using Cascaded Volumetric Fully Convolutional Networks. The approach proposes the use of two Volumetric Fully Convolution Networks (V-Net) in sequence. The first network aims at locating the cardiac region, while the second segment the substructures of the cardiac area and the great vessels. Both networks are trained with the 2016 data set from the MICCAI Workshop on Whole-Heart and Great Vessel Segmentation of 3D Cardiovascular MRI in Congenital Heart Disease (HVSMR). The experimental results show that the proposed method has a promising potential in decision-making in CHD cases (from MR images). The approach obtained on average 98.15 % for Accuracy, 94.89 % for Precision, 98.81 % for Specificity Coefficient, 94.27 % for Sensitivity Coefficient, 93.24 % for Matthews Coefficient, 80.65 % for Jaccard Index, 94.20 % for Dice Coefficient, and 1.61 for the Hausdorff Distance. The proposed method enables the visualization and iteration of the segmented volume in 3D so that the doctor can analyze the entire structure of the heart along with the circulatory network.

**Keywords:** Segmentation, cascaded, V-Net, Congenital Heart Disease.

## LISTA DE FIGURAS

Figura 1 – Exames de (a) ressonância magnética, (b) tomografia computadorizada e (c) ecocardiografia. . . . .	14
Figura 2 – Representação esquemática de um neurônio artificial. . . . .	22
Figura 3 – Função sigmóide (preto) e sua derivada (vermelho). . . . .	22
Figura 4 – Ilustração do efeito da camada de <i>pooling</i> com operador máximo. . . . .	26
Figura 5 – Representação esquemática da arquitetura da rede V-Net 3D. . . . .	27
Figura 6 – Metodologia da abordagem proposta. . . . .	28
Figura 7 – Representação espacial dos três planos anatômicos (coronal, sagital e axial) obtidos a partir de um exame de ressonância magnética cardíaca. . . . .	29
Figura 8 – Representação das etapas para obtenção de uma imagem tridimensional de coração apartir de exame de ressonância magnética. . . . .	30
Figura 9 – Metodologia adotada para teste com o banco de dados <i>MICCAI Workshop, Whole-Heart e Great Vessel Segmentation de 3D Cardiovascular MRI in Congenital Heart Disease (HVSMR) 2016</i> . . . . .	33
Figura 10 – Representação dos componentes da Equação 3.9 em uma imagem HVSMR 2016. . . . .	35
Figura 11 – Segmentação 3D do coração inteiro e grande vaso obtida através do banco de dados (a esquerda), pelo especialista (ao meio) e de forma automática (a direita). . . . .	37
Figura 12 – Regiões de segmentação obtidas através da metodologia proposta, adotando o verde para os <i>pixels</i> do VP, o azul para o FP e o rosa para o FN. . . . .	38
Figura 13 – Comparação 3D das segmentações do coração inteiro e do grande vaso utilizando a abordagem proposta e sua respectiva segmentação por especialista.	39

## **LISTA DE TABELAS**

Tabela 1 – Resumo cronológico dos trabalhos que foram comparados com a abordagem proposta. A tabela evidencia a técnica, as principais vantagens, desvantagens e o modo de segmentação de cada pesquisa. . . . .	16
Tabela 2 – Exemplos de funções de ativação. . . . .	21
Tabela 3 – Acurácia, Precisão, Especificidade, Sensibilidade, coeficiente de Matthews, índice de Jaccard, coeficiente de Dice e valores métricos de distância de Hausdorff. Seus respectivos desvios padrão acompanham todas as métricas.	40
Tabela 4 – Comparação das métricas de abordagem propostas com outros trabalhos da literatura. . . . .	41

## LISTA DE ABREVIATURAS E SIGLAS

HVSMR	<i>MICCAI Workshop, Whole-Heart e Great Vessel Segmentation de 3D Cardiovascular MRI in Congenital Heart Disease</i>
AHA	<i>American Heart Association</i>
DCC	Doença Cardíaca Congênita
RM	Ressonância Magnética
TC	Tomografia Computadorizada
3DE	Ecocardiografia Tridimensional
ROI	Região de Interesse
PDI	Processamento Digital de Imagens
ANN	Redes Neurais Artificiais ( <i>Artificial Neural Networks</i> )
MLP	Perceptron Multicamadas ( <i>Multilayer Perceptron</i> )
CNN	Rede Neural Convolucional ( <i>Convolutional Neural Network</i> )
RNN	Rede Neural Recorrente ( <i>Recurrent Neural Network</i> )
NIFTI	<i>Neuroimaging Informatics Technology Initiative</i>
PReLU	<i>Parametric Rectified Linear Unit</i>
CCM	Coeficiente de Correlação de Matthews
DH	Distância de Hausdorff

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>13</b>
<b>1.1</b>	<b>Estado da arte . . . . .</b>	<b>15</b>
<b>1.2</b>	<b>Objetivos . . . . .</b>	<b>17</b>
<b>1.2.1</b>	<b><i>Objetivos específicos . . . . .</i></b>	<b>17</b>
<b>1.3</b>	<b>Produção científica . . . . .</b>	<b>17</b>
<b>1.4</b>	<b>Organização da dissertação . . . . .</b>	<b>19</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA . . . . .</b>	<b>20</b>
<b>2.1</b>	<b>Redes Neurais Artificiais . . . . .</b>	<b>20</b>
<b>2.1.1</b>	<b><i>Neurônio Artificial . . . . .</i></b>	<b>21</b>
<b>2.1.2</b>	<b><i>Perceptron Multicamadas . . . . .</i></b>	<b>23</b>
<b>2.2</b>	<b>Redes Neurais Convolucionais . . . . .</b>	<b>24</b>
<b>2.2.1</b>	<b><i>Camada convolucional . . . . .</i></b>	<b>25</b>
<b>2.2.2</b>	<b><i>Camada de agrupamento . . . . .</i></b>	<b>26</b>
<b>2.2.3</b>	<b><i>Camada totalmente conectada . . . . .</i></b>	<b>26</b>
<b>2.3</b>	<b>Rede V-Net . . . . .</b>	<b>27</b>
<b>3</b>	<b>METODOLOGIA . . . . .</b>	<b>28</b>
<b>3.1</b>	<b>Conjunto de dados HVMSR 2016 . . . . .</b>	<b>29</b>
<b>3.2</b>	<b>Preparação do conjunto de dados . . . . .</b>	<b>30</b>
<b>3.3</b>	<b>V-Net . . . . .</b>	<b>31</b>
<b>3.4</b>	<b>Treinamento . . . . .</b>	<b>31</b>
<b>3.5</b>	<b>Teste . . . . .</b>	<b>32</b>
<b>3.6</b>	<b>Métricas de avaliação . . . . .</b>	<b>33</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÕES . . . . .</b>	<b>36</b>
<b>4.1</b>	<b>Resultados qualitativos . . . . .</b>	<b>36</b>
<b>4.2</b>	<b>Resultados quantitativos . . . . .</b>	<b>39</b>
<b>4.3</b>	<b>Comparação com outros métodos no conjunto de dados HVMSR 2016 . . . . .</b>	<b>41</b>
<b>5</b>	<b>CONCLUSÃO E TRABALHOS FUTUROS . . . . .</b>	<b>43</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>44</b>
	<b>ANEXOS . . . . .</b>	<b>49</b>

## 1 INTRODUÇÃO

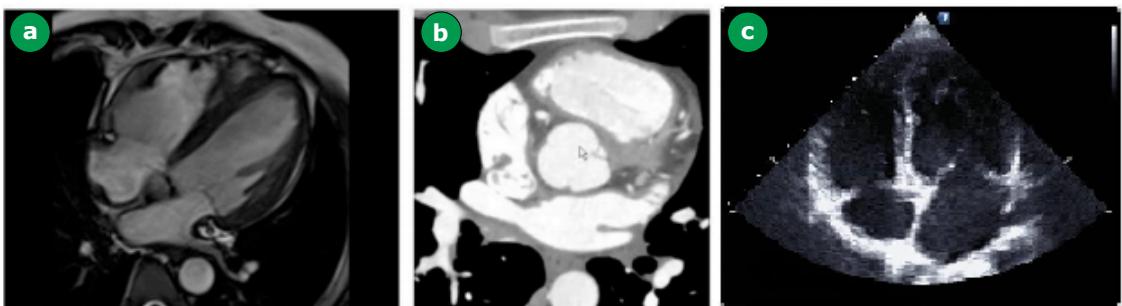
De acordo com a *American Heart Association* (AHA), uma em cada mil crianças nascem no mundo com algum tipo de Doença Cardíaca Congênita (DCC) (STOUT *et al.*, 2019), que corresponde a defeitos anatômicos no coração e/ou na rede circulatória, podendo apresentar gravidades variadas e até comprometer toda a funcionalidade dessas estruturas. Dentro desse cenário, a cada mil casos de DCC diagnosticados, cerca de 6 a 19 estão relacionados a anormalidades moderadas a graves que requerem reparo cirúrgico imediato (STOUT *et al.*, 2019; KWAG *et al.*, 2018).

O formato de imagens 2D são rotineiramente utilizadas para o monitoramento dos pacientes, bem como para o planejamento de eventuais intervenções médicas. No entanto, a falta de informação acerca das relações espaciais da anatomia intracardíaca, em duas dimensões, acaba dificultando procedimentos subsequentes dentro dos casos clínicos (SCHURE; DINARDO, 2019; FAROOQI *et al.*, 2016). Desta forma, a tecnologia 3D no diagnóstico pré e pós-operatório e, no tratamento de pessoas com DCC está ganhando cada vez mais destaque na literatura (BIGLINO *et al.*, 2015; HADEED *et al.*, 2018; SUN *et al.*, 2019; ANWAR *et al.*, 2018; HERMSEN *et al.*, 2017).

Assim, as equipes médicas têm encaminhado os pacientes para exames que permitem visualizações tridimensionais, tais como, Ressonância Magnética (RM) (ZHANG *et al.*, 2018), Tomografia Computadorizada (TC) (RAIMONDI; WARIN-FRESSE, 2016) ou Ecocardiografia Tridimensional (3DE) (JI *et al.*, 2016; SIMPSON; BOSCH, 2019; SPITZER *et al.*, 2017). As varreduras de TC e RM são empregadas para a visualização da anatomia de todo o coração e dos grandes vasos. Já a 3DE é mais adequada para representar estruturas menores, uma vez que, embora exiba um campo de visão menor em comparação com a TC e a RM, apresenta resoluções temporais e espaciais mais altas.

Exemplos de imagens obtidas por exames de RM, TC e 3DE são mostrados na Figura 1. A confiabilidade dos modelos 3D desses exames é totalmente dependente do método de segmentação adotado.

Figura 1 – Exames de (a) ressonância magnética, (b) tomografia computadorizada e (c) ecocardiografia.



Fonte: Adaptado de Capelli e Schievano (2017).

Quando a segmentação da RM é necessária, um especialista deve realizar uma análise cuidadosa de todas as regiões cardíacas. No entanto, a análise manual da RM cardíaca é uma tarefa complexa, subjetiva e sujeita a segmentação errônea. Os potenciais erros são resultados da fadiga e ações repetitivas, dado que, esta análise requer entre 4 a 8 horas. Tais dificuldades podem influenciar adversamente a análise 3D de DCC e, em alguns casos, isso pode afetar a escolha do procedimento médico correto.

A homogeneidade observada entre diferentes regiões do corpo humano e a variabilidade anatômica entre os pacientes com DCC tornam a segmentação das imagens um processo desafiador. Desta forma, muitas pesquisas têm buscado desenvolver tecnologias capazes de analisar até os casos mais complexos, de modo a abranger as necessidades clínicas. Considerando a complexidade envolvida no reconhecimento automatizado de doenças congênitas, melhorias nessa área continuam sendo um grande desafio para os pesquisadores, isto porque, essas adaptações devem apresentar resultados confiáveis, precisos e rápidos, para ser aplicável na rotina clínica.

A fim de contornar esse desafio, o presente estudo propôs o uso de duas redes volumétricas totalmente convolucionais em cascata para segmentar todo o coração e grandes vasos de pessoas com DCC. O conjunto de dados HVSMR 2016 de imagens de RM 3D foi utilizado para verificar a eficiência do método proposto. Em ambas as etapas, as redes foram estruturadas com V-Net. A primeira rede foi construída com o objetivo de localizar a região que contém o coração e os vasos. Já a segunda V-Net foi treinada com base na Região de Interesse (ROI) obtida na primeira rede. Nesse processo, o coração e os grandes vasos foram segmentados em 3D, de modo a possibilitar o diagnóstico, medição de volume, visualização de patologia, planejamento de tratamento, intervenção e geração do mapa de superfície para impressão 3D.

## 1.1 Estado da arte

Pesquisas com enfoque na análise de imagens cardíacas têm recebido atenção significativa nos últimos anos, particularmente, estudos de segmentação da gordura cardíaca (ALBUQUERQUE *et al.*, 2020), átrio (XIONG *et al.*, 2019), vasos cardíacos (GAO *et al.*, 2019) e coração inteiro (LIU *et al.*, 2019; ZHUANG *et al.*, 2010; KHENED *et al.*, 2019). Curiosamente, embora haja necessidade clínica de maiores avanços no que diz respeito a segmentação de todo o coração e dos grandes vasos de pacientes com DCC, poucos estudos têm abordado essas regiões como alvo principal.

Devido a grande variabilidade nos defeitos anatômicos de pessoas com DCC, os métodos encontrados na literatura para segmentação de átrio, ventrículo e artéria podem, por exemplo, não apresentar resultados satisfatórios para a segmentação de todo o coração e grandes vasos (ZHUANG *et al.*, 2019). Os primeiros estudos voltados à segmentação dessas regiões, em pacientes com cardiopatia congênita, surgiram com o desafio da HVSMR em 2016. Essa abordagem é dividida em quatro áreas: Otimização Discreta e Modelagem de Intensidade Probabilística, Estratégias baseadas em Atlas (ZULUAGA *et al.*, 2017; SHAHZAD *et al.*, 2017), *Random Forests* (WANG *et al.*, 2017; MUKHOPADHYAY, 2017) e *Deep Learning* (YU *et al.*, 2017; LI *et al.*, 2017; WOLTERINK *et al.*, 2017).

Zuluaga *et al.* (2017) e Shahzad *et al.* (2017) conduziram uma segmentação usando atlas. Neste caso, considera-se que a forma dos dados do coração é semelhante para as amostras de treinamento e teste. Os valores das métricas obtidas nesses trabalhos (ZULUAGA *et al.*, 2017; SHAHZAD *et al.*, 2017) refletem a dificuldade de tal estratégia para realizar a segmentação do coração de pessoas com deficiência congênita, na qual variações anatômicas são observadas entre os indivíduos. Por outro lado, uma vez que não a relação entre dados de treinamento e dados de teste não é considerada, a confiabilidade da segmentação permanece alta, apesar das diferentes deficiências cardíacas congênitas.

Wang *et al.* (2017) e Mukhopadhyay (2017) usaram o método *Random Forests* para realizar a segmentação cardíaca dos dados do HVSMR 2016. Além do uso de *Random Forests*, Wang *et al.* (2017) realizaram um processamento das imagens para remover truques no resultado do classificador, e por essa razão, um tempo de segmentação maior foi necessário. O estudo de Mukhopadhyay (2017) requereu o uso de Processamento Digital de Imagens (PDI), não se enquadrando no tempo de segmentação necessário. Entretanto, as adversidades acima relatadas foram solucionadas com a utilização da abordagem proposta neste presente estudo,

Tabela 1 – Resumo cronológico dos trabalhos que foram comparados com a abordagem proposta. A tabela evidencia a técnica, as principais vantagens, desvantagens e o modo de segmentação de cada pesquisa.

Trabalhos	Técnica	Vantagens	Desvantagens	Vista
(ZULUAGA <i>et al.</i> , 2017)	Atlas	Os autores avaliaram os pontos fortes e as limitações do uso do Atlas para segmentação de todo o coração.	Requer uma grande variedade de atlas para identificar anomalias cardíacas.	2D
(SHAHZAD <i>et al.</i> , 2017)		Conjuntos combinados de vários atlas e level sets.	O mesmo que Zuluaga <i>et al.</i> (2017).	3D
(WANG <i>et al.</i> , 2017)	Random Forests	Três etapas para segmentação cardíaca: localização, isolamento cardíaco e segmentação miocárdica.	Não forneceu tempo de segmentação.	2D
(MUKHOPADHYAY, 2017)		Apresentou uma segmentação com aspecto mais homogêneo.	Tempo de segmentação: 15-20 min + problema em alguns exames.	2D
(YU <i>et al.</i> , 2017)		Propõe uma nova rede fractal 3D profundamente supervisionada.	Um valor do coeficiente do índice inferior ao alcançado por este trabalho.	3D
(LI <i>et al.</i> , 2017)	Deep Learning	O trabalho comparou várias redes para esta segmentação.	Usou a média de saída de 5 redes para traçar a segmentação 3D.	3D
(WOLTERINK <i>et al.</i> , 2017)		Em casos com variabilidade funcional, a abordagem é satisfatória.	Identificação de grande vaso cardíaco que não está no padrão de referência.	2D

Fonte: Elaborado pelo autor.

que apresentou tempo compatível com as aplicações clínicas, sem a necessidade do emprego de técnicas de processamento de imagem.

Recentemente, as abordagens de aprendizado profundo alcançaram um excelente desempenho na segmentação de objetos (ZHOU *et al.*, 2019), classificação (KHAMPARIA *et al.*, 2019; FERREIRA *et al.*, 2019; NÓBREGA *et al.*, 2018), e detecção (JR *et al.*, 2019; TIWARI *et al.*, 2018), bem como, em problemas de rastreamento (CARNEIRO *et al.*, 2018) e em sistemas de internet das coisas (IoT) (CASOLLA *et al.*, 2019; PICCIALLI *et al.*, 2019; PICCIALLI *et al.*, 2019). Vários modelos de redes neurais convolucionais profundas surgiram nos últimos anos, tais como, AlexNet (KHENED *et al.*, 2019), ZFNet (ZEILER; FERGUS, 2014), VGG (SIMONYAN; ZISSERMAN, 2014), GoogleNet (SZEGEDY *et al.*, 2015), rede residual (HE *et al.*, 2016), rede totalmente convolucional (FCN) (LONG *et al.*, 2015) e U-Net (RONNEBERGER *et al.*, 2015).

Conforme demonstrado na Tabela 1, alguns grupos de pesquisa utilizaram, por exemplo, redes de *Deep Learning* para segmentação 3D do coração inteiro e grandes vasos. Yu *et*

*al.* (2017) e Li *et al.* (2017) usaram redes de convolução 3D e apresentaram o volume segmentado, ao contrário do trabalho de Wolterink *et al.* (2017), que representou apenas a saída em fatias. O grupo de Yu *et al.* (2017) propôs a nova rede fractal 3D profundamente supervisionada, obtendo um coeficiente de Dice de 0,852 - 0,930, que correspondeu ao melhor desempenho dos estudos encontrados na literatura, embora, o tempo de segmentação não tenha sido divulgado. Todavia, nas rotinas clínicas, o tempo é um fator essencial para salvar a vida de uma pessoa, sendo este parâmetro também um dos fatores decisivos para a aplicabilidade de tais propostas.

## 1.2 Objetivos

Segmentar todo o coração e grandes vasos de pessoas com DCC utilizando redes volumétricas totalmente convolucionais em cascata em imagens obtidas a partir de ressonância magnética.

### 1.2.1 *Objetivos específicos*

- Avaliar a viabilidade de uso do conjunto de dados do HVSMR 2016 de imagens de ressonância magnética 3D;
- Localizar a região que contém o coração e os vasos utilizando redes neurais convolucionais nas imagens de ressonância magnética;
- Identificar o coração e grandes vasos nas regiões previamente localizadas;
- Desenvolver uma abordagem de visualização 3D dos resultados obtidos;
- Comparar os resultados obtidos com estudos já publicados para validação da abordagem proposto.

## 1.3 Produção científica

A seguir estão listadas as publicações diretamente relacionadas à pesquisa de mestrado descrita neste documento, sendo o primeiro artigo a produção gerada com o estudo desta dissertação e os demais sendo aplicações relacionadas:

1. HAN, TAO ; IVO, ROBERTO F. ; RODRIGUES, DOUGLAS DE A. ; PEIXOTO, SOLON A. ; DE ALBUQUERQUE, VICTOR HUGO C. ; REBOUÇAS FILHO, PEDRO P. . Cascaded Volumetric Fully Convolutional Networks for Whole-Heart and Great Vessel 3D segmentation. **Future Generation Computer Systems**, v. 108, p. 198-209, 2020. Qualis

- B1 (Engenharia IV).
2. DE ALBUQUERQUE, VICTOR HUGO C.; DE A. RODRIGUES, DOUGLAS ; IVO, ROBERTO F. ; PEIXOTO, SOLON A. ; HAN, TAO ; WU, WANQING ; REBOUÇAS FILHO, PEDRO P. . Fast fully automatic heart fat segmentation in Computed Tomography datasets. **Computerized Medical Imaging and Graphics**, v. 80, p. 101674, 2019. Qualis A1 (Engenharia IV).
  3. CHAGAS, JOÃO V. SOUZA DAS ; IVO, ROBERTO F. ; GUIMARÃES, MATHEUS T. ; DE A. RODRIGUES, DOUGLAS ; DE S. REBOUÇAS, ELIZÂNGELA ; REBOUÇAS FILHO, PEDRO P. . Fast fully automatic skin lesions segmentation probabilistic with Parzen window. **Computerized Medical Imaging and Graphics**, v. 85, p. 101774, 2020. Qualis A1 (Engenharia IV).
  4. CHAGAS, JOÃO V. SOUZA DAS ; RODRIGUES, D. A. ; IVO, R. F. ; HASSAN, M. M. ; ALBUQUERQUE, V. H. C. ; REBOUCAS FILHO, P. P.. A new approach for the detection of pneumonia in children using CXR images based on an real-time IoT system. **Journal of Real-Time Image Processing**, 2021. Qualis A2 (Engenharia IV).
  5. RODRIGUES, DOUGLAS DE A. ; IVO, ROBERTO F. ; SATAPATHY, SURESH CHANDRA ; WANG, SHUIHUA ; HEMANTH, JUDE ; FILHO, PEDRO P. REBOUÇAS . A new approach for classification skin lesion based on transfer learning, deep learning, and IoT system. **Pattern Recognition Letters**, v. 136, p. 8-15, 2020. Qualis A2 (Engenharia IV).
  6. IVO, ROBERTO F.; DE A. RODRIGUES, DOUGLAS ; BEZERRA, GABRIEL M. ; FREITAS, FRANCISCO N.C. ; DE ABREU, HAMILTON FERREIRA GOMES ; REBOUÇAS FILHO, PEDRO P. . Non-grain oriented electrical steel photomicrograph classification using transfer learning. **Journal of Materials Research and Technology-JMRT**, v. 9, p. 8580-8591, 2020. Qualis A1 (Engenharia II).
  7. DE SOUZA, J.W. ; HOLANDA, G. B. ; IVO, R.F. ; ALVER, S.S. ; DA SILVA, S.P. ; NUNES, V.X. ; REBOUÇAS FILHO, PEDRO P.. Predicting body measures from 2D images using Convolutional Neural Networks. **International Joint Conference on Neural Networks (IJCNN)**, p.1-6, 2020.
  8. VIANA, A. A. ; RABELO, S. L. ; SANTOS, J. D. A. ; LIMA FILHO, V. X. ; RODRIGUES, D. A. ; IVO, R. F. ; REBOUÇAS FILHO, PEDRO P.. Automatic Evaluation of the Mechanical Properties of Steels Maraging using Digital Image Processing Techniques.

**Conference on Graphics, Patterns and Images (SIBGRAPI)**, v. 33, p. 211-218, 2020.

#### **1.4 Organização da dissertação**

Esta dissertação de mestrado está organizada em cinco capítulos. O Capítulo 1 introduz aspectos relevantes para o desenvolvimento do presente estudo; o Capítulo 2 fornece fundamentação teórica sobre os métodos utilizados na metodologia proposta; o Capítulo 3 apresenta a metodologia da abordagem desenvolvida; o Capítulo 4 discute os resultados experimentais obtidos nesta pesquisa; e o Capítulo 5, apresenta as principais conclusões e perspectivas futuras.

## 2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta uma revisão das redes neurais convolucionais, constituindo o arcabouço teórico para a compreensão da abordagem proposta e para a avaliação dos resultados obtidos. O aprendizado profundo (*Deep Learning*) é formado por um conjunto de métodos que modelam dados com arquiteturas complexas, combinando diferentes transformações não lineares. A base desse aprendizado são as Redes Neurais Artificiais (*Artificial Neural Networks*) (ANN), que são combinadas para formar as redes neurais profundas (GOODFELLOW *et al.*, 2016; BISHOP, 2006).

Essas técnicas permitiram um progresso significativo nas áreas de processamento de sinais e imagens, incluindo reconhecimento facial, reconhecimento de fala e processamento de linguagem natural. As aplicações potenciais são numerosas. Um exemplo notável é o programa *AlphaGo*, que aprendeu a jogar *Go* pelo método de aprendizado profundo, vencendo o campeão mundial em 2016 (LAPAN, 2018).

Há vários tipos de arquiteturas de redes neurais artificiais que se baseiam em uma cascata profunda de camadas: Perceptron Multicamadas (*Multilayer Perceptron*) (MLP), Rede Neural Convolucional (*Convolutional Neural Network*) (CNN) e Rede Neural Recorrente (*Recurrent Neural Network*) (RNN). Essas arquiteturas precisam de algoritmos de otimização estocásticos, além de uma escolha inteligente de inicialização dos parâmetros para projetar sua estrutura (LAPAN, 2018).

### 2.1 Redes Neurais Artificiais

As ANN baseiam-se em uma coleção de unidades conectadas ou nós, chamados neurônios artificiais, que modelam vagamente os neurônios em um cérebro biológico. Cada conexão, como nas sinapses em um cérebro biológico, pode transmitir um sinal a outros neurônios. Um neurônio artificial que recebe um sinal, o processa e, pode sinalizar os neurônios conectados a ele. O sinal em uma conexão é um número real, e a saída de cada neurônio é calculada por alguma função não linear da soma de suas entradas,  $y = f(\mathbf{x}, \theta)$  (BRAGA *et al.*, 2007).

Como costuma ocorrer no aprendizado estatístico, os parâmetros  $\theta$  são estimados a partir de uma amostra dos dados. A função de minimização não é convexa, levando a minimizadores locais. O sucesso do método veio de um teorema de aproximação universal devido a Cybenko (1989) e Hornik (1991). Além disso, LeCun *et al.* (1988) propuseram uma

maneira eficiente de calcular o gradiente de uma rede neural, chamada *backpropagation* do gradiente, que permite a obtenção facilmente de um minimizador local do critério quadrático.

### 2.1.1 Neurônio Artificial

Um neurônio artificial é uma função  $f_j$  de entrada  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_d)$  ponderada por um vetor de pesos de conexão  $\mathbf{w}_j = (\mathbf{w}_{j,1}, \dots, \mathbf{w}_{j,d})$ , complementada por um viés (*bias*) de neurônio  $b_j$ , e associado a uma função de ativação  $\phi$ ,

$$y_j = f_j(\mathbf{w}) = \phi(\langle \mathbf{w}_j, \mathbf{x} \rangle + b_j). \quad (2.1)$$

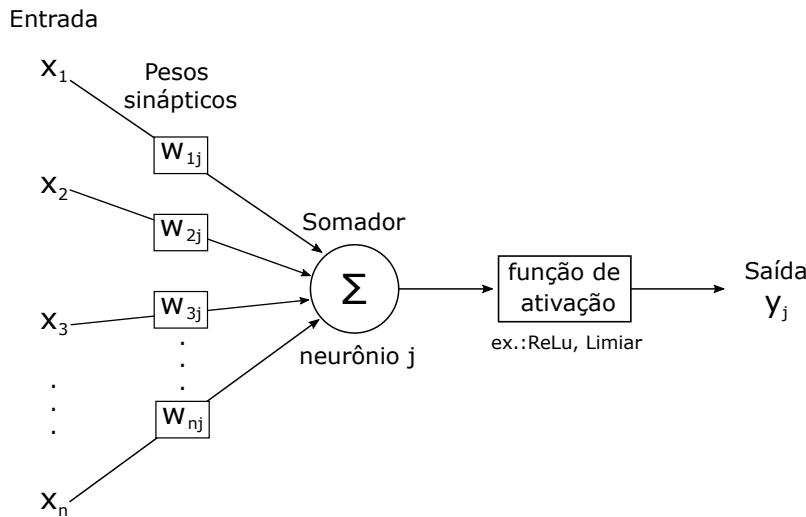
Várias funções de ativação podem ser consideradas. A Tabela 2 mostra as funções mais utilizadas. A representação esquemática de um neurônio artificial é apresentada na Figura 2.

Tabela 2 – Exemplos de funções de ativação.

Função	Definição
<b>Identidade</b>	$\phi(\mathbf{x}) = \mathbf{x}$ ,
<b>Sigmóide (ou logística)</b>	$\phi(\mathbf{x}) = \frac{1}{1 + \exp(-\mathbf{x})}$ ,
<b>Tangente hiperbólica (tanh)</b>	$\phi(\mathbf{x}) = \frac{\exp(\mathbf{x}) - \exp(-\mathbf{x})}{\exp(\mathbf{x}) + \exp(-\mathbf{x})} = \frac{\exp(2\mathbf{x}) - 1}{\exp(2\mathbf{x}) + 1}$ ,
<b>Limiar rígido</b>	$\phi_\beta(\mathbf{x}) = 1_{\mathbf{x} \geq \beta}$ ,
<b>Unidade Linear Retificada (ReLU)</b>	$\phi(\mathbf{x}) = \max(0, \mathbf{x})$ .

Fonte: Elaborado pelo autor.

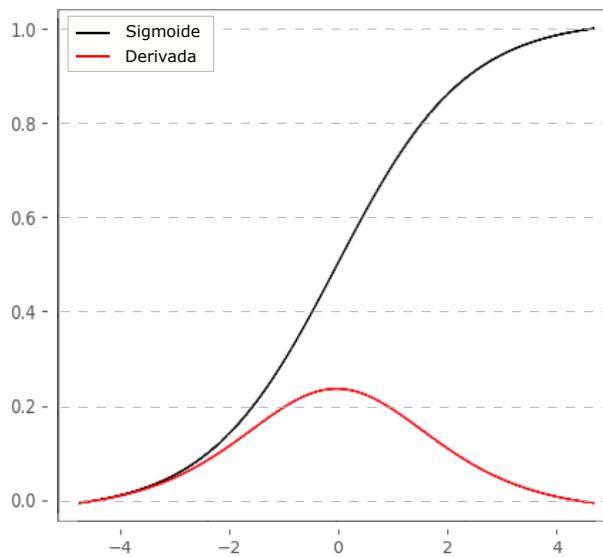
Figura 2 – Representação esquemática de um neurônio artificial.



Fonte: Adaptado de Braga *et al.* (2007).

Historicamente, a sigmóide foi a função de ativação mais utilizada, já que é diferenciável e permite manter os valores no intervalo  $[0, 1]$ . No entanto, seu gradiente é muito próximo de 0 quando  $|x|$  não está próximo a 0. A Figura 3 mostra a função sigmóide e sua derivada.

Figura 3 – Função sigmóide (preto) e sua derivada (vermelho).



Fonte: Elaborado pelo autor.

No caso do aprendizado profundo que é uma rede neural com um grande número de camadas, isso causa problemas para o algoritmo de treinamento estimar os parâmetros. Por conta disto, a função sigmóide foi substituída pela função de Unidade Linear Retificada. A função ReLU não é diferenciável em 0, mas na prática, isso não é realmente um problema, pois a probabilidade de ter uma entrada igual a 0 é geralmente nula. Esse problema costuma

ser contornado adicionando uma pequena polarização positiva para garantir que cada unidade esteja ativa. Diversas variações da função ReLU são considerados para certificar de que todas as unidades tenham um gradiente que não desaparece e que quando  $\mathbf{x} < 0$  a derivada não seja igual a 0. Isto é,

$$\phi(\mathbf{x}) = \max(\mathbf{x}, 0) + \alpha \min(\mathbf{x}, 0), \quad (2.2)$$

onde  $\alpha$  é um parâmetro fixo definido com um pequeno valor positivo ou um parâmetro a ser estimado.

### 2.1.2 Perceptron Multicamadas

A MLP é uma estrutura composta por várias camadas ocultas de neurônios, onde a saída de um neurônio de uma camada torna-se a entrada de um neurônio da próxima camada. Além disso, a saída de um neurônio também pode ser a entrada de um neurônio da mesma camada ou de um neurônio de camadas anteriores (como ocorre nas redes neurais recorrentes). Na última camada, chamada de camada de saída, pode-se aplicar uma função de ativação diferente das camadas ocultas dependendo do tipo de problema: regressão ou classificação.

A MLP têm uma arquitetura básica, pois cada neurônio de uma camada está ligada a todos os da próxima camada, mas não tem nenhuma ligação com os neurônios da mesma camada. Os parâmetros da arquitetura correspondem ao número de camadas ocultas e de neurônios em cada camada. As funções de ativação também são escolhidas pelo usuário. Na camada de saída, a função de ativação é geralmente diferente daquela usada nas camadas ocultas. Em problemas de regressão, nenhuma função de ativação é aplicada na camada de saída. Para classificação binária, a saída fornece uma predição de  $P(Y = 1/X)$ , e uma vez que este valor está em  $[0, 1]$ , a função de ativação sigmóide é geralmente considerada. Para classificação multiclasse, a camada de saída contém um neurônio por classe  $i$ , dando uma predição de  $P(Y = i/X)$ . A soma de todos esses valores deve ser igual a 1. Neste caso, a função multidimensional softmax é geralmente usada:

$$\text{softmax}(z)_i = \frac{\exp(z_i)}{\sum_j \exp(z_j)}. \quad (2.3)$$

Em resumo, a formulação matemática de um Perceptron Multicamadas com  $k =$

$1, \dots, CO$  camadas ocultas, considerando  $h^{(0)}(\mathbf{x}) = \mathbf{x}$  é,

$$\begin{aligned} a^{(k)}(\mathbf{x}) &= b^{(k)} + \mathbf{W}^{(k)} h^{(k-1)}(\mathbf{x}) \\ h^{(k)}(\mathbf{x}) &= \phi(a^{(k)}(\mathbf{x})). \end{aligned} \tag{2.4}$$

Para  $k = CO + 1$  camadas de saída,

$$\begin{aligned} a^{(CO+1)}(\mathbf{x}) &= b^{(CO+1)} + \mathbf{W}^{(CO+1)} h^{(CO)}(\mathbf{x}) \\ h^{(CO+1)}(\mathbf{x}) &= \phi_s(a^{(CO+1)}(\mathbf{x})) := f(\mathbf{x}, \theta), \end{aligned} \tag{2.5}$$

em que  $\phi$  é a função de ativação da camada oculta e  $\phi_s$  é a função de ativação da camada de saída (por exemplo, softmax para classificação multiclasse).

## 2.2 Redes Neurais Convolucionais

As CNNs são uma classe de algoritmo de *Deep Learning* projetadas para processar matrizes estruturadas de dados, normalmente imagens. Ao contrário dos algoritmos de visão computacional clássicos, as CNNs podem operar diretamente em dados brutos com menor pré-processamento. Outra particularidade dessas redes, em relação aos métodos primitivos de visão computacional, é sua capacidade de aprender as características relevantes para cada problema (LECUN *et al.*, 1998).

As camadas convolucionais em uma CNN assemelham-se à estrutura do córtex visual humano, no qual uma série de camadas processam uma imagem de entrada e identificam características progressivamente mais complexas. Assim, cada uma dessas camadas convolucionais empilhadas são capazes de aprender formas mais sofisticadas (BISHOP, 2006).

A principal diferença entre as arquiteturas das redes neurais artificiais convencionais e as CNNs, é o fato de que as CNNs supõem que as entradas são matrizes estruturadas de dados, em geral imagens. Isso torna a implementação do passo de propagação (*forward*) no treinamento mais eficiente e reduz amplamente a quantidade de parâmetros na rede.

Já as ANNs não se adaptam bem com imagens completas. No conjunto de dados *Canadian Institute For Advanced Research* (CIFAR-10) (KRIZHEVSKY *et al.*, 2009), por exemplo, mesmo com o tamanho das imagens sendo  $32 \times 32 \times 3$  (32 canais de largura, 32 de altura e 3 canais de cor), um único neurônio totalmente conectado em uma primeira camada oculta de uma ANN teria  $32 * 32 * 3 = 3072$  pesos. Essa quantidade ainda parece administrável,

mas claramente essa estrutura totalmente conectada não se ajusta a imagens maiores. Por outro lado, uma entrada maior, *e.g.*  $200 \times 200 \times 3$ , exigiria neurônios com  $200 * 200 * 3 = 120.000$  pesos. Notoriamente, essa conectividade é um desperdício e o grande número de parâmetros levaria rapidamente ao sobreajuste (*overfitting*).

Por outro lado, as CNNs exploram o fato da entrada ser uma matriz estruturada de dados para otimizar sua arquitetura. Diferente de uma ANN, os neurônios de suas camadas são organizados em largura, altura e profundidade. Vale ressaltar que o termo profundidade não se refere ao número total de camadas em uma rede. No CIFAR-10, por exemplo, o volume teria dimensões de  $32 \times 32 \times 3$  por possuir 3 canais de cores. Além disso, os neurônios em uma camada estão conectados apenas a uma pequena região da camada anterior, ao invés de em todos os neurônios.

Uma CNN é composta por diversos tipos de camadas, descritas nas próximas seções: camada convolucional, camada de agrupamento (*pooling*) e camada totalmente conectada. O empilhamento dessas camadas formam uma arquitetura CNN (GOODFELLOW *et al.*, 2016).

### 2.2.1 Camada convolucional

A camada convolucional é o bloco de construção central de uma CNN, responsável pela maior parte do trabalho computacional. Seus parâmetros consistem em um conjunto de filtros que podem ser treinados. Cada filtro é espacialmente pequeno (ao longo da largura e da altura), mas se estende por toda a profundidade do volume de entrada. Conforme o filtro é convoluído com esse volume de entrada, um mapa de ativação bidimensional que fornece as respostas desse filtro em cada posição espacial é produzido. Intuitivamente, a rede aprenderá filtros que se ativam com a presença de padrões visuais. Esses mapas de ativação são empilhados ao longo da dimensão de profundidade e produzem o volume de saída (GOODFELLOW *et al.*, 2016). A convolução discreta entre duas funções  $f$  e  $g$  é definida:

$$(f * g)(x) = \sum_t f(t)g(x+t). \quad (2.6)$$

Para sinais bidimensionais, consideramos as convoluções:

$$(\mathbf{W} * \mathbf{I})(i, j) = \sum_{m,n} \mathbf{W}(m, n)\mathbf{I}(i+m, j+n), \quad (2.7)$$

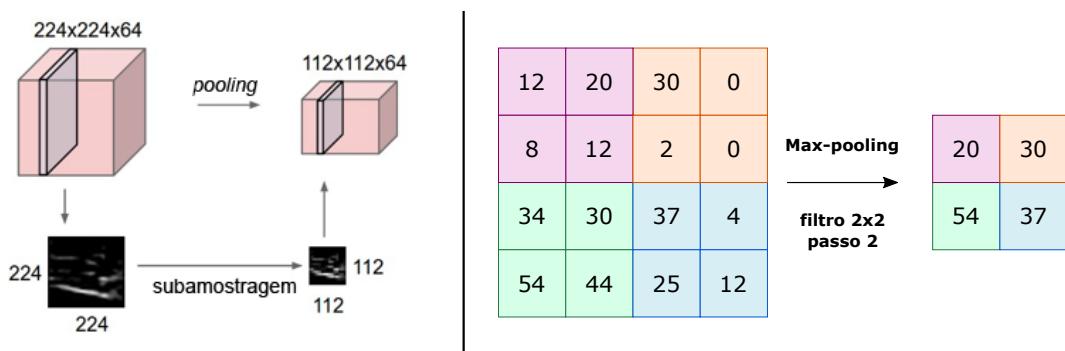
em que  $\mathbf{W}$  é um *kernel* (ou filtro) de convolução aplicado a um sinal bidimensional  $\mathbf{I}$ .

### 2.2.2 Camada de agrupamento

Outra camada importante nas CNNs são as camadas de agrupamento (*pooling*), também conhecidas como subamostragem, que permitem reduzir a dimensão, ao calcular a média ou o máximo em regiões da imagem.

Assim como nas camadas convolucionais, a camada de *pooling* age em pequenos trechos da imagem e podem apresentar passos (*strides*). Se calculado o valor máximo de regiões  $2 \times 2$ , considerando um passo  $s = 2$ , a largura e altura da imagem é então dividida por 2, como mostra a Figura 4. Vale ressaltar que também é possível reduzir a dimensão da camada convolucional com um passo maior do que 1, sem preenchimento de zero. Outra vantagem da camada de *pooling* é tornar a rede menos sensível a pequenas translações das imagens de entrada.

Figura 4 – Ilustração do efeito da camada de *pooling* com operador máximo.



Fonte: Adaptado de *CS231n Convolutional Neural Networks for Visual Recognition*.

### 2.2.3 Camada totalmente conectada

Os neurônios em uma camada totalmente conectada têm conexões para todas as ativações na camada anterior, como visto nas ANNs. Portanto, suas ativações podem ser calculadas com uma multiplicação de matriz juntamente com o *bias*.

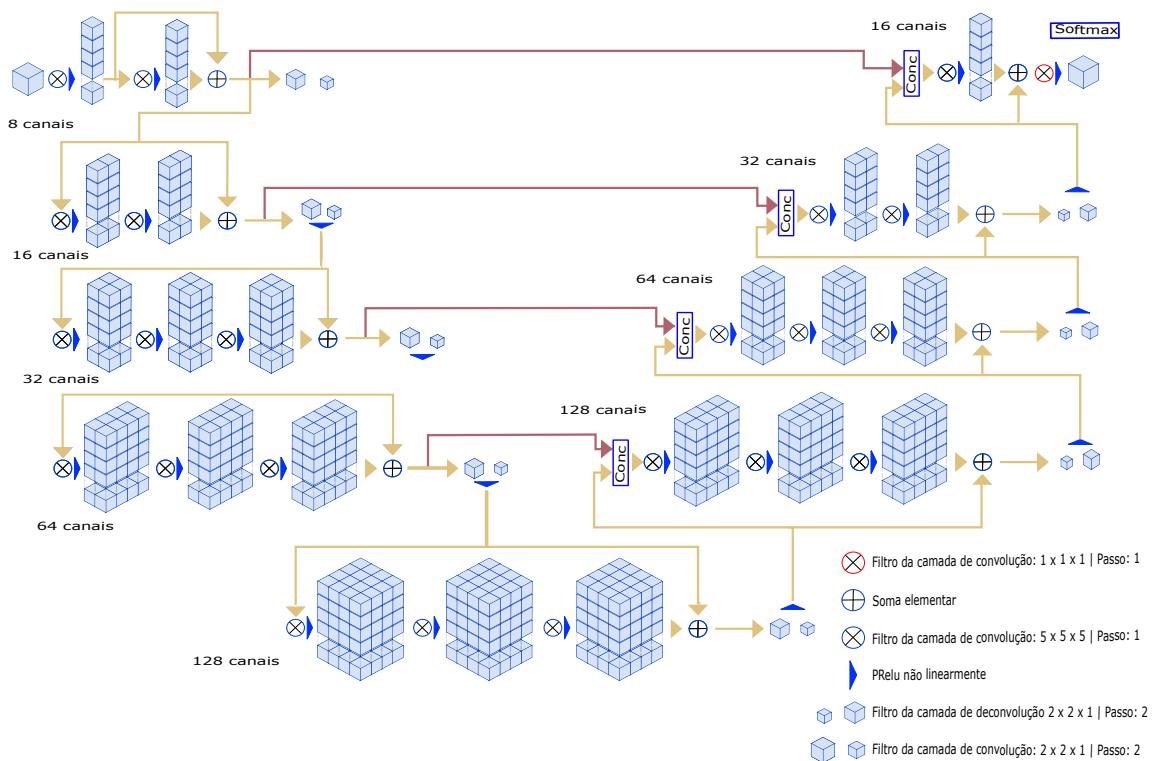
A única diferença entre as camadas totalmente conectadas e as convolucionais é que os neurônios desta camada estão conectados apenas a uma região local na entrada, sendo que muitos deles compartilham parâmetros. No entanto, em ambas as camadas, os neurônios ainda computam produtos escalares, e desta forma, sua forma funcional é idêntica. Neste sentido, uma conversão entre essas camadas é possível (GOODFELLOW *et al.*, 2016).

### 2.3 Rede V-Net

As CNNs tornaram-se famosas pelo processamento de imagens 2D, conforme proposto inicialmente pela rede U-Net (RONNEBERGER *et al.*, 2015). No entanto, nas rotinas clínicas, a maioria dos dados médicos corresponde a volumes 3D. Diante desse cenário, as Redes Volumétricas Totalmente Convolucionais foram introduzidas para segmentar todo o volume de um exame médico.

Uma representação esquemática da arquitetura V-Net é apresentada na Figura 5. Conforme observado, esta rede é semelhante à popular rede U-Net.

Figura 5 – Representação esquemática da arquitetura da rede V-Net 3D.



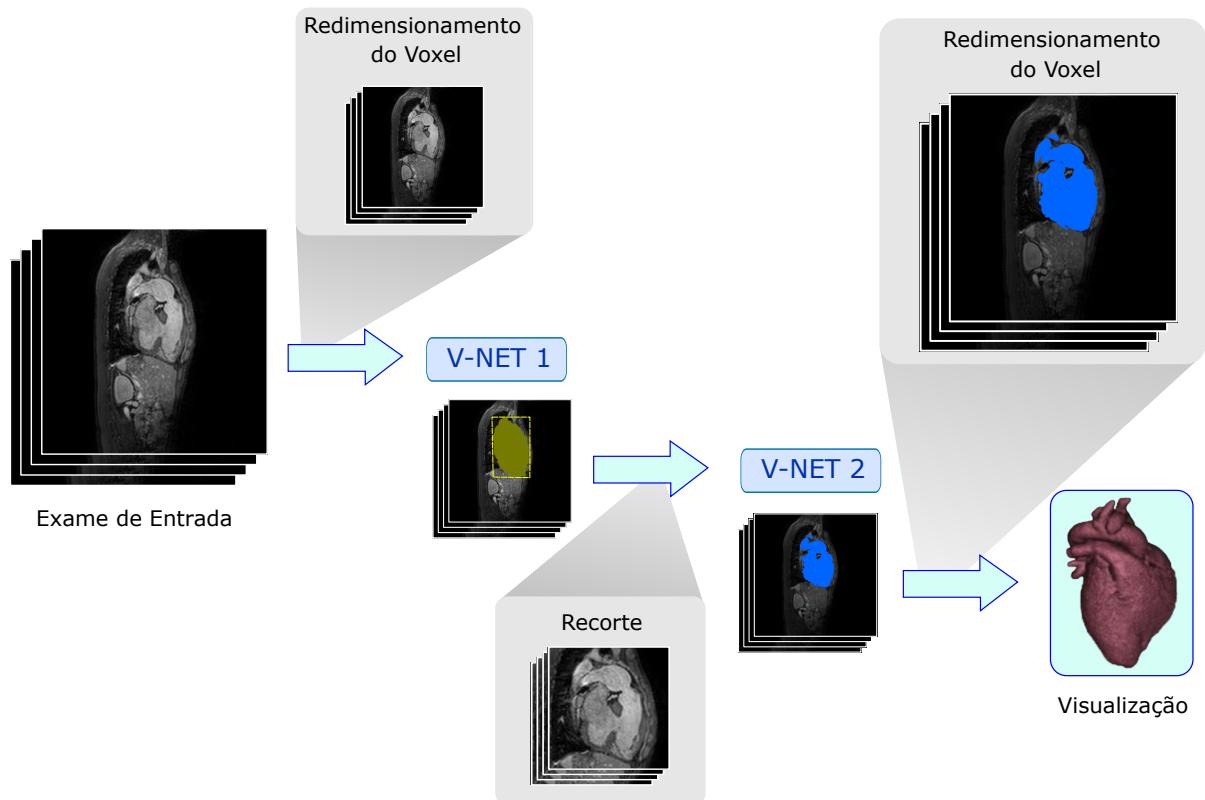
Fonte: Elaborado pelo autor.

Esta arquitetura apresentou um bom desempenho em abordagens envolvendo segmentação de próstata (MILLETARI *et al.*, 2016), atrial (XIA *et al.*, 2019), tumor cerebral (CASAMITJANA *et al.*, 2018), dentre outras. Os estudos de Xia *et al.* (2019) e Casamitjana *et al.* (2018) utilizaram V-Net em cascata para tumores auriculares e cerebrais, apesar de que nenhuma reconstrução 3D dos órgãos segmentados foi realizada. Em suma, esses estudos apontam que o uso de redes V-Net trata-se, sem dúvida, de uma ferramenta que a equipe médica pode utilizar em suas atividades de tomada de decisão.

### 3 METODOLOGIA

Este Capítulo descreve o banco de dados (Seção 3.1), o pré-processamento (Seção 3.2) o método proposto (Seção 3.3), e as métricas empregadas para avaliar a segmentação de todo o coração e grandes vasos (Seção 3.6). A Figura 6 exibe o fluxograma do processo.

Figura 6 – Metodologia da abordagem proposta.



Fonte: Elaborado pelo autor.

Um dos principais desafios ao segmentar o coração é a homogeneidade da intensidade do *pixel* dos grandes vasos na região cardíaca. Conforme mostrado na Figura 6, inicialmente os dados do *voxel* foram redimensionados, para posteriormente, serem utilizados como a entrada da primeira V-Net. A saída da rede obtida definiu a localização da ROI, que tornou-se a entrada para a segunda V-Net. Ao final, um redimensionamento para o tamanho original do exame foi aplicado, e o volume em 3D pôde ser obtido para a visualização e interação do médico.

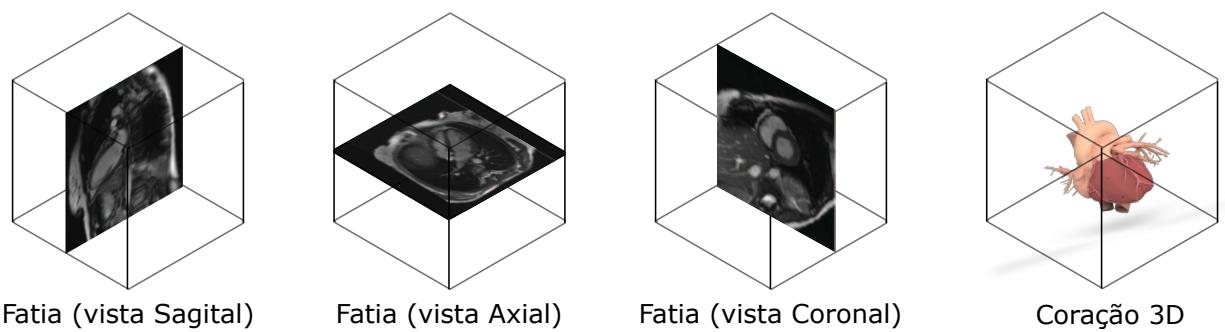
### 3.1 Conjunto de dados HVMSR 2016

O conjunto de dados do HVMSR 2016<sup>1</sup> (PACE *et al.*, 2015) foi utilizado para execução desse estudo. Todos os exames foram adquiridos durante a prática clínica no *Boston Children's Hospital, Boston, MA, EUA*. Dez exames de ressonância magnética cardiovascular 3D foram disponibilizados. Interessantemente, uma variedade de defeitos cardíacos congênitos estava incluída no banco de dados.

As imagens nos exames foram obtidas sob visão axial em um scanner 1.5T (Philips Achieva), sem um agente de contraste. Em média, o volume total encontrado nos exames foi de 390 x 390 x 165 voxels, enquanto que o espaçamento foi de 0,9 x 0,9 x 0,85 mm. Dois especialistas clínicos realizaram as segmentações manuais.

Os exames de RM permitem uma análise nos três planos anatômicos: coronal, sagital e axial como evidenciado na Figura 7.

Figura 7 – Representação espacial dos três planos anatômicos (coronal, sagital e axial) obtidos a partir de um exame de ressonância magnética cardíaca.

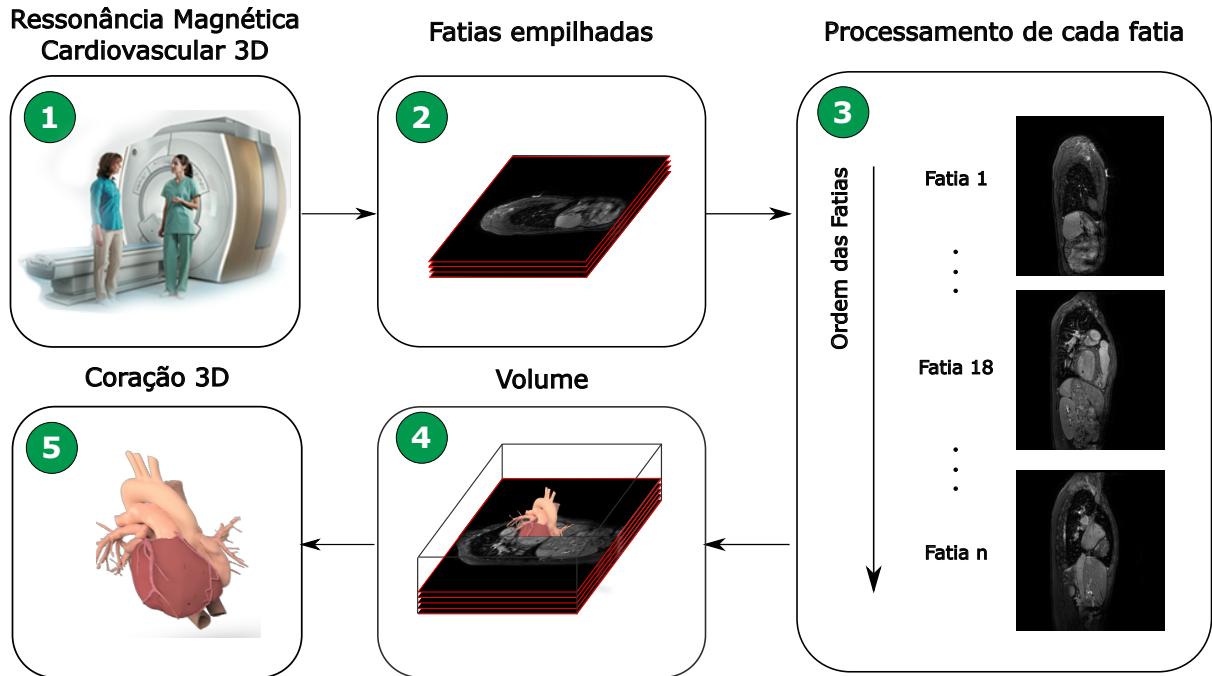


Fonte: Elaborado pelo autor.

A organização HVMSR 2016 forneceu (1) imagens axiais de volume total; (2) imagens axiais cortadas; e (3) imagens recortadas que foram transformadas em uma visão de eixo curto de tamanho aproximado (PACE *et al.*, 2015). Todas as imagens estavam no formato *Neuroimaging Informatics Technology Initiative* (NIFTI), cuja extensão é .nii (LAROBINA *et al.*, 2014). Objetivando uma aplicabilidade em rotinas clínicas, este estudo optou por fazer uso de imagens axiais de volume total, de acordo com a aquisição original disponibilizada através de ressonância magnética. A Figura 8 ilustra o coração em plano axial, as fatias empilhadas das imagens de RM, as imagens da base, meio e topo do coração, volume e coração 3D.

<sup>1</sup> <http://segchd.csail.mit.edu/>

Figura 8 – Representação das etapas para obtenção de uma imagem tridimensional de coração apartir de exame de ressonância magnética.



Fonte: Elaborado pelo autor.

### 3.2 Preparação do conjunto de dados

O banco de dados fornecidos pelo HVSMR 2016 tem diferentes dimensões de volume. Para o plano XY, os diferentes valores são  $256 \times 256$ ,  $384 \times 384$  e  $516 \times 516$  pixels, enquanto o número de divisões (eixo Z) varia de 104 a 174. Uma das características de uma CNN é que as entradas devem possuir um tamanho fixo. Portanto, é necessário padronizar o tamanho dos dados e torná-los adequados para redes V-Net.

Nem todas as fatias dos volumes de RM possuem a ROI, ou seja, a região do coração e dos grandes vasos, e por isso, são removidas do volume. Por exemplo, dado um determinado exame com 174 fatias, em que apenas 150 fatias têm ROI, um novo volume com apenas 150 fatias é então criado. Os volumes de RM ainda podem diferir em tamanho e portanto, para efetuar uma padronização, todos os volumes foram redimensionados para  $256 \times 256 \times 64$  voxels.

Duas interpolações foram utilizadas para esta etapa de padronização. Em ambos os casos, foi utilizada uma interpolação bilinear do volume de dados. Além dos métodos, o volume foi rotacionado  $90^\circ$  no plano x-y para facilitar a visualização. Os dados do volume mostram a ROI em branco e o fundo em preto.

### 3.3 V-Net

Inicialmente, a V-Net comprime o volume do exame (ilustrados a esquerda - Figura 5). Em seguida, esse volume é descomprimido até que as dimensões originais do exame clínico sejam atingidas. Além disso, existe ainda conexões horizontais entre as duas partes. A Figura 5 ilustra todos esses recursos.

Os primeiros estágios operam em diferentes resoluções. O valor das resoluções de cada estágio é resultado das convoluções com núcleos volumétricos de tamanho  $2 \times 2 \times 2$  voxels. Portanto, as dimensões dos volumes são divididas por dois, enquanto o número de canais de recursos aumenta na mesma proporção. Em cada estágio, existe de uma a três camadas convolucionais, que usam núcleos de tamanho  $5 \times 5 \times 5$  voxel. O V-Net usa a *Parametric Rectified Linear Unit* (PReLU) (HE *et al.*, 2015) como sua função de ativação e uma função residual é aprendida em cada um de seus estágios, ao contrário do que é observado na rede U-Net.

A segunda parte da V-Net (ilustrada a direita - Figura 5) expande os recursos anteriormente reduzidos para informações que restauram o volume. Assim como mencionado anteriormente, uma função residual também é aprendida nesta etapa. Na última camada da rede, um núcleo volumétrico  $1 \times 1 \times 1$  é usado para calcular os dois mapas de recursos finais e assim, atingir o tamanho inicial. Esses dois mapas correspondem às segmentações probabilísticas que indicam as regiões de primeiro e segundo plano usando a função *Softmax*. Já na deconvolução, as conexões horizontais fornecem as informações de localização. Além de melhorar a qualidade da previsão, essas conexões reduzem o tempo de convergência.

### 3.4 Treinamento

Em redes em cascata, a saída da primeira rede é a entrada para a segunda. Nesse caso, a primeira rede visa localizar a região do coração, enquanto que a segunda, segmenta a região do coração e quaisquer subestruturas significativas de vasos.

Para realizar o treinamento das duas redes, os pesos foram atualizados com uma taxa de aprendizagem fixa de 0,001, sendo que nove em cada dez exames foram selecionados como amostras de treinamento. O número de épocas para a primeira e segunda redes, respectivamente, foi 10 e 300 e o modelo com a pontuação mais alta foi salvo. O tamanho do lote utilizado foi 1 devido à restrição de memória. Além disso, o modelo V-Net proposto nesta dissertação usa a função de perda binária *Crossentropy* (Equação 3.1), em que o rótulo é representado por  $y$ , e  $\hat{y}$  é

a probabilidade esperada de pertencer ou não à região de interesse.

$$L(y, \hat{y}) = -\frac{1}{N} \sum_{i=0}^N (y \cdot \log(\hat{y}_i) + (1-y) \cdot \log(1-\hat{y}_i)) \quad (3.1)$$

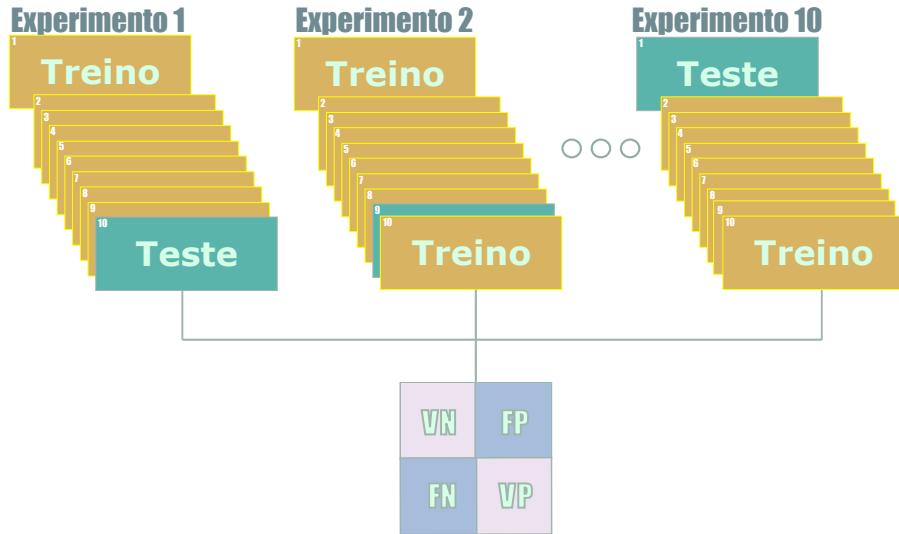
Inicialmente, a rede V-Net possui 16 canais no primeiro estágio, depois 32 canais e assim por diante (MILLETARI *et al.*, 2016). Nesta dissertação, a V-Net foi adaptada para que o primeiro estágio tenha oito canais, que resulta em um segundo estágio que possui 16 canais e assim por diante. Esta adaptação requer menos poder computacional.

A saída da primeira rede V-Net apresenta apenas a região onde todo o coração e os grandes vasos estão presentes no volume, implicando numa redução no número de cortes. O critério empregado é o máximo e o mínimo da máscara de saída. Por sua vez, o treinamento da segunda rede é baseado apenas na saída da primeira V-Net. Já a saída desta segunda rede corresponde à segmentação do coração e dos grandes vasos, onde há, em seguida, o redimensionamento do *voxel*, que retorna ao tamanho original do volume. Todos esses processos são mostrados detalhadamente na Figura 6.

### 3.5 Teste

A metodologia para a realização dos testes de segmentação 3D de todo o coração e grandes vasos é exemplificada na Figura 9. No estágio de teste, houve um processo de validação cruzada de 10 vezes em um conjunto de dados aleatório. Na primeira validação, um exame foi separado para as provas e os outros foram destinados para treinamento. Na segunda validação, um exame diferente foi encaminhado para o conjunto de testes e o restante para o treinamento. As demais validações se seguiram conforme esse procedimento, e assim, os parâmetros foram obtidos para constituir a matriz de confusão em cada processo de validação cruzada.

Figura 9 – Metodologia adotada para teste com o banco de dados HVSMR 2016.



Fonte: Elaborado pelo autor.

### 3.6 Métricas de avaliação

A avaliação da segmentação é baseada no número de *pixels* e *voxels* corretamente e/ou incorretamente alocados para uma região da imagem. As métricas medem a discrepância ou semelhança entre a região segmentada e a região da imagem de referência.

As seguintes métricas de avaliação foram calculadas para comparar o desempenho da segmentação de *pixels* quantitativamente: Acurácia (Equação 3.2), Precisão (Equação 3.3), Especificidade (Equação 3.4), Sensibilidade (Equação 3.5), Coeficiente de Matthews (Equação 3.6), Índice de Jaccard (Equação 3.7), coeficiente de Dice (Equação 3.8) e Distância de Hausdorff (Equação 3.9).

Para cada imagem, há uma contagem de verdadeiro positivo (VP), verdadeiro negativo (VN), falso positivo (FP) e falso negativo (FN). Os valores de VP e FN representam, respectivamente, o número de *pixels* de uma determinada região alocados correta e incorretamente. Já VN corresponde ao número de *pixels* que não pertencem a uma determinada região classificada como não pertencente a esta região. Os valores de FP representam o número de *pixels* classificados incorretamente como pertencentes a uma determinada região.

Acurácia (Acc) significa a proporção de *pixels* do coração e grandes vasos que estão corretamente segmentados por todos os *pixels* na imagem. Precisão (Pre) relaciona todos os verdadeiros positivos a todas as previsões positivas. A especificidade (Sp) reflete o quanto eficaz é a identificação correta de *pixels* que não fazem parte do ROI. Sp é a porcentagem de *pixels*

adequadamente excluídos dos resultados de segmentação em relação a todos os *pixels* que não sejam de coração e vasos. Sensibilidade (Sen) é a taxa na qual os *pixels* pertencentes à região cardíaca são identificados corretamente em comparação com o número total de *pixels*.

Considerando que há menos *pixels* para classificar a região de interesse do que o fundo, o Coeficiente de Correlação de Matthews (CCM) torna-se uma métrica relevante neste estudo para o cálculo da correlação de predição independentemente do tamanho das regiões.

$$Acc = \frac{VP + VN}{VP + VN + FP + FN} \quad (3.2)$$

$$Pre = \frac{VP}{VP + FN} \quad (3.3)$$

$$Sp = \frac{VN}{VN + FP} \quad (3.4)$$

$$Sen = \frac{VP}{VP + FN} \quad (3.5)$$

$$CCM = \frac{VP \cdot VN - FP \cdot FN}{\sqrt{(VP + FP)(VP + FN)(VN + FP)(VN + FN)}} \quad (3.6)$$

Nas Equações 3.7 e 3.8, os termos  $S_{gt}$  e  $S_a$ , respectivamente, representam a região segmentada manualmente pelo especialista e a região segmentada pelo método.

O índice de Jaccard (Jac) é uma medida estatística de similaridade entre conjuntos de amostras. Para dois conjuntos de dados, essa métrica é a cardinalidade de sua interseção dividida pela cardinalidade da soma entre a segmentação prevista e especializada. O coeficiente de Dice é usado para medir a similaridade entre regiões segmentadas de forma automática e manual. O valor de sobreposição dos intervalos com taxa 0 caracteriza a não sobreposição entre duas regiões, enquanto um valor igual a 1 representa a sobreposição perfeita.

$$Jac = \frac{S_{gt} \cap S_a}{S_{gt} \cup S_a} = \frac{VP}{VP + FP + FN} \quad (3.7)$$

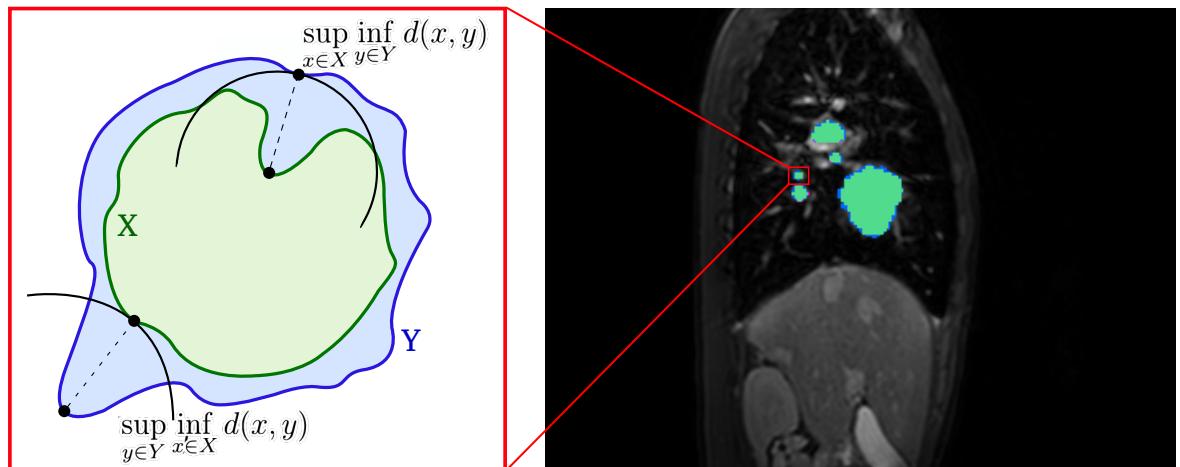
$$Dice = \frac{S_{gt} \cap S_a}{S_{gt} + S_a} = \frac{2VP}{(VP + FP)(VP + FN)} \quad (3.8)$$

A Distância de Hausdorff (DH) é a métrica usada para verificar o alcance entre dois objetos (HAUSDORFF, 1914; MULQUEEN *et al.*, 2001). Como X e Y correspondem a dois contornos de dois objetos, a métrica é avaliada pela Equação 3.9.

$$DH(X, Y) = \max\{A, B\} \quad (3.9)$$

O termo A da Equação 3.9 é  $\sup_{x \in X} \inf_{y \in Y} d(x, y)$ , enquanto o parâmetro B é  $\sup_{y \in Y} \inf_{x \in X} d(x, y)$ , no qual  $\sup$ ,  $\inf$  e  $d(x, y)$  indicam os operadores supremos, operadores infinitos e a distância euclidiana entre os dois pontos x e y, respectivamente. A Figura 10 ilustra os componentes da Equação 3.9 (FILHO *et al.*, 2019). Nesta dissertação, foi utilizado o método para cálculo de DH conforme Taha e Hanbury (2015). O algoritmo apresenta um desempenho eficiente para um grande número de pontos, além de ter uma complexidade quase linear, superando o algoritmo DH padrão da Biblioteca ITK (TAHA; HANBURY, 2015).

Figura 10 – Representação dos componentes da Equação 3.9 em uma imagem HVSMR 2016.



Fonte: Adaptado de Filho *et al.* (2019).

## 4 RESULTADOS E DISCUSSÕES

Este Capítulo apresenta os resultados obtidos pela abordagem proposta, categorizados em qualitativos (Seção 4.1) e quantitativos (Seção 4.2), e apresenta posteriormente uma comparação entre esses resultados com aqueles obtidos a partir de estudos que utilizaram a mesma base de dados (Seção 4.3).

As metodologias aplicadas foram implementadas na linguagem *python* usando o IDE *PyCharme*, juntamente com a biblioteca gratuita *OpenCV 3.0* integrada com *Tensorflow* (ABADI *et al.*, 2016). Todo o processo computacional foi realizado em um computador com sistema operacional Ubuntu, processador Intel Core i5 com 3,7 GHz, 16 GB de RAM e GPU Nvidia RTX 2060 6G.

### 4.1 Resultados qualitativos

A comparação qualitativa da segmentação da abordagem proposta com a segmentação por especialistas é demonstrada nas Figuras 11, 12 e 13.

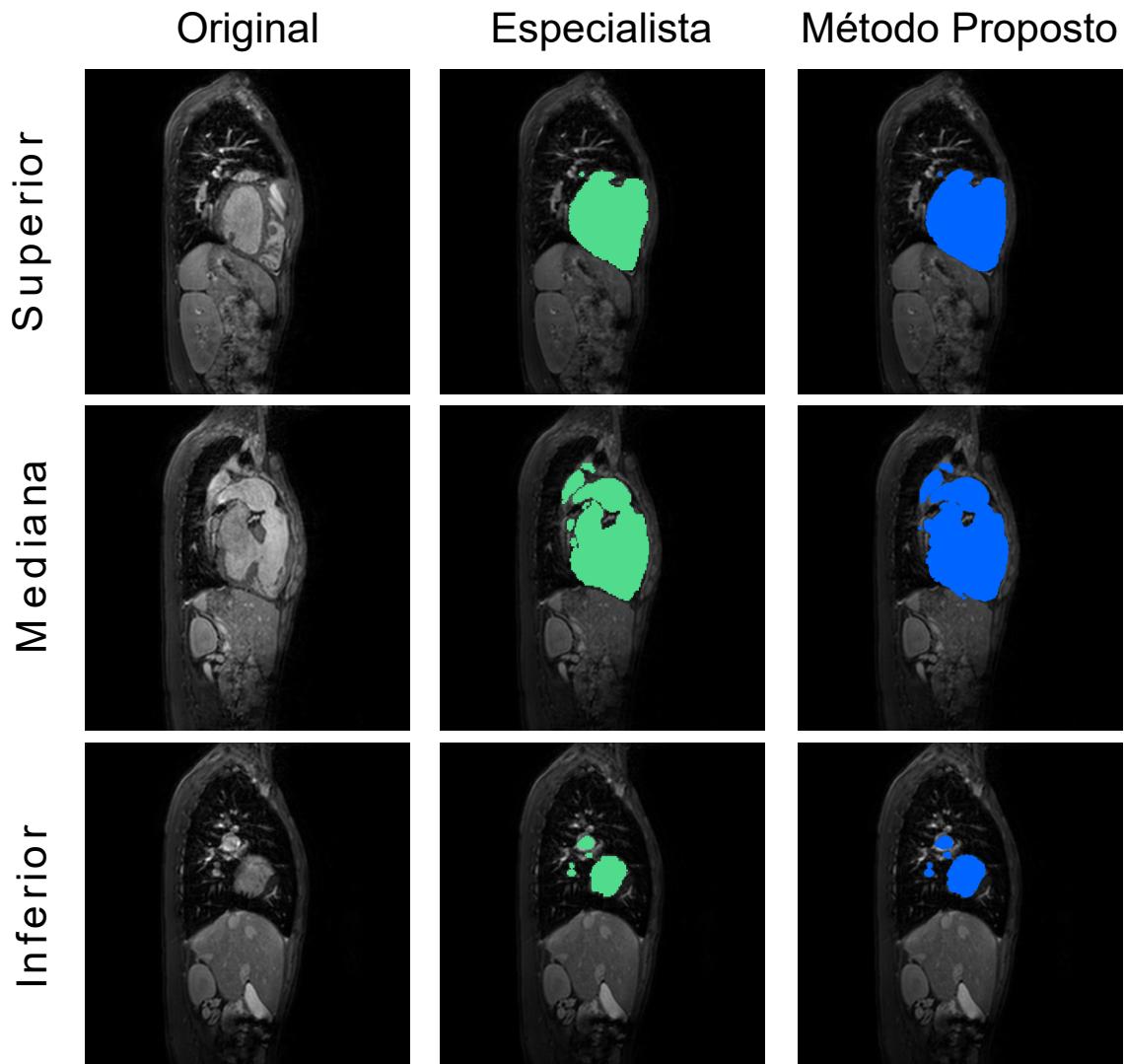
Conforme ilustra a Figura 11, alguns cortes do mesmo exame de ressonância magnética cardíaca, onde são retratadas a imagem original resultante do exame, a imagem de segmentação realizada pelo especialista e a imagem de segmentação da abordagem propostam, abrangendo as fatias correspondentes à parte superior do coração, à parte média do volume do coração e uma fatia da parte inferior do coração.

A sobreposição da imagem segmentada pelo método proposto com a imagem segmentada do especialista é mostrada na Figura 12. A cor verde indica *pixels* de verdadeiro positivo (VP), branco *pixels* de verdadeiro negativo (VN), azul falso positivo (FP) e rosa falso negativo (FN). A diferença nessas cores está presente devido à segmentação excessiva ou subsegmentação. A supersegmentação corresponde ao ROI não marcado pelo especialista e a sub-segmentação corresponde a detecção de apenas parte da região de interesse. A abordagem proposta foi capaz de segmentá-lo o ROI, e nos três exames da Figura 11, a região verde foi maior do que as áreas de outras cores; demonstrando assim, a precisão na segmentação do método proposto.

A Figura 13 exemplifica três exames clínicos provenientes de HVSMR 2016, junto com suas respectivas verdades fundamentais em diferentes ângulos que variaram horizontalmente:  $60^\circ$ ,  $120^\circ$ ,  $180^\circ$ ,  $240^\circ$ ,  $300^\circ$  e  $360^\circ$ .

Todos os resultados abaixo retratados 11, 12 e 13 indicaram que as segmentações

Figura 11 – Segmentação 3D do coração inteiro e grande vaso obtida através do banco de dados (a esquerda), pelo especialista (ao meio) e de forma automática (a direita).

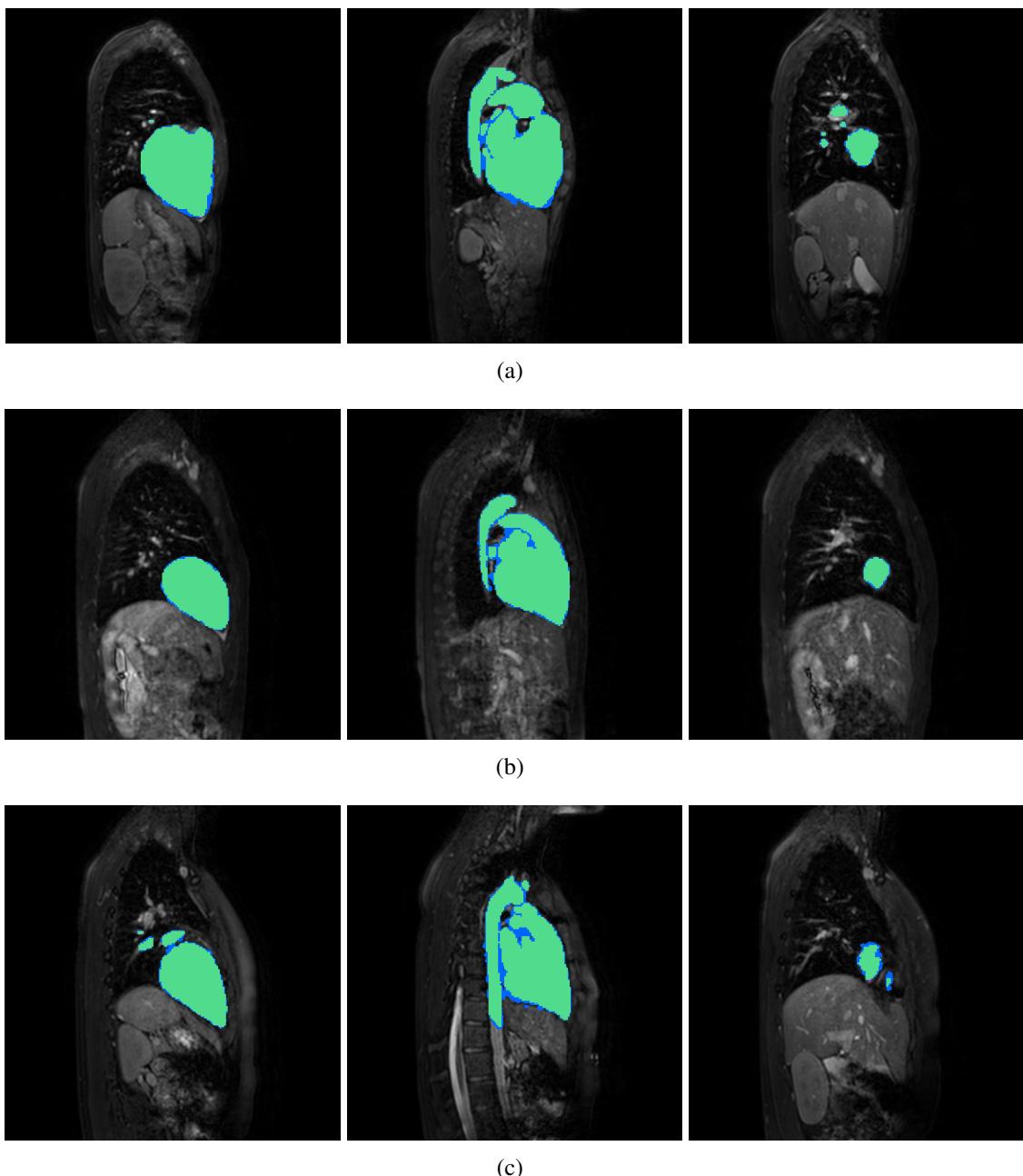


Fonte: Elaborado pelo autor.

realizadas pelas redes volumétricas totalmente convolucionais em cascata podem ser usados para a finalidade proposta.

A segmentação de todo o coração e dos grandes vasos é resultado, muitas vezes, da soma de várias escolhas pessoais de um profissional. Por exemplo, conforme observado na primeira coluna da Figura 11, há um limite mínimo entre o que faz parte da região do coração e as outras regiões do corpo humano, e isso pode gerar dúvidas para o analista quanto à determinação de uma área fazer ou não parte da ROI, sendo possível da mesma incerteza acontecer em diversas regiões. Consequentemente, existem localizações de ROI que acabam sendo perdidas pelo especialista e por sua vez, resultam em uma análise clínica incorreta. Nesse caso, considerando que a abordagem proposta utiliza segmentação semântica, uma redução dos erros de segmentação

Figura 12 – Regiões de segmentação obtidas através da metodologia proposta, adotando o verde para os *pixels* do VP, o azul para o FP e o rosa para o FN.

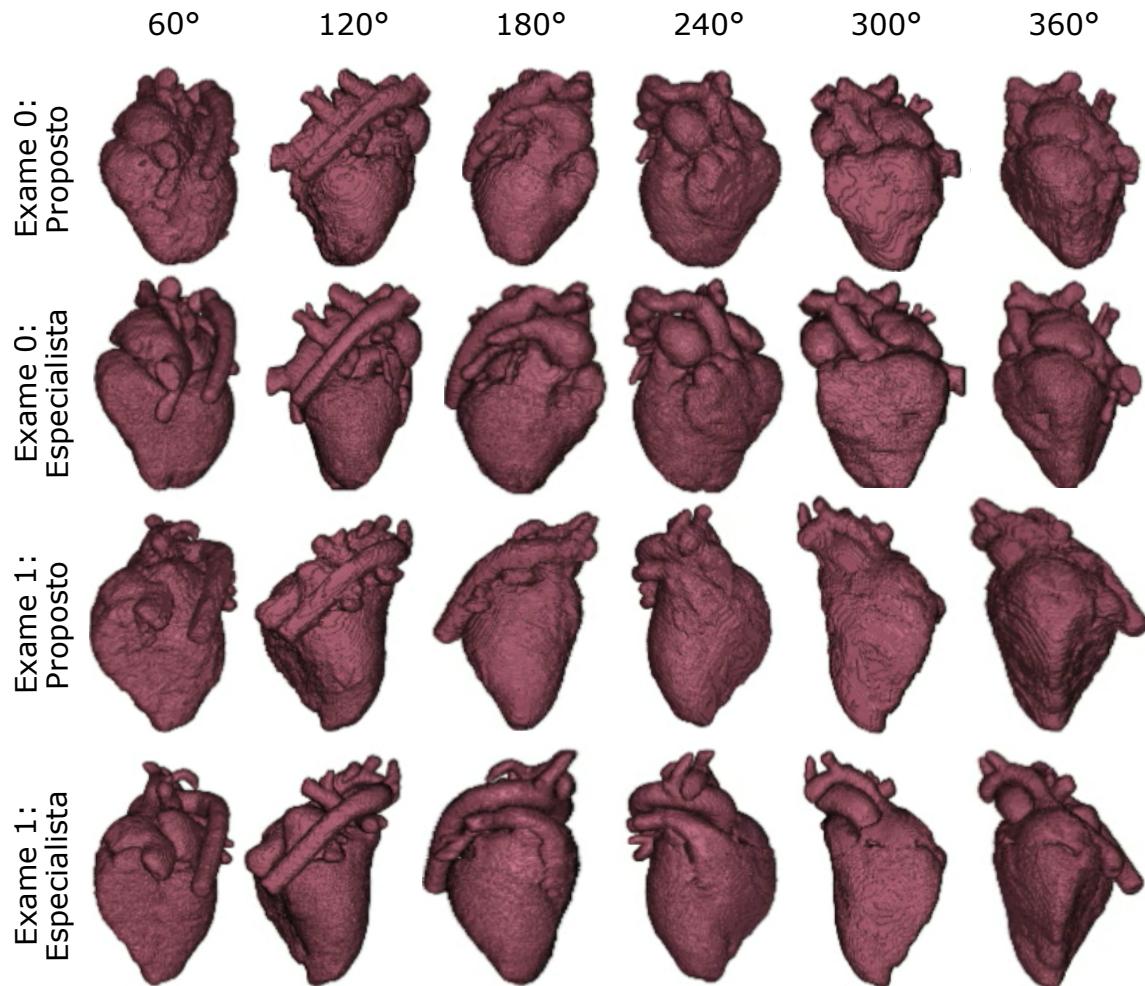


Fonte: Elaborado pelo autor.

pode ser observada.

A partir dessa segmentação, uma equipe médica teria menor dificuldade em tomar decisões sobre o tratamento ou procedimento cirúrgico adequado para pessoas com cardiopatia congênita. A análise qualitativa mostrou claramente os detalhes das estruturas do coração, podendo o volume do órgão ser calculado, e isso possibilitaria uma impressão 3D do coração, o

Figura 13 – Comparação 3D das segmentações do coração inteiro e do grande vaso utilizando a abordagem proposta e sua respectiva segmentação por especialista.



Fonte: Elaborado pelo autor.

que permitiria simulações cirúrgicas.

## 4.2 Resultados quantitativos

O tempo utilizado para o treinamento da primeira rede V-Net foi de 6 minutos, consumindo em média 5,4 GB de memória da GPU, enquanto que a segunda rede V-Net precisou de 35 minutos para treinamento com consumo semelhante de cerca de 5,4 GB de memória GPU.

Para estimar a segmentação de uma varredura de ressonância magnética, isto é, na fase de teste, a segmentação conseguiu ser estimada em apenas 4 segundos com o uso de cerca de 1,8 GB de memória GPU. Esse curto tempo de segmentação de um coração e grandes vasos reforça o potencial da abordagem para uso em rotinas clínicas.

Os valores das métricas avaliativas de cada exame com seus respectivos desvios

padrão estão indicados na Tabela 3, que mostra as métricas de exatidão, precisão, especificidade, sensibilidade, coeficiente de Matthews, índice de Jaccard, coeficiente de Dice e distância de Hausdorff.

Tabela 3 – Acurácia, Precisão, Especificidade, Sensibilidade, coeficiente de Matthews, índice de Jaccard, coeficiente de Dice e valores métricos de distância de Hausdorff. Seus respectivos desvios padrão acompanham todas as métricas.

Exame	Acurácia (%)	Precisão (%)	Especificidade (%)	Sensibilidade (%)	Matthews (%)	Jaccard (%)	Dice (%)	Hausdorff (mm)
1	98.39 ± 0.01	94.14 ± 0.06	98.21 ± 0.03	97.38 ± 0.02	94.61 ± 0.04	91.79 ± 0.06	95.55 ± 0.04	1.58 ± 0.05
2	98.31 ± 0.01	92.49 ± 0.04	97.99 ± 0.00	98.99 ± 0.03	94.60 ± 0.03	91.66 ± 0.05	95.58 ± 0.03	1.61 ± 0.25
3	98.16 ± 0.01	93.51 ± 0.03	97.85 ± 0.02	95.69 ± 0.03	93.05 ± 0.05	89.33 ± 0.09	93.90 ± 0.06	1.56 ± 0.03
4	98.12 ± 0.01	96.89 ± 0.09	99.73 ± 0.02	92.00 ± 0.05	93.02 ± 0.06	89.19 ± 0.09	93.96 ± 0.06	1.61 ± 0.30
5	98.09 ± 0.01	96.95 ± 0.09	99.50 ± 0.02	91.27 ± 0.06	92.69 ± 0.05	88.67 ± 0.09	93.63 ± 0.07	1.61 ± 0.29
6	98.08 ± 0.01	97.19 ± 0.08	99.74 ± 0.02	91.27 ± 0.06	92.83 ± 0.06	88.87 ± 0.08	93.80 ± 0.06	1.63 ± 0.29
7	98.01 ± 0.01	97.01 ± 0.08	99.73 ± 0.02	90.60 ± 0.06	92.42 ± 0.06	88.08 ± 0.08	93.36 ± 0.06	1.60 ± 0.28
8	98.12 ± 0.01	91.21 ± 0.07	98.00 ± 0.02	96.92 ± 0.06	92.83 ± 0.06	88.99 ± 0.08	93.91 ± 0.06	1.58 ± 0.27
9	98.06 ± 0.01	92.57 ± 0.05	97.71 ± 0.02	97.28 ± 0.04	93.54 ± 0.05	90.37 ± 0.06	94.78 ± 0.04	1.67 ± 0.27
10	98.14 ± 0.01	97.00 ± 0.07	99.71 ± 0.02	91.30 ± 0.06	92.81 ± 0.07	88.74 ± 0.08	93.76 ± 0.06	1.61 ± 0.28
Média	98.15	94.89	98.81	94.27	93.24	80.65	94.2	1.61

Fonte: Elaborado pelo autor.

Os resultados apresentados na Tabela 3 (referentes ao exame 1) mostram que o método proposto obteve o maior valor de acurácia com 98,39% e o menor desvio padrão de 0,01%, sendo que a acurácia média foi de 98,15%. Esses resultados podem ser confirmados pelos altos níveis de precisão, sensibilidade e especificidade que acompanham os valores de exatidão. O valor médio dessas métricas foram de 94,89%, 98,81% e 94,27%, respectivamente. Desvios padrão baixos indicam que o percentual dessas métricas tende a ficar próximo da média de cada exame.

Acurácia, Precisão, Especificidade e Sensibilidade são métricas amplamente utilizadas na literatura para avaliar a segmentação. Entretanto, a área do coração e dos grandes vasos é menor do que a região apresentada no exame de ressonância magnética, e por isso, essas métricas não mostram claramente as pequenas diferenças nos resultados obtidos com a segmentação. Assim, torna-se fundamental a análise de outras métricas que privilegiam a avaliação apenas da região de destino, ou seja, a região de interesse. Essas métricas são coeficiente de Matthews, índice de Jaccard e coeficiente de Dice.

Os resultados obtidos para CCM, Índice de Jaccard e coeficiente de Dice apresentaram valores médios iguais a 93,24%, 80,65% e 94,20%. Além disso, no exame 1 foi obtido um valor de CCM de 94,61%, indicando que a abordagem proposta usando redes volumétricas totalmente convolucionais em cascata forneceu uma melhor previsão para o coração inteiro e

o grande vaso. No exame 1, o valor do Índice de Jaccard foi 91,79% e o valor do coeficiente de Dice foi 95,55%. Desta forma, a região segmentada pela abordagem proposta possui grande semelhança com a segmentação dos especialistas. Ademais, os valores de DH representam qual pode ser a separação máxima entre os contornos em um determinado ponto. Aqui, os resultados foram próximos a 1,61 mm, que indica uma distância pequena entre as curvas. A menor distância foi de 1,56 mm obtida no exame 3.

#### **4.3 Comparação com outros métodos no conjunto de dados HVMRS 2016**

Os resultados da segmentação 3D do coração e dos grandes vasos pelo método proposto foram comparados com outros estudos da literatura utilizando estratégias baseadas em Atlas, *Random Forests* e *Deep Learning*. Neste caso, a segmentação de especialistas foi considerada a verdade fundamental. Na abordagem aqui proposta oito métricas foram utilizadas, além de uma avaliação qualitativa em 2D e 3D, ao contrário dos trabalhos na literatura que restringiram a análise relacionadas ao coeficiente de Dice e DH. A Tabela 4 retrata os valores das métricas percentuais do coeficiente de Dice e DH em milímetros.

Tabela 4 – Comparação das métricas de abordagem propostas com outros trabalhos da literatura.

Trabalhos	Coeficiente de Dice	Distância de Hausdorff
<b>Método Proposto</b>	<b>94.20 %</b>	<b>1.61 mm</b>
(ZULUAGA <i>et al.</i> , 2017)	61.00 %	-
(SHAHZAD <i>et al.</i> , 2017)	82.00 %	7.25 mm
(WANG <i>et al.</i> , 2017)	76.00 %	9.22 mm
(MUKHOPADHYAY, 2017)	72.80 %	-
(YU <i>et al.</i> , 2017)	89.35 %	4.05 mm
(LI <i>et al.</i> , 2017)	80.10 %	-
(WOLTERINK <i>et al.</i> , 2017)	86.00 %	6.60 mm

Fonte: Elaborado pelo autor.

Com exceção da pesquisa de Li *et al.* (2017), os trabalhos utilizando *Deep learning* apresentavam resultados superiores para as métricas avaliadas. Porém, ao realizar a comparação com o método aqui proposto, nota-se que um maior valor de coeficiente de Dice foi obtido, além de um melhor valor de DH, isto é, a menor distância. Neste caso, o coeficiente de Dice obtido foi de 94,20% contra 89,35% obtido através de outra rede de aprendizagem profunda (YU *et al.*, 2017). Além disso, a DH é cerca de 2,5 vezes menor que a distância relatada por Yu *et al.* (2017).

A principal limitação do método proposto é a necessidade de um computador com alta alocação de memória, o que justifica a realização do redimensionamento do *voxel* antes da primeira V-Net e o retorno ao tamanho inicial após a segunda rede. Embora este redimensionamento cause a perda de informações de alguns exames, conforme observado pelos resultados, nenhum prejuízo significativo foi observado. A metodologia, portanto, é a mais eficiente para segmentação.

Uma das principais implicações práticas deste método é a sua aplicabilidade nas rotinas clínicas devido a sua segmentação 3D eficiente, rápida e totalmente automática para diferentes casos de pacientes com DCC. Além disso, os excelentes resultados observados foram alcançados com um número reduzido de testes disponibilizados pelo HVSMR 2016. A mudança no número de canais de convoluções volumétricas considerado neste trabalho abre a possibilidade de usar V-Nets em outras aplicações sem a necessidade de um extenso uso de consumo de memória computacional.

Ademais, a alta precisão e o tempo reduzido na segmentação 3D do coração e grandes vasos de pacientes com DCC reforçam a empregabilidade desta metodologia para fins de assistência médica. Além disso, outros impactos positivos que essa metodologia traz para a área da saúde podem ser citados, tais como, visualização 3D da patologia para planejamento do tratamento e/ou intervenção cirúrgica, medição do volume cardíaco, mapeamento de superfície para impressão 3D do órgão e utilização como ferramenta para fins educacionais. Além de que, outro fator relevante a ser considerado é a possibilidade de aplicação desta metodologia eficiente a outros órgãos humanos, que de um modo geral, forneceria informações cruciais para auxiliar no diagnóstico e tratamentos de outras patologias.

## 5 CONCLUSÃO E TRABALHOS FUTUROS

O estudo do desempenho da rede V-Net em cascata aplicada para segmentação 3D do coração e grandes vasos em imagens de RM apresentou resultados promissores no que concerne as métricas avaliadas de acurácia, precisão, coeficiente de especificidade, coeficiente de sensibilidade, coeficiente de Matthews, índice de Jaccard, coeficiente de Dice e distância de Hausdorff. Esses resultados confirmam a possibilidade de integração da abordagem proposta com as rotinas clínicas, em que essa metodologia poderia funcionar como ferramenta de auxílio as equipes médicas em processos de tomada de decisão envolvendo casos de DCC.

Dentre as principais contribuições desta dissertação, podem ser citados: (1) o desenvolvimento de um método de segmentação totalmente automático, rápido e eficaz em nível de voxel para as diferentes variabilidades anatômicas do coração e dos grandes vasos cardíacos, (2) a apresentação de uma visão tridimensional (3D) do Coração Inteiro e do Grande Vaso para auxiliar a equipe médica em casos clínicos de DCC, (3) o uso de redes V-Net em cascata para resolver a segmentação de um problema complexo na região cardíaca de pacientes com DCC, além da (4) redução do número de canais de convolução volumétrica, reduzindo o consumo de memória.

De um modo geral, o desenvolvimento desta dissertação, abre portas para novas pesquisas com enfoque em redes convolucionais aplicadas a segmentação 3D do coração. Nesse sentido, a utilização da rede V-Net para distinguir as múltiplas estruturas cardíacas; a obtenção do cálculo do volume do coração a partir dos resultados da segmentação 3D e a aplicação de diferentes redes convolucionais para desenvolver a segmentação de vasos sanguíneos e miocárdio são algumas das propostas que contribuirão para o desenvolvimento de sistemas computacionais que visam o auxílio ao diagnóstico clínico.

## REFERÊNCIAS

- ABADI, M.; BARHAM, P.; CHEN, J.; CHEN, Z.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; IRVING, G.; ISARD, M.; KUDLUR, M.; LEVENBERG, J.; MONGA, R.; MOORE, S.; MURRAY, D. G.; STEINER, B.; TUCKER, P.; VASUDEVAN, V.; WARDEN, P.; WICKE, M.; YU, Y.; ZHENG, X. Tensorflow: A system for large-scale machine learning. In: **Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation**. Berkeley, CA, USA: USENIX Association, 2016. (OSDI'16), p. 265–283. ISBN 978-1-931971-33-1.
- ALBUQUERQUE, V. H. C. de; RODRIGUES, D. de A.; IVO, R. F.; PEIXOTO, S. A.; HAN, T.; WU, W.; FILHO, P. P. R. Fast fully automatic heart fat segmentation in computed tomography datasets. **Computerized Medical Imaging and Graphics**, v. 80, p. 101674, 2020. ISSN 0895-6111.
- ANWAR, S.; SINGH, G. K.; MILLER, J.; SHARMA, M.; MANNING, P.; BILLADELLO, J. J.; EGHTESADY, P.; WOODARD, P. K. 3d printing is a transformative technology in congenital heart disease. **JACC: Basic to Translational Science**, v. 3, n. 2, p. 294 – 312, 2018. ISSN 2452-302X.
- BIGLINO, G.; CAPELLI, C.; WRAY, J.; SCHIEVANO, S.; LEAVER, L.-K.; KHAMBADKONE, S.; GIARDINI, A.; DERRICK, G.; JONES, A.; TAYLOR, A. M. 3d-manufactured patient-specific models of congenital heart defects for communication in clinical practice: feasibility and acceptability. **BMJ open**, British Medical Journal Publishing Group, v. 5, n. 4, p. e007165, 2015.
- BISHOP, C. M. **Pattern recognition and machine learning**. Secaucus, NJ, USA: Springer, 2006. ISBN 0387310738.
- BRAGA, A. de P.; FERREIRA, A. C. P. de L.; LUDEMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro, RJ, Brasil: LTC Editora, 2007.
- CAPELLI, C.; SCHIEVANO, S. 4 - computational analyses and 3d printed models: A combined approach for patient-specific studies. In: KALASKAR, D. M. (Ed.). **3D Printing in Medicine**. [S. l.]: Woodhead Publishing, 2017. p. 73 – 90. ISBN 978-0-08-100717-4.
- CARNEIRO, T.; Medeiros Da Nóbrega, R. V.; Nepomuceno, T.; Bian, G.; De Albuquerque, V. H. C.; Filho, P. P. R. Performance analysis of google colaboratory as a tool for accelerating deep learning applications. **IEEE Access**, v. 6, p. 61677–61685, 2018. ISSN 2169-3536.
- CASAMITJANA, A.; CATÀ, M.; SÁNCHEZ, I.; COMBALIA, M.; VILAPLANA, V. Cascaded v-net using roi masks for brain tumor segmentation. In: CRIMI, A.; BAKAS, S.; KUIJF, H.; MENZE, B.; REYES, M. (Ed.). **Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries**. Cham: Springer International Publishing, 2018. p. 381–391. ISBN 978-3-319-75238-9.
- CASOLLA, G.; Cuomo, S.; Di Cola, V. S.; Piccialli, F. Exploring unsupervised learning techniques for the internet of things. **IEEE Transactions on Industrial Informatics**, p. 1–1, 2019. ISSN 1941-0050.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of control, signals and systems**, Springer, v. 2, n. 4, p. 303–314, 1989.

- FAROOQI, K. M.; UPPU, S. C.; NGUYEN, K.; SRIVASTAVA, S.; KO, H. H.; CHOUEITER, N.; WOLLSTEIN, A.; PARNESS, I. A.; NARULA, J.; SANZ, J.; NIELSEN, J. C. Application of virtual three-dimensional models for simultaneous visualization of intracardiac anatomic relationships in double outlet right ventricle. **Pediatric Cardiology**, v. 37, n. 1, p. 90–98, Jan 2016. ISSN 1432-1971.
- FERREIRA, D. S.; Ramalho, G. L. B.; Medeiros, F. N. S.; Bianchi, A. G. C.; Carneiro, C. M.; Ushizima, D. M. Saliency-driven system with deep learning for cell image classification. In: **2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)**. [S. l.: s. n.], 2019. p. 1284–1287. ISSN 1945-7928.
- FILHO, P. P. R.; BARROS, A. C. da S.; ALMEIDA, J. S.; RODRIGUES, J.; ALBUQUERQUE, V. H. C. de. A new effective and powerful medical image segmentation algorithm based on optimum path snakes. **Applied Soft Computing**, v. 76, p. 649 – 670, 2019. ISSN 1568-4946.
- GAO, Z.; CHUNG, J.; ABDELRAZEK, M.; LEUNG, S.; HAU, W. K.; XIAN, Z.; ZHANG, H.; LI, S. Privileged modality distillation for vessel border detection in intracoronary imaging. **IEEE transactions on medical imaging**, IEEE, 2019.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge, Massachusetts: MIT Press, 2016.
- HADEED, K.; ACAR, P.; DULAC, Y.; CUTTONE, F.; ALACOQUE, X.; KARSENTY, C. Cardiac 3d printing for better understanding of congenital heart disease. **Archives of cardiovascular diseases**, v. 111, n. 1, p. 1, 2018.
- HAUSDORFF, F. Bemerkung über den inhalt von punktmengen. **Mathematische Annalen**, v. 75, n. 3, p. 428–433, 1914.
- HE; ZHANG; REN; SUN. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: **2015 IEEE International Conference on Computer Vision (ICCV)**. [S. l.: s. n.], 2015. p. 1026–1034.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Identity mappings in deep residual networks. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 9908 LNCS, p. 630–645, 2016.
- HERMSEN, J. L.; BURKE, T. M.; SESLAR, S. P.; OWENS, D. S.; RIPLEY, B. A.; MOKADAM, N. A.; VERRIER, E. D. Scan, plan, print, practice, perform: Development and use of a patient-specific 3-dimensional printed model in adult cardiac surgery. **The Journal of Thoracic and Cardiovascular Surgery**, v. 153, n. 1, p. 132 – 140, 2017. ISSN 0022-5223.
- HORNIK, K. Approximation capabilities of multilayer feedforward networks. **Neural Networks**, v. 4, n. 2, p. 251 – 257, 1991. ISSN 0893-6080.
- JI; GU; SUN; YANG; QIAO. Congenital heart disease (chd) discrimination in fetal echocardiogram based on 3d feature fusion. In: **2016 IEEE International Conference on Image Processing (ICIP)**. [S. l.: s. n.], 2016. p. 3419–3423.
- JR, C. M. D.; SILVA, S. P. P. da; NÓBREGA, R. V. M. da; BARROS, A. C. da S.; FILHO, P. P. R.; ALBUQUERQUE, V. H. C. de. Deep learning iot system for online stroke detection in skull computed tomography images. **Computer Networks**, v. 152, p. 25 – 39, 2019. ISSN 1389-1286.

KHAMPARIA, A.; SAINI, G.; GUPTA, D.; KHANNA, A.; TIWARI, S.; ALBUQUERQUE, V. H. C. de. Seasonal crops disease prediction and classification using deep convolutional encoder network. **Circuits, Systems, and Signal Processing**, Jan 2019.

KHENED, M.; KOLLERATHU, V. A.; KRISHNAMURTHI, G. Fully convolutional multi-scale residual densenets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. **Medical Image Analysis**, v. 51, p. 21 – 45, 2019. ISSN 1361-8415.

KRIZHEVSKY, A.; HINTON, G. *et al.* Learning multiple layers of features from tiny images. Citeseer, 2009.

KWAG, E. M.; LEE, J. S.; KIM, S. H. The incidentally diagnosed adult congenital heart disease during routine medical health checkups in 27,897 koreans at a single center over seven years. **BMC cardiovascular disorders**, BioMed Central, v. 18, n. 1, p. 223, 2018.

LAPAN, M. **Deep Reinforcement Learning Hands-On: Apply modern RL methods, with deep Q-networks, value iteration, policy gradients, TRPO, AlphaGo Zero and more.** Birmingham, United Kingdom: Packt Publishing Ltd, 2018.

LAROBINA; MICHELE; MURINO; LOREDANA. Medical image file formats. **Journal of digital imaging**, Springer, v. 27, n. 2, p. 200–206, 2014.

LECUN, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278–2324, 1998.

LECUN, Y.; TOURESKY, D.; HINTON, G.; SEJNOWSKI, T. A theoretical framework for back-propagation. In: CMU, PITTSBURGH, PA: MORGAN KAUFMANN. **Proceedings of the 1988 connectionist models summer school**. San Matedo, Ca, 1988. v. 1, p. 21–28.

LI, J.; ZHANG, R.; SHI, L.; WANG, D. Automatic whole-heart segmentation in congenital heart disease using deeply-supervised 3d fcn. In: ZULUAGA, M. A.; BHATIA, K.; KAINZ, B.; MOGHARI, M. H.; PACE, D. F. (Ed.). **Reconstruction, Segmentation, and Analysis of Medical Images**. Cham: Springer International Publishing, 2017. p. 111–118. ISBN 978-3-319-52280-7.

LIU, T.; Tian, Y.; Zhao, S.; Huang, X.; Wang, Q. Automatic whole heart segmentation using a two-stage u-net framework and an adaptive threshold window. **IEEE Access**, v. 7, p. 83628–83636, 2019. ISSN 2169-3536.

LONG, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In: **2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S. l.: s. n.], 2015. p. 3431–3440.

MILLETARI, F.; NAVAB, N.; AHMADI, S. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: **Fourth International Conference on 3D Vision, 3DV 2016, Stanford, CA, USA, October 25-28, 2016**. [S. l.: s. n.], 2016. p. 565–571.

MUKHOPADHYAY, A. Total variation random forest: Fully automatic mri segmentation in congenital heart diseases. In: ZULUAGA, M. A.; BHATIA, K.; KAINZ, B.; MOGHARI, M. H.; PACE, D. F. (Ed.). **Reconstruction, Segmentation, and Analysis of Medical Images**. Cham: Springer International Publishing, 2017. p. 165–171.

- MULQUEEN, C. M.; STETZ, T. A.; BEAUBIEN, J.; O'CONNELL, B. Developing dynamic work roles using jaccard similarity indices of employee competency data. **Ergometrika**, v. 2, p. 26–37, 2001.
- NÓBREGA, R. V. M. d.; Peixoto, S. A.; Silva, S. P. P. d.; Filho, P. P. R. Lung nodule classification via deep transfer learning in ct lung images. In: **2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)**. [S. l.: s. n.], 2018. p. 244–249. ISSN 2372-9198.
- PACE, D. F.; DALCA, A. V.; GEVA, T.; POWELL, A. J.; MOGHARI, M. H.; GOLLAND, P. Interactive whole-heart segmentation in congenital heart disease. In: NAVAB, N.; HORNEGGER, J.; WELLS, W. M.; FRANGI, A. F. (Ed.). **Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015**. Cham: Springer International Publishing, 2015. p. 80–88. ISBN 978-3-319-24574-4.
- PICCIALLI, F.; Casolla, G.; Cuomo, S.; Giampaolo, F.; Schiano di Cola, V. Decision making in iot environment through unsupervised learning. **IEEE Intelligent Systems**, p. 1–1, 2019. ISSN 1941-1294.
- PICCIALLI, F.; CUOMO, S.; COLA, V. S. di; CASOLLA, G. A machine learning approach for iot cultural data. **Journal of Ambient Intelligence and Humanized Computing**, Springer, p. 1–12, 2019.
- RAIMONDI, F.; WARIN-FRESSE, K. Computed tomography imaging in children with congenital heart disease: Indications and radiation dose optimization. **Archives of Cardiovascular Diseases**, v. 109, n. 2, p. 150 – 157, 2016. ISSN 1875-2136.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: NAVAB, N.; HORNEGGER, J.; WELLS, W. M.; FRANGI, A. F. (Ed.). **Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015**. Cham: Springer International Publishing, 2015. p. 234–241. ISBN 978-3-319-24574-4.
- SCHURE, A. Y.; DINARDO, J. A. 18 - cardiac physiology and pharmacology. In: COTÉ, C. J.; LERMAN, J.; ANDERSON, B. J. (Ed.). **A Practice of Anesthesia for Infants and Children (Sixth Edition)**. Sixth edition. Philadelphia: Content Repository Only!, 2019. p. 424 – 457.e17. ISBN 978-0-323-42974-0.
- SHAHZAD, R.; GAO, S.; TAO, Q.; DZYUBACHYK, O.; GEEST, R. van der. Automated cardiovascular segmentation in patients with congenital heart disease from 3d cmr scans: Combining multi-atlases and level-sets. In: ZULUAGA, M. A.; BHATIA, K.; KAINZ, B.; MOGHARI, M. H.; PACE, D. F. (Ed.). **Reconstruction, Segmentation, and Analysis of Medical Images**. Cham: Springer International Publishing, 2017. p. 147–155. ISBN 978-3-319-52280-7.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **International Conference on Learning Representations**, 2014.
- SIMPSON, J. M.; BOSCH, A. van den. Educational series in congenital heart disease: Three-dimensional echocardiography in congenital heart disease. **Echo Research and Practice**, Bioscientifica Ltd, Bristol, UK, v. 6, n. 2, 2019.

- SPITZER, E.; REN, B.; ZIJLSTRA, F.; MIEGHEM, N. M. V.; GELEIJNSE, M. L. The role of automated 3d echocardiography for left ventricular ejection fraction assessment. **Cardiac failure review**, Radcliffe Cardiology, v. 3, n. 2, p. 97, 2017.
- STOUT, K. K.; DANIELS, C. J.; ABOULHOSN, J. A.; BOZKURT, B.; BROBERG, C. S.; COLMAN, J. M.; CRUMB, S. R.; DEARANI, J. A.; FULLER, S.; GURVITZ, M. *et al.* 2018 aha/acc guideline for the management of adults with congenital heart disease: a report of the american college of cardiology/american heart association task force on clinical practice guidelines. **Circulation**, Am Heart Assoc, v. 139, n. 14, p. e698–e800, 2019.
- SUN, Z.; LAU, I.; WONG, Y. H.; YEONG, C. H. Personalized three-dimensional printed models in congenital heart disease. **Journal of clinical medicine**, Multidisciplinary Digital Publishing Institute, v. 8, n. 4, p. 522, 2019.
- SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCKE, V.; RABINOVICH, A. Going deeper with convolutions. In: . [S. l.: s. n.], 2015. v. 07-12-June-2015, p. 1–9.
- TAHA; HANBURY. An efficient algorithm for calculating the exact hausdorff distance. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 37, n. 11, p. 2153–2163, Nov 2015. ISSN 1939-3539.
- TIWARI, P.; QIAN, J.; LI, Q.; WANG, B.; GUPTA, D.; KHANNA, A.; RODRIGUES, J. J.; ALBUQUERQUE, V. H. C. de. Detection of subtype blood cells using deep learning. **Cognitive Systems Research**, Elsevier, v. 52, p. 1036–1044, 2018.
- WANG, C.; WANG, Q.; SMEDBY, Ö. Automatic heart and vessel segmentation using random forests and a local phase guided level set method. In: ZULUAGA, M. A.; BHATIA, K.; KAINZ, B.; MOGHARI, M. H.; PACE, D. F. (Ed.). **Reconstruction, Segmentation, and Analysis of Medical Images**. Cham: Springer International Publishing, 2017. p. 159–164. ISBN 978-3-319-52280-7.
- WOLTERINK, J. M.; LEINER, T.; VIERGEVER, M. A.; ISGUM, I. Dilated convolutional neural networks for cardiovascular MR segmentation in congenital heart disease. **CoRR**, abs/1704.03669, 2017.
- XIA, Q.; YAO, Y.; HU, Z.; HAO, A. Automatic 3d atrial segmentation from ge-mris using volumetric fully convolutional networks. In: POP, M.; SERMESANT, M.; ZHAO, J.; LI, S.; MCLEOD, K.; YOUNG, A.; RHODE, K.; MANSI, T. (Ed.). **Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges**. Cham: Springer International Publishing, 2019. p. 211–220. ISBN 978-3-030-12029-0.
- XIONG, Z.; Fedorov, V. V.; Fu, X.; Cheng, E.; Macleod, R.; Zhao, J. Fully automatic left atrium segmentation from late gadolinium enhanced magnetic resonance imaging using a dual fully convolutional neural network. **IEEE Transactions on Medical Imaging**, v. 38, n. 2, p. 515–524, Feb 2019. ISSN 1558-254X.
- YU, L.; YANG, X.; QIN, J.; HENG, P.-A. 3d fractalnet: Dense volumetric segmentation for cardiovascular mri volumes. In: ZULUAGA, M. A.; BHATIA, K.; KAINZ, B.; MOGHARI, M. H.; PACE, D. F. (Ed.). **Reconstruction, Segmentation, and Analysis of Medical Images**. Cham: Springer International Publishing, 2017. p. 103–110. ISBN 978-3-319-52280-7.

- ZEILER, M.; FERGUS, R. Visualizing and understanding convolutional networks. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 8689 LNCS, n. PART 1, p. 818–833, 2014.
- ZHANG, L.; GOOYA, A.; PEREANEZ, M.; DONG, B.; PIECHNIK, S.; NEUBAUER, S.; PETERSEN, S.; FRANGI, A. F. Automatic assessment of full left ventricular coverage in cardiac cine magnetic resonance imaging with fisher discriminative 3d cnn. **IEEE Transactions on Biomedical Engineering**, IEEE, 2018.
- ZHOU, T.; RUAN, S.; CANU, S. A review: Deep learning for medical image segmentation using multi-modality fusion. **Array**, v. 3-4, p. 100004, 2019. ISSN 2590-0056.
- ZHUANG, X.; LI, L.; PAYÉ, C.; ŠTERN, D.; URSCHLER, M.; HEINRICH, M. P.; OSTER, J.; WANG, C.; SMEDBY Örjan; BIAN, C.; YANG, X.; HENG, P.-A.; MORTAZI, A.; BAGCI, U.; YANG, G.; SUN, C.; GALISOT, G.; RAMEL, J.-Y.; BROUARD, T.; TONG, Q.; SI, W.; LIAO, X.; ZENG, G.; SHI, Z.; ZHENG, G.; WANG, C.; MACGILLIVRAY, T.; NEWBY, D.; RHODE, K.; OURSELIN, S.; MOHIADDIN, R.; KEEGAN, J.; FIRMIN, D.; YANG, G. Evaluation of algorithms for multi-modality whole heart segmentation: An open-access grand challenge. **Medical Image Analysis**, v. 58, p. 101537, 2019. ISSN 1361-8415.
- ZHUANG, X.; RHODE, K. S.; RAZAVI, R. S.; HAWKES, D. J.; OURSELIN, S. A registration-based propagation framework for automatic whole heart segmentation of cardiac mri. **IEEE transactions on medical imaging**, IEEE, v. 29, n. 9, p. 1612–1625, 2010.
- ZULUAGA, M. A.; BIFFI, B.; TAYLOR, A. M.; SCHIEVANO, S.; VERCAUTEREN, T.; OURSELIN, S. Strengths and pitfalls of whole-heart atlas-based segmentation in congenital heart disease patients. In: ZULUAGA, M. A.; BHATIA, K.; KAINZ, B.; MOGHARI, M. H.; PACE, D. F. (Ed.). **Reconstruction, Segmentation, and Analysis of Medical Images**. Cham: Springer International Publishing, 2017. p. 139–146.