



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

**Segmentação do ventrículo esquerdo em imagens de  
ressonância magnética usando rede neural  
convolucional e modelo Chan-Vese**

Marcelo Áxel Chiapinotto de Nazaré

Monografia apresentada como requisito parcial  
para conclusão do Curso de Engenharia da Computação

Orientadora  
Prof.a Dr.a Roberta Barbosa Oliveira

Brasília  
2021



# Dedicatória

*Dedico este trabalho a minha família, em especial aos meus pais, que sempre me deram apoio incondicional e estiveram ao meu lado nos momentos mais difíceis e ao meu irmão que me acompanhou durante toda a trajetória, estudando juntos e fortalecendo um ao outro.*

# Agradecimentos

Agradeço a minha família por sempre me apoiar, em especial aos meus pais Edmar e Marizete e ao meu irmão Geovanne, por sempre estarem comigo em todos os momentos.

Agradeço a minha professora orientadora Roberta Barbosa de Oliveira, por me guiar durante toda a realização desde trabalho, desde a escolha do tema até a resolução do trabalho, desde o primeiro momento se mostrou disposta a ajudar.

Agradeço aos meus amigos de curso que conheci ao longo dessa jornada e que se tornaram membros importantes na minha vida. Os inúmeros desafios da vida acadêmica não seriam ultrapassados sem os momentos de estudo e descontração que tivemos juntos.

# Resumo

Uma das principais causas de morte em todo o mundo são as doenças cardiovasculares. Um dos exames utilizados para o diagnóstico dessas doenças é a ressonância magnética, que é um exame de imageamento cardíaco não invasivo. A partir desse exame é possível visualizar a estrutura do coração. Um passo fundamental do exame é a segmentação do ventrículo esquerdo para posteriormente identificar alguma anomalia no funcionamento do coração. Esse procedimento, geralmente, é realizado manualmente por um especialista, o que exige tempo e esforço e pode ocasionar falhas no diagnóstico. O objetivo deste trabalho é comparar métodos de *deep learning*, com um método tradicional de segmentação e um método híbrido para identificar o ventrículo esquerdo em imagens de ressonância magnética cardíaca. Com essa finalidade, os modelos de CNNs U-net tradicional e *Attention* U-net foram aplicados, bem como o modelo de contorno ativo Chan-Vese. Além disso, foi avaliado um modelo híbrido considerando a U-net tradicional para melhorar o resultado do modelo Chan-Vese. Foram considerados conjuntos de treinamento, validação e teste, utilizando a base de dados Sunnybrook. Os métodos foram comparados sem a aplicação de pré-processamento de imagens, e aplicando técnicas de realce de imagens e *data augmentation* para avaliar o efeito de cada uma das técnicas sobre os métodos. O método de segmentação que obteve o melhor desempenho foi a U-net tradicional sem aplicação de pré-processamento, com 0,890 de coeficiente Dice e 0,997 de acurácia. Os resultados do método híbrido também se mostraram promissores em relação a utilização individual do modelo Chan-Vese.

**Palavras-chave:** Segmentação de imagens, imagens médicas, aprendizado profundo, coração, ressonância magnética

# Abstract

One of the leading causes of death worldwide is cardiovascular disease. Every year, more people die from these diseases than from any other cause. One of the tests to diagnose these diseases is magnetic resonance imaging, which is a non-invasive cardiac imaging test. From this exam, it is possible to visualize the structure of the heart. A fundamental step in the exam is the segmentation of the left ventricle to later identify any abnormality in the functioning of the heart. This procedure is usually performed manually by a specialist, which requires time and effort, which can lead to diagnostic failures. The goal of this work is to compare deep learning methods, with a traditional segmentation method and a hybrid method to identify the left ventricle in cardiac magnetic resonance images. For this purpose, the traditional U-net and Attention U-net CNN models were applied, as well as the Chan-Vese active contour model. Furthermore, a hybrid model considering the traditional U-net was evaluated to improve the result of the Chan-Vese model. Training, validation, and testing sets were considered, using the Sunnybrook database. The methods were compared without applying image pre-processing and applying image enhancement and data augmentation techniques to assess the effect of each technique on the methods. The segmentation method that obtained the best performance was the traditional U-net without pre-processing application, with 0.890 of Dice coefficient and 0.997 of accuracy. The results of the hybrid method were also promising in relation to the individual use of the Chan-Vese model.

**Keywords:** image segmentation, medical images, deep learning, heart, magnetic resonance

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Objetivos . . . . .	3
1.2	Estrutura do documento . . . . .	3
<b>2</b>	<b>Fundamentação Teórica</b>	<b>4</b>
2.1	Imagens de Ressonância Magnética . . . . .	4
2.2	Pré-processamento de Imagens . . . . .	7
2.2.1	Histograma de Imagens . . . . .	7
2.2.2	Data Augmentation . . . . .	8
2.3	Segmentação de imagens baseada no Modelo Chan-Vese . . . . .	8
2.4	Rede Neural Convolutiva . . . . .	10
2.4.1	Convolução . . . . .	12
2.4.2	Pooling . . . . .	13
2.4.3	Camada Totalmente Conectada . . . . .	13
2.5	Arquiteturas de CNN . . . . .	13
2.5.1	U-net . . . . .	14
2.5.2	<i>Attention</i> U-net . . . . .	14
2.6	Métricas de Avaliação de Desempenho . . . . .	16
<b>3</b>	<b>Revisão Bibliográfica</b>	<b>18</b>
<b>4</b>	<b>Metodologia</b>	<b>21</b>
4.1	Base de imagens . . . . .	21
4.1.1	Separação das imagens . . . . .	22
4.1.2	Leitura do Contorno . . . . .	23
4.1.3	Normalização dos <i>pixels</i> . . . . .	23
4.2	Pré-Processamento das imagens . . . . .	23
4.2.1	Equalização do histograma . . . . .	24
4.2.2	Data Augmentation . . . . .	24

4.3	Segmentação . . . . .	25
4.3.1	U-net tradicional . . . . .	25
4.3.2	<i>Attention</i> U-net . . . . .	26
4.3.3	Modelo tradicional Chan-Vese . . . . .	26
4.3.4	Modelo Híbrido . . . . .	26
4.4	Extrair Contorno da imagem binária . . . . .	26
<b>5</b>	<b>Resultados Experimentais</b>	<b>29</b>
5.1	Implementação . . . . .	29
5.2	Segmentação tradicional . . . . .	29
5.3	Redes Neurais Convolucionais . . . . .	30
5.3.1	Segmentação sem pré-processamento . . . . .	32
5.3.2	Aumento da quantidade de imagens . . . . .	32
5.3.3	Dados com realce da imagem . . . . .	34
5.4	Modelo Híbrido . . . . .	35
<b>6</b>	<b>Conclusão</b>	<b>38</b>
	<b>Referências</b>	<b>40</b>



# Lista de Figuras

1.1	Fatias do coração. . . . .	2
2.1	Orientações dos principais planos do corpo em relação ao paciente e suas imagens. . . . .	5
2.2	Orientações dos planos cardíacos em relação ao coração e suas respectivas imagens. . . . .	6
2.3	Histograma de uma imagem de ressonância magnética: (a) Imagem original da base de imagens Sunnybrook (b) Histograma da imagem (a). . . . .	8
2.4	Possíveis posições da curva. . . . .	9
2.5	Neurônio biológico. . . . .	10
2.6	Modelo de um neurônio artificial. . . . .	10
2.7	Exemplo de funcionamento da camada de convolução. . . . .	12
2.8	Aplicação de <i>max pooling</i> . . . . .	13
2.9	Arquitetura da rede U-net. . . . .	14
2.10	Arquitetura da rede <i>Attention</i> U-net. A imagem de entrada é progressivamente filtrada e reduzida por um fator de 2 em cada fase da parte de codificação da rede. <i>Attention gates</i> (AGs) filtra os recursos propagados pelas conexões. . . . .	15
2.11	Esquemático do <i>Attention gate</i> (AG). . . . .	16
4.1	Diagrama da metodologia empregada. O modelo híbrido é descrito na Figura 4.4. . . . .	22
4.2	Exemplo de equalização do histograma em uma imagem original da base. . . . .	24
4.3	Exemplos de <i>data augmentation</i> : (a) Imagem original retirada do conjunto de dados Sunnybrook [1], (b) Imagem rotacionada 180 graus, (c) Imagem invertida verticalmente, e (d) Imagem invertida horizontalmente. . . . .	25
4.4	Diagrama da rede híbrida proposta. . . . .	27

4.5	Exemplo do teorema de Green [2] em uma imagem binária gerada pela U-net tradicional no conjunto de imagens (a) a linha em vermelho são os pontos do contorno do maior objeto, (b) é o resultado após a aplicação do teorema de Green. . . . .	28
5.1	Exemplo de segmentação Chan-Vese com objeto de interesse dentro do contorno inicial. Sendo o contorno inicial do modelo em amarelo, em verde o <i>ground-truth</i> e em vermelho o resultado da segmentação. . . . .	30
5.2	Exemplo de segmentação Chan-Vese com objeto de interesse fora do contorno inicial. Sendo o contorno inicial do modelo em amarelo, em verde o <i>ground-truth</i> e em vermelho o resultado da segmentação. . . . .	31
5.3	Coeficiente Dice em relação as épocas da U-net tradicional sem pré-processamento. . . . .	32
5.4	Coeficiente Dice da <i>Attention</i> U-net sem pré-processamento. . . . .	33
5.5	Coeficiente Dice da U-net tradicional com <i>data augmentation</i> . . . . .	33
5.6	Coeficiente Dice da <i>Attention</i> U-net com <i>data augmentation</i> . . . . .	34
5.7	Coeficiente Dice da U-net tradicional com realce da imagem. . . . .	35
5.8	Coeficiente Dice da <i>Attention</i> U-net com realce da imagem. . . . .	35
5.9	Segmentação do método híbrido (a) Imagem original, (b) máscara <i>ground-truth</i> , (c) resultado da segmentação, (d) contorno sobreposto à imagem original (verde (b) e vermelho (c)). . . . .	37

# Lista de Tabelas

5.1 Resultado dos experimentos. . . . .	36
---	----

# Lista de Abreviaturas e Siglas

**ACR** *American College of Radiology.*

**CLAHE** *Contrast Limited Adaptive Histogram Equalization.*

**CMRI** *Ressonância Magnética Cardíaca.*

**CNN** *Rede Neural Convolutacional.*

**DICOM** *Digital Imaging and Communications in Medicine.*

**FFR** *fluxo fracionado de reserva do miocárdio.*

**FPN** *Feature Pyramid Network.*

**LBF** *Labeled-Based Forecasting.*

**MLP** *Multilayer Perceptron.*

**NEMA** *National Electrical Manufacturers Association.*

**ReLU** *Rectified Linear Units.*

**RNA** *Rede Neural Artificial.*

**ROI** *Região de interesse.*

**SVM** *Support Vector Machines.*

# Capítulo 1

## Introdução

Segundo a Organização Mundial da Saúde (OMS) [3], a principal causa de morte em todo o mundo entre os anos de 2000 e 2019 foram doenças cardiovasculares, chegando a quase 9 milhões de mortes em 2019. Esse número corresponde a 16% de todas as mortes no mundo.

Para o diagnóstico dessas doenças existem vários exames não invasivos de imageamento cardíaco, tais como ecocardiograma, tomografia computadorizada por emissão de fóton único, tomografia computadorizada e ressonância magnética cardíaca. Um dos exames mais utilizados é o exame de ressonância magnética, no qual é possível visualizar a estrutura do coração sem precisar realizar procedimentos cirúrgicos e não utiliza radiação ionizante [4]. Cada exame de ressonância magnética divide o coração em diversas fatias (slices), representadas na Figura 1.1, o que gera um volume grande de imagens para o médico analisar.

O ventrículo esquerdo é responsável por bombear sangue para todo o corpo. Além disso, o ventrículo esquerdo é uma das partes mais analisadas nesse tipo de exame, pois sua dimensão, volume e outros dados servem de indício de um mal funcionamento cardíaco do paciente. A identificação do ventrículo esquerdo nessas fatias depende da avaliação de um médico, o que se torna uma análise exaustiva podendo levar a falhas no diagnóstico devido a grande quantidade de imagens por exame, fato pelo qual existe uma necessidade de automatizar esse processo.

Essa identificação de forma automatizada geralmente é feita pelo processo conhecido como segmentação que consiste em dividir a imagem em um conjunto de *pixels* ou objetos de acordo com suas propriedades em comum, como cor e textura. Existem vários métodos de segmentação de imagens como, por exemplo, detecção de borda, limiarização e crescimento de região [6].

Um método utilizado nas pesquisas realizadas utilizando crescimento de região foi o algoritmo tradicional de contorno ativo Chan-Vese [7] que se mostra eficiente na segmen-

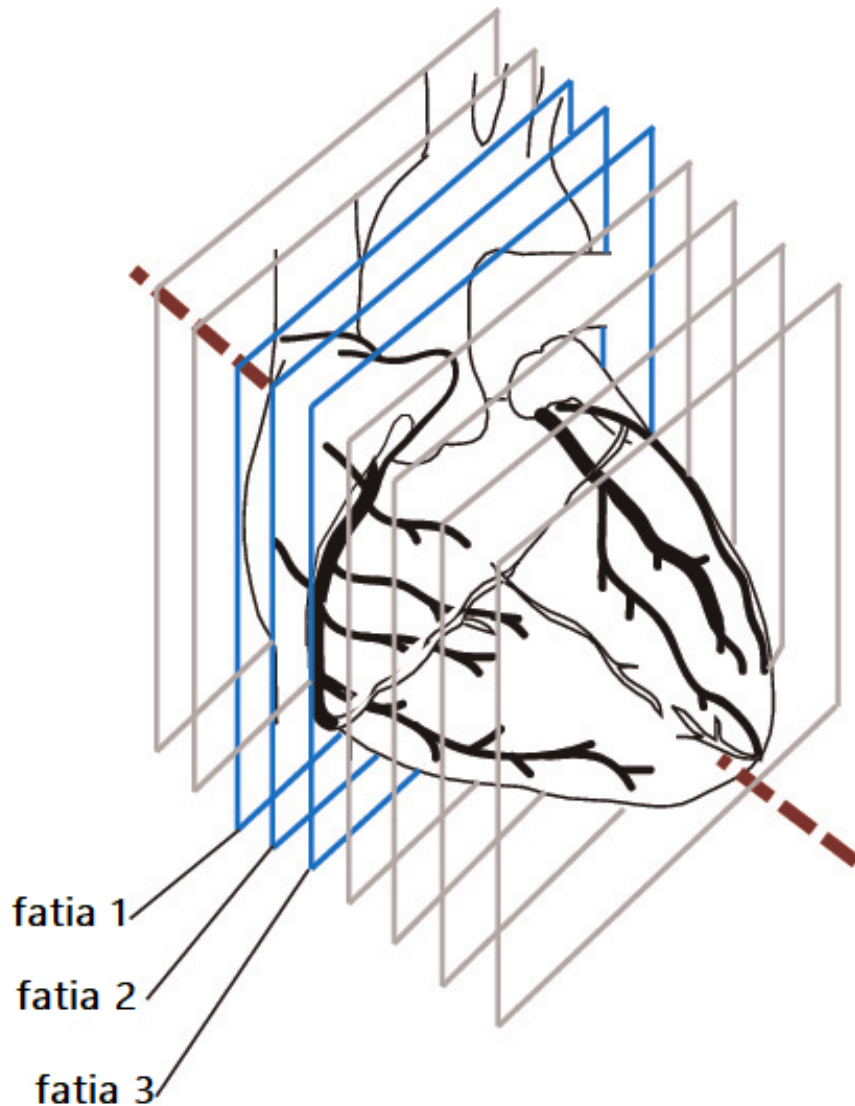


Figura 1.1: Fatias do coração (Fonte: [5]).

tação do ventrículo esquerdo [8]. Uma abordagem com aprendizado de máquina que vem se destacando na segmentação de imagens médicas [9] é a rede neural convolucional U-net, devido a sua capacidade de aprendizado com uma quantidade menor de imagens rotuladas [10]. A rede neural *Attention* U-net, inicialmente utilizada em imagens do pâncreas [11], mostra-se interessante para ser analisada para o problema de segmentação do ventrículo esquerdo, apesar de até o momento desta pesquisa não ter sido encontrada nenhuma publicação para o problema apresentado. Essa rede aprende a focar em estruturas de interesse para segmentação e atribui um menor peso as estruturas que não são de interesse, aprendendo ao máximo suas características.

## 1.1 Objetivos

O objetivo geral deste trabalho é realizar experimentos relacionados a segmentação do ventrículo esquerdo em exames de ressonância magnética cardíaca para auxiliar no diagnóstico de doenças como isquemia cardíaca [12] e hipertrofia ventricular. Para alcançar o objetivo geral, foi definido os seguintes objetivos específicos:

- Analisar modelos tradicional e de *deep learning* para a segmentação do ventrículo esquerdo;
- Verificar se a utilização das técnicas de *data augmentation* e equalização de histograma, promovem um aumento no desempenho das redes neurais convolucionais;
- Verificar se a utilização de um modelo híbrido, promove um aumento no desempenho do modelo tradicional Chan-Vese.

## 1.2 Estrutura do documento

O presente trabalho está dividido nos seguintes capítulos: O Capítulo 2 apresenta toda a fundamentação teórica necessária para um melhor entendimento dos conceitos utilizados no desenvolvimento desse trabalho. O Capítulo 3 traz os artigos do estado da arte relevantes para o desenvolvimento dos experimentos realizados. O Capítulo 4 descreve a metodologia desenvolvida para que seja possível reproduzir os experimentos realizados. O Capítulo 5 mostra os resultados obtidos com o método híbrido e a comparação com o uso de diferentes métodos. O Capítulo 6 traz conclusões acerca dos métodos analisados e considerações de trabalhos futuros.

# Capítulo 2

## Fundamentação Teórica

Este capítulo apresenta os conceitos teóricos para um melhor entendimento dos métodos empregados durante a realização deste trabalho. São apresentados conceitos fundamentais sobre a aquisição das imagens e o formato DICOM, informações sobre histograma e *data augmentation*, além de conceitos sobre o método de segmentação tradicional Chan-Vese e, por fim, conceitos sobre aprendizado de máquina, arquiteturas de rede neural e métricas utilizadas para avaliar o desempenho da rede.

### 2.1 Imagens de Ressonância Magnética

A Ressonância magnética é um método de obtenção de imagens computadorizadas que, diferente de outras modalidades como tomografia e raio X, não utiliza radiação ionizante que é uma forma de radiação que carrega energia suficiente para remover elétrons que estão ligados a átomos criando os íons. Os efeitos da radiação dependem da quantidade e qualidade da radiação incidente, pacientes expostos a altas doses podem apresentar efeitos agudos como náuseas, perda de cabelo, fraqueza, etc chegando a casos com efeitos crônicos como alteração no DNA e câncer [13].

Os dois principais sistemas de coordenadas usados para a realização do Ressonância Magnética Cardíaca (CMRI) são os planos do corpo e os planos cardíacos [14]. Os planos do corpo são orientados ortogonalmente ao eixo longo do corpo e são utilizados para obter as imagens que fornecem uma descrição qualitativa da morfologia cardíaca [14]. Os principais planos do corpo são: coronal, sagital e axial, e suas orientações são apresentadas na Figura 2.1.

O plano coronal pode ser usado para avaliar as veias pulmonares que carregam sangue rico em oxigênio dos pulmões até o átrio esquerdo, também pode-se avaliar a câmara inferior e superior à esquerda do coração. São elas o átrio esquerdo e o ventrículo esquerdo. O átrio esquerdo é responsável por coletar o sangue oxigenado que volta ao coração pelas



veias pulmonares e passa esse volume de sangue para o ventrículo esquerdo. O ventrículo esquerdo por sua vez recebe o sangue oxigenado do átrio esquerdo, através da válvula mitral, e em seguida bombeia esse volume de sangue em alta pressão para todo o corpo [15]. O plano sagital mostra os grandes vasos surgindo em continuidade com os ventrículos. O plano axial captura, simultaneamente, as quatro câmaras do coração (dois átrios e dois ventrículos) e o pericárdio<sup>1</sup>. Entretanto, a obliquidade (aproximadamente 45°) destes planos com as paredes do coração impede uma caracterização anatômica e funcional precisa. Para obter essas informações é preciso o auxílio dos planos cardíacos.

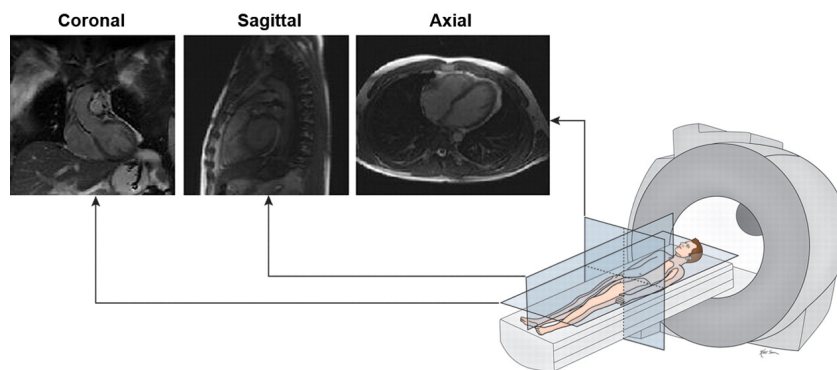


Figura 2.1: Orientações dos principais planos do corpo em relação ao paciente e suas imagens (Fonte: [14]).

Os planos cardíacos são estabelecidos utilizando o topograma, uma imagem digital utilizada pelo técnico para selecionar a área de aquisição de imagens do exame, ao longo da linha que se estende do ápice cardíaco ao centro da válvula mitral (eixo longo do coração). O eixo longo do coração é obtido através da imagem do plano axial do corpo descrito anteriormente e incluem eixo curto (*short axis*), eixo longo horizontal (*horizontal long axis*) (vista de quatro camadas) e eixo longo vertical (*vertical long axis*) (vista de duas camadas), conforme a Figura 2.2.

O eixo curto é o plano perpendicular ao eixo longo do coração e ortogonal ao ventrículo esquerdo, considerado o eixo que alinha a base do coração e o ápice. Esse eixo fornece uma visão em corte transversal dos ventrículos esquerdo e direito e geralmente exhibe o esqueleto cardíaco e os anéis valvulares. O eixo longo horizontal é o plano horizontal perpendicular ao eixo curto que mostra as quatro câmaras cardíacas (ventrículo esquerdo, átrio esquerdo, ventrículo direito, átrio direito) permitindo fazer a avaliação do tamanho e da posição valvular delas. Enquanto o eixo longo horizontal é perpendicular ao eixo curto o eixo longo vertical é o plano vertical ortogonal ao eixo curto, este eixo pode ser adquirido em duas direções, esquerda ou direita, mostrando as duas camadas (átrio e ventrículo)

<sup>1</sup>Membrana serosa que envolve externamente o coração.

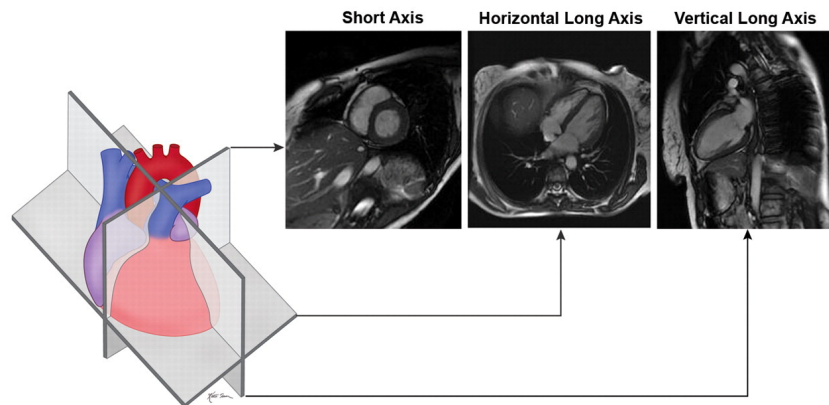


Figura 2.2: Orientações dos planos cardíacos em relação ao coração e suas respectivas imagens (Fonte: [14]).

da direção escolhida. Com o eixo longo vertical pode-se identificar a obliquidade no eixo cardíaco no plano coronal.

O método utilizado denominado ressonância magnética cinemática consiste na aquisição de múltiplas imagens ao longo do intervalo de um batimento cardíaco para cada posição de corte, através desse método é possível visualizar o funcionamento do coração em fatias e pelas fases do ciclo cardíaco.

Para a capturar uma imagem do coração sem que esta seja afetada pelo movimento do coração a aquisição das imagens deve ser sincronizada com o sinal de eletrocardiograma através de eletrodos colocados no tórax do paciente. Além disso, o movimento da respiração pode afetar a qualidade das imagens, por isso a captura dessas imagens deve ser feita em pausas respiratórias feitas pelo paciente [16]. Ao final desse processo é gerado um conjunto de imagens em formato DICOM contendo diversas informações do paciente.

DICOM, que significa *Digital Imaging and Communications in Medicine* (Comunicação de Imagens Digitais na Medicina), é um conjunto de normas para padronizar o arquivamento e comunicação de imagens e dados obtidos em exames. O conjunto de normas foi criado em 1983, pelo comitê formado pelo *American College of Radiology* (ACR) em parceria com National Electrical Manufacturers Association (NEMA) [17].

Desde então, o conjunto de normas passou por três versões. A versão 1.0 publicada em 1985, foi a primeira forma de arquivar dados em mídia com comunicação de forma não proprietária. Em 1988 a versão 1.0 foi sucedida por uma revisão e tornou-se publica a ACR-NEMA versão 2.0. Essa versão incluiu as principais definições sobre terminologia, estrutura de dados e codificação. Sua última versão foi apresentada como DICOM versão 3.0 em 1993, cujo a diferença essencial foi a especificação de um protocolo de rede baseado no modelo ISO/OSI e o uso de TCP/IP para permitir a independência de soluções

específicas do fornecedor do equipamento.

Os arquivos em formato DICOM possui a imagem codificada e um cabeçalho com informações sobre o exame, que incluem informações do paciente, número de fatias que o coração foi dividido, dimensões da imagem capturada, qual a técnica utilizada na aquisição da imagem entre outros. Dependendo da técnica de aquisição escolhida (Ressonância magnética, tomografia computadorizada ou raio X), algumas informações são adicionadas ao cabeçalho, para a ressonância magnética o tempo de repetição e tempo de eco são adicionados.

## 2.2 Pré-processamento de Imagens

O pré-processamento de imagens é o procedimento inicial geralmente realizado nas imagens originais com o intuito de melhorar a qualidade das imagens para as etapas posteriores do processo.

Um procedimento de pré-processamento que pode ser realizado é o realce de contraste da imagem, feito através da análise da distribuição de ocorrências de níveis de cinza através do histograma [18]. Outro procedimento é o aumento da quantidade de imagens feito através da adição de cópias ligeiramente modificadas, sendo conhecida como *data augmentation* [19].

### 2.2.1 Histograma de Imagens

Cada *pixel* de uma imagem digital é representado por um número que, em imagens em escala de cinza, representa a intensidade de cinza e varia de 0 a 255 em imagens de 24 bits. O histograma é a representação gráfica da quantidade de cada *pixel* em cada nível de intensidade de cinza [20]. A Figura 2.3 apresenta uma imagem com seu respectivo histograma.

Através do histograma é possível obter informações sobre o contraste, qualidade e brilho médio da imagem. O percentual de *pixels* para cada nível de intensidade de cinza de cada *pixel* é calculado por:

$$p_r(r_k) = \frac{n_k}{n}, \quad (2.1)$$

sendo:

- $0 \leq r_k \leq 1$ ;
- $k = 0, 1, \dots, L - 1$ , onde  $L$  é o número de níveis de cinza;
- $n$  = número total de *pixels* da imagem;

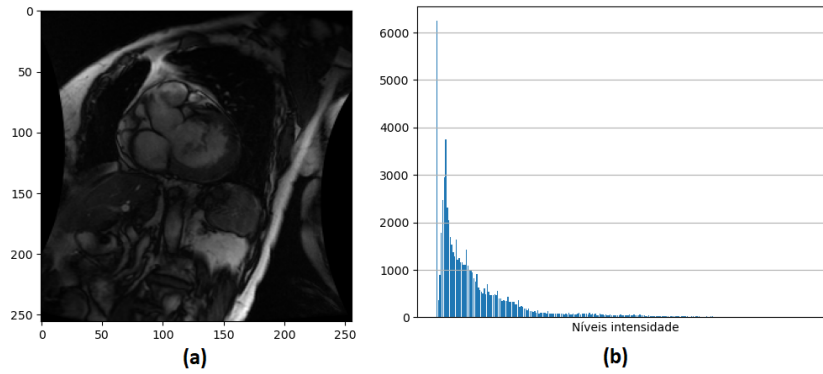


Figura 2.3: Histograma de uma imagem de ressonância magnética: (a) Imagem original da base de imagens Sunnybrook (b) Histograma da imagem (a).

- $p_r(r_k)$  = percentual de *pixels* no  $K$  –ésimo nível de cinza;
- $n_k$  = número de *pixels* cujo nível de cinza corresponde a  $k$ .

### 2.2.2 Data Augmentation

*Data Augmentation* é uma técnica de regularização utilizada para aumentar o número de imagens em um banco de dados adicionando cópias dos dados levemente modificadas. Estas modificações são feitas aplicando técnicas como variação no brilho, saturação, zoom, rotação, reflexão horizontal ou vertical [19].

## 2.3 Segmentação de imagens baseada no Modelo Chan-Vese

Para extrair informações de uma imagem digital é necessário dividi-la em partes ou objetos que a compõem. Este processo é chamado de segmentação de imagem e pode ser formalmente definido como o processo de identificação das regiões homogêneas de uma determinada imagem e das extremidades ou limites dessa região.

Os métodos de segmentação podem ser classificadas em duas categorias, baseado em borda e baseado em região [6]. Os métodos baseados em borda procuram detectar as bordas do objeto de interesse na imagem, enquanto os métodos baseados em região procuram identificar todos os *pixels*. Além disso, os métodos baseados em região visam agrupar em um conjunto *pixels* adjacentes que atendem a um critério de heterogeneidade.

O modelo proposto por Chan e Vese [7] é um modelo de contorno ativo sem bordas para segmentação de imagem baseado em região. O modelo Chan-Vese se baseia na minimização da função de energia que leva em conta a intensidade da imagem.

Definindo a curva em evolução  $C$  como o limite do subconjunto  $\omega$  no domínio da imagem  $\Omega$ , ou seja,  $\omega \subset \Omega$  e  $C = \partial\omega$ , em que  $\bar{\omega}$  representa a região dentro da curva  $C$  e  $\omega$  representa a região de fora da curva  $C$ . Este modelo considera uma energia funcional para a minimização da energia, que faz com que a curva  $C$  se deforme no sentido da fronteira  $C_0$  de uma imagem  $u_0$  com regiões de intensidade  $u_0^i$  e  $u_0^o$  constantes. Considerando a energia funcional formada por dois termos temos:

$$F_1(C) + F_2(C) = \int_{dentro(C)} |u_0 - c_1|^2 dx dy + \int_{fora(C)} |u_0 - c_2|^2 dx dy, \quad (2.2)$$

onde  $c_1$  é a média da imagem  $u_0$  dentro da curva  $C$  e  $c_2$  é a média da imagem  $u_0$  fora da curva  $C$ . A fronteira do objeto  $\eta$  é o minimizador da energia funcional dado por:

$$\inf_C \{F_1(C) + F_2(C)\} \approx 0 \approx F_1(\eta) + F_2(\eta). \quad (2.3)$$

Se a curva  $C$  está fora do objeto, então  $F_1(C) > 0$  e  $F_2(C) \approx 0$ . Se a curva  $C$  está dentro do objeto, então  $F_1(C) \approx 0$  e  $F_2(C) > 0$ . A energia funcional será minimizada se  $C = \eta$ , logo a curva  $C$  estará na fronteira do objeto. Todas as possibilidades são apresentadas na Figura 2.4.

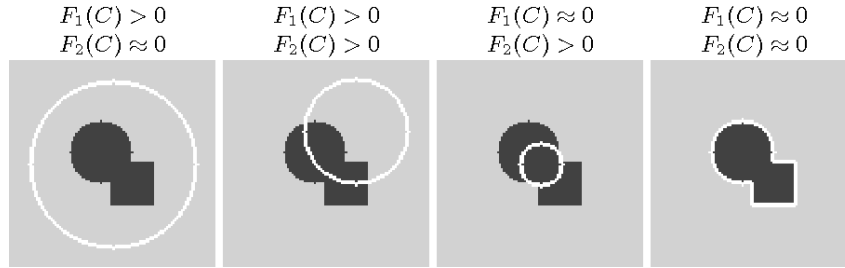


Figura 2.4: Possíveis posições da curva (Fonte: [7]).

Para minimizar o termo de energia  $F$  são utilizados termos de regularização, como o comprimento da curva  $Comprimento(C)$  e a área dentro da curva  $Area(dentro(C))$ . Então a energia funcional considerando os termos de regularização é dada por:

$$F(c_1, c_2, C) = \mu \cdot Comprimento(C) + v \cdot Area(dentro(C)) + \lambda_1 \int_{dentro(C)} |u_0 - c_1|^2 dx dy + \lambda_2 \int_{fora(C)} |u_0 - c_2|^2 dx dy, \quad (2.4)$$

sendo  $\mu \geq 0$ ,  $v \geq 0$ ,  $\lambda_1, \lambda_2 > 0$  parâmetros fixos que auxiliam no resultado.

## 2.4 Rede Neural Convolutacional

O sistema nervoso é formado por uma rede interligada de neurônios que tem papel essencial no funcionamento e comportamento do corpo humano. Os neurônios são constituídos de três partes principais conforme Figura 2.5: (i) dendritos, (ii) corpo celular e (iii) axônio. Dendritos que recebem os impulsos nervosos, corpo celular onde encontram-se o núcleo e é responsável pelo processamento das informações dos dendritos e o axônio que transmite a informação ao neurônio seguinte ou a outra estrutura.

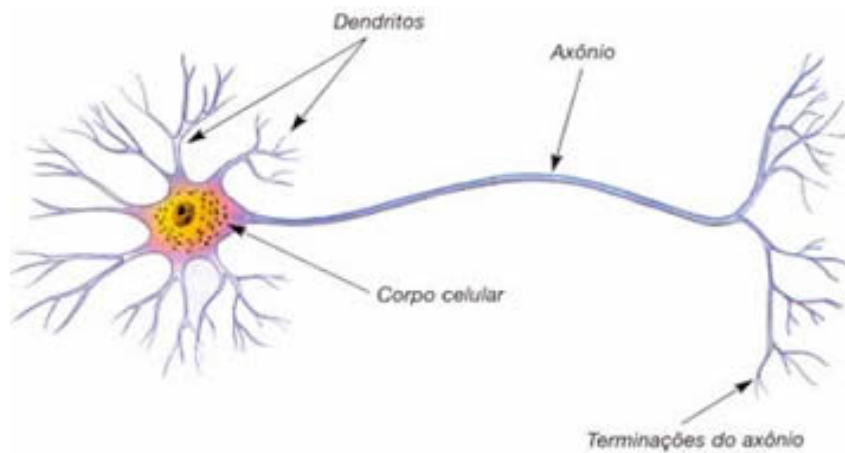


Figura 2.5: Neurônio biológico (Fonte: [21]).

Inspirada neste sistema, a Rede Neural Artificial é uma estrutura interligada por elementos de processamento simples chamados de neurônios artificiais apresentado na Figura 2.6. Os neurônios recebe um vetor de entradas  $x$  e aplica uma função  $f$  tal que:

$$f(x_1, x_2, \dots, x_i) = \sum_{n=1}^i W_n x_n. \quad (2.5)$$

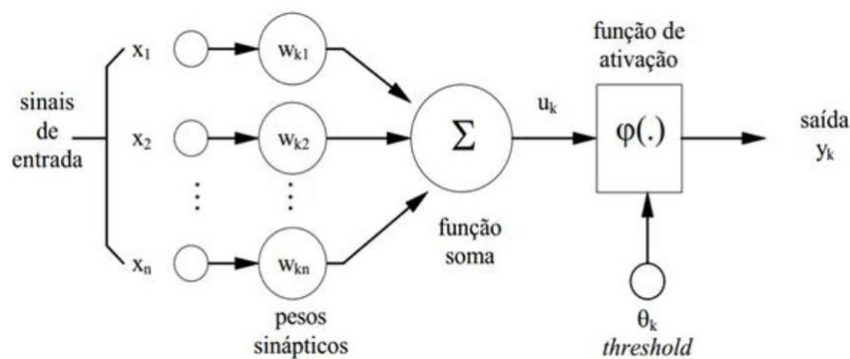


Figura 2.6: Modelo de um neurônio artificial (Fonte: [22]).

Dentre os sinais recebidos  $x$ , como alguns deles são mais estimulantes do que outros, um peso  $W$  é atribuído para cada um deles. Em seguida, é feita a soma das entradas multiplicada pelos seus respectivos pesos. O resultado da função dada pela Equação 2.5 é posteriormente usado como entrada em uma função de ativação  $g()$  que gera uma saída do neurônio. Abaixo são descritos alguns exemplos de funções de ativação:

- Função *Rectified Linear Units* (ReLU): Dada uma entrada  $x$ , se o valor de  $x$  for maior que 0 retorna o valor de  $x$  caso contrário retorna 0, de acordo com:

$$ReLU(x) = \max(0, x); \quad (2.6)$$

- Função de etapa binária (*Binary Step Function*): Dado um limiar  $l$  e uma entrada  $x$ , se o valor de entrada  $x$  for inferior à este limite a saída será zero, senão um:

$$\begin{cases} BSF(x) = 0 & \text{se } x < l \\ BSF(x) = 1 & \text{se } x \geq l \end{cases}; \quad (2.7)$$

- Função sigmóide: Função não linear que dada uma entrada  $x$  retorna um valor real entre  $[0, 1]$ . Pela Equação 2.8, quanto maior a entrada  $x$  mais próximo de 1 está a saída:

$$Sigmoid(x) = \frac{1}{1 + e^{-x}}; \quad (2.8)$$

- Função *Leaky ReLU*: Uma versão melhorada da função ReLU. Na função ReLU o retorno é 0 quando a entrada  $x$  for negativa o que fez os neurônios morrerem nessa região de ativação. Para resolver esse problema, a função Leaky ReLU define um linear  $a$ :

$$\begin{cases} Leaky(x) = ax & \text{se } x < 0 \\ Leaky(x) = x & \text{se } x \geq 0 \end{cases}. \quad (2.9)$$

Uma classe de RNA é a Rede Neural Convolutacional (*Convolutional Neural Network*, CNN). Uma CNN é um algoritmo de Aprendizado Profundo desenvolvido de modo que suas camadas realizem operações de forma linear e não linear, recebendo uma imagem de entrada, atribuindo peso a aspectos ou objetos da imagem e diferenciando um objeto do outro. A primeira aplicação de uma CNN foi feita por LeCun et al. [23] em 1998, com sete camadas entre camadas de convoluções e camadas totalmente conectadas. Nas próximas subseções são apresentadas uma descrição mais detalhada das estruturas de uma CNN.

### 2.4.1 Convolução

A convolução é uma parte vital para a operação da CNN, sendo composta por diversos neurônios, cada neurônio está conectado a um conjunto de *pixels* da camada anterior e é responsável por aplicar um filtro neste conjunto e atribuir um peso. Esses pesos atribuídos às conexões de um neurônio podem ser interpretadas como uma matriz que representa o filtro de uma convolução (também conhecido como *kernel*) e são atualizados a cada nova entrada durante o processo de *backpropagation*. A região na imagem onde o filtro foi aplicado pelo neurônio é chamada de *receptive field* e o agrupamento das saídas dos neurônios da mesma região é chamado de mapa.

A otimização da saída de uma camada de convolução se dá através de três hiperparâmetros, profundidade, saltos e *padding*. A profundidade da saída de uma convolução é igual a quantidade de filtros aplicados, ou seja, quanto mais profunda são as camadas de convolução, mais detalhados são os traços identificados no mapa. Os saltos ou *stride* representam quantos *pixels* serão pulados entre cada *receptive field*, informando qual será o tamanho da camada seguinte na mesma unidade. O *padding* serve para que as camadas não diminuam mais rápido que o necessário para o aprendizado. Pode ser definido como nenhum ou *zero-padding*, quando a escolha é nenhum, a saída da convolução ficará no seu tamanho original, enquanto no *zero-padding* uma borda é adicionada e preenchida com zeros.

Um exemplo do funcionamento da camada de convolução é apresentado na Figura 2.7, considerando uma entrada com dimensão  $28 \times 28$  e *receptive field* de área  $5 \times 5$ .

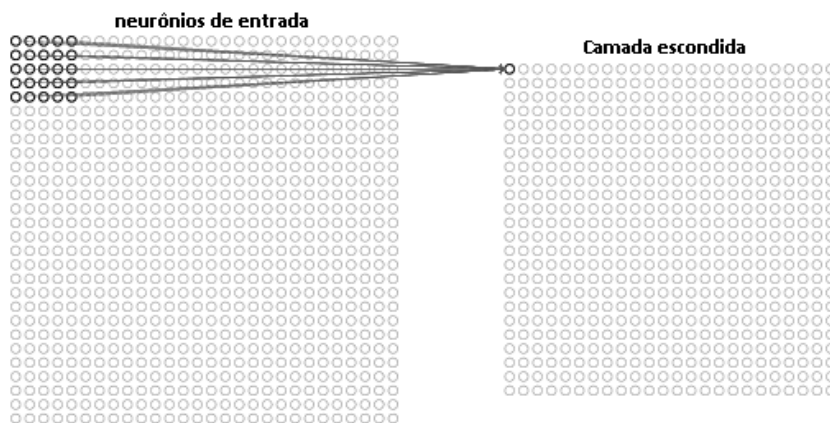


Figura 2.7: Exemplo de funcionamento da camada de convolução (Fonte: [24]).



## 2.4.2 Pooling

A camada de *pooling* tem o objetivo de simplificar a informação da camada anterior, e assim diminuir o número de parâmetros e a complexidade computacional do modelo. Assim como na convolução da Seção 2.4.1, escolhe-se uma região para percorrer toda a saída da camada anterior.

A camada de *pooling* opera sobre cada um dos mapas, o método mais comum, *max pooling*, consiste no dimensionamento utilizando apenas o maior número da unidade passado para a saída, com *stride* de 2, redimensionando assim o mapa com 25% do tamanho original apresentado na Figura 2.8, em que cada cor representa um filtro 2×2, de forma que dos 4 elementos da área selecionada pelo filtro é escolhido o elemento de maior valor para a saída.

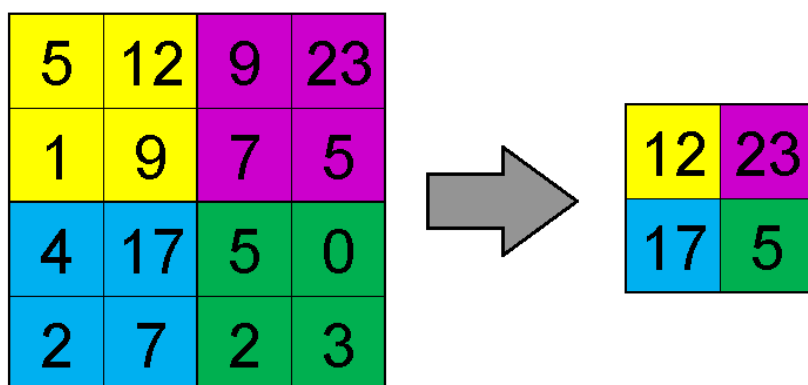


Figura 2.8: Aplicação de *max pooling*.

## 2.4.3 Camada Totalmente Conectada

A camada totalmente conectada tem como entrada a todos os neurônios de saída da camada de *pooling* anterior que são transformados em um vetor unidimensional onde sua saída será este vetor multiplicado por um peso.

## 2.5 Arquiteturas de CNN

A maneira como os neurônios artificiais de uma rede neural são agrupados é uma característica que define o tipo de arquitetura da rede neural.

## 2.5.1 U-net

A U-net é uma Rede Neural Convolutacional desenvolvida por Ronneberger et al. [10] para segmentação em imagens biomédicas no *Computer Science Department and BIOS Centre for Biological Signalling Studies* da *University of Freiburg* na Alemanha.

A rede foi baseada em uma *fully convolutional network* inicialmente proposta por Long et al. [25], onde os autores propõem uma adaptação de redes já utilizadas para segmentação de imagens, transformando as camadas totalmente conectadas em camadas de convolução permitindo a geração de mapas de características de segmentação para cada imagem, podendo ser de qualquer tamanho.

A U-net consiste em um caminho de contração e um de expansão, sua simetria da o formato de U da arquitetura visto na Figura 2.9. O caminho de contração, parte a esquerda, é uma rede convolutacional típica que consiste na aplicação repetida de convoluções, cada uma seguida por uma função ReLU e um operador *max pooling*. Durante a contração, as informações espaciais são reduzidas enquanto as *features* são incrementadas. O caminho de expansão, parte a direita, combina *features* e informações espaciais por meio de uma sequência de convoluções e concatenações de *features* de alta resolução do caminho de contração.

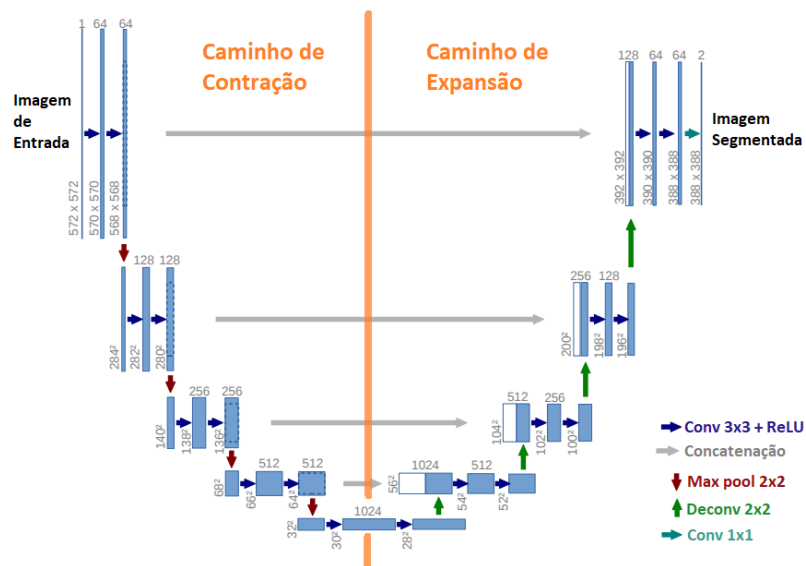


Figura 2.9: Arquitetura da rede U-net (Fonte: [10]).

## 2.5.2 Attention U-net

Uma variação na arquitetura da U-net é a *attention U-net* que utiliza mecanismos de atenção para guiar o modelo enfatizando as características mais importantes e evitando

as características menos importantes para a segmentação durante o treinamento.

Existem dois tipos de mecanismos de atenção:

- *Hard Attention*: É um processo estocástico, em que a rede obtém informações apenas de uma pequena parte da imagem. Essa parte é escolhida de forma iterativa pela rede neural por meio de um mecanismo de seleção de atenção;
- *Soft Attention*: É um processo probabilístico, em que uma CNN gera um mapa de característica aplicando um filtro na camada anterior (*pooling*). Na *soft attention* mapas de característica diferentes tem pesos diferentes e são ponderados de forma diferente.

Com a adição desse mecanismo a arquitetura da U-net deve ser modificada, adicionando o *attention gate* no caminho de expansão da rede, visto na Figura 2.10. Nessa arquitetura a imagem de entrada é progressivamente filtrada e reduzida por um fator de 2 em cada fase do caminho de contração da rede.

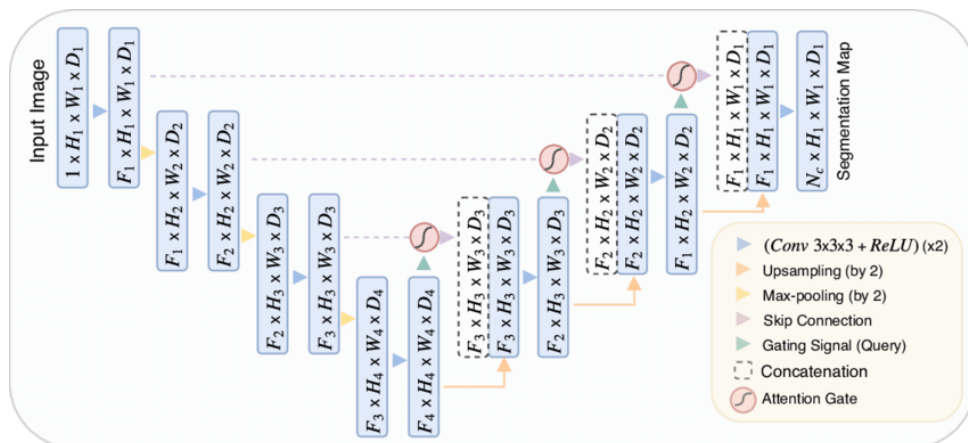


Figura 2.10: Arquitetura da rede *Attention* U-net. A imagem de entrada é progressivamente filtrada e reduzida por um fator de 2 em cada fase da parte de codificação da rede. *Attention gates* (AGs) filtra os recursos propagados pelas conexões (Fonte: [11]).

Os *attention gates* filtram as *features* propagados pelo caminho de contração. O esquemático dos *attention gates* é apresentado na Figura 2.11. As *features* de entrada ( $x^l$ ) são escalonadas com coeficiente de atenção ( $\alpha$ ) calculados no AG. As regiões espaciais são selecionadas avaliando as ativações e as informações contextuais fornecidas pelo sinal de *gating* ( $g$ ), que foi coletado anteriormente de forma mais grosseira, por último é feito a reamostragem do coeficiente de atenção usando interpolação trilinear.

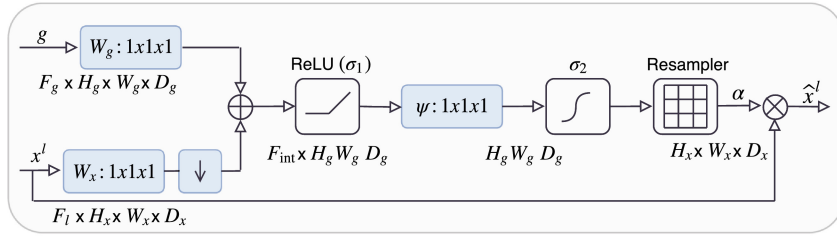


Figura 2.11: Esquemático do *Attention gate* (AG) (Fonte: [11]).

## 2.6 Métricas de Avaliação de Desempenho

Métricas de avaliação de desempenho são utilizadas para comparar modelos de segmentação, cada uma avaliando um aspecto diferente do modelo. A principal métrica é a matriz de confusão, uma vez que as demais métricas, são obtidas através de combinações de seus valores.

A matriz de confusão é formada pela classificação dos *pixels* gerados pelo método de segmentação, comparando-os com os *pixels* do *ground-truth*. Sendo a classe positiva o conjunto de *pixels* que pertencem ao objeto a ser segmentado e classe negativa o conjunto de *pixels* que pertencem ao fundo, ou seja, não pertencem ao objeto a ser segmentado.

A comparação é feita separando em quatro possíveis classificações:

- Verdadeiro Positivo (VP): *Pixel* que pertence a classe positiva e foi classificado como positivo;
- Falso Positivo (FP): *Pixel* que pertence a classe negativa e foi classificado como positivo;
- Verdadeiro Negativo (VN): *Pixel* que pertence a classe negativa e foi classificado como negativo;
- Falso Negativo (FN): *Pixel* que pertence a classe positiva e foi classificado como negativo.

Da matriz de confusão podemos obter as seguintes métricas:

- Acurácia: Taxa de acerto do modelo, ou seja, a quantidade de segmentos classificados corretamente. Para isso mede-se a relação de VP e VN em comparação com o conjunto total:

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN}; \quad (2.10)$$

- **Precisão:** Quantidade de segmentos classificados como positivo que realmente pertencem a classe positiva. Para isso mede-se a relação de VP em comparação com o conjunto classificado como positivo (VP e FP):

$$Precisão = \frac{VP}{VP + FP}; \quad (2.11)$$

- **Revocação:** Quantidade de segmentos que pertencem a classe positiva e que foram classificados corretamente. Para isso mede-se a relação de VP em comparação com o conjunto que pertence a classe positiva (VP e FN):

$$Revocação = \frac{VP}{VP + FN}; \quad (2.12)$$

- **Coefficiente Dice:** Utilizado para medir a similaridade entre dois objetos. Calculado pela média harmônica entre a precisão e a revocação. Seu valor varia entre 0 e 1, sendo mais próximo de 1 significa que os valores aferidos da matriz de confusão não tem grandes distorções:

$$Dice = 2 * \frac{Precisão * Revocação}{Precisão + Revocação}. \quad (2.13)$$

# Capítulo 3

## Revisão Bibliográfica

As tecnologias para obtenção de imagens cardiovasculares vêm aumentando cada vez mais sua capacidade de gerar grandes quantidades de dados [26], despertando assim interesse no uso de técnicas de aprendizado profundo para auxiliar no diagnóstico de doenças cardíacas. Neste capítulo serão apresentados alguns trabalhos sobre essas técnicas.

Geralmente, para um treinamento bem sucedido de uma rede neural profunda é necessário uma grande quantidade de imagens rotuladas. No entanto, conjuntos de imagens médicas tendem a ser pequenos. Considerando esse fato, Ronneberger et al. [10] propõem uma rede neural denominada U-net, baseada em uma rede neural convolucional tradicional com estratégias de treinamento que contam com forte uso de *data augmentation* para utilizar as amostras disponíveis mais eficientemente.

Ammar et al. [27] também considerou o problema da quantidade de imagens, em sua base com inicialmente 1902 imagens, aplicando técnica de *data augmentation* gerando a mesma quantidade de imagens que o conjunto de dados original oferece, totalizando 3804 imagens. A geração dessas novas imagens foi feita por meio de deslocamento na altura e largura da imagem, pequenas rotações e inversão horizontal. Os autores fazem a segmentação de três estruturas do coração: ventrículo esquerdo, ventrículo direito e miocárdio<sup>1</sup>. Na segmentação do ventrículo esquerdo, a técnica de *data augmentation* obteve um aumento de 0,01 no coeficiente dice, passando de 0,90 para 0,91.

Um artigo que inspirou o desenvolvimento do presente trabalho foi o artigo feito por Zreik et al. [28], nele é apresentado um processo de segmentação através da medição do fluxo fracionado de reserva do miocárdio (*fractional flow reserve* - FFR), um procedimento invasivo que permite identificar isquemia cardíaca<sup>2</sup>. Com o intuito de diminuir o número de procedimentos invasivos, Zreik et al. [28], apresentam um método utilizando *deep learning*, através da análise do ventrículo esquerdo em tomografia, para isso o projeto foi

---

<sup>1</sup>Músculo cardíaco responsável por ejetar sangue que se encontra no interior do coração.

<sup>2</sup>Dano ou doença nos principais vasos sanguíneos do coração

separado em etapas. Inicialmente, o ventrículo esquerdo foi segmentado utilizando uma CNN e as *deep features* retiradas foram adicionadas a um *autoencoder* convolucional. Em seguida, essas *deep features* foram utilizadas para identificar os pacientes com doenças cardíacas utilizando um classificador *Support Vector Machines* (SVM) com base nas medidas do FFR. Para examinar a viabilidade de se utilizar essa metodologia os autores utilizaram o método em 20 imagens resultando em um coeficiente Dice de 0,91, outras métricas foram apresentadas somente para o processo de classificação como a acurácia de 94%.

O método proposto por Binbin et al. [29] para segmentação ventricular em imagens de ressonância magnética nuclear, foi o Mask R-CNN. Consiste em um aperfeiçoamento do modelo Faster R-CNN, que além de retornar a classe e a caixa limitadora, também retorna a máscara do objeto detectado na imagem. Para o aperfeiçoamento do Faster R-CNN foi utilizado a arquitetura ResNet para resolver o problema de perda de informações do gradiente e *Feature Pyramid Network* (FPN), pois com o aumento do número de camadas o reconhecimento de pequenos objetos diminui. Os autores apresentam coeficiente Dice de 0,91 para o ventrículo esquerdo e 0,89 para o ventrículo direito e confirmam experimentalmente que o método pode ser usado para determinar se o paciente tem disfunção sistólica<sup>3</sup>.

Isensee et al. [30] apresenta o mesmo problema citado anteriormente por Ammar et al. [27]. Ambos os trabalhos foram feitos sobre a base *Automated Cardiac Diagnosis Challenge 2017* [31] composto por 150 pacientes com imagens 3D de ressonância magnética, com a diferença que o pré-processamento feito por Isensee et al. [30] foi na escala de cinza das imagens, enquanto Ammar et al. [28], aumentou a base fazendo ajustes no tamanho, rotação e invertendo as imagens e utilizando U-net na resolução do problema, porém o artigo de Isensee et al. [30] fez *cross-validation* obtendo um coeficiente Dice de 0,945 comparado a 0,93 que foi obtido por Ammar et al.

A classificação feita por Isensee et al. [30] utilizou MLP com quatro camadas escondidas que contêm 32 unidades, ReLU não linear e camada de regularização Gaussiana com acurácia de classificação de 93% e fez a comparação com o método floresta aleatória chegando a uma acurácia de 94%, um resultado mais preciso quando comparada a 92% obtidos por Zreik et al. [28].

Em seu artigo, Liao et al. [32] apresentam comparações de diversos métodos para resolver a proposta apresentada no *Second Annual Data Science Bowl* organizado pelo Kaggle<sup>4</sup>. A base de dados é dividida em treinamento, validação e teste contendo 500, 400 e 440 pacientes, respectivamente. Com o objetivo de calcular o volume de sangue bombeado

---

<sup>3</sup>Incapacidade do coração bombear sangue, em outras palavras, o coração fica "fraco"

<sup>4</sup><https://www.kaggle.com/c/second-annual-data-science-bowl>

pelo ventrículo esquerdo, foi feito um pré-processamento nas imagens para definir quais frames serão utilizados, além disso, em alguns casos a resolução, a orientação e o tamanho dos frames são inconsistentes entre os frames de um mesmo paciente, para resolver este problema estes casos foram alterada para  $48 \times 48$ . Em seguida, foi extraída a Região de Interesse (*Region of Interest*, ROI) e aplicada CNN para pontuar as ROIs e escolher as que tiveram maior pontuação e por fim calcular a área dessas ROIs com base em uma *Labeled-Based Forecasting* (LBF). A acurácia do método foi 97,2%.

A partir do embasamento teórico levantado nos artigos apresentados neste capítulo, foi possível perceber que a arquitetura U-net se mostrou bastante utilizada para segmentação de imagens médicas. Um fato em comum entre os trabalhos é a baixa quantidade de imagens presentes nas bases de dados, sendo assim, a aplicação de técnicas de *data augmentation* se mostra uma solução viável. Além disso, a *Attention U-net* [11] é uma arquitetura que apresentou bons resultados na segmentação do pâncreas de imagens de tomografia, sendo assim é interessante avaliar o seu resultado para o problema abordado neste trabalho. O modelo de segmentação tradicional Chan-Vese [7] apresenta vantagens em relação aos contornos ativos, sendo uma motivação para a sua comparação com modelos de *deep learning* utilizados atualmente. Com isso, foram definidos os métodos descritos no Capítulo 4 a serem utilizados nos experimentos.



# Capítulo 4

## Metodologia

Este capítulo apresenta a metodologia empregada neste trabalho. O trabalho consiste na comparação de modelos de Rede Neural Convolutiva e um método tradicional para segmentação do ventrículo esquerdo a partir de imagens de ressonância magnética, bem como um modelo híbrido baseado nas duas abordagens. Todos os experimentos realizados são apresentados na Figura 4.1. Para avaliar o desempenho de cada uma das técnicas de pré-processamento citadas no Capítulo 2 os experimentos com modelos de CNNs foram divididos em três partes: (i) sem nenhum pré-processamento nas imagens de entrada, (ii) com realce das imagens através da equalização de histograma (CLAHE) e (iii) com aumento na quantidade de imagens (*data augmentation*). Em seguida, foi feito um experimento com o método tradicional Chan-Vese, em seguida. Por último foi elaborado um modelo híbrido com o intuito de melhorar a segmentação do modelo Chan-Vese.

### 4.1 Base de imagens

Para realizar os experimentos foi utilizada a base de imagens *Sunnybrook Cardiac Data* (SCD) [1], também conhecido como *2009 Cardiac MR Left Ventricle Segmentation Challenge Data*. A base consiste em conjuntos de imagens de ressonância magnética de 45 pacientes divididos em quatro patologias: (1) saudáveis, (2) hipertrofia, (3) insuficiência cardíaca com infarto e (4) insuficiência cardíaca sem infarto. As imagens estão em formato *Digital Imaging and Communications in Medicine* (DICOM).

O conjunto total de dados possui 32000 imagens monocromáticas com dimensões distintas ( $64 \times 64$ ,  $128 \times 128$ ,  $256 \times 256$  e  $512 \times 512$ ). Além destas imagens, a base contém um conjunto de contornos (*ground-truth* da segmentação), em arquivos de texto, definidos manualmente pelo especialista Perry Radau do *Sunnybrook Health Sciences Centre*, Toronto, Canadá. Os contornos definidos manualmente estão disponíveis apenas para algumas imagens de dimensão  $256 \times 256$ . Considerando um total de 12235 imagens com

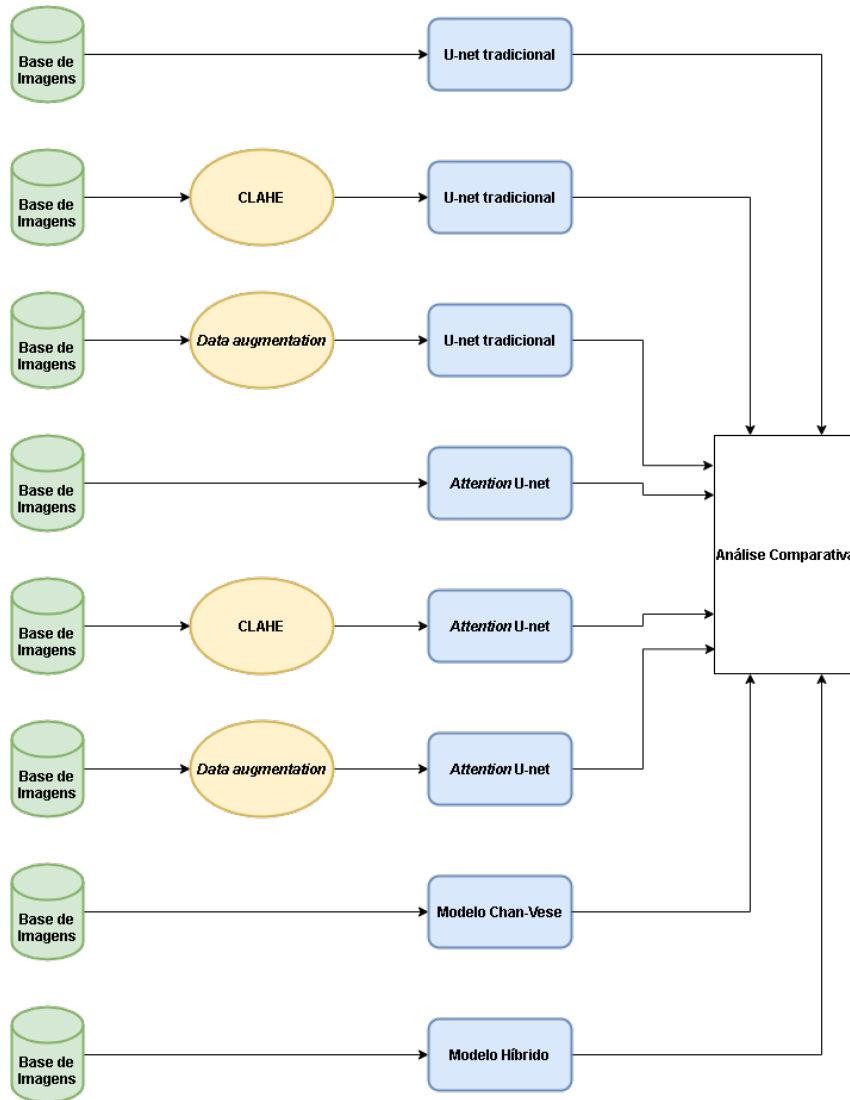


Figura 4.1: Diagrama da metodologia empregada. O modelo híbrido é descrito na Figura 4.4.

dimensão  $256 \times 256$ , apenas 805 dessas imagens possuem a definição do contorno. Com o intuito de treinar e validar os métodos é necessário que as imagens utilizadas contenham o *ground-truth*. Portanto, para os experimentos foram usadas apenas as respectivas 805 imagens do conjunto de dados.

#### 4.1.1 Separação das imagens

Para a escolha dessas imagens, primeiramente foi feita a leitura do conjunto de *ground-truth* disponível em arquivo texto e gerada uma imagem para cada um deles. Estes arquivos estão separados por pastas para cada paciente e o nome do arquivo contém um número de identificação.

Em seguida, as 12235 imagens de dimensão  $256 \times 256$  foram agrupadas pelo código de identificação do paciente e pelo número de identificação da imagem no exame presente nos metadados da imagem. Então, essas identificações foram cruzadas com as identificações do *ground-truth* gerando um conjunto de possíveis imagens correspondentes para cada *ground-truth*. Por fim, através da validação visual foi feita a correspondência do *ground-truth* com sua imagem, resultando no total de 805 imagens com *ground-truth*.

### 4.1.2 Leitura do Contorno

Foi necessário converter os contornos manuais para um formato que seja compatível com os tensores de entrada da rede neural. Os contornos manuais inicialmente se encontram em formato de arquivo texto (.txt) contendo um conjunto de coordenadas cartesianas referentes as coordenadas de cada pixel do contorno. As coordenadas foram lidas e transformadas em uma máscara no formato *numpy array* com dimensão de  $256 \times 256$  e os *pixels* internos ao contorno foram preenchidos.

### 4.1.3 Normalização dos *pixels*

Fazendo uma análise dos metadados das imagens do conjunto de dados, observa-se que as imagens são do tipo *MONOCHROME2* com 16 bits de resolução, sendo assim possui uma resolução maior com valores do *pixel* chegando até 65536. Devido a essa discrepância de valores é importante fazer a normalização para evitar problemas de estabilidade numérica. Portanto, os *pixels* foram normalizados de acordo com a Equação 4.1.

$$x_{normalized} = \frac{x}{max(x)}. \quad (4.1)$$

$x$  representa o *pixel* e  $max(x)$  representa o maior valor de todos os *pixels* da imagem. Desta forma, todos os valores da imagem pertencem ao intervalo de  $[0, 1]$ .

## 4.2 Pré-Processamento das imagens

A etapa de pré-processamento consiste em melhorar a qualidade das imagens para permitir uma melhor discriminação dos objetos da imagem. Portanto, algumas técnicas de pré-processamento foram utilizadas neste trabalho com o intuito de analisar o desempenho dos algoritmos de segmentação testados na resolução do problema. Abaixo são explicadas quais técnicas de pré-processamento foram adotadas.

### 4.2.1 Equalização do histograma

O conjunto de dados é formado por imagens de ressonância magnética que possuem baixa luminosidade e contraste, o que dificulta a delimitação dos contornos dos objetos na imagem, sendo assim grande parte dos *pixels* no histograma se encontram nos menores valores de intensidade. Com o intuito de melhorar o contraste e brilho destas imagens [33], foi feita a equalização do histograma.

A técnica aplicada para tal foi a CLAHE, que significa *Contrast Limited Adaptive Histogram Equalization* (Equalização de Histograma Adaptativo Limitado por Contraste) [34], na qual a imagem é dividida em  $N \times N$  blocos de equalização, o histograma é calculado para cada bloco e é feita sua equalização com limite para a amplitude definido pelo parâmetro *ClipLimit*. A Figura 4.2 é um exemplo de uma imagem da base [1], sendo dividida em 64 blocos de equalização ( $8 \times 8$ ) com *ClipLimit* de 2.0.

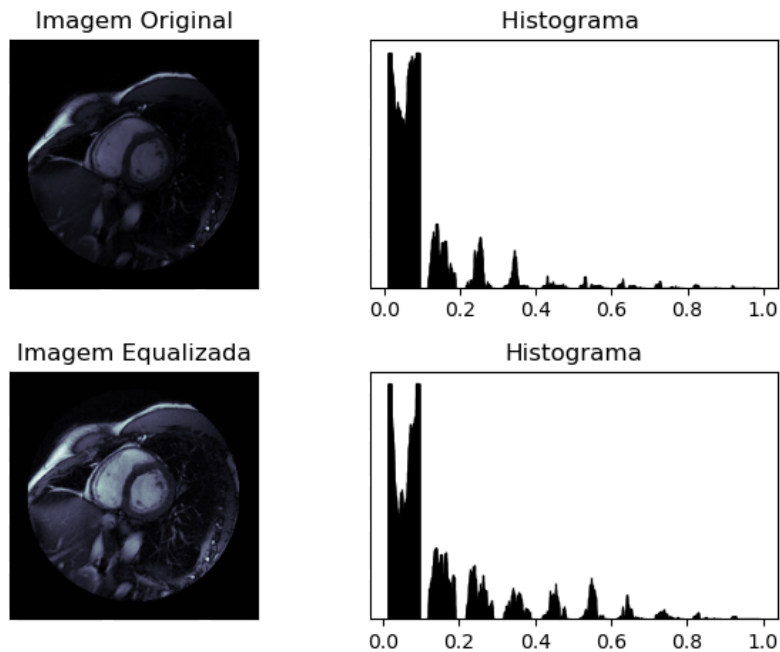


Figura 4.2: Exemplo de equalização do histograma em uma imagem original da base.

### 4.2.2 Data Augmentation

Considerando que a base de imagens contém poucas imagens, com o *data augmentation* é possível aumentar a capacidade do modelo sem precisar coletar novos dados, reduzindo a possibilidade de *overfitting*. As estratégias utilizadas neste trabalho foram: rotacionar a imagem em 180 graus, inversão da imagem tanto horizontal e verticalmente, aplicação de

*zoom* de até 20%. A Figura 4.3 apresenta tais estratégias, que foram aplicadas nos conjuntos de validação e treinamento, duplicando a quantidade de imagens nesses conjuntos, conforme descritos no Capítulo 5.

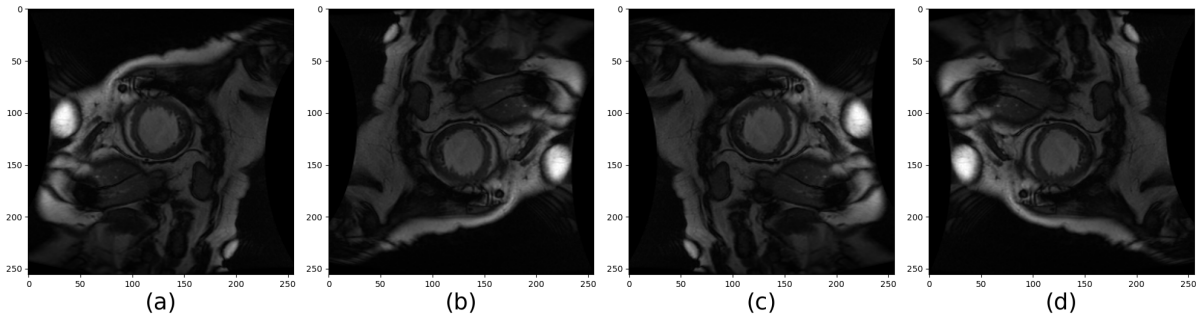


Figura 4.3: Exemplos de *data augmentation*: (a) Imagem original retirada do conjunto de dados Sunnybrook [1], (b) Imagem rotacionada 180 graus, (c) Imagem invertida verticalmente, e (d) Imagem invertida horizontalmente.

## 4.3 Segmentação

Para fazer a segmentação do ventrículo esquerdo foram utilizadas as imagens que possuem o *ground-truth* descritas na Seção 4.1.1. Foram utilizados quatro métodos de segmentação: U-net tradicional, *Attention* U-net, Modelo Chan-Vese e modelo híbrido. A U-net [10] é uma das mais utilizadas atualmente conforme a revisão de literatura. Já a *Attention* U-net [11] é uma arquitetura que apresentou bons resultados na segmentação do pâncreas de imagens de tomografia, sendo assim é interessante avaliar o seu resultado para o problema abordado neste trabalho e o modelo Chan-Vese [7] se mostrou promissor em experimentos preliminares.

### 4.3.1 U-net tradicional

A U-net apresenta forte uso de *data augmentation* para aumentar a quantidade de imagens virtualmente [10], sem necessidade de mais imagens rotuladas para treinamento. As técnicas de *data augmentation* são feitas em sua arquitetura por meio de deformações que alteram a forma dos objetos na imagem e atuam como se os objetos representados estivessem em posições diferentes.

Outro fator importante da arquitetura é a simetria da rede devido as conexões de salto entre o caminho de contração e expansão que combinam as informações de localização do caminho de contração com as informações contextuais do caminho de expansão.

O objetivo consiste em dada uma imagem de entrada gerar uma imagem binária de saída onde 1 representa o ventrículo esquerdo e 0 representa o fundo, ou seja, *pixels* que não são parte do ventrículo esquerdo.

### 4.3.2 *Attention* U-net

Na implementação da *Attention* U-net [11] foi utilizada a técnica de *Soft Attention*, que consiste em atribuir pesos maiores a partes mais relevantes da imagem e pesos menores a partes menos relevantes. Portanto, foi atribuído um peso maior aos *pixels* pertencentes ao ventrículo esquerdo, gerando uma imagem binária de saída como resultado da segmentação do ventrículo.

### 4.3.3 Modelo tradicional Chan-Vese

Foi implementado um modelo de detecção Chan-Vese para a segmentação do ventrículo esquerdo, utilizando como entrada as imagens com o melhoramento de contraste que contêm *ground-truth* e como contorno inicial uma máscara binária contém um círculo no centro da máscara.

### 4.3.4 Modelo Híbrido

O modelo híbrido feito consiste na combinação do modelo tradicional Chan-Vese com a U-net tradicional. Para fazer a segmentação com este modelo, primeiramente a U-net tradicional recebe como entrada as imagens com melhoramento de contraste que contêm *ground-truth* gerando uma imagem binária. Essa imagem então é utilizada como contorno inicial no modelo tradicional Chan-Vese gerando assim uma nova segmentação. Os passos descritos estão presentes na Figura 4.4.

## 4.4 Extrair Contorno da imagem binária

Para construir o modelo híbrido é preciso refinar a imagem binária gerada pela arquitetura U-net tradicional, uma vez que aconteceu da U-net tradicional segmentar dois objetos na imagem, para isso utilizado a função *findContours* da biblioteca OpenCV<sup>1</sup> para extrair a linha que une todos os pontos ao longo do limite da imagem que, por se tratar de

---

<sup>1</sup><https://opencv.org/>

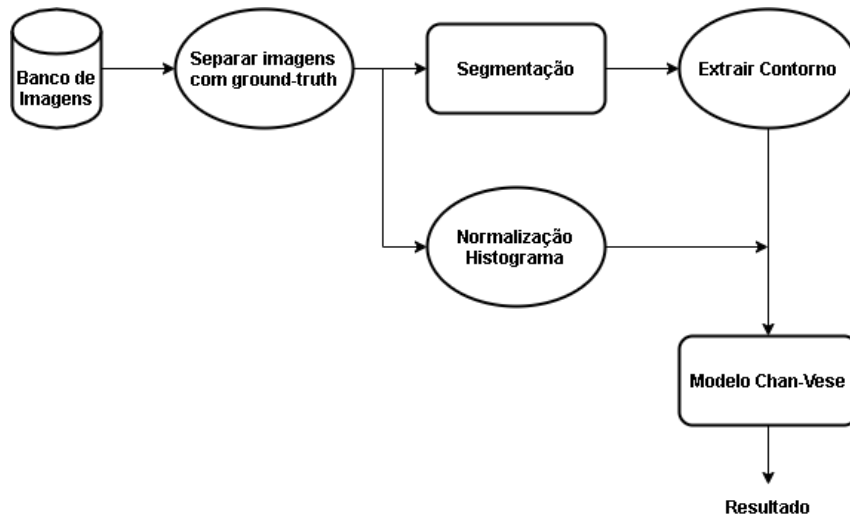


Figura 4.4: Diagrama da rede híbrida proposta.

uma imagem binária, possuem intensidade igual a 1. Para isso, foram definidos dois parâmetros:

- O modo de recuperação foi definido como *RETR\_EXTERNAL* para recuperar apenas os contornos externos do objeto.
- O modo de aproximação de contorno foi definido como *CHAIN\_APPROX\_NONE* que armazena todos os pontos do contorno.

O retorno da função é um vetor contendo as coordenadas  $x$  e  $y$  de todos os pontos do contorno. Porém, em alguns casos a segmentação feita anteriormente resultou em mais de um objeto, e portanto, a função retornou mais de um contorno. Esse problema foi resolvido utilizando o teorema de Green [2] para o cálculo da área dos objetos com base no contorno, assim, visto que na segmentação feita pela U-net os outros objetos segmentados de forma errônea são menores que o ventrículo esquerdo, o contorno de maior área foi selecionado. Com esse vetor de pontos foi utilizado uma segunda função da biblioteca OpenCV<sup>2</sup>, a *drawContours* para criar uma nova imagem binária contendo apenas a linha que define o contorno do objeto conforme Figura 4.5.

---

<sup>2</sup><https://opencv.org/>

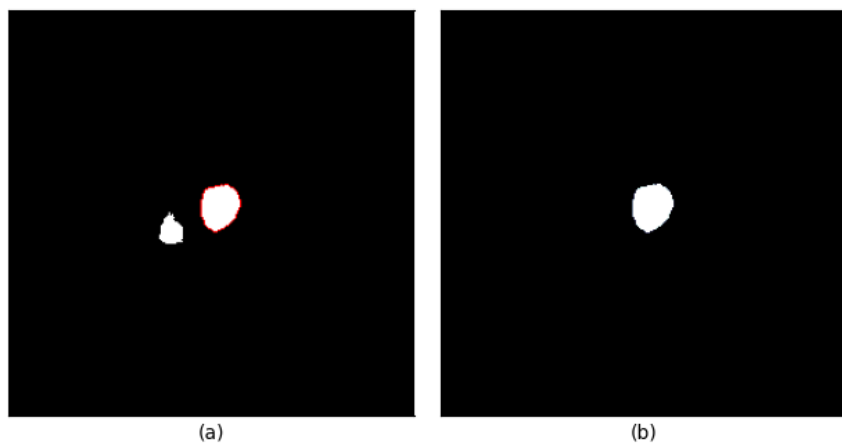


Figura 4.5: Exemplo do teorema de Green [2] em uma imagem binária gerada pela U-net tradicional no conjunto de imagens (a) a linha em vermelho são os pontos do contorno do maior objeto, (b) é o resultado após a aplicação do teorema de Green.



# Capítulo 5

## Resultados Experimentais

Neste capítulo são apresentados os resultados obtidos nos experimentos descritos no Capítulo 4 além da análise dos resultados dos métodos, com o intuito de comparar seus desempenhos na segmentação do ventrículo esquerdo em imagens de ressonância magnética. Primeiramente, são apresentados os recursos computacionais e bibliotecas utilizados para o desenvolvimento deste trabalho. Em seguida, são apresentados os resultados obtidos na segmentação tradicional utilizando o modelo Chan-Vese. Na sequência, são analisados os resultados da segmentação utilizando as redes U-net tradicional e *Attention* U-net, e por último, são analisados os resultados do modelo híbrido.

### 5.1 Implementação

As rede neural convolucionais e todo o processamento e testes foram implementados no sistema operacional Windows 10 utilizando a linguagem Python 3.8 com as bibliotecas Scikit-learn, Kera, NumPy e OpenCV na plataforma Jupyter. Todos os experimentos foram realizados em uma máquina com CPU Intel(R) Core(TM) i7-7700k @ 4.2 GHz, GPU NVIDIA GeForce GTX 1070 8 GB Vram e 8 GB 2400 MHz de memória RAM. Para diminuir o tempo computacional de execução e otimizar as tarefas, foi utilizado processamento paralelo com a GPU, utilizando as bibliotecas CUDA e cuDNN.

### 5.2 Segmentação tradicional

Primeiramente foi realizado um experimento com técnicas de segmentação por contorno ativo. Para isso, foi utilizado o modelo Chan-Vese [7] com 100 iterações, escolhido de forma empírica, sendo o contorno inicial um círculo de raio 20, escolhido de forma visual, com centro sobre com o centro da imagem, conforme Figura 5.1. A posição do contorno

inicial foi considerada central, pois foi observado que em grande parte das imagens o ventrículo esquerdo está localizado mais próximo do centro da imagem.

As Figuras 5.1 a 5.2 destacam dois exemplos de uma segmentação feita com o modelo Chan-Vese [7] sendo o contorno em verde é o *ground-truth*, em vermelho o contorno detectado pela modelo e o contorno inicial passado para a função em amarelo. Pode-se observar na Figura 5.1 que com o ventrículo dentro do contorno inicial, o modelo Chan-Vese apresenta um resultado satisfatório pois consegue segmentar o objeto de interesse de forma eficiente. Por outro lado, quando o objeto de interesse está fora do contorno inicial (Figura 5.2) o modelo pode perder o controle e segmentar uma região que não seja a região de interesse. Com isso, um melhor contorno inicial deve ser selecionado.

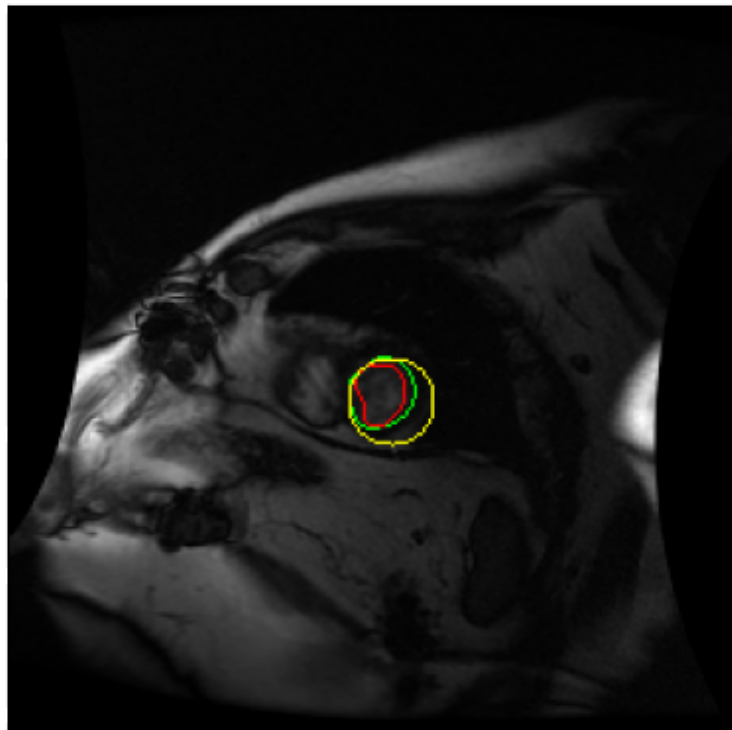


Figura 5.1: Exemplo de segmentação Chan-Vese com objeto de interesse dentro do contorno inicial. Sendo o contorno inicial do modelo em amarelo, em verde o *ground-truth* e em vermelho o resultado da segmentação.

### 5.3 Redes Neurais Convolucionais

O conjunto de imagens foi dividido em 70% das imagens no conjunto de treinamento, 20% no conjunto de teste e 10% no conjunto de validação correspondendo a 563, 161 e 81

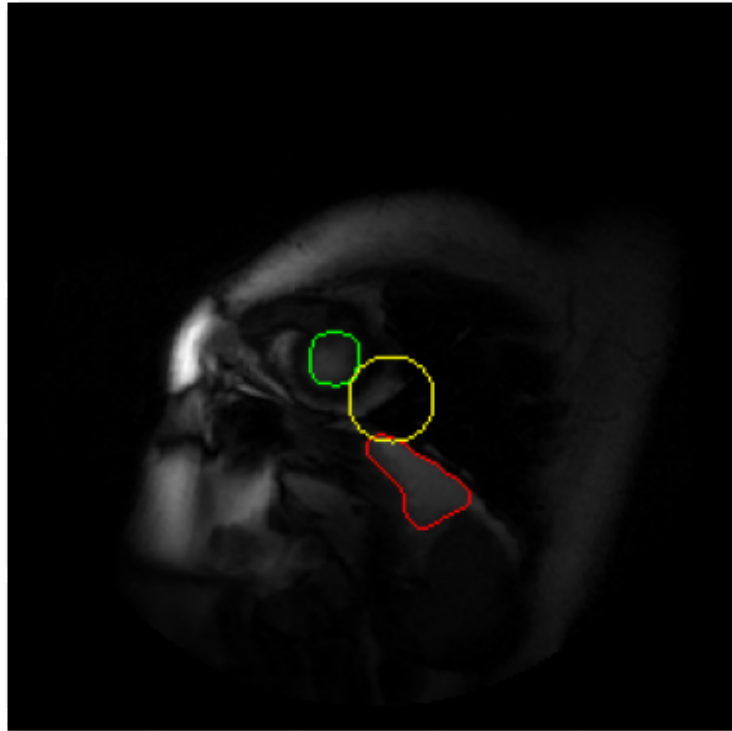


Figura 5.2: Exemplo de segmentação Chan-Vese com objeto de interesse fora do contorno inicial. Sendo o contorno inicial do modelo em amarelo, em verde o *ground-truth* e em vermelho o resultado da segmentação.

imagens, respectivamente.

Para o treinamento das redes neurais convolucionais foram utilizados os conjuntos de treinamento e validação. Foram aplicadas as técnicas de realce da imagem e *data augmentation* separadamente com método de otimização *Adam* e fixando o número de épocas em 100 com passos por época dado pela  $T/batch\_size$ , sendo  $T$  o tamanho do conjunto de treinamento.

Com o objetivo de avaliar o desempenho de cada uma das redes, os experimentos foram divididos em três partes para que assim, seja possível avaliar o efeito de cada etapa de pré-processamento nas redes. As três partes são:

- Sem nenhum pré-processamento nas imagens de entrada;
- Apenas com *data augmentation*;
- Apenas com realce da imagem.

### 5.3.1 Segmentação sem pré-processamento

O primeiro experimento realizado com rede neural convolucional foi utilizar as imagens sem nenhum pré-processamento obtidas na separação das imagens com *ground-truth* (Seção 4.1.1) como entrada nas arquiteturas U-net tradicional e *Attention* U-net.

Nas Figuras 5.3 a 5.4, o eixo X representa o número de épocas e o eixo Y representa o coeficiente Dice. Pode-se observar que, com os dados sem nenhum pré-processamento, a U-net tradicional obteve uma melhora significativa nas primeiras 40 épocas enquanto a *Attention* U-net manteve uma pequena melhora ao decorrer das 100 épocas.

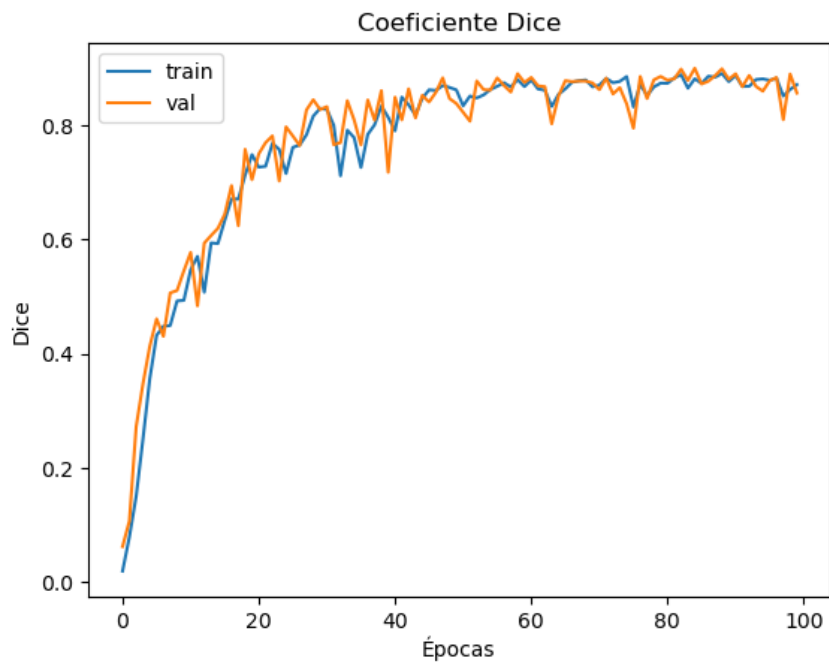


Figura 5.3: Coeficiente Dice em relação às épocas da U-net tradicional sem pré-processamento.

Com o decorrer das épocas os pesos da rede vão se reajustando de forma que o algoritmo de otimização vai convergindo, isto é, a rede aprende os padrões do objeto a ser segmentado. Vale ressaltar que a curva da Figura 5.3 cresce exponencialmente e na 30<sup>a</sup> começa a se estabilizar e segue constante até a última época o que indica que a rede teve um bom treinamento, sem presença de *overfitting*.

### 5.3.2 Aumento da quantidade de imagens

Em seguida, foi utilizado o *data augmentation* com o intuito de analisar seu impacto nas arquiteturas. Para isso, foi duplicada a quantidade de imagens nos conjuntos de

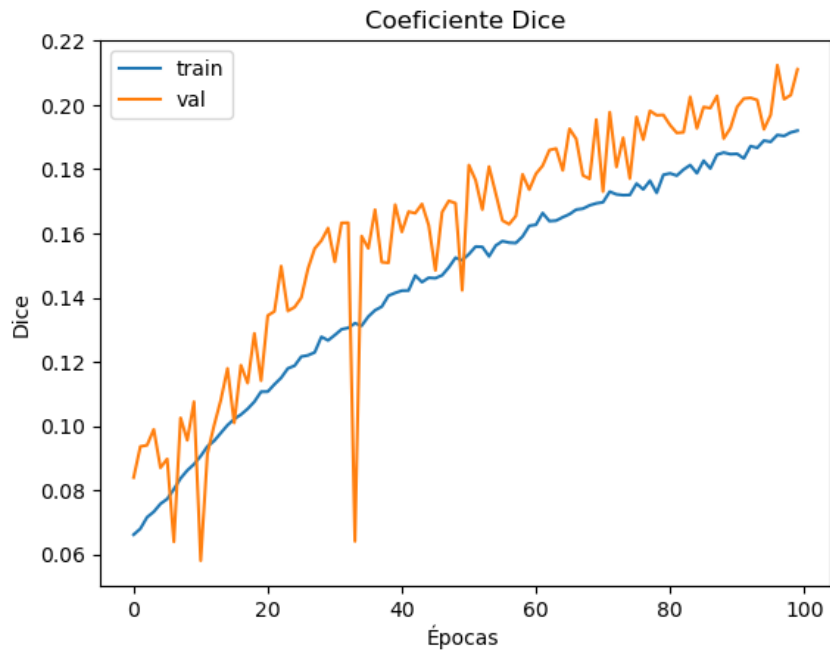


Figura 5.4: Coeficiente Dice da *Attention* U-net sem pré-processamento.

treinamento e validação totalizando 1127 imagens para treinamento e 162 imagens para validação.

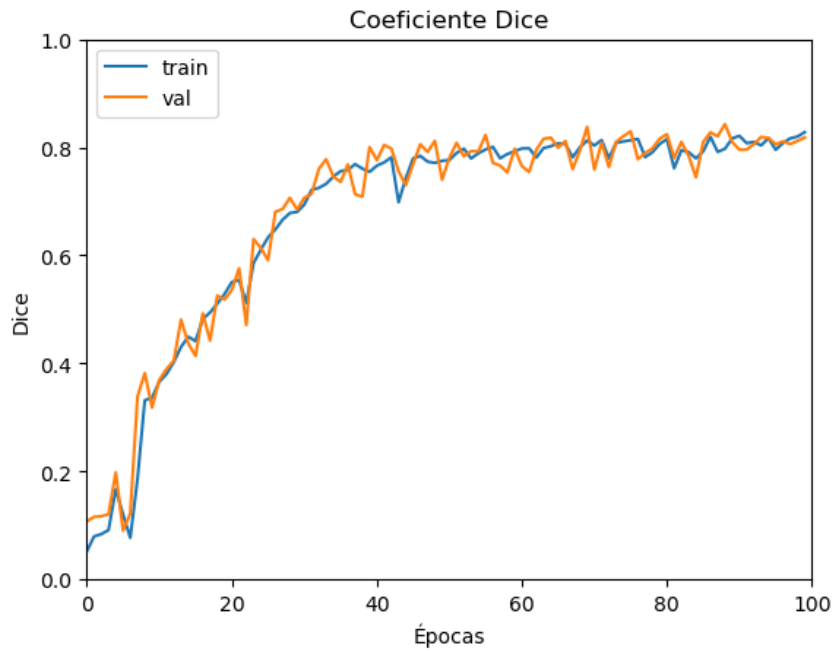


Figura 5.5: Coeficiente Dice da U-net tradicional com *data augmentation*.

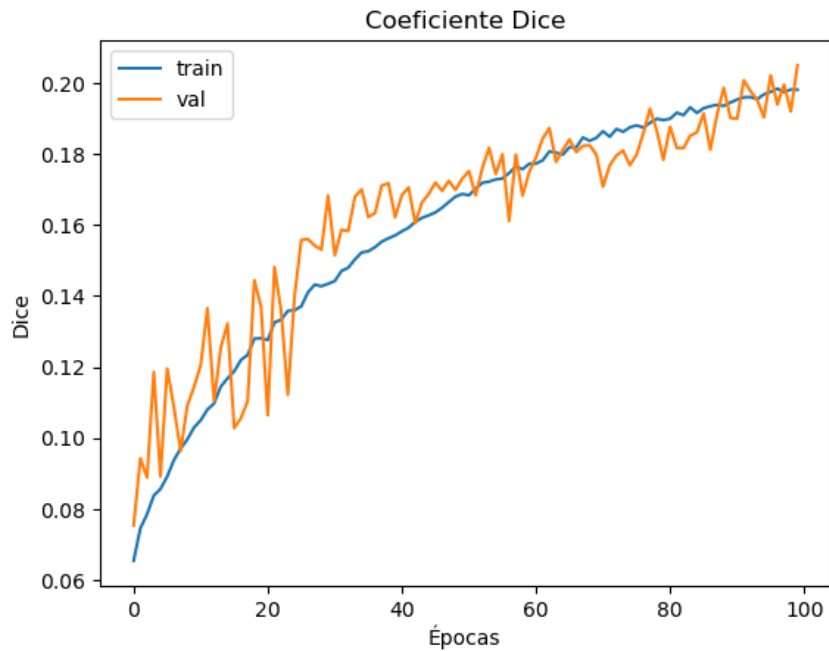


Figura 5.6: Coeficiente Dice da *Attention* U-net com *data augmentation*.

Analisando as Figuras 5.5 a 5.6, percebe-se que o aumento de dados teve um efeito negativo no desempenho de ambas as arquiteturas levando a um coeficiente Dice muito inferior aos resultados obtidos sem nenhum processamento. Uma possível causa seria a ocorrência de *overfitting* devido a duplicação das imagens pois na *Attention* U-net percebe-se uma diferença maior entre o conjunto de validação e treinamento.

### 5.3.3 Dados com realce da imagem

Outro experimento realizado consistiu em utilizar como entrada das redes as imagens com o método de equalização de histograma (Seção 4.2.1). Para analisar o efeito que o aumento no brilho e contraste fizeram nas redes.

Com esse método, a rede neural *Attention* U-net apresenta uma curva crescente até a 100<sup>a</sup> época, isso indica que a rede poderia estar aprendendo quando o treinamento foi encerrado. No que diz respeito a U-net tradicional, os valores do coeficiente Dice foram altos, porém não superior aos valores obtidos na segmentação sem pré-processamento. Logo, a técnica não obteve um aumento no desempenho da rede. Apesar de realçar o brilho e contraste da imagem, o método pode ter sido prejudicado pela presença de ruídos nas imagens o que foi refletido no aprendizado da rede neural.

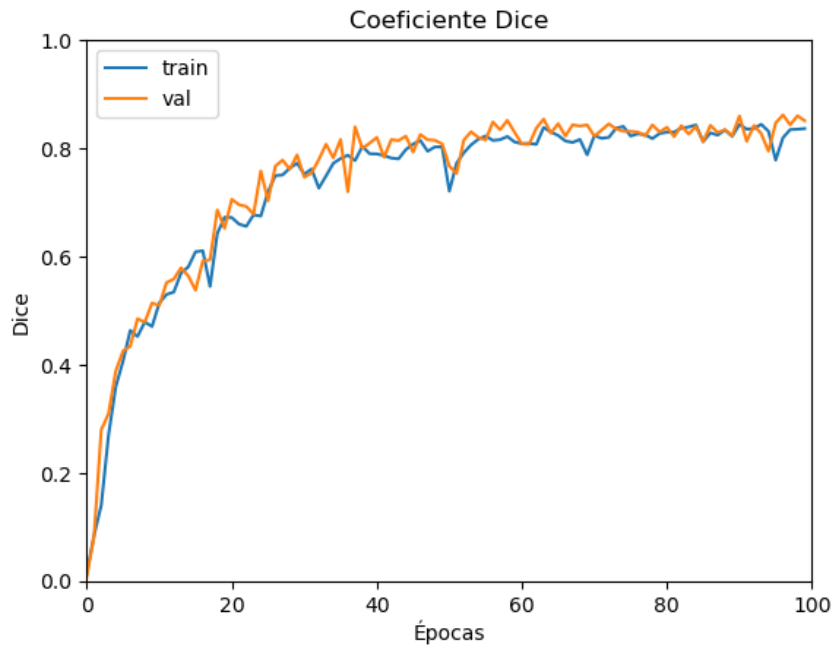


Figura 5.7: Coeficiente Dice da U-net tradicional com realce da imagem.

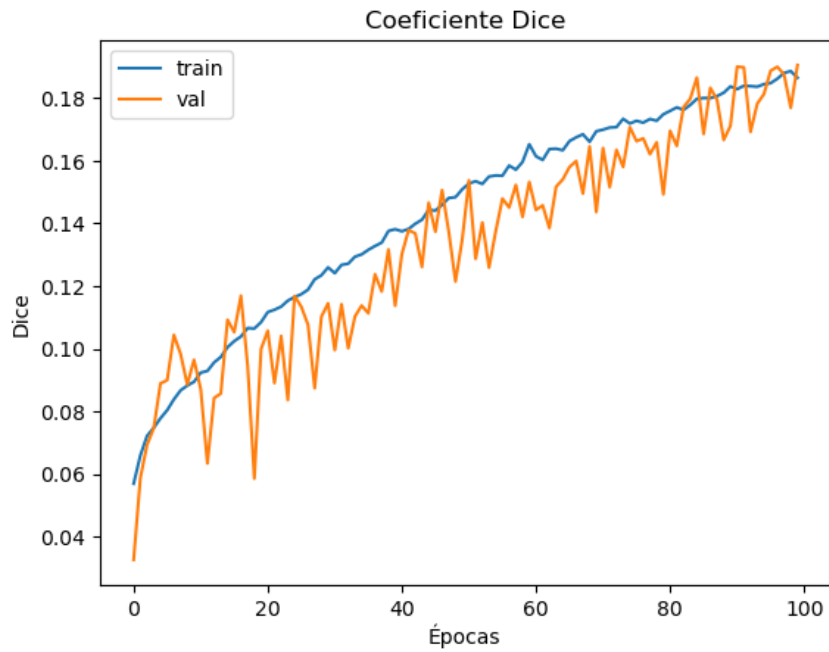


Figura 5.8: Coeficiente Dice da *Attention* U-net com realce da imagem.

## 5.4 Modelo Híbrido

A Tabela 5.1 apresenta a média do coeficiente Dice de todos os experimentos descritos nas seções anteriores. Observa-se na Tabela 5.1 que o melhor resultado foi obtido utilizando

Tabela 5.1: Resultado dos experimentos.

<b>Experimento</b>	<b>Coefficiente Dice</b>	<b>Acurácia</b>
U-net tradicional sem pré-processamento	<b>0,890</b>	<b>0,997</b>
U-net tradicional com realce de imagem	<b>0,871</b>	<b>0,998</b>
U-net tradicional com <i>data augmentation</i>	0,836	0,993
<i>Attention</i> U-net sem pré-processamento	0,760	0,994
<i>Attention</i> U-net com realce da imagem	0,158	0,983
<i>Attention</i> U-net com <i>data augmentation</i>	0,172	0,983
Modelo Chan-Vese com realce de imagem	0,485	0,974
Modelo híbrido	0,690	0,994

a U-net sem nenhum pré-processamento. Vale também destacar o desempenho da U-net com pré-processamento que ficou muito próximo. Outro ponto importante é que em ambas as arquiteturas, U-net tradicional e *Attention* U-net, o uso de *data augmentation* teve um impacto muito negativo, reduzindo o desempenho do modelo drasticamente. O objetivo de melhorar o modelo Chan-Vese com o uso do modelo híbrido proposto foi alcançado, uma vez que houve uma melhoria no modelo, resultando em um coeficiente dice de 0,485 para 0,690. Essa melhora ocorreu devido ao melhor posicionamento do contorno inicial do modelo Chan-Vese, uma vez que o contorno inicial tem possibilidade de estar mais próximo ao ventrículo esquerdo, de acordo com o desempenho da U-net.

Além disso, é importante salientar que conforme visto nas Figuras 5.4, 5.6 e 5.8, a curva da rede *Attention* U-net ainda estava crescente e para obter um melhor resultado uma opção seria aumentar o número de épocas para esta arquitetura, por outro lado, na rede U-net a curva já estava estabilizada ao atingir a 100<sup>a</sup> época.

Portanto, a U-net tradicional sem pré-processamento foi considerada para compor o método híbrida da Figura 4.4, juntamente com a segmentação tradicional pelo modelo Chan-Vese. O método híbrido foi implementado com todos os métodos descritos no Capítulo 4 resultando em um coeficiente Dice de 0,690 com acurácia de 0,994. Podemos observar na Tabela 5.1 que o modelo híbrido teve um melhora no desempenho do modelo de segmentação tradicional Chan-Vese, porém uma piora no desempenho da rede U-net tradicional.

Com o objetivo de ter uma análise visual, foram geradas imagens da saída do método híbrido e comparado com suas respectivas máscaras do conjunto de teste conforme a Figura 5.9. A partir das imagens da Figura 5.9, pode-se observar que, quando o ventrículo esquerdo (objeto de interesse a ser segmentado), é uma parte muito pequena da imagem (linhas 2 e 3) o método híbrido não obteve um bom desempenho, fazendo a segmentação de área vizinha e até mesmo segmentando outra região da imagem, o que ocasionou uma redução na média do coeficiente Dice.



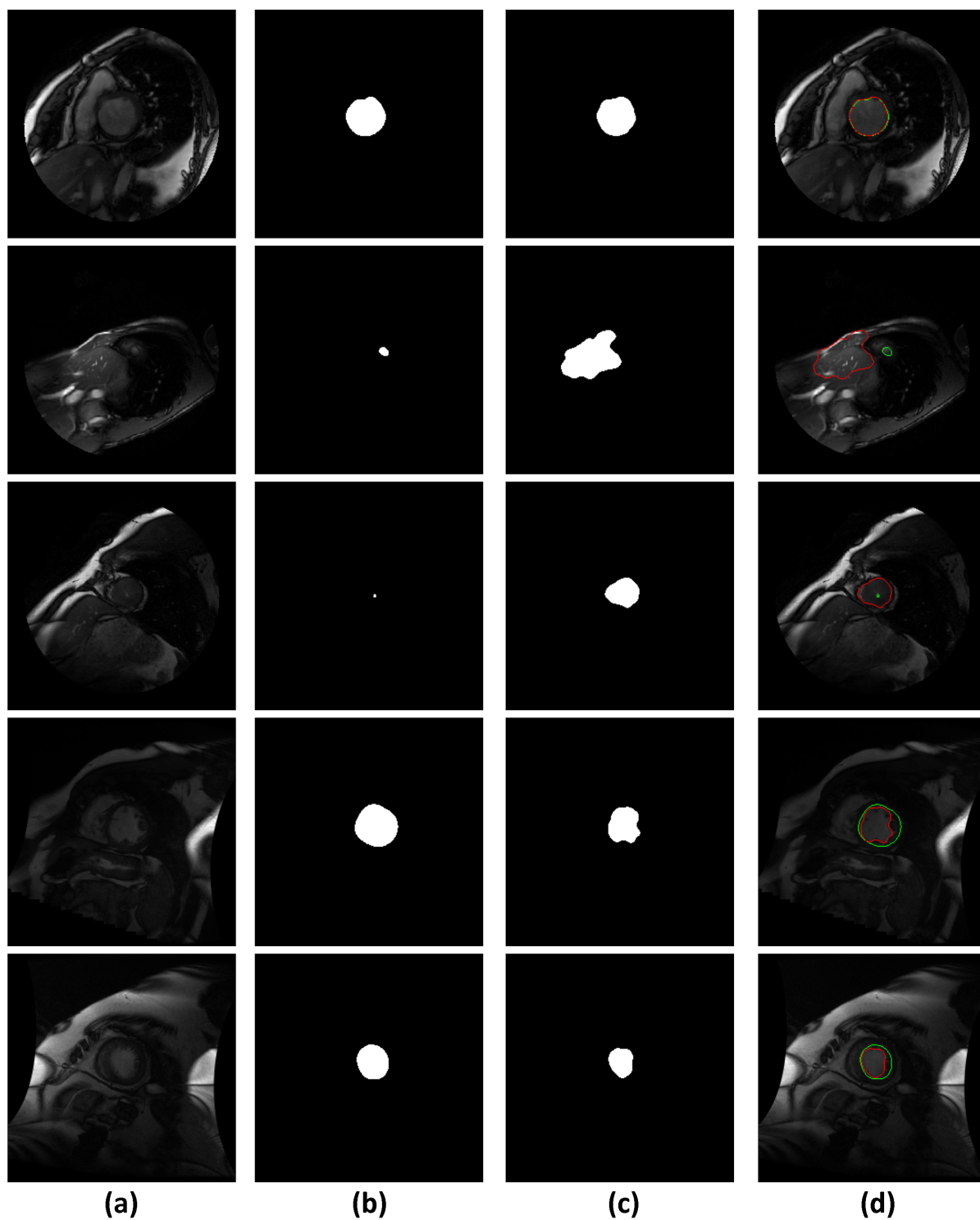


Figura 5.9: Segmentação do método híbrido (a) Imagem original, (b) máscara *ground-truth*, (c) resultado da segmentação, (d) contorno sobreposto à imagem original (verde (b) e vermelho (c)).

# Capítulo 6

## Conclusão

Este trabalho analisou diferentes métodos para auxiliar no diagnóstico de doenças cardíacas em exames de ressonância magnética através da segmentação do ventrículo esquerdo uma vez que esse processo feito manualmente é cansativo e pode levar a falhas. Foi testado um método de segmentação tradicional de contorno ativo denominado Chan-Vese e foram testadas duas arquiteturas de redes neurais: U-net tradicional e *Attention* U-net, sendo cada uma testada usando as imagens de entrada sem pré-processamento, com realce de imagem e com *data augmentation* para assim determinar qual combinação seria mais apropriada a ser utilizada para compor o método híbrido juntamente com a segmentação tradicional Chan-Vese.

A métrica utilizada para tais testes foi o coeficiente Dice, uma vez que a acurácia não representaria o resultado real esperado pois o objeto a ser segmentado representa uma pequena parte da imagem, sendo assim a acurácia sempre seria elevada. O modelo híbrido obteve sucesso em melhorar o modelo Chan-veese resultando em um aumento no coeficiente Dice. Passando de 0,485 no modelo Chan-Vese com realce de imagem para 0,690 no modelo híbrido, entretanto, o melhor resultado foi obtido com a rede neural U-net tradicional sem pré-processamento, resultando em um coeficiente Dice de 0,890. O valor do modelo híbrido se mostrou abaixo, uma vez que em algumas imagens o ventrículo esquerdo era muito pequeno, sendo assim o modelo segmentou outra parte do coração.

A U-net tradicional foi a arquitetura que mais se mostrou consistente, mantendo seu resultado em uma média de 0,8656 em todos os experimentos realizados. Outro ponto importante foi o desempenho da *Attention* U-net nos experimentos utilizando algum método de pré-processamento seu desempenho caiu drasticamente que poderia ter sido melhor caso fossem utilizados um número maior de épocas, pois sua curva ainda estava crescente, mesmo que aos poucos.

Uma proposição de trabalho futuro seria a utilização de técnicas de recorte nas imagens para lidar com o problema do ventrículo ser muito pequeno na imagem de entrada.

Outra proposta seria fazer o uso de classificadores com a saída do modelo para classificar patologias gerando um possível diagnóstico.

# Referências

- [1] Project, Cardiac Atlas: *Sunnybrook cardiac data - cardiac atlas project*. <https://www.cardiacatlas.org/studies/sunnybrook-cardiac-data/>, acesso em 2021-08-07. ix, 21, 24, 25
- [2] Yang, Luren e Fritz Albregtsen: *Fast and exact computation of cartesian geometric moments using discrete green's theorem*. *Pattern Recognition*, 29(7):1061–1073, 1996. x, 27, 28
- [3] Organization, World Health: *The top 10 causes of death*. <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>, acesso em 2021-06-16. 1
- [4] Amaro Júnior, Edson e Helio Yamashita: *Aspectos básicos de tomografia computadorizada e ressonância magnética*. *Brazilian Journal of Psychiatry*, 23:2–3, 2001. 1
- [5] Wong, Kelvin, Zhonghua Sun e Jiyuan Tu: *Medical imaging and computer-aided flow analysis of a heart with atrial septal defect*. *Journal of Mechanics in Medicine and Biology*, 12:1–28, dezembro 2012. 2
- [6] Gonzalez, Rafael e Richard Woods: *Digital Image Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2002. 1, 8
- [7] Chan, T.F. e L.A. Vese: *Active contours without edges*. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001. 1, 8, 9, 20, 25, 29, 30
- [8] Zhou, Yan, Wei Ren Shi, Wei Chen, Yong lin Chen, Ying Li, Li Wen Tan e Dai Qiang Chen: *Active contours driven by localizing region and edge-based intensity fitting energy with application to segmentation of the left ventricle in cardiac ct images*. *Neurocomputing*, 156:199–210, 2015. 2
- [9] Sharma, Kiran, Abeer Alsadoon, PWC Prasad, Thair Al-Dala'in, Tran Quoc Vinh Nguyen e Duong Thu Hang Pham: *A novel solution of using deep learning for left ventricle detection: Enhanced feature extraction*. *Computer Methods and Programs in Biomedicine*, 197:105751, 2020. 2
- [10] Ronneberger, Olaf, Philipp Fischer e Thomas Brox: *U-net: Convolutional networks for biomedical image segmentation*. Em *International Conference on Medical image computing and computer-assisted intervention*, páginas 234–241. Springer, 2015. 2, 14, 18, 25

- [11] Oktay, Ozan, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, Ben Glocker e Daniel Rueckert: *Attention u-net: Learning where to look for the pancreas*, 2018. 2, 15, 16, 20, 25, 26
- [12] Ishida, Masaki, Shingo Kato e Hajime Sakuma: *Cardiac mri in ischemic heart disease*. Circulation Journal, páginas 0908060454–0908060454, 2009. 3
- [13] Câncer, Instituto Nacional de. <https://www.inca.gov.br/exposicao-no-trabalho-e-no-ambiente/radiacoes/radiacoes-ionizantes>, acesso em 2021-11-05. 4
- [14] Ginat, Daniel T., Michael W. Fong, David J. Tuttle, Susan K. Hobbs e Rajashree C. Vyas: *Cardiac imaging: Part 1, mr pulse sequences, imaging planes, and basic anatomy*. American Journal of Roentgenology, 197(4):808–815, 2011. <https://doi.org/10.2214/AJR.10.7231>, PMID: 21940567. 4, 5, 6
- [15] Tortora, Gerard J e Bryan Derrickson: *Corpo Humano-: Fundamentos de Anatomia e Fisiologia*. Artmed Editora, 2016. 5
- [16] Ridgway, John P: *Cardiovascular magnetic resonance physics for clinicians: part i*. Journal of cardiovascular magnetic resonance, 12(1):1–28, 2010. 6
- [17] Mildemberger, Peter, Marco Eichelberg e Eric Martin: *Introduction to the dicom standard*. European radiology, 12(4):920–927, 2002. 6
- [18] Zimmerman, J.B., S.M. Pizer, E.V. Staab, J.R. Perry, W. McCartney e B.C. Brenton: *An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement*. IEEE Transactions on Medical Imaging, 7(4):304–312, 1988. 7
- [19] Eaton-Rosen, Zach, Felix Bragman, Sebastien Ourselin e M Jorge Cardoso: *Improving data augmentation for medical image segmentation*. 2018. 7, 8
- [20] Marques Filho, Ogê e Hugo Vieira Neto: *Processamento digital de imagens*. Brasport, 1999. 7
- [21] *Células nervosas*. <https://www.sobiologia.com.br/conteudos/FisiologiaAnimal/nervoso2.php>, acesso em 2021-10-15. 10
- [22] Haykin, Simon: *Redes neurais: princípios e prática*. Bookman Editora, 2007. 10
- [23] Lecun, Y., L. Bottou, Y. Bengio e P. Haffner: *Gradient-based learning applied to document recognition*. Proceedings of the IEEE, 86(11):2278–2324, 1998. 11
- [24] [https://miro.medium.com/max/628/1\\*0j4LnMvcbPRw2P9rf7TAga.png](https://miro.medium.com/max/628/1*0j4LnMvcbPRw2P9rf7TAga.png), acesso em 2021-10-15. 12
- [25] Long, Jonathan, Evan Shelhamer e Trevor Darrell: *Fully convolutional networks for semantic segmentation*. Em *Proceedings of the IEEE conference on computer vision and pattern recognition*, páginas 3431–3440, 2015. 14

- [26] Silva Neto, Benedito Rodrigues da: *A inteligência artificial na medicina clínica: Principais avanços e aplicações*. A Inteligência Artificial na medicina clínica: Principais Avanços e aplicações, páginas 1–388. 18
- [27] Ammar, Abderazzak, Omar Bouattane e Mohamed Youssfi: *Automatic cardiac cine mri segmentation and heart disease classification*. Computerized Medical Imaging and Graphics, 88:101864, 2021, ISSN 0895-6111. <https://www.sciencedirect.com/science/article/pii/S0895611121000124>. 18, 19
- [28] Zreik, Majd, Nikolas Lessmann, Robbert W van Hamersvelt, Jelmer M Wolterink, Michiel Voskuil, Max A Viergever, Tim Leiner e Ivana Išgum: *Deep learning analysis of the myocardium in coronary ct angiography for identification of patients with functionally significant coronary artery stenosis*. Medical image analysis, 44:72–85, 2018. 18, 19
- [29] Yong, Binbin, Chen Wang, Jun Shen, Fucun Li, Hang Yin e Rui Zhou: *Automatic ventricular nuclear magnetic resonance image processing with deep learning*. Multimedia Tools and Applications, Apr 2020, ISSN 1573-7721. <https://doi.org/10.1007/s11042-020-08911-9>. 19
- [30] Isensee, Fabian, Paul F Jaeger, Peter M Full, Ivo Wolf, Sandy Engelhardt e Klaus H Maier-Hein: *Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features*. Em *International workshop on statistical atlases and computational models of the heart*, páginas 120–129. Springer, 2017. 19
- [31] *Automated cardiac diagnosis challenge*. <https://www.creatis.insa-lyon.fr/Challenge/acdc/>, acesso em 2021-11-07. 19
- [32] Liao, Fangzhou, Xi Chen, XXiaolin Hu e Sen Song: *Estimation of the volume of the left ventricle from mri images using deep neural networks*. IEEE Transactions on Cybernetics, 49(2):495–504, Feb 2019, ISSN 2168-2275. 19
- [33] Chang, Dah Chung e Wen Rong Wu: *Image contrast enhancement based on a histogram transformation of local standard deviation*. IEEE transactions on medical imaging, 17(4):518–531, 1998. 24
- [34] Pisano, Etta D, Shuquan Zong, Bradley M Hemminger, Marla DeLuca, R Eugene Johnston, Keith Muller, M Patricia Braeuning e Stephen M Pizer: *Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms*. Journal of Digital imaging, 11(4):193, 1998. 24