

UNIVERSIDADE FEDERAL DO MARANHÃO - UFMA
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ITALO FRANCYLES SANTOS DA SILVA

**SEGMENTAÇÃO AUTOMÁTICA DE ESTRUTURAS CARDÍACAS EM
IMAGENS DE CINE-RESSONÂNCIA MAGNÉTICA DE EIXO CURTO
USANDO REDES NEURAIAS CONVOLUCIONAIS**

SÃO LUÍS – MA

2023

ITALO FRANCYLES SANTOS DA SILVA

**SEGMENTAÇÃO AUTOMÁTICA DE ESTRUTURAS CARDÍACAS EM
IMAGENS DE CINE-RESSONÂNCIA MAGNÉTICA DE EIXO CURTO
USANDO REDES NEURAIIS CONVOLUCIONAIS**

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da UFMA como parte dos requisitos necessários para obtenção do grau de Doutor em Engenharia Elétrica.

Área de concentração: Ciência da Computação.

Orientador: Prof. Dr. Aristófanés Corrêa Silva

Coorientador: Prof. Dr. Anselmo Cardoso de Paiva

SÃO LUÍS – MA

2023

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

Silva, Italo Francyles Santos da.

Segmentação automática de estruturas cardíacas em imagens de cine-ressonância magnética de eixo curto usando redes neurais convolucionais / Italo Francyles Santos da Silva. - 2023.

126 p.

Coorientador(a): Anselmo Cardoso de Paiva.

Orientador(a): Aristóphanes Corrêa Silva.

Tese (Doutorado) - Programa de Pós-graduação em Engenharia Elétrica/ccet, Universidade Federal do Maranhão, São Luís, Maranhão, 2023.

1. Cine-ressonância magnética. 2. Doenças cardiovasculares. 3. Redes totalmente convolucionais. 4. Segmentação de estruturas cardíacas. I. Paiva, Anselmo Cardoso de. II. Silva, Aristóphanes Corrêa. III. Título.

Tese de autoria de Italo Francyles Santos da Silva, sob o título “**Segmentação automática de estruturas cardíacas em imagens de cine-ressonância magnética de eixo curto usando redes neurais convolucionais**”, apresentado ao Programa de Pós-Graduação em Engenharia Elétrica da UFMA, como parte dos requisitos para obtenção do título de Doutor em Engenharia Elétrica, na área de concentração Ciência da Computação, aprovada em 03 de março de 2023 pela comissão julgadora constituída pelos doutores:

Prof. Dr. Aristófanês Corrêa Silva
Orientador

Prof. Dr. Anselmo Cardoso de Paiva
Coorientador

Prof. Dr. Vicente Leonardo Paucar Casas
Examinador Interno

Prof. Dr. João Dallyson Sousa de Almeida
Examinador Interno

Prof. Dr. Pedro de Alcantara dos Santos Neto
Examinador Externo - Universidade Federal do Piauí (UFPI)

Prof. Dr. João Pedrosa
Examinador Externo - Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência (INESC TEC - Portugal)

À minha família, namorada e amigos

AGRADECIMENTOS

A Deus. Até aqui, sempre me ajudou o Senhor.

À minha família. Em especial, à minha avó, Maria Santana, e à minha mãe, Elza Helena, pelo incentivo a continuar a jornada acadêmica. E também ao meu pai, Paulo Henrique, pelo apoio diuturno.

À minha namorada, Rayana, pelo amor, cumplicidade, e por me ajudar a manter o controle mediante as adversidades.

Ao meu orientador, Aristófanés, e ao meu coorientador, Anselmo, pela confiança, pelo direcionamento, pelas oportunidades oferecidas, e, sobretudo, pela amizade.

À professora Simara e aos professores Geraldo e João Dallyson pela amizade e ensinamentos. Inspirações no âmbito pessoal e profissional.

À UFMA e ao PPGEE pela oportunidade que me foi concedida.

À CAPES pelo importante apoio financeiro durante o doutorado.

A todos que me acompanharam nesta jornada, e que contribuíram direta ou indiretamente em minha formação acadêmica.

*“Eu não tenho nada pra dizer, por isso eu digo.
Eu não tenho muito o que perder, por isso jogo.
Eu não tenho hora pra morrer, por isso sonho.”*

Rita Lee

RESUMO

Doenças cardiovasculares são responsáveis por milhões de óbitos todos os anos. Nesse cenário, exames não invasivos como o de cine-ressonância magnética (cine-RM) tem favorecido uma melhor compreensão acerca dessas patologias, auxiliando no diagnóstico precoce e tratamentos prévios essenciais para melhorar a qualidade de vida dos indivíduos. Por meio desse exame, os especialistas conseguem obter dados mais precisos das estruturas cardíacas, entre elas o miocárdio (Mio), a cavidade ventricular esquerda (CV) e o ventrículo direito (VD). Dado esse contexto, este trabalho apresenta métodos automáticos para a segmentação dessas estruturas cardíacas em imagens de cine-RM. São propostos dois métodos. O primeiro, denominado de Segmentação em Cascata com Reconstrução, é dividido em três etapas principais. A primeira consiste na extração de uma região de interesse (ROI) para reduzir o escopo de processamento. A segunda aplica uma rede totalmente convolucional (FCN) proposta para gerar a segmentação inicial do Mio, CV e VD. Essas segmentações iniciais são passadas para a terceira etapa, chamada de refinamento, na qual é usado um módulo de reconstrução de máscaras baseado em U-Net capaz de regenerar as segmentações geradas. Além disso, nessa etapa também são aplicadas algumas técnicas de pós-processamento específicas para cada estrutura de interesse. O segundo método, chamado de Segmentação Especializada por Contexto, é similarmente dividido em três etapas, sendo que a primeira delas também foca na extração de uma ROI; a segunda etapa compreende numa combinação de FCNs para a segmentação do Mio e da CV; e a terceira etapa utiliza outra FCN proposta, baseada na X-Net, para a segmentação do VD. Os métodos desenvolvidos alcançam resultados promissores em testes com as bases disponibilizadas pelo desafio ACDC, tanto em âmbito local, quanto na avaliação feita pela plataforma do próprio desafio, na qual os métodos apresentam resultados aproximados das melhores abordagens. Resultados expressivos também são obtidos em testes com a base M&Ms, que contém mais exames que a ACDC, sendo estes obtidos por meio de *scanners* de diferentes fabricantes, tendo portanto uma maior variabilidade de características, de modo semelhante ao cenário real da aplicação.

Palavras-chave: Doenças cardiovasculares, Segmentação de estruturas cardíacas, Redes totalmente convolucionais, Cine-ressonância magnética.

ABSTRACT

Cardiovascular diseases are responsible for millions of deaths every year. In this scenario, non-invasive exams such as cine-magnetic resonance imaging (cine-RM) have favored a better understanding of these pathologies, helping early diagnosis and previous treatments essential to improve the quality of life of individuals. Through this exam, specialists can obtain more accurate information about cardiac structures, including the myocardium (Myo), the left ventricular cavity (LVC) and the right ventricle (RV). Given this context, this work presents automatic methods for the segmentation of those cardiac structures in short-axis cine-MRI images. Two methods are proposed. The first, called Cascaded Segmentation with Reconstruction, is divided into three main steps. The first step consists in extracting a region of interest (ROI) to reduce the scope of processing. The second applies a fully convolutional network (FCN) proposed to generate the initial Myo, LVC and RV segmentations. These initial segmentations are passed on to the third step, called refinement, in which a mask reconstruction module based on U-Net is used to restore the generated segmentations. In addition, in this step some specific post-processing techniques are also applied for each structure of interest. The second method, called Specialized Segmentation by Context, is similarly divided into three steps, the first of which also focuses on extracting an ROI; the second step comprises a combination of FCNs for Myo and LCV segmentation; and the third step uses another proposed FCN, based on X-Net, for RV segmentation. The methods developed reach promising results in tests with the dataset made available by the ACDC challenge, both at the local level and in the evaluation made by the challenge's online platform, in which the proposed methods present results with little difference to the best approaches. Expressive results are also obtained in tests with the M&Ms dataset, which contains more exams than the ACDC, which are acquired by scanners from different vendors, therefore having a greater variability of features, similar to the real application scenario.

Keywords: Cardiovascular diseases, Cardiac structures segmentation, Fully convolutional network, Cine-MRI.

LISTA DE ILUSTRAÇÕES

Figura 1 – O Coração Humano: Visão externa (A) e interna (B).	33
Figura 2 – Sequências do exame de cine-RM e categorização dos seguintes planos anatômicos: (A) Coronal; (B) Sagital; (C) Axial; (D) Visão de duas câmaras; (E) Visão de quatro câmaras; (F) Planejamento de aquisição do eixo curto; e (G) Eixo curto.	35
Figura 3 – Exemplo da operação de dilatação sobre uma imagem (A) resultando em (B). O elemento estruturante (E) é um quadrado com dimensões 10 x 10.	38
Figura 4 – Exemplo da operação de erosão sobre uma imagem (A) resultando em (B). O elemento estruturante (E) é circular e tem dimensões 15 x 15.	39
Figura 5 – Operações aplicadas sobre a imagem (A). Abertura resulta em (B). Fechamento, em (C). O elemento estruturante (E) é um quadrado de dimensões 15 x 15.	40
Figura 6 – Arquitetura básica de uma CNN contendo convoluções, subamostragens, mapas de característica e a camada completamente conectada.	42
Figura 7 – Camadas de convolução e subamostragem.	42
Figura 8 – Ilustração da arquitetura de uma rede completamente convolucional.	43
Figura 9 – Ilustração da arquitetura de uma rede U-Net.	44
Figura 10 – Ilustração da arquitetura de uma rede X-Net.	45
Figura 11 – Bloco MBCConv, com uma visão ampliada do bloco Squeeze-and-Excitation (SAE).	47
Figura 12 – Bloco Attention.	48
Figura 13 – Bloco <i>Inception</i>	49
Figura 14 – Visão geral do método SCR e suas etapas: extração da ROI (1), segmentação inicial (2) e refinamento (3).	53
Figura 15 – Arquitetura U-Net utilizada para a geração das segmentações de referência.	55
Figura 16 – Visibilidade das estruturas de interesse nas fatias basal, medial e apical.	56
Figura 17 – EAIS-Net: FCN proposta para a segmentação inicial das estruturas cardíacas no método SCR.	57
Figura 18 – Estruturas internas do Decoder Block proposto.	58

Figura 19 – Visão geral do processo de reconstrução.	60
Figura 20 – Refinamento de uma segmentação inicial da CV através do pós-processamento.	61
Figura 21 – Exemplos de alongamentos e descontinuidades.	61
Figura 22 – Eliminação de alongamentos: exemplo de fatia recortada (A), <i>ground-truth</i> (B), máscara antes (C) e depois (D) da eliminação de alongamentos.	62
Figura 23 – Correção de descontinuidades: máscara de entrada (A), detecção de circunferência (B), máscara transformada para o espaço polar (C), detecção de descontinuidades (D), conexão entre pontos (E) e resultado final (F).	62
Figura 24 – Pós-processamento de máscaras VD: exemplo de fatia recortada (A), seu <i>ground-truth</i> (B), máscara antes (C) e depois (D) o pós-processamento.	63
Figura 25 – Exemplos do dataset ACDC com suas respectivas marcações.	65
Figura 26 – Exemplo de reconstrução do miocárdio: (A) fatia de entrada, (B) <i>ground-truth</i> , (C) segmentação inicial com Dice de 0,7368 e (D) segmentação após o módulo de reconstrução, com Dice de 0,7528.	73
Figura 27 – Resultados qualitativos do método SCR, sendo (A) a entrada, (B) o <i>ground-truth</i> , (C) a segmentação inicial e (D) a segmentação final.	76
Figura 28 – Quantidade de exames por tipo de scanner contida nos conjuntos de treinamento, validação e teste da base M&Ms.	77
Figura 29 – Base M&Ms: exemplos de fatias mediais obtidas pelos diferentes scanners utilizados na elaboração da base.	77
Figura 30 – Resultados qualitativos do método SCR nos experimentos com a base M&Ms, sendo (A) a entrada, (B) o <i>ground-truth</i> , (C) o resultado do experimento SCR_ACDC e (D) o resultado do experimento SCR_MMS.	79
Figura 31 – Visão geral das etapas do método SEC: (1) Extração da ROI, (2) Segmentação do Mio e CV, e (3) Segmentação do VD.	82
Figura 32 – Processo de agregação de máscaras para a extração de uma nova ROI.	84
Figura 33 – Arquitetura da IRAX-Net proposta para a segmentação do VD no método SEC.	86
Figura 34 – Resultados qualitativos do método SEC aplicado em fatias basais, mediais e apicais, sendo (A) a entrada, (B) o <i>ground-truth</i> , e (C) o resultado final.	91

Figura 35 – Exemplos de casos de erros de segmentação produzidos pelo método SEC em fatias na fase de ES, sendo (A) a entrada, (B) o <i>ground-truth</i> , e (C) o resultado final.	92
Figura 36 – Resultados qualitativos do método SEC nos experimentos com a base M&Ms, sendo (A) a entrada, (B) o <i>ground-truth</i> , (C) o resultado do experimento SEC_ACDC e (D) o resultado do experimento SEC_MMS.	97
Figura 37 – Exemplo comparativo entre os métodos SCR e SEC no caso da segmentação do VD, sendo (A) a entrada, (B) o <i>ground-truth</i> , (C) o resultado gerado pelo método SCR, e (D) o resultado do método SEC.	100
Figura 38 – Casos em que o método SCR apresenta mais falhas que o método SEC, sendo (A) a entrada, (B) o <i>ground-truth</i> , (C) o resultado gerado pelo método SCR, e (D) o resultado do método SEC.	101
Figura 39 – Exemplo comparativo entre os métodos SCR e SEC, com a segmentação das três estruturas de interesse. Neste caso, (A) é a entrada, (B) é o <i>ground-truth</i> , (C) é o resultado gerado pelo método SCR, e (D) é o resultado do método SEC.	102
Figura 40 – Exemplo comparativo entre os métodos SCR e SEC, com foco na segmentação do Mio e da CV. Neste caso, (A) é a entrada, (B) é o <i>ground-truth</i> , (C) é o resultado gerado pelo método SCR, e (D) é o resultado do método SEC.	102
Figura 41 – Exemplo comparativo entre os métodos SCR e SEC no experimento com a base M&Ms, sendo (A) a entrada, (B) o <i>ground-truth</i> , (C) o resultado gerado pelo método SCR, e (D) o resultado do método SEC.	108
Figura 42 – Outro exemplo comparativo dos métodos desenvolvidos no experimento com a base M&Ms, sendo (A) a entrada, (B) o <i>ground-truth</i> , (C) o resultado gerado pelo método SCR, e (D) o resultado do método SEC.	108
Figura 43 – Exemplo de falha de ambos os métodos, em que (A) representa a entrada, (B) o <i>ground-truth</i> , (C) o resultado gerado pelo método SCR, e (D) o resultado do método SEC.	109

LISTA DE TABELAS

Tabela 1 – Visão geral dos trabalhos relacionados	32
Tabela 2 – Comparação entre os experimentos realizados para a Extração da ROI	66
Tabela 3 – Resultados de MAE e mAP para a etapa de extração da ROI.	66
Tabela 4 – Segmentação das estruturas cardíacas: cavidade ventricular (CV), mio- cárdio (Mio) e ventrículo direito (VD). Comparação entre os resultados obtidos pela EAIS-Net e por outros métodos de segmentação.	69
Tabela 5 – Análise comparativa entre abordagens de uso da EAIS-Net.	70
Tabela 6 – Experimentos com a EAIS-Net em dois conjuntos de teste: (C1) com as ROIs extraídas manualmente e (C2) ROIs extraídas pela etapa 1 do método SCR.	70
Tabela 7 – Resultados dos testes de hipóteses: valores p encontrados para as métricas Dice, IoU, SEN, e PRC obtidos pelas abordagens semi- supervisionada e completamente automática. Nível $\alpha = 0,05$	71
Tabela 8 – Análise comparativa entre o método SCR (A) que faz segmentação dividida em etapas e métodos de segmentação em etapa única via EAIS-Net (B) e U-Net (C).	71
Tabela 9 – Resultados dos experimentos realizados com diferentes métodos para o módulo de reconstrução.	72
Tabela 10 – Resultados alcançados pelo método SCR em dois cenários: execução das etapas 1 e 2; e a execução completa.	74
Tabela 11 – Avaliação do método SCR a partir das métricas Dice e IoU para as fatias nas fases de fim da diástole (ED) e fim da sístole (ES).	74
Tabela 12 – Método SCR: resultados dos experimentos realizados com a base M&Ms.	78
Tabela 13 – Resultados da aplicação do método SCR na base M&Ms separados por fase cardíaca.	79
Tabela 14 – Valores p encontrados para as métricas de avaliação do método SCR nos experimentos SCR_ACDC e SCR_MMS.	80
Tabela 15 – Comparação entre os resultados obtidos nos experimentos para a seg- mentação inicial do Mio e da CV.	87
Tabela 16 – Comparação entre os resultados obtidos nos experimentos para a seg- mentação final do Mio e da CV.	88

Tabela 17 – Análise comparativa entre as redes usadas no processo de segmentação do VD.	89
Tabela 18 – Abordagens escolhidas para compor o método SEC e seus respectivos resultados de segmentação da CV, Mio, e VD.	90
Tabela 19 – Comparativo entre os resultados obtidos pelo método SEC e outras abordagens de segmentação de etapa única.	93
Tabela 20 – Experimentos realizados com o método SEC seguindo duas abordagens: a extração semi-supervisionada da ROI, e a proposta completamente automatizada.	94
Tabela 21 – Valores p encontrados para as métricas Dice e IoU resultantes das abordagens semi-supervisionada e automática de extração da ROI no método SEC.	94
Tabela 22 – Método SEC: resultados dos experimentos realizados com a base M&Ms.	95
Tabela 23 – Resultados da aplicação do método SEC na base M&Ms separados por fase cardíaca.	96
Tabela 24 – Valores p encontrados para as métricas de avaliação do método SEC nos experimentos SEC_ACDC e SEC_MMS.	97
Tabela 25 – Base ACDC: resultados finais obtidos pelos métodos propostos. . . .	99
Tabela 26 – Base ACDC: resultados finais separados por fase cardíaca (ED e ES).	103
Tabela 27 – Resultados dos testes de hipóteses: valores p encontrados para as métricas Dice, IoU, SEN, e PRC produzidos pelos métodos SCR e SEC para cada estrutura de interesse. Nível $\alpha = 0,05$	103
Tabela 28 – Resultados dos métodos propostos para a segmentação da cavidade ventricular (CV) na base de teste e comparação com os melhores métodos elencados no desafio ACDC.	104
Tabela 29 – Resultados dos métodos propostos para a segmentação do miocárdio (Mio) na base de teste e comparação com os melhores métodos elencados no desafio ACDC.	105
Tabela 30 – Resultados dos métodos propostos para a segmentação do ventrículo direito (VD) na base de teste e comparação com os melhores métodos elencados no desafio ACDC.	105
Tabela 31 – Base M&Ms: resultados finais obtidos pelos métodos propostos. . . .	107

Tabela 32 – Base M&Ms: resultados finais dos métodos SCR e SEC separados por fase cardíaca (ED e ES).	109
Tabela 33 – Resultados dos testes de hipóteses aplicados na comparação entre os métodos SCR e SEC nos experimentos com a base M&Ms.	110
Tabela 34 – Resultados de Dice Geral (G) e por fase cardíaca (ED e ES) dos métodos propostos para a segmentação da CV, Mio, e VD na base de teste e comparação com os trabalhos relacionados.	111
Tabela 35 – Artigo relacionado à segmentação de estruturas cardíacas.	118
Tabela 36 – Publicações envolvendo processamento de imagens e visão computacional.	118

LISTA DE ABREVIATURAS E SIGLAS

CAD	<i>Computer-Aided Detection</i>
CADx	<i>Computer-Aided Diagnosis</i>
CNN	<i>Convolutional Neural Network</i>
CV	Cavidade Ventricular Esquerda
DCV	Doenças cardiovasculares
ED	<i>End Diastole</i> - Fim da Diástole
ES	<i>End Systole</i> - Fim da Sístole
FCN	<i>Fully Convolutional Network</i>
FN	Falso Negativo
FP	Falso Positivo
HD	Distância de Hausdoff
IOU	<i>Intersection Over Union</i>
Mio	Miocárdio
PRC	Precisão
RNA	Rede Neural Artificial
ROI	Região de Interesse
RM	Ressonância Magnética
SCR	Segmentação em Cascata com Reconstrução
SEC	Segmentação Especializada por Contexto
SEN	Sensibilidade
VD	Ventrículo Direito
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

SUMÁRIO

1	INTRODUÇÃO	19
1.1	Hipóteses do Trabalho	21
1.2	Objetivos	22
1.3	Contribuições	23
1.4	Organização do Trabalho	24
2	TRABALHOS RELACIONADOS	25
2.1	Base <i>Sunnybrook Cardiac Dataset</i>	25
2.2	Base LVSC	26
2.3	Bases Privadas	26
2.4	Base ACDC	27
2.5	Base M&Ms	29
2.6	Considerações Finais	30
3	FUNDAMENTOS TEÓRICOS	33
3.1	O Coração Humano	33
3.2	Cine-Ressonância Magnética Cardíaca	34
3.3	Processamento de Imagens Digitais	36
3.3.1	Suavização e Normalização	37
3.3.2	Morfologia Matemática	37
3.4	Aprendizado Profundo	40
3.4.1	Redes Neurais Convolucionais	41
3.4.2	Redes Totalmente Convolucionais	42
3.4.3	EfficientNet	45
3.4.4	Arquiteturas de Blocos Convolucionais	47
3.4.4.1	Bloco <i>Attention</i>	47
3.4.4.2	Blocos <i>Inception</i>	48
3.5	Métricas de Avaliação	49
3.6	Testes de Hipóteses	51
3.7	Considerações Finais	52

4	MÉTODO DE SEGMENTAÇÃO EM CASCATA COM RE- CONSTRUÇÃO	53
4.1	Extração da ROI	54
4.2	Segmentação Inicial	56
4.3	Refinamento	58
4.3.1	Módulo de Reconstrução	59
4.3.2	Pós-Processamento	60
4.4	Experimentos e Resultados	63
4.4.1	Base de Imagens ACDC	64
4.4.2	Resultados com a Base ACDC	65
4.4.2.1	Extração da ROI	65
4.4.2.2	Segmentação Inicial	67
4.4.2.3	Refinamento: Reconstrução e Pós-Processamentos	72
4.4.3	Base de Imagens M&Ms	76
4.4.4	Resultados com a Base M&Ms	78
4.5	Considerações Finais	81
5	MÉTODO DE SEGMENTAÇÃO ESPECIALIZADA POR CONTEXTO	82
5.1	Segmentação do Mio e CV	83
5.1.1	Segmentação Inicial	83
5.1.2	Agregação de Máscaras para a Extração de uma Nova ROI	84
5.1.3	Segmentação Final do Mio e da CV	85
5.2	Segmentação do VD	85
5.3	Experimentos e Resultados	86
5.3.1	Resultados com a base ACDC	86
5.3.1.1	Segmentação do Mio e da CV	87
5.3.1.2	Segmentação do VD	89
5.3.1.3	Resultado Final e Estudos de Caso	90
5.3.1.4	Análise de Impacto na Segmentação	93
5.3.2	Resultados com a Base M&Ms	94
5.4	Considerações Finais	98

6	ANÁLISE COMPARATIVA DOS MÉTODOS PROPOSTOS	99
6.1	Base ACDC	99
6.1.1	Conjunto Local de Testes	99
6.1.2	Avaliação na Plataforma Online e Comparação com Trabalhos Relacionados	104
6.2	Base M&Ms	106
6.2.1	Comparação entre os Métodos SCR e SEC	107
6.2.2	Comparação com os Trabalhos Relacionados	111
6.3	Discussão	113
7	CONCLUSÃO	116
7.1	Trabalhos Futuros	117
7.2	Produções Científicas	117
	REFERÊNCIAS	119

1 INTRODUÇÃO

As doenças cardiovasculares (DCV) correspondem a um grupo de problemas relacionados ao coração e aos vasos sanguíneos. Elas surgem conforme a idade e estão ligadas a uma série de fatores como maus hábitos alimentares, sedentarismo, ou até mesmo fatores genéticos. No mundo, as DCV são responsáveis por 17,5 milhões de óbitos por ano (HAZRA et al., 2017). No Brasil, segundo a Sociedade Brasileira de Cardiologia (SBC), mais de 336 mil pessoas faleceram entre os meses de janeiro e novembro de 2022 por causa de DCV (SBC, 2022). Só no Estado do Maranhão, entre 2016 e 2019, foram registrados um total de 42.558 óbitos por DCV entre a população, o que indica um elevado índice de mortalidade (NETO et al., 2022). Essas constatações mostram a necessidade de ações estratégicas para conter o agravamento desse cenário, visto que, quanto mais cedo as DCV forem diagnosticadas e tratadas, maiores serão as chances de melhorar a qualidade de vida dos indivíduos.

O avanço contínuo de tecnologias tem favorecido uma melhor compreensão acerca das patologias que afetam o ser humano, principalmente por meio de exames não invasivos como o de ressonância magnética (RM) (MURDOCH; DETSKY, 2013). No contexto cardíaco, esse exame também é chamado de cine-ressonância magnética (cine-RM), e é considerado um dos métodos de maior segurança para o diagnóstico cardiovascular através da análise da anatomia e morfologia do coração, assim como para a identificação de DCV de natureza congênita (AZIZ et al., 2013; SARA et al., 2014).

Através da cine-RM, os especialistas podem analisar o funcionamento das estruturas cardíacas, entre elas o miocárdio (Mio), a cavidade ventricular esquerda (CV), e ventrículo direito (VD), bem como quantificar valores de massa e volume. Esse passo é importante também no acompanhamento pós-operatório, com a avaliação das funções cardíacas e da anatomia ventricular (SECHTEM et al., 1987; MYERSON et al., 2002). O cálculo desses valores requer, a princípio, a delimitação dos contornos internos e externos dessas estruturas ao longo das fatias do exame (LI et al., 2010). No entanto, esse processo é bastante demorado quando feito manualmente pelo especialista. Além disso, pode ser fadigante, tornando-se suscetível a erros que podem levar a um diagnóstico impreciso.

Em processamento de imagens, a delimitação de uma região de interesse é chamada de segmentação. No contexto da identificação de DCVs, métodos computacionais podem

auxiliar os especialistas na avaliação das estruturas cardíacas provendo uma segmentação automática, de modo que a análise dos exames seja mais célere.

Dentre esses métodos, é possível destacar aqueles baseados em Aprendizado Profundo, como as redes totalmente convolucionais (FCN), que têm apresentado resultados expressivos em diversas aplicações, como identificação de objetos em imagem e vídeo (LONG et al., 2015; CAELLES et al., 2017) e também na segmentação de imagens médicas (RONNEBERGER et al., 2015; WANG; LIU, 2021). As FCN têm como fundamentos a extração de características do nível mais amplo ao mais profundo, analisando padrões latentes para, a partir disso, identificar quais pixels fazem parte das regiões de interesse.

Ainda no âmbito das imagens médicas, essas técnicas podem ser aplicadas em exames com visualização em duas ou mais dimensões. No entanto, dependendo da solução proposta, o processamento desses exames pode demandar alto custo computacional. Por isso, desenvolver um método preciso e computacionalmente viável torna-se uma tarefa desafiadora.

Nesse cenário, o presente trabalho apresenta dois métodos para a segmentação de imagens de cine-RM cardíaca de eixo curto. Esses métodos têm como característica comum a divisão do processo de segmentação em etapas. O primeiro método apresenta uma abordagem de segmentação em cascata com reconstrução. Ele se inicia com a extração da região de interesse (ROI), seguida pela segmentação inicial e, depois, o refinamento. Na primeira etapa, é usada uma abordagem baseada em U-Net (RONNEBERGER et al., 2015) para localizar a ROI na fatia original. Na segunda, a ROI extraída é submetida a uma FCN proposta, chamada EAIS-Net, que usa a Efficient-Net (TAN; LE, 2019), blocos Attention (OKTAY et al., 2018), Inception (SZEGEDY et al., 2015), e Squeeze-and-Excitation (HU et al., 2018). A EAIS-Net gera uma segmentação inicial que é passada para a terceira etapa, responsável por melhorar os resultados alcançados através de pós-processamentos, entre eles um módulo de reconstrução de máscaras baseado em U-Net.

O segundo método segue uma abordagem de segmentação especializada por contexto. O referido método também se inicia com uma etapa de extração da ROI. Em seguida, as ROIs extraídas são submetidas a outras duas etapas, cada uma relacionada a um contexto definido: segmentação do Mio e da CV; e segmentação do VD. Na primeira delas, a ROI extraída é passada para a EAIS-Net a fim de gerar uma segmentação inicial para o Mio e a CV. Esse resultado é utilizado para gerar uma nova ROI, que serve de entrada para uma segunda U-Net, responsável por produzir a segmentação final dessas

estruturas. Na etapa de segmentação do VD, a ROI extraída é passada a uma outra FCN proposta, chamada IRAX-Net, para segmentar o VD. Essa rede, por sua vez, é baseada na X-Net (BULLOCK et al., 2019). Por fim, para ambos os métodos, o resultado final é a combinação das segmentações geradas para o ventrículo direito, esquerdo e o miocárdio.

1.1 Hipóteses do Trabalho

Conforme mencionado anteriormente, o exame de cine-RM é importante para o diagnóstico de DCVs bem como faz parte de um protocolo de acompanhamento de pacientes em tratamento. E a segmentação do miocárdio, cavidade ventricular esquerda e do ventrículo direito é uma tarefa importante para fornecer dados necessários para a assistência médica. Essa tarefa é desafiadora tanto no contexto da aquisição das imagens, quanto na sua composição, uma vez que as estruturas de interesse apresentam variações não só ao longo do volume, mas no decorrer do tempo em virtude das fases cardíacas de sístole e diástole. Além disso, há o desbalanceamento entre as classes de pixels correspondentes às estruturas de interesse e às demais partes aparentes nas fatias do exame. Com base nessas observações, são consideradas as seguintes hipóteses:

Hipótese 1: A segmentação das estruturas cardíacas através de uma abordagem em cascata apresenta melhor resultado comparada ao processo em passo único.

Essa hipótese surge com base nas características observadas nos exames de cine-RM, tais quais o amplo desbalanceamento entre as classes de pixels das estruturas de interesse e do resto do exame; além das alterações de forma e textura dessas estruturas em virtude dos movimentos cardíacos (ABDELTAWAB et al., 2020). Para lidar com essas características, um processo de segmentação em passo único deve ser composto por uma técnica robusta, capaz de produzir resultados precisos. Uma abordagem em cascata, por sua vez, divide o processo em etapas menores, podendo ser constituída por técnicas mais simples cuja elaboração considera as especificidades de cada etapa. Com isso, há a possibilidade de que sejam obtidos resultados melhores em comparação com o processo em passo único.

Hipótese 2: O uso de mecanismos baseados em conexões densas, ponderação de mapas de característica e análise multi-escala na composição de uma rede totalmente convolucional implica na obtenção de resultados melhores em comparação aos mecanismos de convolução tradicionais.

Conforme mencionado, os exames de cine-RM cardíaca apresentam uma certa complexidade uma vez que as características das estruturas de interesse se alteram por causa dos movimentos cardíacos. Com base nisso, pressupõe-se que mecanismos tradicionais de convolução podem não ser suficientes para a obtenção de resultados promissores, sendo, portanto, necessária a utilização de mecanismos mais robustos cujo desempenho tem sido exitoso em tarefas de classificação e segmentação, como as conexões densas (TAN; LE, 2019), a análise multi-escala (SZEGEDY et al., 2015) e a ponderação de mapas de características (OKTAY et al., 2018; HU et al., 2018).

Hipótese 3: O uso de um módulo de reconstrução de máscaras é relevante para alcançar melhores resultados.

O módulo de reconstrução de máscaras integrado a uma etapa de refinamento pode garantir a correção da segmentação gerada inicialmente, ocasionando melhoramentos aos resultados.

Hipótese 4: Abordagens semi-supervisionadas para a extração de uma ROI como um passo inicial dos métodos propostos alcançarão resultados melhores do que abordagens automáticas.

Devido ao conhecimento prévio, a intervenção humana em etapas iniciais dos métodos propostos pode acarretar resultados significativamente melhores em comparação às suas execuções de modo totalmente automático.

Hipótese 5: Um cenário com uma elevada quantidade de exames obtidos por meio de *scanners* de diferentes fabricantes, pode acarretar perdas para os resultados de segmentação.

Uma simulação possível do cenário real de aplicação ocorre por meio de experimentos com um conjunto de exames obtidos por diferentes *scanners*, cada um com características próprias de aquisição das imagens de cine-RM. Essa variabilidade pode impactar negativamente os resultados da segmentação, ainda que essa permaneça em um nível satisfatório.

1.2 Objetivos

O objetivo geral deste trabalho é propor métodos para a segmentação das estruturas cardíacas (miocárdio, cavidade ventricular esquerda e ventrículo direito) em imagens de cine-ressonância magnética de eixo curto, dividindo esse processo em etapas e usando

combinações de redes neurais convolucionais.

A realização do objetivo geral é possibilitada com o cumprimento dos seguintes objetivos específicos:

- Desenvolver um método computacional para a localização e extração de uma região de interesse (ROI);
- Desenvolver um método computacional para geração de mapas de segmentação iniciais das estruturas de interesse;
- Desenvolver processos de refinamento capazes de corrigir falhas presentes nas segmentações iniciais, melhorando os resultados;
- Realizar uma análise comparativa entre o desempenho de arquiteturas de rede propostas e outras utilizadas na literatura no âmbito da segmentação das estruturas cardíacas.
- Fazer uma análise comparativa entre os métodos propostos para identificar pontos fortes e limitações no contexto de sua aplicabilidade.

1.3 Contribuições

Este trabalho apresenta algumas contribuições, dentre as quais podem ser destacadas:

- A proposição de métodos para a segmentação das estruturas cardíacas, que dividem esse processo em etapas, sendo que cada uma delas usa técnicas específicas conforme características do domínio da aplicação;
- Uma abordagem para a localização e extração da ROI baseada em U-Net que busca a redução de escopo e processamento;
- A proposta de uma arquitetura de FCN, chamada de EAIS-Net, que combina Efficient-Net, Blocos Inception, Blocos Attention, Blocos Squeeze-and-Excitation e conexões salteadas;
- A reconstrução baseada em U-Net para o refinamento das segmentações iniciais das estruturas cardíacas.

- A proposta de uma arquitetura de FCN, chamada de IRAX-Net, composta por estruturas da X-Net combinada com Blocos Residuais Inception e Attention;

1.4 Organização do Trabalho

Este trabalho está organizado da seguinte forma: no **Capítulo 2** são descritos os trabalhos relacionados ao tema investigado; no **Capítulo 3**, são apresentados os fundamentos teóricos importantes para a compreensão do presente trabalho; no **Capítulo 4**, é apresentado o método de segmentação em cascata com reconstrução, e os resultados obtidos por este; no **Capítulo 5**, é discorrido sobre o método de segmentação especializada por contexto, bem como a sua avaliação por meio dos experimentos; no **Capítulo 6**, é apresentada uma análise comparativa dos métodos propostos, confrontando-os, também, com os trabalhos relacionados; e, por fim, no **Capítulo 7**, são relatadas as conclusões obtidas, as contribuições da pesquisa, bem como sugestões para trabalhos futuros.

2 TRABALHOS RELACIONADOS

Este capítulo apresenta os trabalhos relacionados utilizados como referência para o desenvolvimento deste estudo. Os trabalhos aqui mencionados dizem respeito ao processo de segmentação das estruturas cardíacas no âmbito do exame de cine-RM de eixo curto utilizando abordagens baseadas em redes neurais convolucionais.

Alguns desses trabalhos visam segmentar somente os contornos do ventrículo esquerdo, outros consideram também a cavidade ventricular esquerda e o ventrículo direito, executando essa tarefa em uma ou mais etapas. Alguns deles realizam experimentos com bases privadas, e outros com as bases públicas *Sunnybrook Cardiac Dataset* (RADAU et al., 2009), LVSC (SUINESIAPUTRA et al., 2014), ACDC (BERNARD et al., 2018) e a M&Ms (CAMPELLO et al., 2021a). Com base nisso, neste capítulo, os trabalhos apresentados a seguir estão organizados por base de dados.

2.1 Base *Sunnybrook Cardiac Dataset*

Os trabalhos envolvendo a base *Sunnybrook Cardiac Dataset* visam a segmentação dos contornos do epicárdio e endocárdio. Entre eles, a abordagem de Tran (2016) foi a pioneira em usar FCN para a segmentação em RMC de eixo curto. A arquitetura de FCN implementada é a mesma proposta por Long et al. (2015), sem relevantes modificações. O método assume a priori que a cavidade ventricular está sempre no centro das imagens e usa essa informação para extrair uma ROI que engloba os ventrículos. Contudo, essa estrutura não está centralizada em todos os casos, sendo essa a principal justificativa para as falhas de segmentação apresentadas pelo método. Os resultados obtidos foram 0,92 e 0,96 de Dice, respectivamente, para a segmentação do endocárdio e epicárdio.

Hu et al. (2019) implementam uma abordagem dividida em duas etapas, sendo que a primeira consiste na extração de uma ROI a partir de uma segmentação primária via SegNet (BADRINARAYANAN et al., 2017). Em seguida, a ROI e a máscara são transformadas do espaço cartesiano para o espaço polar, e passadas como entrada a um módulo de refinamento, no qual é aplicado um processo de otimização baseado em mapa de bordas e programação dinâmica. Nos experimentos, esse método obteve 0,90 e 0,93 de Dice para a segmentação do endocárdio e do epicárdio, respectivamente.

2.2 Base LVSC

Os trabalhos desenvolvidos no âmbito da base LVSC têm como principal objetivo a segmentação dos contornos do ventrículo esquerdo. Destaca-se também que a métrica principal de avaliação, diferentemente dos demais cenários, é o índice de Jaccard.

Tan et al. (2017) propuseram um método semi-supervisionado que usa uma CNN para prever o centro do ventrículo esquerdo (VE), a partir do qual é feita a extração da região de interesse (ROI) que é passada para uma segunda etapa de segmentação. Neste processo, a ROI é transformada para o espaço das coordenadas polares e serve de entrada a uma rede neural convolucional (CNN) que delimita os contornos do VE via regressão. Apesar da robustez e dos resultados acurados, trata-se de uma proposta que necessita de intervenção humana em seus passos iniciais, o que pode deixar sua execução mais lenta e suscetível a falhas. Esse método alcançou 0,74 de índice de Jaccard nos experimentos.

O método proposto por Cui et al. (2021) também é validado com a base LVSC. Este método utiliza a U-Net com blocos Attention incorporados para reforçar o aprendizado das estruturas de interesse sem elevar muito a quantidade de parâmetros treináveis. Essa rede recebe uma entrada multi-escala de fatias, e é combinada com a função de perda Focal Tverky (ABRAHAM; KHAN, 2019) para lidar com o desbalanceamento de classes. Como resultado, essa abordagem obtém índice de Jaccard de 0,75.

2.3 Bases Privadas

Os trabalhos a seguir realizam experimentos com bases privadas as quais apresentam outras estruturas cardíacas, além do ventrículo esquerdo, como objeto de interesse da segmentação.

O método proposto por Abdeltawab et al. (2020) realiza um processo em duas etapas, inicialmente encontrando a cavidade ventricular esquerda (CV) para extrair uma ROI que contenha essa estrutura e o miocárdio (Mio). Em seguida, essa região é usada como entrada para a etapa de segmentação baseada em FCN. Os autores utilizaram uma base privada nos experimentos que avaliam a segmentação considerando as fases cardíacas de fim da sístole (ES) e da diástole (ED). O método obteve 0,96 e 0,92 de Dice para a segmentação da CV respectivamente em ES e ED. No caso do Mio, os resultados de Dice foram 0,88 (ES) e 0,89 (ED).

A abordagem desenvolvida por (SHI et al., 2021) propõe um modelo baseado na U-Net chamado MIFNet, que é aplicado na segmentação dos ventrículos direito, esquerdo, e átrio esquerdo. Nessa abordagem, as características são extraídas de entrada com diferentes escalas e depois combinadas, gerando novos mapas de características. Como resultado, um Dice médio de 0.97 foi obtido em experimentos com uma base adquirida para testes locais, o que dificulta a sua reprodutibilidade.

2.4 Base ACDC

A seguir, serão apresentadas abordagens direcionadas para a segmentação da CV, do Mio, e do VD, cujos métodos foram avaliados em experimentos com a base ACDC. As métricas também foram calculadas considerando as fases cardíacas ES e ED. Ressalta-se que esta foi a base utilizada durante o desenvolvimento dos métodos propostos nesta tese (Capítulos 4 e 5).

Isensee et al. (2017) propuseram um ensemble de U-Net 2D e 3D para segmentar os ventrículos direito, esquerdo e a cavidade ventricular esquerda. Os modelos 2D e 3D recebem na entrada exemplos com dimensões elevadas, exigindo, portanto, um alto custo computacional que nem sempre é viável. Como resultado, os autores obtiveram respectivamente para ED e ES coeficientes Dice de 0,967 e 0,928 na segmentação da CV; 0,904 e 0,923 para o Mio; 0,951 e 0,904 para o VD.

Baumgartner et al. (2017) também fizeram experimentos com as arquiteturas U-Net (2D e 3D) e a FCN-8s, cada qual com dimensões específicas para a entrada, assumindo que a redução de tamanho poderia acarretar em perda de informação. Esse método obteve os seguintes valores de Dice em ED e ES: 0,963 e 0,911 para a segmentação da CV; 0,892 e 0,901 para o Mio; 0,932 e 0,883 para o VD.

Calisto e Lai-Yuen (2020) propuseram um ensemble adaptativo de FCNn 2D e 3D aplicado à segmentação de imagens médicas, inclusive de cine-RM cardíaca, chamado de AdaEn-Net. O processo de adaptação ocorre por meio de um algoritmo de otimização multiobjetivo que determina tamanhos de kernel, número de filtros e quantidade de blocos, e com isso definindo a largura e profundidade da rede. Os resultados de Dice obtidos por essa abordagem nas fases ED e ES foram de 0,958 e 0,903 (CV); 0,873 e 0,895 (Mio); e 0,936 e 0,884 (VD).

A abordagem proposta por Zotti et al. (2017) apresenta a arquitetura Grid-Net.

Trata-se de uma modificação na U-Net em que as conexões salteadas são substituídas por camadas convolucionais a fim de ampliar a quantidade de mapas com características relevantes. Esse método alcançou nas fases ED e ES coeficientes Dice de 0,964 e 0,912 (CV); 0,886 e 0,902 (Mio); e 0,941 e 0,882 (VD).

Khened et al. (2019) propuseram uma abordagem de segmentação baseada em Dense-Nets Residuais em Multi-escala. Nessa proposta, utiliza-se uma FCN densamente conectada com o intuito de ampliar a complexidade do modelo, garantindo uma maior quantidade de mapas de características, e evitar o problema da explosão do gradiente. Os resultados obtidos por esse método para as fases ED e ES em relação ao Dice foram de 0,964 e 0,912 (CV); 0,886 e 0,902 (Mio); e 0,941 e 0,882 (VD).

Simantiris e Tziritas (2020) apresentam um método baseado em uma FCN mais simples que executa convoluções dilatadas para preservar informações de localização. A segmentação ocorre em duas etapas, sendo que a primeira é inspirada por Grinias e Tziritas (2017) para extração de uma ROI menor, reduzindo o escopo da imagem para submetê-la ao processo de segmentação. Essa abordagem obteve valores de Dice nas fases ED e ES de 0,964 e 0,912 (CV); 0,886 e 0,902 (Mio); e 0,941 e 0,882 (VD).

A abordagem proposta por Ammar et al. (2021) também utiliza uma U-Net para a segmentação das estruturas cardíacas. As fatias são submetidas a um processo de extração da ROI baseado no uso da Transformada Rápida de Fourier (FFT) para localizar, nas fatias, a região do coração. As ROIs extraídas são passadas à uma U-Net padrão. Essa abordagem também conta com um processo de refinamento da segmentação do Mio feito no espaço polar. Os resultados de obtidos no desafio ACDC, respectivamente, para as fases ED e ES, foram 0,968 e 0,911 (CV); 0,891 e 0,901 (Mio); 0,929 e 0,886 (VD).

Li et al. (2022) apresentam uma arquitetura de rede convolucional chamada de DCNet, composta por blocos convolucionais chamados de *Diversity Convolution Blocks*, cuja finalidade é a extração de características de diferentes escalas ao longo do volume através da combinação entre convoluções dilatadas e um mecanismo de atenção chamado de *Dual-path Channel Attention* (DCAM). Na entrada, são passados recortes de tamanho 160×160 obtidos a partir do centro das imagens originais. Esse método alcançou Dices ED e ES, respectivamente, de 0,970 e 0,930 (CV); 0,890 e 0,910 (Mio); 0,940 e 0,890 (VD).

2.5 Base M&Ms

Nesta seção, serão apresentadas abordagens cujo desenvolvimento e validação foram realizados com a base M&Ms. Assim como nos trabalhos relatados nas seções anteriores, aqueles apresentados na sequência usam métricas de avaliação calculadas com base nas fases cardíacas ES e ED.

O método desenvolvido por Scannell et al. (2020) apresenta uma abordagem chamada *Domain-adversarial learning*, que visa melhorar a generalização das redes neurais, incentivando-as a aprender características independentemente do domínio da entrada. Nesse processo, durante o treinamento de uma U-Net 2D para a segmentação das estruturas cardíacas, seus mapas de características intermediários são então passados para uma CNN que é otimizada para classificar os domínios de entrada. Por fim, os parâmetros da U-Net são atualizados via *adversarial back-propagation*, o que, segundo os autores, serve para incentivar a invariância de domínio. Como resultado, essa abordagem alcançou Dices ED e ES, respectivamente, de 0,905 e 0,848 (CV); 0,772 e 0,820 (Mio); 0,876 e 0,809 (VD).

Huang et al. (2020) também usam a U-Net como método de segmentação em conjunto com o processo de transferência de estilo usando uma rede neural proposta chamada de rede ST, cujo treinamento visa, ao receber uma imagem de entrada, gerar, na saída, uma nova imagem obedecendo a um novo padrão de aparência. A U-Net, por sua vez, é treinada com imagens que obedecem esse novo padrão. Dessa forma, segundo os autores, o método pode ser aplicado em imagens de cine-RM oriundas de diferentes scanners. Os resultados obtidos de Dice ED e ES, respectivamente, foram de 0,896 e 0,772 (CV); 0,761 e 0,721 (Mio); 0,820 e 0,698 (VD).

A abordagem desenvolvida por Full et al. (2020) utiliza um *framework* chamado nnU-Net, desenvolvido para realizar o treinamento automatizado de redes de segmentação, facilitando a configuração de hiperparâmetros e a experimentação de diversas operações de *Data Augmentation*. Como resultado, foram obtidos os respectivos valores de Dice ED e ES: 0,939 e 0,886 (CV); 0,839 e 0,867 (Mio); 0,910 e 0,860 (VD).

Li et al. (2021) propuseram um método que combina a U-Net e uma *Conditional GAN* (CGAN) como um autoencoder. Nesse processo, as imagens são previamente recortadas em tamanho 160×160 e registradas tendo uma imagem escolhida como referência de modo que os centros das CV sejam alinhados. Isso caracteriza uma abordagem semi-supervisionada. A U-Net produz inicialmente as máscaras, as quais são submetidas à

rede geradora no contexto da CGAN. Os autores propõem também uma função de perda para guiar o aprendizado geral da combinação de redes. Dessa forma, o método alcança os seguintes resultados de Dices ED e ES: 0,930 e 0,894 (CV); 0,764 e 0,828 (Mio); 0,883 e 0,822 (VD).

Lin et al. (2022) desenvolveram um método em cascata que usa duas redes propostas chamadas de AEM-Net. A primeira rede é usada para segmentar toda a região cardíaca que envolve as três estruturas de interesse (CV, Mio e VD). Esse resultado é passado para a segunda rede que gerará simultaneamente as máscaras para cada estrutura. Essa abordagem obtém os seguintes resultados de Dices ED e ES: 0,924 e 0,889 (CV); 0,827 e 0,859 (Mio); 0,878 e 0,843 (VD).

Huang et al. (2022) propuseram uma rede chamada de RegU-Net. Ela possui três módulos: codificador, decodificador e localização. O codificador extrai as características semânticas das imagens de entrada. A sua saída é passada simultaneamente para o módulo de localização da ROI e para o decodificador. O módulo de localização identifica uma ROI que envolve CV, Mio e VD. E, por fim, o decodificador gera a segmentação com as características extraídas pelos dois módulos anteriores. Os autores utilizaram como métrica avaliadora o Dice Geral, sem separar por fase cardíaca. Portanto, como resultado, foram obtidos os seguintes valores: 0,945 (CV); 0,869 (Mio); e 0,912 (VD).

2.6 Considerações Finais

Assim como alguns dos trabalhos mencionados, os métodos propostos pelo presente trabalho também são divididos em etapas, iniciando com a extração de uma ROI que envolva as estruturas de interesse (CV, Mio e VD). No entanto, os trabalhos relacionados usam essa ROI como entrada para um processo de segmentação que gera, em passo único, as máscaras dessas estruturas. Diferentemente deles, as etapas dos métodos propostos foram desenvolvidas considerando características específicas das fatias dos exames e das estruturas de interesse, utilizando técnicas visando produzir resultados expressivos sem ampliar consideravelmente o custo computacional.

Como dito anteriormente, este trabalho propõe dois métodos. O primeiro deles, chamado de Segmentação em Cascata com Reconstrução (SCR), usa a U-Net para extração da ROI, que é passada para a segmentação inicial das estruturas de interesse. Para essa etapa, é proposta uma arquitetura de FCN que combina a EfficientNet, blocos Inception,

Attention e Squeeze-And-Excitation. E como um passo posterior de refinamento dos resultados, propõe-se o uso de um módulo de reconstrução baseado em U-Net.

O segundo método, por sua vez, também aplica uma abordagem baseada em U-Net para a extração da ROI. Essas ROIs são submetidas a uma etapa de segmentação da CV e Mio, que produz inicialmente uma segmentação usando a FCN proposta para o primeiro método, sendo esse resultado refinado através de uma segunda segmentação via U-Net. Essa etapa também possui um processo de redução de escopo para mitigar o desbalanceamento entre classes de pixel. E, para a segmentação do VD, é proposta outra FCN com base na X-Net combinada com blocos Inception e Attention. Esse método é chamado de Segmentação Especializada por Contexto (SEC).

A Tabela 1 traz um resumo dos trabalhos relacionados mencionados neste capítulo, apontando os autores e as técnicas utilizadas.

Tabela 1 – Visão geral dos trabalhos relacionados

Trabalhos relacionados	Métodos desenvolvidos	Bases Utilizadas
Tran (2016)	FCN	Sunnybrook Cardiac Dataset
Hu et al. (2019)	SegNet + refinamento baseado em otimização	Sunnybrook Cardiac Dataset
Tan et al. (2017)	CNNs + segmentação do VE no espaço polar	LVSC
Cui et al. (2021)	U-Net com blocos <i>Attention</i>	LVSC
Abdeltawab et al. (2020)	FCN e U-Net modificada para segmentação	Base privada
Shi et al. (2021)	MIFNet	Base privada
Isensee et al. (2017)	Ensemble de U-Nets 2D e 3D	ACDC
Baumgartner et al. (2017)	U-Nets 2D e 3D	ACDC
Zotti et al. (2017)	Grid-Net	ACDC
Khened et al. (2019)	Multi-Scale Residual Dense-Net	ACDC
Calisto e Lai-Yuen (2020)	AdaEn-Net	ACDC
Simantiris e Tziritas (2020)	FCN com convoluções dilatadas	ACDC
Ammar et al. (2021)	Extração de ROI baseada em FFT + U-Net	ACDC
Li et al. (2022)	DCNet	ACDC
Scannell et al. (2020)	<i>Domain-adversarial learning</i>	M&Ms
Huang et al. (2020)	U-Net + Rede ST	M&Ms
Full et al. (2020)	nnU-Net	M&Ms
Li et al. (2021)	U-Net + CGAN	M&Ms
Lin et al. (2022)	AEM-Net	M&Ms
Huang et al. (2022)	RegU-Net	M&Ms
Método Proposto nº 1	Segmentação em Cascata com Reconstrução	ACDC/ M&Ms
Método Proposto nº 2	Segmentação Especializada por Contexto	ACDC/ M&Ms

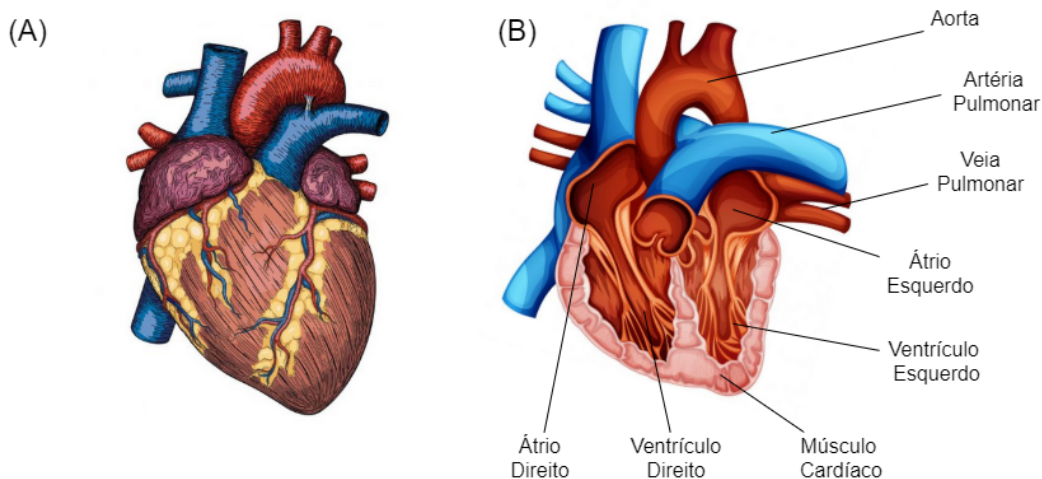
3 FUNDAMENTOS TEÓRICOS

Este capítulo apresenta os fundamentos teóricos utilizados no desenvolvimento deste trabalho. Esses conceitos são importantes para a compreensão do contexto no qual o trabalho se insere, bem como as técnicas adotadas para alcançar os objetivos. Na sequência, serão abordados os aspectos estruturais do coração humano, contextualizando-se também o exame de cine-ressonância magnética cardíaca. O capítulo também explana sobre processamento de imagens digitais, aprendizado profundo e as métricas utilizadas para a avaliação dos resultados.

3.1 O Coração Humano

O coração humano é o principal órgão do sistema circulatório. Ele é responsável pelo bombeamento do sangue para as diferentes partes do corpo por meio dos vasos sanguíneos, garantindo assim o abastecimento das células no que se refere aos nutrientes e oxigênio, além de permitir que os resíduos do metabolismo celular sejam carregados até os órgãos excretórios. A Figura 1 apresenta uma visão geral das estruturas do coração.

Figura 1 – O Coração Humano: Visão externa (A) e interna (B).



Fonte: Elaborada pelo autor a partir de imagens disponíveis em <<https://br.depositphotos.com/>>

O coração possui quatro câmaras dispostas em duas séries de bombas funcionais: uma bomba direita que compreende átrio e ventrículo direitos que recebem sangue venoso da circulação sistêmica e o bombeia para os pulmões para oxigenação; e uma bomba esquerda caracterizada pelo átrio e ventrículo esquerdos que recebem sangue oxigenado

da circulação pulmonar, bombeando-o para a circulação sistêmica (BAZIRA, 2021). Os ventrículos possuem uma parede muscular mais espessa que os auxilia na função de bombear o sangue para fora do coração. No caso do ventrículo esquerdo, essa composição muscular é cerca de três vezes mais espessa que a do direito por causa da alta pressão hemodinâmica sustentada por essa estrutura. O músculo cardíaco é caracterizado por uma parte medial (miocárdio) e seus revestimentos interno (endotélio) e externo (pericárdio).

O coração possui um ciclo de movimento em que os músculos se contraem e relaxam. A fase de contração é chamada de sístole e a de relaxamento é denominada diástole. A primeira garante o bombeamento de sangue, enquanto que a segunda permite que o sangue preencha as cavidades ventriculares. Quando os ventrículos se contraem, cerca de 70 ml de sangue é transmitido para o corpo por meio da artéria Aorta, sendo propagado por aproximadamente 100.000 Km de vasos sanguíneos (KHAN, 2005).

Por ter uma importância vital para o corpo humano, é salutar uma rotina de consultas médicas, e conseqüentemente de exames, para a avaliação das funções cardíacas em prol da manutenção da qualidade de vida a partir das orientações do especialista.

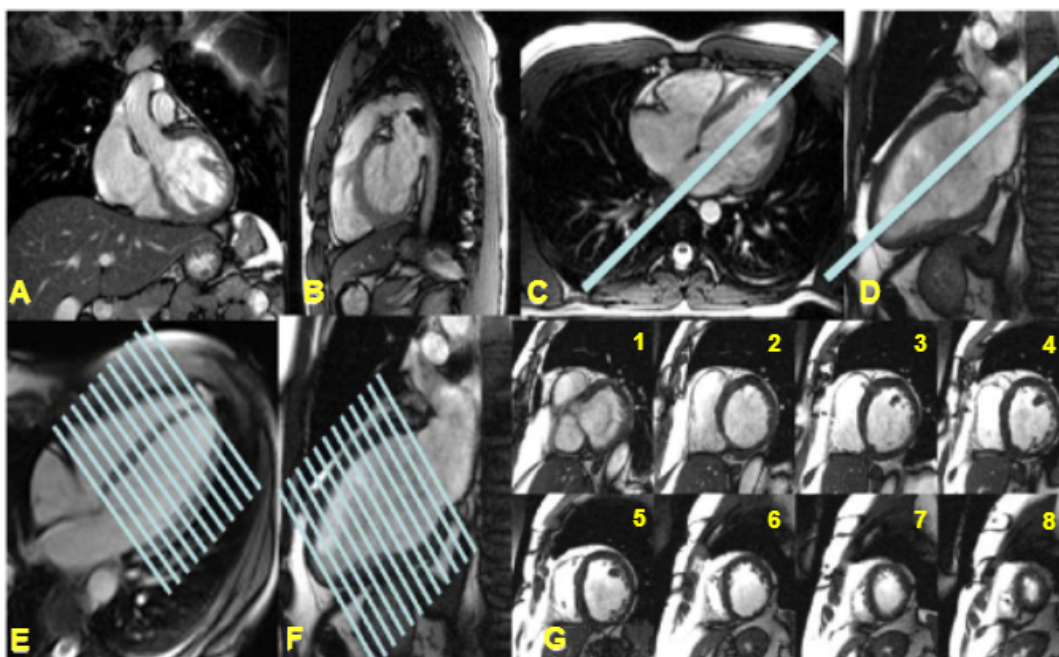
3.2 Cine-Ressonância Magnética Cardíaca

A Ressonância Magnética Cardíaca (RMC) é uma técnica baseada no uso da radiofrequência em tecidos submetidos a um campo magnético e pode ser considerado como um dos métodos de maior segurança no diagnóstico cardiovascular (HIGGINS, 1992). Essa técnica não utiliza radiação ionizante ou contrastes nefrotóxicos. Trata-se de um método aplicável clinicamente em um espectro de doenças cardiovasculares (DCVs) bastante extenso, como no diagnóstico das doenças congênitas, avaliação das alterações ventriculares, além do seguimento pós-operatório de adultos com estas doenças (ROCHITTE et al., 2006).

Segundo Bezerra et al. (2006), no âmbito da RMC, que também pode ser chamada de cine-ressonância magnética (cine-RM), a aquisição de imagens via sequências de equilíbrio estável segmentadas e associadas ao eletrocardiograma oferecem uma precisa avaliação da função cardíaca, apontando também que isso se deve ao alto contraste entre o sangue e o miocárdio. Esse protocolo requisita que o paciente expire o ar e segure a respiração por alguns segundos (apneia ou *breath-hold*), mas há outros que dispensam esse procedimento, bem como a aplicação de contraste (SAKUMA et al., 1993).

Os exames de ressonância magnética em geral consideram três planos anatômicos: coronal, sagital e o axial, como mostram as Figuras 2A, 2B e 2C, respectivamente. Esses são eixos ortogonais ao tórax e não respeitam a orientação cardíaca. Contudo, no exame de cine-RM cardíaca são prescritos outros planos que fornecem diferentes visualizações e, por conseguinte, uma melhor análise. São eles (NACIF et al., 2010): eixo longo vertical; eixo longo horizontal; e eixo curto.

Figura 2 – Sequências do exame de cine-RM e categorização dos seguintes planos anatômicos: (A) Coronal; (B) Sagital; (C) Axial; (D) Visão de duas câmaras; (E) Visão de quatro câmaras; (F) Planejamento de aquisição do eixo curto; e (G) Eixo curto.



Fonte: Adaptada de (NACIF et al., 2010).

O eixo longo vertical é obtido a partir do eixo axial por meio de um plano que abrange o ventrículo esquerdo (VE) e o átrio esquerdo (AE) mediano à cavidade ventricular (Figura 2C). Ele também é denominado de visualização em duas câmaras (Figura 2D e 2F) e é usado no planejamento da aquisição do eixo longo horizontal ou quatro câmaras (Figura 2E). Este irá identificar o VE, o AE, o ventrículo direito (VD) e o átrio direito (AD). Por fim, a partir dessa visualização é possível prescrever o eixo curto, traçado de forma ortogonal ao VE, como visto na Figura 2G. O eixo curto se caracteriza pelas fatias basais (Figura 2G.1), mediais (Figura 2G.2-5), e apicais (Figura 2G.6-8).

Segundo Filho et al. (2004), uma vez adquiridas as imagens de cine-RM do eixo curto, que evidenciam os ventrículos direito e esquerdo, os especialistas podem avaliá-las por meio da observação, e de um método chamado planimetria digital, que consiste no

uso de um software para delimitar manualmente, nas imagens, os contornos referentes às estruturas de interesse da análise. Com esses contornos, é possível quantificar os volumes ventriculares sistólico e diastólico através da multiplicação entre a soma das áreas planimetradas e a espessura de cada fatia do exame. Tais volumes são importantes para o cálculo de outras medidas, como a fração de ejeção de sangue, que é determinante para o diagnóstico da insuficiência cardíaca (PINHEIRO et al., 2013).

Para o profissional que realiza a cine-RM cardíaca, é condição *sine qua non* o sólido embasamento acerca dos aspectos anatômicos e fisiopatológicos, além das diferentes ferramentas de estudo que o exame oferece. Pois, o desconhecimento ou eventuais equívocos podem reduzir a capacidade diagnóstica do método. Isso reforça o desenvolvimento de tecnologias e métodos computacionais para auxiliar os profissionais na obtenção das imagens e em sua análise.

3.3 Processamento de Imagens Digitais

De acordo com Gonzalez e Woods (2010), processamento de imagens pode ser definido como um conjunto de técnicas computacionais que transformam uma imagem de entrada em uma saída desejada. O uso dessas técnicas viabiliza a extração e a identificação de informações nas imagens, melhorando também a qualidade visual de aspectos estruturais, facilitando a percepção humana e a interpretação automática por meio de máquinas (PEDRINI; SCHWARTZ, 2007).

Muitos algoritmos podem ser incorporados a um determinado método de processamento de imagens, cada um para determinado propósito. Mas, apesar de suas especificidades, esses métodos geralmente obedecem uma sequência de passos fundamentais.

A aquisição das imagens digitais é o passo inicial. Nessas imagens, podem ser aplicados pré-processamentos em busca de melhoramentos de visualização. Outra etapa é a segmentação, que consiste na extração das regiões de interesse da análise. Além desses passos, há também a representação, descrição, reconhecimento e interpretação dos resultados. Ressalta-se que o resultado final de cada passo pode ser utilizado como entrada para outro subsequente em conformidade com o objetivo a ser alcançado. Ademais, em processamento de imagens, as metodologias podem seguir somente um subconjunto destas etapas.

Dado esse contexto, a presente seção mostra um detalhamento das principais técnicas de processamento de imagens utilizadas no desenvolvimento deste trabalho.

3.3.1 Suavização e Normalização

A suavização de imagens é um processo no qual são utilizadas técnicas, denominadas filtros, para causar um certo grau de borramento e, com isso, eliminar alguns ruídos e uniformizar regiões das imagens. No presente trabalho, a técnica de suavização recorrentemente utilizada em virtude das características visuais dos objetos de análise é chamada de filtro da mediana.

O filtro da mediana faz parte do conjunto de filtros estatísticos. Eles são filtros espaciais não lineares que produzem respostas com base na ordenação ou classificação dos pixels pertencentes à área da imagem coberta pelo filtro e na substituição do valor do pixel central pelo resultado da classificação. No caso do filtro da mediana, esse valor do pixel central é substituído pela mediana dos valores classificados. Essa técnica é bastante útil para a remoção de ruídos do tipo "sal e pimenta", produzindo um borramento consideravelmente menor em relação a outros filtros (GONZALEZ; WOODS, 2010).

Outra técnica utilizada no presente trabalho é a normalização de imagens. Esse processo consiste na padronização dos valores de pixel em uma determinada faixa, geralmente entre 0 e 1. O uso dessa técnicas é importante uma vez que as imagens de cine-RM cardíaca têm uma patente variação de intensidades em virtude dos diferentes protocolos de aquisição e *scanners* utilizados para realizar os exames (BAI et al., 2015). Com a normalização esse problema é reduzido, garantindo um escopo de análise mais definido.

No âmbito do presente trabalho, a normalização das imagens é feita através do cálculo apresentado pela Equação 3.1.

$$N(x, y) = \frac{M(x, y) - \min(M)}{\max(M) - \min(M)} \quad (3.1)$$

onde $M(x,y)$ é o valor do pixel na imagem M e $N(x,y)$ o seu correspondente na nova imagem normalizada.

3.3.2 Morfologia Matemática

Morfologia matemática é o estudo da estrutura geométrica das entidades presentes em uma imagem (SERRA, 1986). Tais conceitos podem ser aplicados em várias tarefas e

com diferentes objetivos, tais como melhoramento de imagens, segmentação, entre outros.

Através da morfologia matemática, é possível extrair informações geométricas e topológicas de uma imagem através de um subconjunto previamente definido, chamado de elemento estruturante, de modo que os resultados obtidos para cada transformação aplicada estão diretamente ligados a ele. Esse elemento estruturante é uma entidade que pode assumir vários formatos, a depender do processamento desejado. As formas mais comuns são a cruz, a elipse e o retângulo.

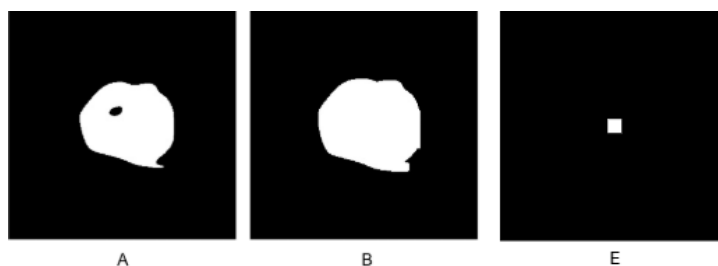
As operações básicas do processamento morfológico são a dilatação e a erosão. A compreensão do seu funcionamento, portanto, auxilia o entendimento das outras operações a serem explanadas posteriormente.

Dados os conjuntos C e D definidos no espaço Z^2 . A dilatação de C por D é definida como:

$$C \oplus D = \{x | (\hat{D})_x \cap C \neq \emptyset\} \quad (3.2)$$

D é o elemento estruturante. O processo consiste em rotacionar D em torno da origem a fim de que se obtenha o seu conjunto reflexão e depois fazer um deslocamento de x . A dilatação então consiste em todos os x deslocamentos para os quais a interseção entre a reflexão de D e C inclui pelo menos um elemento diferente de zero (GONZALEZ; WOODS, 2010). Como resultado, a dilatação produz o inchaço da imagem. A Figura 3 ilustra esse processo.

Figura 3 – Exemplo da operação de dilatação sobre uma imagem (A) resultando em (B). O elemento estruturante (E) é um quadrado com dimensões 10 x 10.



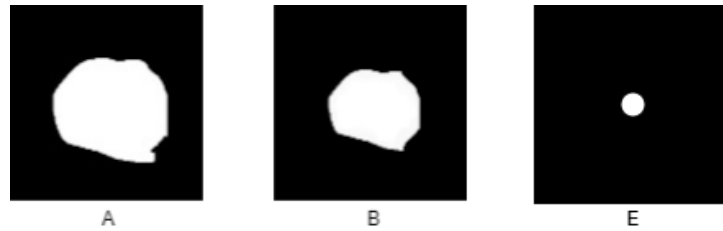
Fonte: Adaptada de (SILVA, 2019).

Sejam C e D conjuntos definidos no espaço Z^2 , a erosão de C por D é definida como:

$$C \ominus D = \{x | (D)_x \subseteq C\} \quad (3.3)$$

D é o elemento estruturante. A erosão de C por D resulta no conjunto de pontos x de modo que D , transladado de x , está contido em C . Em contrapartida ao inchaço resultante da dilatação, a erosão ocasiona a diminuição do tamanho das estruturas, como ilustrado pela Figura 4.

Figura 4 – Exemplo da operação de erosão sobre uma imagem (A) resultando em (B). O elemento estruturante (E) é circular e tem dimensões 15 x 15.



Fonte: Adaptada de (SILVA, 2019).

Quando aplicadas repetidamente na mesma imagem, dilatação e erosão têm seus efeitos ampliados. É importante mencionar que os efeitos produzidos por essas operações dependem diretamente da definição do elemento estruturante.

As outras operações a serem explanadas e que integram parte do método proposto são a abertura e o fechamento. Seja o conjunto C com o elemento estruturante D , pode-se dizer que a abertura consiste em uma erosão seguida de dilatação conforme mostra a Equação 3.4.

$$C \circ D = (C \ominus D) \oplus D \quad (3.4)$$

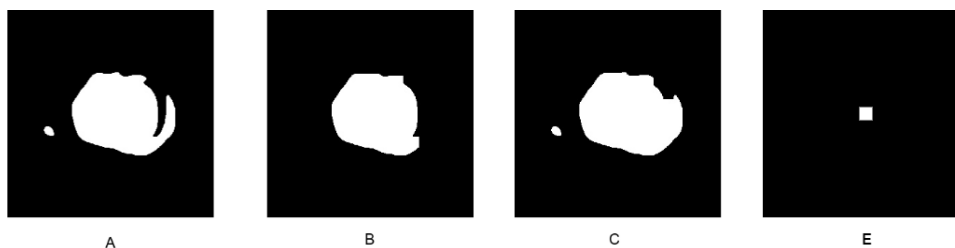
Por sua vez, o fechamento de um conjunto C com o elemento estruturante D consiste em uma dilatação seguida de erosão como define a Equação 3.5.

$$C \circ D = (C \oplus D) \ominus D \quad (3.5)$$

A abertura, em geral, ocasiona a suavização de contornos bem como a separação de estruturas ligadas por linhas tênues. O fechamento, por outro lado, pode conectar estruturas separadas por pequenas lacunas e eliminar pequenos orifícios. A Figura 5 mostra os resultados produzidos por essas operações, todas fazendo uso do mesmo elemento estruturante.

Gonzalez e Woods (2010) destacam que a abertura e o fechamento satisfazem uma propriedade chamada idempotência. Isso significa que elas não produzem efeitos

Figura 5 – Operações aplicadas sobre a imagem (A). Abertura resulta em (B). Fechamento, em (C). O elemento estruturante (E) é um quadrado de dimensões 15 x 15.



Fonte: Adaptada de (SILVA, 2019).

acumulativos na imagem, logo, a aplicação sucessiva dessas operações sobre uma mesma imagem não irá alterá-la, a menos que haja outros processamentos entre as suas aplicações ou mudanças no elemento estruturante.

3.4 Aprendizado Profundo

Aprendizado profundo é uma área a qual representa uma tentativa de aprendizado de abstrações de alto nível em um conjunto de dados. Essa área tem sido amplamente estudada em visão computacional e, conseqüentemente, muitas propostas têm sido desenvolvidas e aplicadas em diversos cenários, principalmente na análise de imagens (REDMON et al., 2015; SZEGEDY et al., 2016).

O conceito de aprendizado profundo se origina no estudo das redes neurais artificiais (RNA). Elas são constituídas por unidades de processamento chamadas de neurônios, os quais implementam mecanismos denominados funções de ativação. Essas redes são capazes de aproximar qualquer função mensurável para qualquer grau de precisão desejado (HORNIK et al., 1989). Ademais, dependendo do problema, uma rede pode precisar de muitas camadas para alcançar resultados satisfatórios, retardando, por conseguinte, o seu treinamento e a tornando-a mais suscetível ao *overfitting* (LIU et al., 2017).

Os algoritmos de aprendizado profundo também se apresentam como arquiteturas hierárquicas contendo múltiplas camadas, cada qual composta por sua unidade de processamento. Diferente das RNAs tradicionais, esses algoritmos permitem a extração automática de características a partir do conjunto de treinamento, sem a necessidade de uma série de técnicas de processamento de imagens e reconhecimento de padrões (HUA et al., 2015). Isso implica na redução da intervenção humana no processo.

Neste trabalho, será dado foco para duas abordagens de aprendizado profundo: redes neurais convolucionais (CNN - *Convolutional Neural Networks*) e redes totalmente convolucionais (FCN - *Fully Convolutional Networks*). Essas técnicas são utilizadas pelo método proposto ao longo de suas etapas percorridas nos Capítulos 4 e 5. As seções a seguir apresentam mais detalhes sobre o funcionamento dessas abordagens.

3.4.1 Redes Neurais Convolucionais

As redes neurais convolucionais (CNN - *Convolutional Neural Networks*) são modelos biologicamente inspirados capazes de construir um aprendizado hierárquico de características (LECUN et al., 1998). CNNs são um tipo de arquitetura profunda discriminativa (AREL et al., 2010), mostrando desempenho satisfatório no processamento de imagens e vídeos.

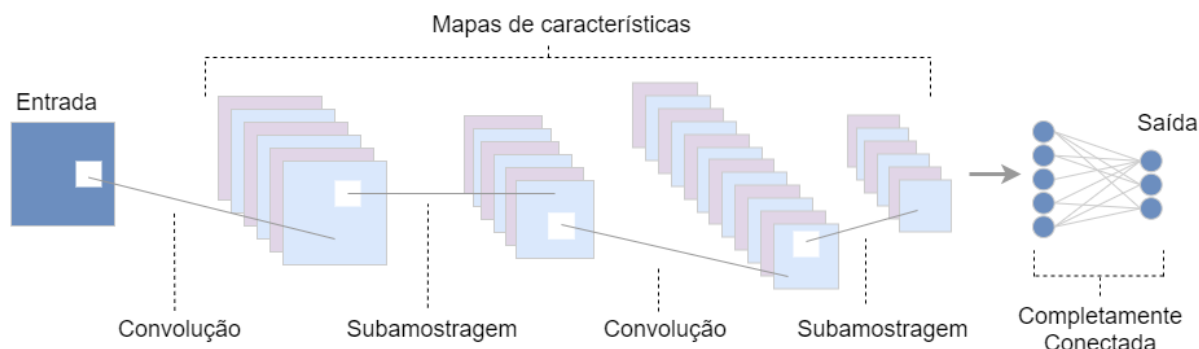
Goodfellow et al. (2016) enumeram três conceitos chave para o entendimento das CNNs: interação esparsa, compartilhamento de parâmetros e representação equivariante. A interação esparsa por meio de *kernels* de dimensões pequenas aplicados sobre toda a imagem de entrada numa operação chamada de convolução, em substituição às multiplicações de matriz inerentes às RNAs tradicionais. O compartilhamento de parâmetros é uma habilidade que permite o aprendizado de somente parte do total de parâmetros de uma CNN, permitindo que a outra parte permaneça "congelada", flexibilizando o treinamento. Essa característica possibilita a identificação de outra propriedade chamada equivariância, na qual, dadas as mudanças no conjunto de entrada, a saída mudará da mesma forma.

A CNN geralmente utiliza em sua arquitetura três tipos de camadas: convolução, subamostragem e completamente conectada (GUO et al., 2016). As camadas de convolução e subamostragem compõem as camadas iniciais da rede, e são implicitamente responsáveis pela extração de características. A camada completamente conectada, por fim, localiza-se no final da rede, e é responsável pela classificação dos padrões. A Figura 6 ilustra uma arquitetura básica de CNN.

A camada de convolução é composta por filtros treináveis aplicados sobre a imagem de entrada, gerando, portanto, vários mapas de características (GUO et al., 2016). Ao fim do treinamento, cada filtro será responsável pela detecção de características específicas da imagem, ou de parte dela. As camadas de subamostragem são subsequentes às de convolução, e podem ser utilizadas para reduzir as dimensões dos mapas de características.

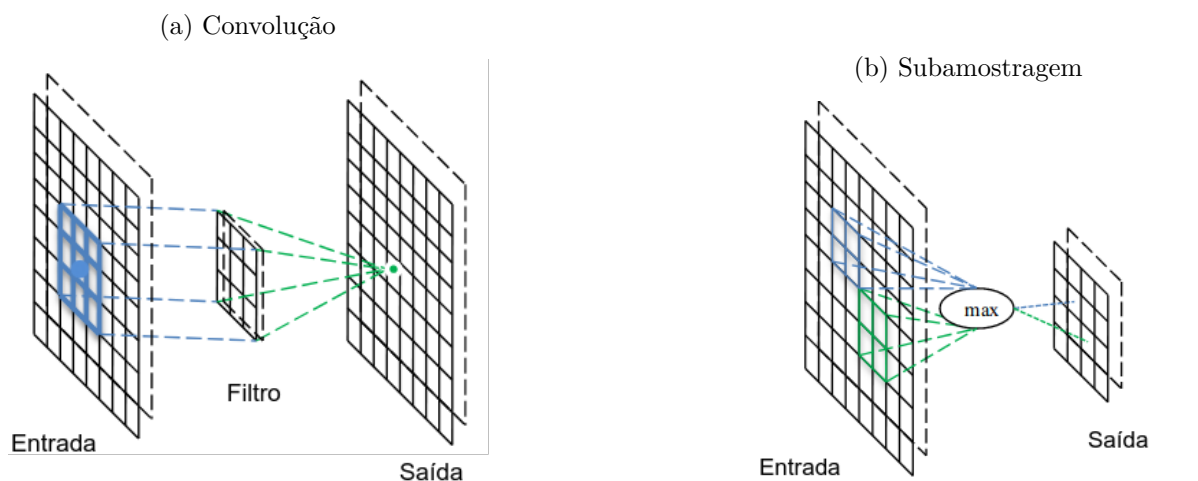
Nelas, os filtros são aplicados sobre os mapas extraíndo um valor que represente cada sub-região. Isso se dá através de operações como média aritmética (*Average Pooling*) ou valor máximo (*Max Pooling*). A Figura 7 ilustra o funcionamento das camadas supracitadas.

Figura 6 – Arquitetura básica de uma CNN contendo convoluções, subamostragens, mapas de característica e a camada completamente conectada.



Fonte: Elaborada pelo autor

Figura 7 – Camadas de convolução e subamostragem.



Fonte: Adaptado de (GUO et al., 2016).

Por fim, a camada completamente conectada consiste em uma estrutura de neurônios que, na entrada, recebem os *pixels* dos mapas de característica gerados pela camada anterior e produzem a classificação dos padrões extraídos.

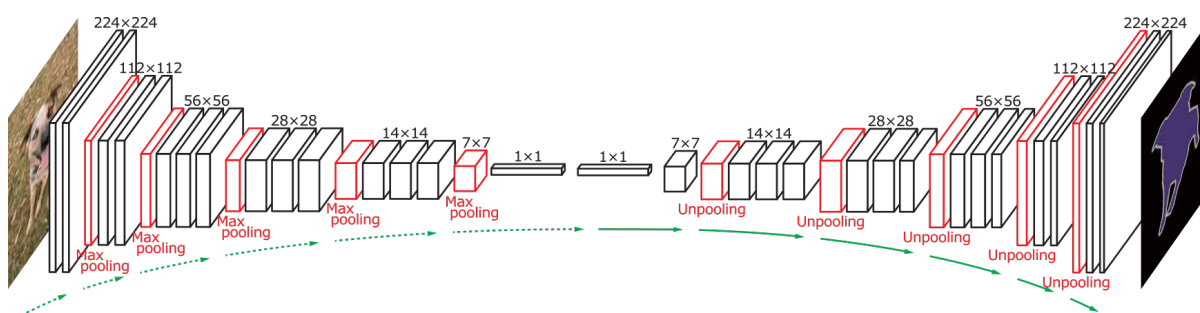
3.4.2 Redes Totalmente Convolucionais

As CNNs são utilizadas tipicamente em tarefas de classificação. Dada uma imagem de entrada, as camadas de convolução e subamostragem irão gerar mapas de características

que, hierarquicamente, representarão os padrões identificados e servirão de insumo ao processo de classificação realizado pela camada completamente conectada.

Para a tarefa de segmentação, no entanto, são necessárias informações de localização, ou seja, indicar quais *pixels* compõem ou não o objeto de interesse na imagem. Para essa finalidade, Long et al. (2015) desenvolveram o conceito de Redes totalmente convolucionais (FCN - *Fully Convolutional Networks*), que são arquiteturas projetadas para a segmentação de imagens. A Figura 8 ilustra uma arquitetura básica de FCN.

Figura 8 – Ilustração da arquitetura de uma rede completamente convolucional.



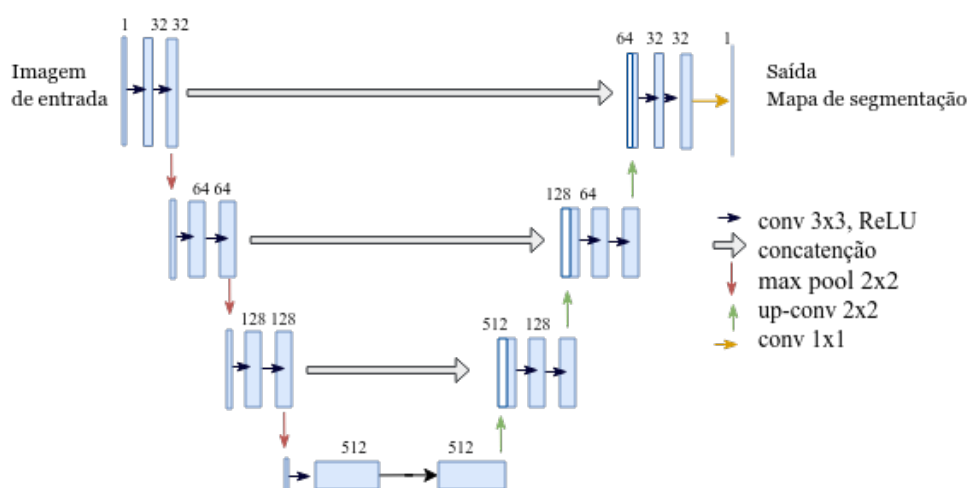
Fonte: Adaptada de (NOH et al., 2015).

As FCNs são compostas por dois estágios: contração e expansão. No primeiro, os mapas de características são gerados igualmente às CNNs por meio das camadas de convolução e subamostragem. Remove-se a camada completamente conectada, portanto, os mapas gerados servem como entrada para o estágio de ampliação. Nessa fase, novos mapas de características são produzidos através de novas camadas convolucionais e ampliados pelos operadores de *upsampling* (ou *unpooling*). Na última camada, os mapas gerados são submetidos a uma função de ativação que realizará a classificação pixel a pixel, produzindo, desta feita, a máscara de segmentação.

Os métodos apresentados por este trabalho utilizam três arquiteturas de FCN em suas etapas, cada qual para uma tarefa específica. Uma delas se chama U-Net. Proposta por Ronneberger et al. (2015), ela se diferencia das FCNs comuns porque combina os mapas de característica da fase de contração com seu correspondente simétrico na etapa de ampliação por meio de conexões salteadas. Dessa forma, a U-Net possui forma similar à letra "U", como pode ser visto pela Figura 9.

A etapa de contração (descida do "U") contém blocos convolucionais compostos por duas camadas de convolução com filtros de dimensões 3 x 3 e função de ativação ReLU. Cada bloco é sucedido de uma camada de subamostragem (*Max Pooling*) com

Figura 9 – Ilustração da arquitetura de uma rede U-Net.



Fonte: Adaptada de (RONNEBERGER et al., 2015).

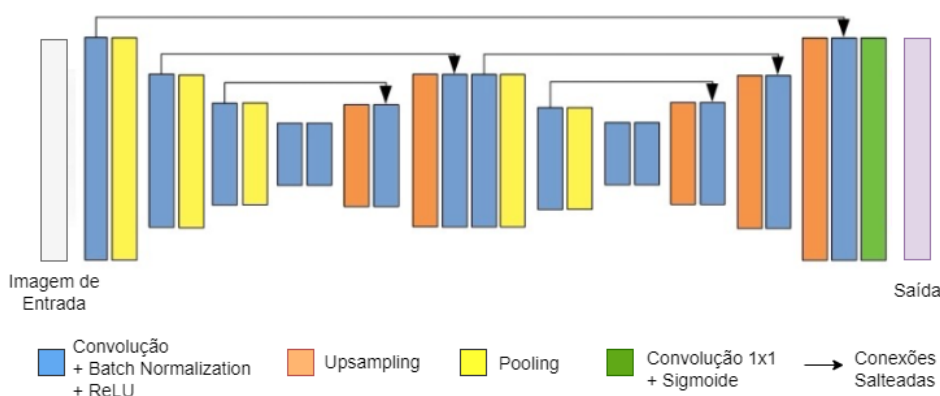
janelamento 2×2 . Ressalta-se que, a partir do segundo bloco, são gerados o dobro de mapas de característica do bloco anterior. Após o último bloco de convolução, no nível mais profundo da rede, não haverá camada completamente conectada. Logo, os mapas gerados nesse nível serão a entrada da etapa de ampliação.

A fase de expansão (subida do "U") contém blocos convolucionais similares aos da fase de contração. Eles são precedidos da operação *upsampling*, a qual amplia o tamanho das saídas produzidas pelo bloco que o antecede. Nessa fase, concatena-se os mapas de características dos blocos simétricos por meio das conexões salteadas. O último bloco se diferencia dos demais, pois é composto por três camadas convolucionais. As duas primeiras, contendo filtros de dimensão 3×3 e função de ativação ReLU. A última, entretanto, possui filtros de tamanho 1×1 e função de ativação sigmoide. Essa função atua como um classificador pixel-a-pixel, gerando valores entre 0 e 1 os quais indicam, respectivamente, fundo e região de interesse.

Uma variação da U-Net também utilizada neste trabalho é chamada de X-Net. Proposta por Bullock et al. (2019), essa rede possui dois estágios, cada um composto por um caminho de contração e outro de ampliação, como mostra a Figura 10. Segundo os autores, a composição dos dois caminhos de contração provê uma extração de característica em vários níveis. E a utilização de dois caminhos de expansão permite vários graus de reconstrução de características.

Em ambos os caminhos de contração são utilizados apenas blocos convolucionais (kernel 3×3) e Max Pooling (kernel 2×2). Operadores Upsampling (kernel 2×2) são

Figura 10 – Ilustração da arquitetura de uma rede X-Net.



Fonte: Adaptada de (BULLOCK et al., 2019)

usados nos dois caminhos de expansão. Em todos os estágios, ocorre o compartilhamento de mapas de característica entre as camadas através das conexões salteadas. A última camada da X-Net usa convoluções 1×1 e função de ativação sigmoide para a classificação pixel a pixel.

Os métodos propostos pelo presente trabalho utilizam FCNs baseadas na U-Net e X-Net, aplicando os mesmos princípios acima mencionados. Logo, elas contêm as fases de contração e expansão, e também usam as conexões salteadas. Entretanto, neste trabalho, foram propostas modificações em suas estruturas por meio da combinação de outros modelos de CNN com outras arquiteturas de blocos convolucionais, os quais serão explanados nas seções seguintes.

3.4.3 EfficientNet

Proposta por Tan e Le (2019), a família EfficientNet é conhecida pelas altas taxas de acerto alcançadas nos desafios ImageNet e ImageNet-V2 (RECHT et al., 2019). Essas arquiteturas surgiram a partir do estudo de um método de escalonamento de modelo que usa uma combinação de coeficientes para alterar a estrutura as redes em diversos aspectos, como número de filtros e de camadas convolucionais, diferenciando-as de abordagens tradicionais em que, comumente, esses parâmetros são definidos de maneira arbitrária ou através de testes. Esse escalonamento é dado por meio de um método heurístico baseado em *Grid Search* responsável pelo dimensionamento uniforme da largura e profundidade da rede, bem como dos mapas de características a partir de um conjunto fixo de coeficientes de escala, como mostra a Equação 3.6.

$$\begin{cases} d = \alpha^\phi \\ w = \beta^\phi \\ r = \gamma^\phi \end{cases} \quad (3.6)$$

onde d é a profundidade, w representa a largura, r a resolução; ϕ é um coeficiente definido pelo usuário, permitindo controlar quantos recursos estão disponíveis; α, β, γ especificam como atribuir esses recursos, respectivamente, à profundidade, largura e resolução da rede, devendo assumir valores maiores ou iguais a 1, respeitando a condição mostrada na Equação 3.7.

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \quad (3.7)$$

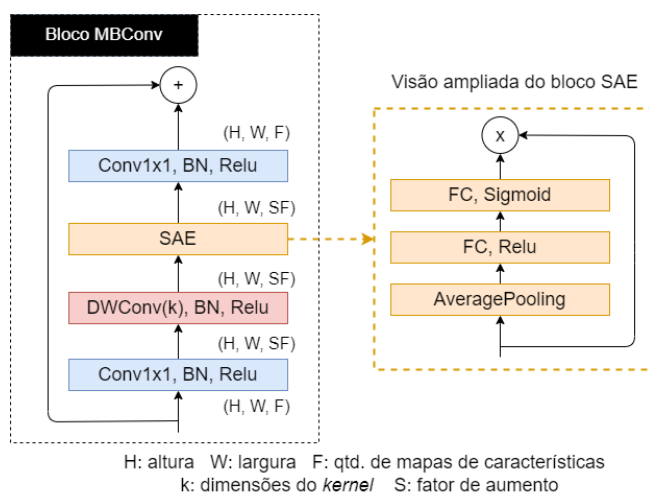
A estimação desses parâmetros pode ocorrer em dois passos principais: primeiro, definindo-se $\phi = 1$ e estimando-se α, β , e γ ; ou, considerando os valores α, β , e γ como constantes e estimar diferentes valores de ϕ . Através desses processos, surgem as redes da família EfficientNet, nomeadas de B0 até B7 (TAN; LE, 2019).

Em relação à estrutura das EfficientNets, elas são compostas por blocos denominados MBConv caracterizados por uma combinação entre blocos chamados de *Inverted Bottlenecks* (IB) (SANDLER et al., 2018) e blocos Squeeze-and-Excitation (SAE) (HU et al., 2018). Os blocos IB fazem uso de convoluções em profundidade (DWConv - *Depth-wise convolutions*) como uma alternativa à convolução padrão para reduzir custo computacional, visto que DWConv possuem menos parâmetros a serem ajustados (HOWARD et al., 2017). A Figura 11 mostra a composição de um bloco MBConv.

Inspirando-se nos blocos IB, os blocos MBConv recebem entradas de dimensões $H \times W \times F$, representando, respectivamente, altura, largura e o número de mapas de características. Através da primeira convolução 1×1 , é produzida uma quantidade maior de mapas de características. Esse aumento leva em consideração um fator de aumento S que multiplica F . Em seguida, usa-se a DWConv que gera também uma nova quantidade aumentada de mapas, os quais servem de entrada ao bloco SAE.

Blocos SAE visam a ponderação dos mapas de características. Dessa forma, atribui-se pesos maiores ou menores aos referidos mapas de acordo com sua relevância ao processo de aprendizado. Isso se torna uma estratégia importante, haja vista a quantidade ampliada de mapas produzidos internamente nos blocos MBConv. Por fim, a última

Figura 11 – Bloco MBConv, com uma visão ampliada do bloco Squeeze-and-Excitation (SAE).



Fonte: Elaborada pelo autor.

convolução 1x1 diminui o número de mapas para o seu valor inicial.

Então, por meio da composição dos blocos MBConv, assim como do processo de estimação dos parâmetros, a família EfficientNet alcançou resultados promissores, e se tornou uma referência no campo do aprendizado profundo.

3.4.4 Arquiteturas de Blocos Convolucionais

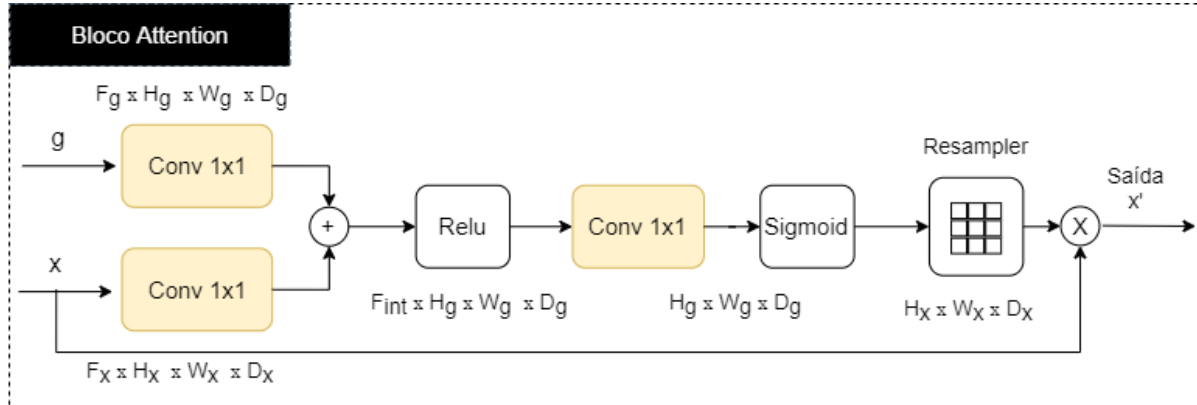
Além dos modelos U-Net, X-Net e EfficientNet apresentados anteriormente, neste trabalho também são utilizadas outras duas arquiteturas de blocos convolucionais como mecanismos para o aprimoramento dos modelos em busca de melhores resultados. São eles os blocos *Attention* e *Inception*.

3.4.4.1 Bloco *Attention*

Os blocos *Attention* foram inicialmente propostos por Oktay et al. (2018). Esses blocos têm como ideia principal a atribuição de pesos para as regiões internas dos mapas de características que apresentam maior relevância no processo de aprendizado. O funcionamento desses blocos é mostrado pela Figura 12. Neles, as entradas g e x são, respectivamente, a saída de uma camada anterior da rede, e a conexão saltada. Ambas passam por convoluções 1x1 o que faz com que elas tenham o mesmo número de canais. Depois, elas são somadas e passam pela ativação ReLU, seguidas de outra convolução 1x1 com ativação Sigmoid. Isso gera um único canal, composto de valores entre 0 e 1, que ponderam as características de acordo com a relevância que elas têm em cada parte do

mapa. Esse mapa ponderado, chamado de *attention map*, é multiplicado com a entrada x oriúnda de uma conexão salteada, produzindo, portanto, a saída do bloco Attention.

Figura 12 – Bloco Attention.



Fonte: Adaptada de (OKTAY et al., 2018)

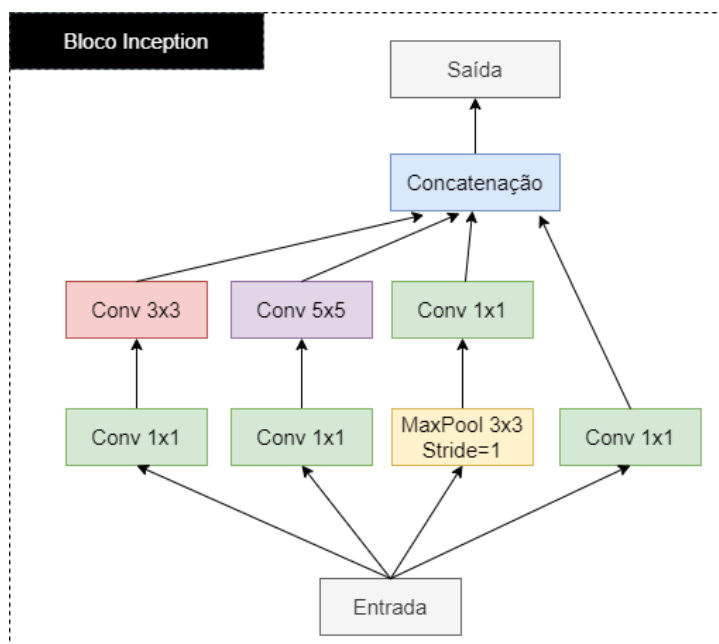
3.4.4.2 Blocos *Inception*

Propostos por Szegedy et al. (2015), os blocos *Inception* visam a geração de mapas de características em diferentes escalas. A estrutura interna desses blocos *Inception* pode ser conferida na Figura 13. As entradas são processadas paralelamente por convoluções com *kernels* de tamanho 1×1 , 3×3 e 5×5 , além de MaxPooling 3×3 . O resultado é concatenado na saída do bloco e passado adiante.

A estrutura composta por *kernels* de diferentes dimensões proporciona aprendizado de características em várias escalas, uma vez que os *kernels* pequenos são capazes de identificá-las em regiões menores, assim como os de tamanho maior focam em áreas maiores, podendo, dessa forma, analisar as características espaciais de modo mais amplo.

Um aprimoramento desses blocos foi proposto por Szegedy et al. (2016), e consiste na combinação da arquitetura *Inception* com conexões residuais (HE et al., 2016). Esses blocos são chamados de *Residual Inception*, e a principal diferença entre estes e o modelo tradicional é que a concatenação de filtros antes da saída é substituída pela conexão residual. Com isso, os autores observaram que não só o problema de degradação (ou desaparecimento) do gradiente causado pela profundidade das redes profundas é evitado, mas também o tempo de treinamento é reduzido.

Os blocos *Residual Inception* constituem a família de CNNs chamadas de InceptionResnet, dentre as quais a InceptionResnet-v2 obteve os resultados mais promissores

Figura 13 – Bloco *Inception*.

Fonte: Elaborada pelo autor

em experimentos com imagens do desafio ImageNet. Essa rede possui 164 camadas, e seu desempenho no referido desafio mostrou um aprendizado de características bastante amplo, com tempo de treinamento inferior às redes *Inception* tradicionais (SZEGEDY et al., 2016).

3.5 Métricas de Avaliação

A avaliação do desempenho de um método objetiva promover a validação dos resultados obtidos. Esta seção, portanto, apresenta as métricas utilizadas para quantificar o desempenho obtido pelos métodos propostos.

A avaliação da similaridade entre a segmentação gerada e as anotações manuais presentes na base de dados é feita por meio das seguintes métricas: Coeficiente Dice, Índice de Jaccard, Sensibilidade e Precisão. O coeficiente de similaridade Dice é o mais utilizado para a validação de segmentação em imagens médicas (TAHA; HANBURY, 2015). Ele é definido conforme Equação 3.8, onde X e Y representam as máscaras comparadas para cada imagem de entrada, no caso, a anotação manual e o mapa de segmentação gerado pelo processo de segmentação automática.

$$DICE = \frac{2 * |X \cap Y|}{|X| + |Y|} \quad (3.8)$$

O índice de Jaccard é dado pela razão entre a interseção dos conjuntos X e Y e a união desses (Equação 3.9). Esse índice também é conhecido como Interseção sobre a União (IoU, sigla em inglês). Dice e Jaccard são métricas comumente recomendadas para a avaliação métodos de segmentação de imagens médicas (BLAND, 2015).

$$IoU = \frac{|X \cap Y|}{|X \cup Y|} \quad (3.9)$$

A avaliação através das métricas de Sensibilidade e Precisão considera o cenário da classificação em nível de pixel. Neste caso, os pixels podem assumir um dos quatro estados possíveis listados a seguir:

- Verdadeiro Positivo (VP): o pixel pertence à estrutura de interesse e teve classificação correta;
- Falso Positivo (FP): o pixel não pertence à estrutura de interesse, mas foi classificado como se pertencesse;
- Verdadeiro Negativo (VN): o pixel não pertence à estrutura de interesse e foi classificado corretamente;
- Falso Negativo (FN): o pixel pertence à estrutura de interesse porém foi classificado como se não pertencesse.

Posto isso, a sensibilidade (Equação 3.10) é a proporção de pixels corretamente classificados como pertencentes à estrutura de interesse. E a precisão (Equação 3.11) relaciona verdadeiros positivos e falsos positivos.

$$SEN = \frac{VP}{VP + FN} \quad (3.10)$$

$$PRC = \frac{VP}{VP + FP} \quad (3.11)$$

Por fim, para avaliar os processos de extração de regiões de interesse, são usadas as métricas Erro Médio Absoluto (MAE) (SAMMUT; WEBB, 2010) e *Mean Average Precision* (mAP) (LIU; ÖZSU, 2009). Neste trabalho, o MAE (Equação 3.12) é usado para calcular a similaridade entre *bounding boxes*. No caso, a comparação é feita entre as *bounding boxes* geradas automaticamente pelos métodos e aquelas obtidas manualmente (*ground-truth*).

$$MAE = \left(\frac{1}{n}\right) \sum_{i=1}^n |y_i - x_i| \quad (3.12)$$

onde x_i e y_i são, respectivamente, as i -ésimas *bounding boxes* do *ground-truth* e dos conjuntos gerados pelos métodos.

Já o cálculo do mAP utiliza o conceito de IoU para estabelecer um conjunto de limiares, de modo que, se o IoU entre duas bounding boxes for igual ou superior a um limiar, será definido como positivo para aquele limiar; do contrário, será negativo. No final, serão totalizados os positivos e negativos para cada limiar do conjunto, e, assim, serão calculadas (também para cada limiar) as sensibilidades e precisões médias (AP) como mostra a Equação 3.13. O mAP consiste, portanto, na média das precisões médias.

$$AP = \sum_n (SEN_n - SEN_{n-1}) PRC_n \quad (3.13)$$

onde SEN_n e PRC_n são as sensibilidades e precisões médias para um n -ésimo limiar.

3.6 Testes de Hipóteses

Hipóteses podem surgir a partir da observação dos dados produzidos em um determinado experimento. Essas hipóteses podem ser aceitas ou rejeitadas a partir de algum fator determinante. Os testes de hipóteses, portanto, são procedimentos estatísticos que permitem a tomada de decisão sobre duas ou mais hipóteses levantadas, isto é, sua aceitação ou rejeição.

Neste trabalho, os testes de hipóteses realizados usam um método chamado de "teste Z" (ROMANO; LEHMANN, 2005). Através deste, é possível determinar se a diferença entre duas proporções é significativa. Essas proporções, no contexto do trabalho, referem-se às métricas calculadas para avaliar o desempenho dos métodos propostos. A abordagem, portanto, consiste em quatro etapas principais: (1) declarar as hipóteses, (2) formular um plano de análise, (3) analisar os dados da amostra, e (4) interpretar os resultados.

A declaração de hipóteses é o passo em que afirmações são feitas sobre a diferença entre as proporções analisadas. Nesse cenário, surge a hipótese nula, que deve ser aceita ou rejeitada, e a hipótese alternativa, que é decorrente da rejeição da hipótese nula. Por exemplo, seja a hipótese nula "há diferença significativa entre duas proporções". Caso ela

seja rejeitada, a interpretação, por conseguinte, é de que "não há diferença significativa", sendo esta a hipótese alternativa.

A formulação do plano de análise serve para descrever como usar os dados da amostra para rejeitar ou aceitar a hipótese nula. Nesse passo, é definido o nível de significância α , que pode assumir valores entre 0 e 1. O nível recomendado é $\alpha = 0,05$. O método de teste usa os dados da amostra para calcular a proporção da amostra combinada (PAC), o erro padrão (EP), o escore z (EZ), e o valor p (p) (Equações 3.14 a 3.17, respectivamente). Outros parâmetros são as duas proporções a serem analisadas (P1 e P2) e os tamanhos de cada amostra (N1 e N2).

$$PAC = \frac{P1 \times N1 + P2 \times N2}{N1 + N2} \quad (3.14)$$

$$EP = \sqrt{PAC \times (1 - PAC) \times \left(\frac{1}{N1} + \frac{1}{N2}\right)} \quad (3.15)$$

$$EZ = \frac{P1 - P2}{EP} \quad (3.16)$$

$$p = 2 \times (1 - Norm.Dist(EZ)) \quad (3.17)$$

onde *Norm.Dist* é a distribuição normal.

Por fim, usa-se o valor p e o nível α na etapa de interpretação dos resultados. Portanto, se $p < \alpha$, a hipótese nula é aceita, indicando há uma diferença significativa entre as proporções analisadas. Senão, essa hipótese será rejeitada, levando à aceitação da hipótese alternativa.

3.7 Considerações Finais

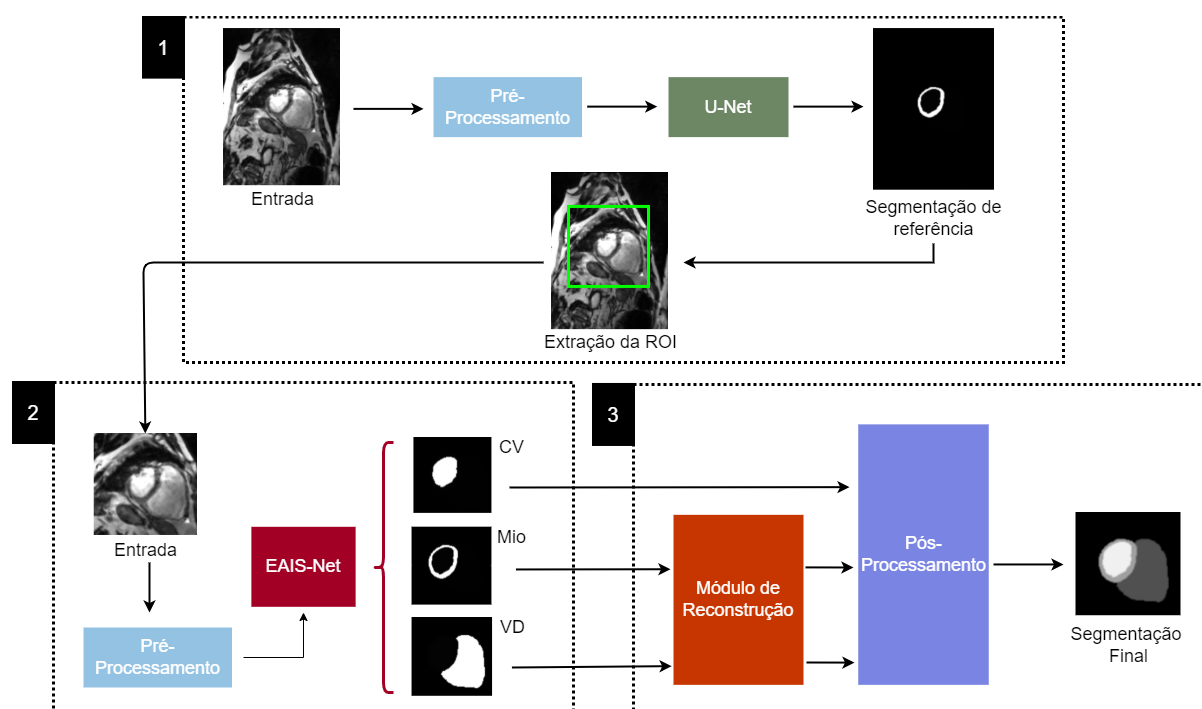
Neste capítulo, foram apresentados os conceitos e demais fundamentos teóricos imprescindíveis para a compreensão do método proposto e de sua aplicação. Abordaram-se tópicos como o exame de cine-RM cardíaca e sua potencial utilidade para a obtenção mais precisa de diagnósticos; técnicas de processamento e análise de imagens digitais; tópicos em aprendizado profundo, como as FCN, a Efficient Net, e arquiteturas de blocos convolucionais; e, por fim, as métricas de avaliação usadas nos experimentos, e uma explanação sobre os testes de hipóteses. No próximo capítulo, portanto, serão apresentados os materiais utilizados e os métodos propostos por este trabalho.

4 MÉTODO DE SEGMENTAÇÃO EM CASCATA COM RECONSTRUÇÃO

Este capítulo descreve o primeiro método proposto para a segmentação de estruturas cardíacas em imagens de cine-RM de eixo curto. Primeiro, será feita uma explanação sobre o referido método, e, em seguida, sobre os experimentos realizados com a apresentação dos resultados obtidos.

Este método pode ser visualizado de forma mais ampla pela Figura 14. Ele segue uma abordagem em cascata que é composta por três etapas principais. A primeira delas visa reduzir o escopo da imagem através do recorte de uma região menor que envolva as estruturas de interesse que são a cavidade ventricular (CV), miocárdio (Mio) e ventrículo direito (VD). A segunda tem como objetivo produzir as segmentações iniciais dessas estruturas que serão passadas para a terceira etapa, responsável pelo processo de refinamento com o uso do módulo de reconstrução e de pós-processamentos. Em virtude dessas características, esse método é doravante chamado de Segmentação em Cascata com Reconstrução (SCR). Uma melhor explicação de suas etapas é feita na sequência.

Figura 14 – Visão geral do método SCR e suas etapas: extração da ROI (1), segmentação inicial (2) e refinamento (3).



Fonte: Elaborada pelo autor.

4.1 Extração da ROI

A primeira etapa do método SCR visa a extração de uma região menor (ROI) que envolva os objetos de interesse da segmentação (CV, Mio e VD). Pretende-se, dessa forma, reduzir o escopo de cada volume, eliminando regiões de fundo cujas características podem interferir no aprendizado do modelo de segmentação.

Esse processo procura também diminuir o custo computacional da solução proposta, uma vez que evita lidar com as fatias em seu tamanho original. Outra vantagem é que ao eliminar mais regiões de fundo, mitiga-se o desbalanceamento entre as classes de pixels, problema bastante comum no processamento de imagens médicas (GAO et al., 2020).

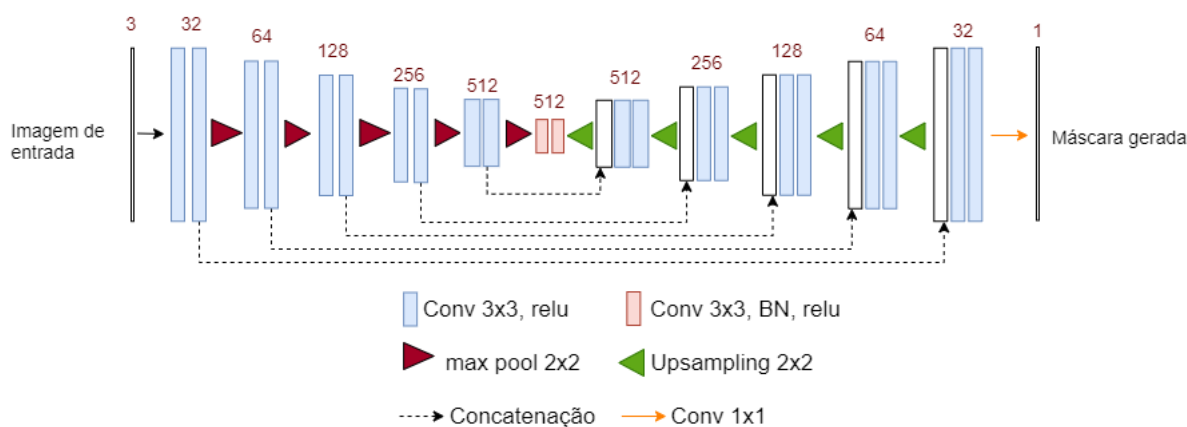
Para a extração da ROI, submete-se cada fatia do exame a uma U-Net (RONNEBERGER et al., 2015). Essa rede produzirá máscaras que servirão como referência para a localização da *bounding box* que melhor delimite os objetos de interesse, a partir da qual será extraída a ROI. Esses processos serão detalhados na sequência.

Antes de serem submetidas à U-Net, as fatias passam por um pré-processamento que consiste em três procedimentos: redimensionamento, eliminação de picos de intensidade e a normalização. Primeiro, as fatias são redimensionadas para 160×160 pixels por questões de limitações de hardware. Em seguida, eliminam-se os picos de intensidade, que são ruídos inerentes à cine-RM. Para isso, aplica-se uma técnica similar à que foi proposta por Nasr-Esfahani et al. (2018). Esse processo consiste em definir, para cada fatia, um teto correspondente a 70% da maior intensidade. Os pixels que ultrapassam esse teto recebem como valor o do referido teto. Por fim, o último procedimento é a normalização (Seção 3.3.1), que busca padronizar os valores de pixel entre 0 e 1.

A U-Net utilizada para gerar as segmentações iniciais possui uma arquitetura simples (Figura 15), tendo o seu caminho de contração baseado na VGG16 (SIMONYAN; ZISSERMAN, 2014), com 5 níveis de blocos convolucionais seguidos por operações de *Max Pooling* 2D com *kernel* 2×2 . Os blocos convolucionais são compostos por duas camadas de convolução 2D com função de ativação relu (NAIR; HINTON, 2010). *Batch Normalization* é utilizado nos blocos mais profundos por ser um recurso importante para melhorar a generalização do aprendizado e acelerar a convergência do modelo (SANTURKAR et al., 2018).

O caminho de expansão também possui 5 níveis de convolução, sendo que os 4

Figura 15 – Arquitetura U-Net utilizada para a geração das segmentações de referência.



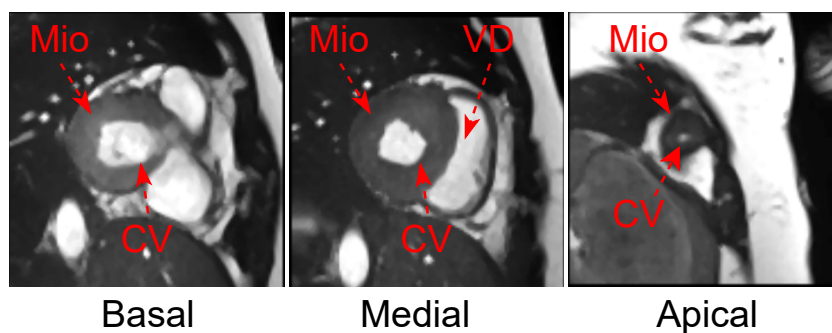
Fonte: Elaborada pelo autor.

primeiros são iniciados por camadas de ampliação (*Upsampling*) de $kernel\ 2 \times 2$, seguidos por duas camadas de convolução 3×3 com ativação ReLU. Os mapas de características de cada bloco do caminho de expansão são concatenados ao do nível correspondente no caminho de contração. Dessa forma, é possível recuperar características possivelmente perdidas por causa do *Max Pooling* (e.g informações espaciais) e ampliar a quantidade de mapas nas últimas camadas do modelo. A camada final é uma convolução 1×1 que usa a função de ativação Sigmoid para classificar os pixels entre 0 e 1, esse último representando o objeto de interesse.

Para o treinamento da U-Net, foram selecionadas apenas as fatias mediais dos exames de cine-RM. Essa escolha se deve pelo fato de as estruturas de interesse sempre estarem visíveis nas fatias mediais, o que nem sempre ocorre entre as do tipo apical e basal. Um exemplo pode ser visto na Figura 16. Além da visibilidade, as estruturas se apresentam de forma mais definida em relação às demais regiões da imagem, possibilitando que a segmentação produza resultados mais acurados nestes casos. Ademais, a rede foi treinada para gerar máscaras somente do miocárdio ao invés das três estruturas, porque, nas fatias mediais, o Mio e a CV estão mais centralizados em comparação ao ventrículo direito. Ressalta-se também que o processo de aprendizado se torna menos complexo visto que será segmentada somente uma região.

Sendo assim, para um dado volume, todas as fatias serão submetidas à U-Net. E como resultado, essa rede gerará máscaras somente para as fatias mediais. Essas máscaras são chamadas de segmentações de referência, pois, a partir delas, será definida a ROI a ser extraída.

Figura 16 – Visibilidade das estruturas de interesse nas fatias basal, medial e apical.



Fonte: Elaborada pelo autor.

Como essas segmentações podem apresentar alguns ruídos, é necessário removê-los. Para isso, aplica-se uma erosão morfológica de elemento estruturante retangular de tamanho 3×3 . Após esse procedimento, busca-se entre as segmentações de referência geradas aquela que possui a maior área, pois, a partir do seu centro, delimita-se uma *bounding box* cujas dimensões variam de acordo com o tamanho original da fatia antes do redimensionamento aplicado para a entrada na U-net. Por meio de testes, foram estabelecidos os seguintes tamanhos de *bounding box*: 120×120 para fatias com tamanho original inferior a 200×200 ; 140×140 para aquelas cujo tamanho é menor que 250×250 ; e 160×160 quando não se enquadram nesses casos. Por fim, a *bounding box* delimitada será replicada por todo o volume, servindo como base para a extração da ROI.

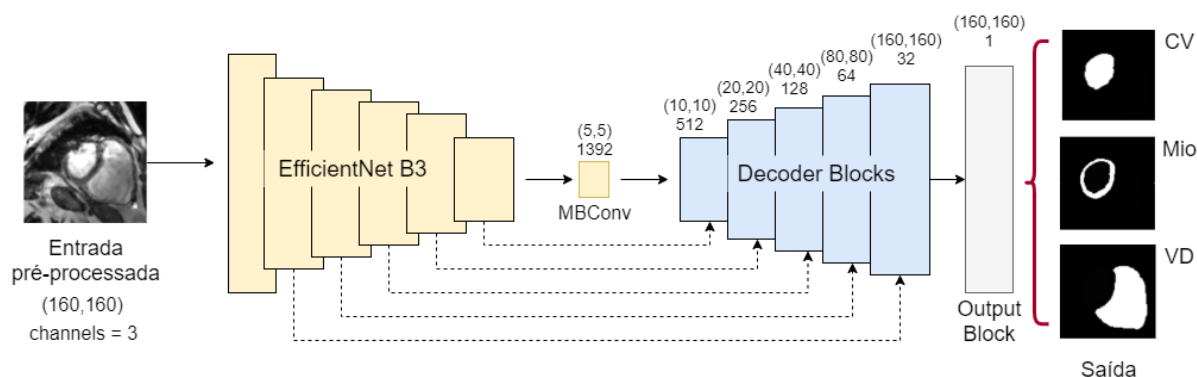
Ao término desse processo, as ROIs extraídas são passadas como entrada para a segunda etapa, que é responsável pela geração da segmentação inicial das estruturas cardíacas.

4.2 Segmentação Inicial

As ROIs extraídas das fatias pelo processo anterior são submetidas a uma FCN para gerar as máscaras referentes à cavidade ventricular (CV), miocárdio (Mio) e ventrículo direito (VD). Essa rede, chamada EAIS-Net, possui os caminhos de contração e expansão, além de conexões salteadas assim como a U-Net. Contudo, ela apresenta uma estrutura modificada, que é baseada na EfficientNet, e nos blocos Attention, Inception e SAE (o que dá origem à sigla EAIS). Uma visão geral da arquitetura proposta é apresentada pela Figura 17.

A rede recebe como entrada as ROIs extraídas pela etapa anterior. Essas fatias

Figura 17 – EAIS-Net: FCN proposta para a segmentação inicial das estruturas cardíacas no método SCR.



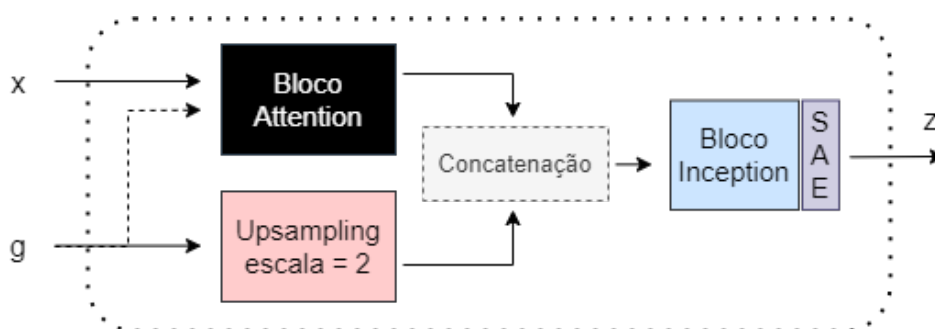
Fonte: Elaborada pelo autor.

recortadas são submetidas inicialmente a um pré-processamento. Primeiro, aplica-se *zero padding* para padronizar em 160×160 o tamanho das ROIs. Em seguida, são aplicados o janelamento e a normalização anteriormente mencionados na Seção 4.1. Por fim, as entradas pré-processadas terão dimensões 160×160 com três canais.

O processo de extração de características no caminho de contração é feito por uma EfficientNet B3 (TAN; LE, 2019). Utiliza-se essa rede em virtude das dimensões das fatias de entrada (160×160) e suas características. Como as dimensões de entrada não são muito grandes, o problema não exige uma rede muito profunda. Mas as fatias têm características complexas, sendo necessário usar uma rede robusta para extrair padrões de forma mais refinada. Isso foi observado através de testes com algumas redes da família EfficientNet, dentre as quais a B3 foi a que melhor se adequou ao problema.

No caminho de expansão da EAIS-Net, são utilizados 5 *decoder blocks* compostos por blocos Inception (Seção 3.4.4.2) com injeção de blocos SAE (Seção 3.4.3), seguidos por *Upsampling* 2×2 , sempre reduzindo pela metade a quantidade de filtros a serem gerados (de 512 a 32). De maneira similar à U-Net, a EAIS-Net utiliza conexões salteadas de modo a concatenar características extraídas no caminho de contração às novas geradas no caminho de expansão. Os *decoder blocks* contam também com a utilização de blocos *Attention* (Seção 3.4.4.1). A Figura 18 ilustra a estrutura interna de um *decoder block*, bem como o referido processo de concatenação. Seja g a saída de um bloco anterior e x a conexão salteada oriunda do caminho de expansão. Ambas são enviadas para o bloco *Attention*, e g é submetida ao *Upsampling* de fator de escala igual a 2. Por fim, a saída desses blocos é concatenada e passada como entrada para o bloco Inception.

Figura 18 – Estruturas internas do Decoder Block proposto.



Fonte: Elaborada pelo autor.

Como mencionado na Seção 3.4.4.1, os blocos *Attention* permitem que as regiões dos mapas de características transmitidos pelas conexões saltadas sejam ponderadas. Dessa forma, são atribuídos pesos maiores para as regiões mais relevantes dos mapas, enquanto que as de menor relevância para o processo de aprendizado da rede receberão pesos menores. Em relação aos blocos *Inception*, o seu uso visa o aprendizado de características em várias escalas. Assim, busca-se identificar as características das áreas menores pertencentes às estruturas de interesse, assim como de áreas maiores, auxiliando na redução de falsos positivos no caso de regiões com características espaciais similares.

No final do caminho de expansão, a saída do último bloco *Inception* é submetida a convolução 1×1 com função de ativação *Sigmoid*, gerando, portanto, a máscara de segmentação para cada estrutura de interesse.

É importante ressaltar que nesta etapa do método, cada fatia de entrada é submetida a três modelos da EAIS-Net. Esses modelos foram treinados individualmente para gerar separadamente as máscaras da CV, VD e Mio. A partir de experimentos, foi verificado que essa abordagem apresenta resultados melhores em comparação ao uso de um modelo único treinado para gerar simultaneamente as três máscaras.

4.3 Refinamento

A terceira etapa do método SCR é chamada de Refinamento, e consiste no uso de técnicas para melhorar o resultado inicialmente produzido na etapa anterior com o uso da EAIS-Net para a geração de mapas de segmentação.

Como CV, Mio e VD possuem características diferentes entre si, principalmente em relação à forma e tamanho, a busca por um conjunto único de técnicas capazes de melhorar as três segmentações se torna uma tarefa exaustiva. Sendo assim, foram

aplicadas técnicas específicas para cada uma dessas estruturas, de modo a melhorá-las individualmente.

As técnicas utilizadas na etapa de refinamento são agrupadas em dois módulos: reconstrução e pós-processamento. As segmentações iniciais do Mio e VD são submetidas primeiro à reconstrução e depois ao pós-processamento. Já a segmentação inicial da CV é direcionada apenas para este último. A explicação de como cada módulo é executado é apresentada a seguir.

4.3.1 Módulo de Reconstrução

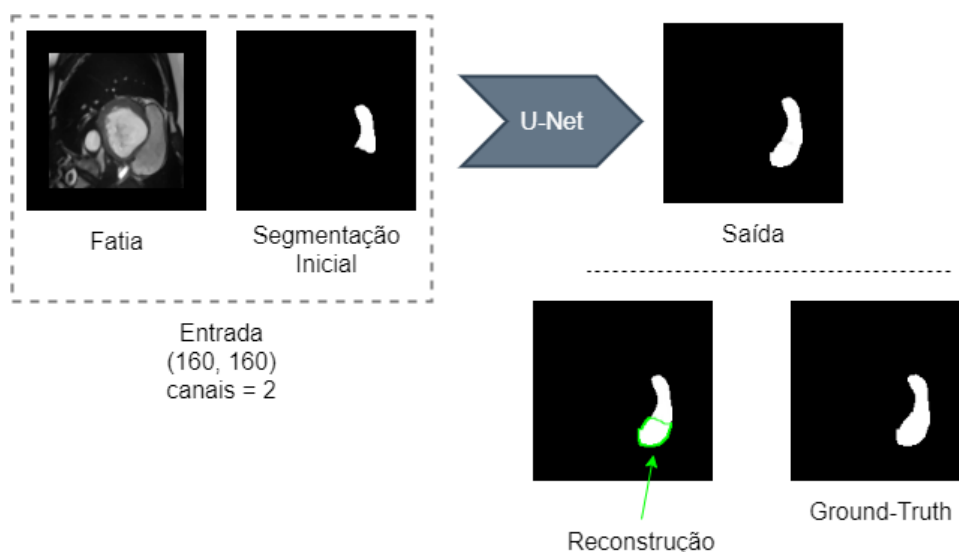
O módulo de reconstrução foi desenvolvido em busca do melhoramento das segmentações iniciais do Mio e VD. Ao percorrer o volume, observa-se que essas máscaras possuem formato e tamanhos diferentes, sendo grandes e bem definidas nas fatias mediais e mais irregulares nas fatias basais e apicais. Assim, o uso de técnicas tradicionais em processamento de imagens se torna mais complexo uma vez que muitas delas trabalham com limiares fixos e muitos parâmetros cuja estimação pode ser demorada e exaustiva. Além disso, é possível alcançar bons resultados em certas fatias, mas, ao mesmo tempo, comprometer a segmentação de outras.

Considerando esse cenário, o módulo de reconstrução então consiste no uso de uma rede convolucional profunda para produzir melhorias às segmentações iniciais. Inspirado por Souza et al. (2019), a ideia é que a rede aprenda as características das máscaras originais para corrigir as falhas apresentadas pelas predições anteriores. Dessa forma, os casos de erro podem ser melhorados com o cuidado para que os casos de sucesso não sejam piorados. O processo de reconstrução é apresentado pela Figura 19.

A rede de reconstrução recebe uma entrada com dimensões 160×160 e dois canais, representados pela fatia e sua segmentação inicial. A saída será o novo mapa de segmentação reconstruído. A arquitetura usada nesse processo é a mesma U-Net utilizada para na etapa de extração da ROI (Seção 4.1). Para o treinamento dessa rede, foi necessário construir uma base composta pelas fatias, suas máscaras originais e as máscaras preditas inicialmente. Para isso, os conjuntos de treino, validação e teste foram submetidos para o modelo inicial de segmentação (EAIS-Net) para a geração das máscaras que compõem a entrada da rede de reconstrução.

Assim sendo, foram treinados modelos específicos para cada estrutura (Mio e VD)

Figura 19 – Visão geral do processo de reconstrução.



Fonte: Elaborada pelo autor.

de modo que não houvesse a interferência entre as características, reduzindo também a complexidade do processo de aprendizado. Portanto, cada rede foca na reconstrução da sua respectiva estrutura de interesse.

4.3.2 Pós-Processamento

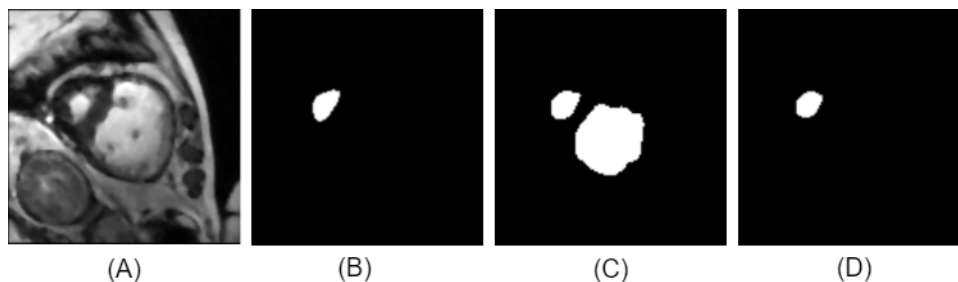
O módulo de pós-processamento consiste em conjuntos de técnicas aplicadas de forma específica para o melhoramento das segmentações iniciais da CV produzidas pela EAIS-Net, e também das segmentações geradas pelo módulo de reconstrução nos casos do Mio e VD.

Iniciando-se pela CV, verificou-se nos testes que essa estrutura obteve resultados consistentes de segmentação inicial, de modo que poucas técnicas foram suficientes para o seu melhoramento. As principais falhas detectadas foram a predição de falsos positivos em regiões das fatias com textura similar à CV, assim como algumas lacunas. Então, o pós-processamento das máscaras CV consiste na aplicação de: (1) filtro de área e (2) fechamento. O filtro de área é útil para remover regiões de falsos positivos. Em geral, eles possuem uma área maior que a CV, ocorrendo com maior frequência nas fatias apicais. Dessa forma, baseando-se no conhecimento *a priori* de que a CV é a menor área, o filtro é usado para mantê-la. Em seguida, aplica-se o fechamento para preencher lacunas internas. Essa operação usa elemento estruturante elíptico com dimensões 15×15 .

A Figura 20 mostra o processo de refinamento para uma segmentação inicial da

CV através do pós-processamento. Na Figura 20(A) e (B), são mostrados, respectivamente, a fatia e o *ground-truth*. Em 20(C) podem ser vistas duas regiões preditas, sendo que a menor corresponde à CV e a maior contém falsos positivos. Na Figura 20(D), pode ser visto o resultado do pós-processamento com a eliminação da região predita erroneamente.

Figura 20 – Refinamento de uma segmentação inicial da CV através do pós-processamento.



Fonte: Elaborada pelo autor.

Em relação às estruturas Mio e VD, as suas segmentações iniciais passam pelo módulo de reconstrução a fim de que sejam gerados melhores resultados. Esse processo é importante para reduzir a quantidade de casos de erro. Entretanto, mesmo após a reconstrução, algumas falhas oriundas da segmentação inicial podem persistir. Portanto, faz-se necessário aplicar o pós-processamento para que melhoramentos possam ser realizados.

Ao analisar as segmentações do Mio geradas pela reconstrução, foi verificado que as principais falhas remanescentes foram os alongamentos e as descontinuidades (Figura 21). Os alongamentos são extensões da máscara gerada composta por falsos positivos que ocorrem devido à similaridade de textura entre esses pixels e aqueles pertencentes ao Mio. E as descontinuidades são lacunas causadas por falsos negativos preditos, que desconectam regiões preditas corretamente.

Figura 21 – Exemplos de alongamentos e descontinuidades.

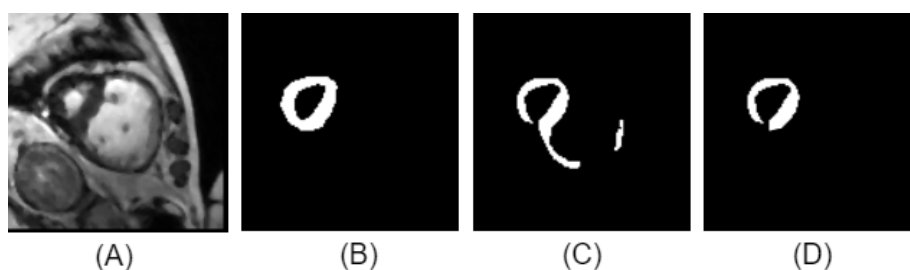


Fonte: Elaborada pelo autor.

Sendo assim, o pós-processamento aplicado para a correção dessas falhas consiste na execução dos seguintes procedimentos: (1) eliminação de alongamentos, (2) filtro de área e (3) correção de descontinuidades no espaço polar.

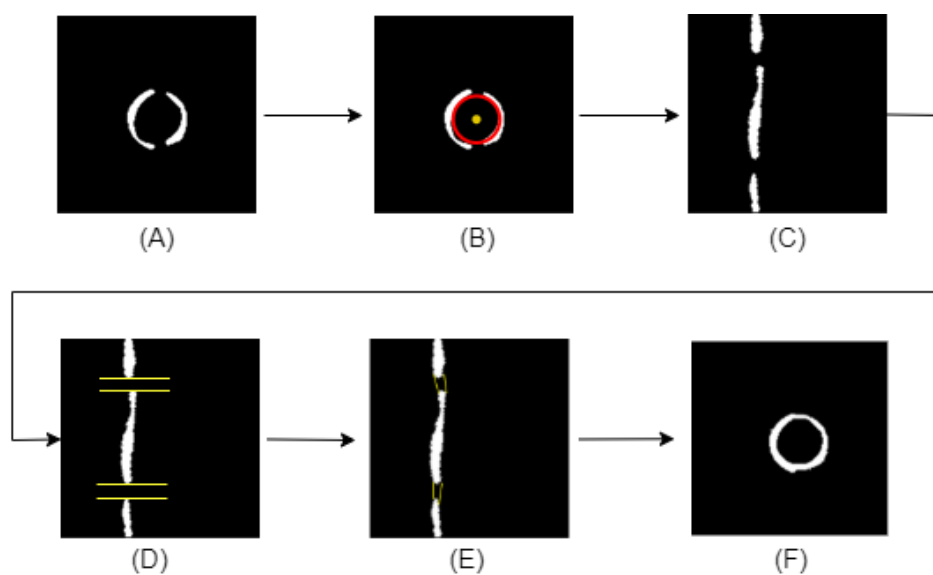
A eliminação de alongamentos consiste em, dada a máscara binária, calcular histogramas de projeção horizontal e vertical para eliminar as linhas ou colunas que contém menos de cinco pixels com intensidade igual a 1. Esse limiar foi definido através de experimentos. Em seguida, usa-se o filtro de área para recuperar o maior objeto. A Figura 22 demonstra esse processo.

Figura 22 – Eliminação de alongamentos: exemplo de fatia recortada (A), *ground-truth* (B), máscara antes (C) e depois (D) da eliminação de alongamentos.



Fonte: Elaborada pelo autor.

Figura 23 – Correção de descontinuidades: máscara de entrada (A), detecção de circunferência (B), máscara transformada para o espaço polar (C), detecção de descontinuidades (D), conexão entre pontos (E) e resultado final (F).



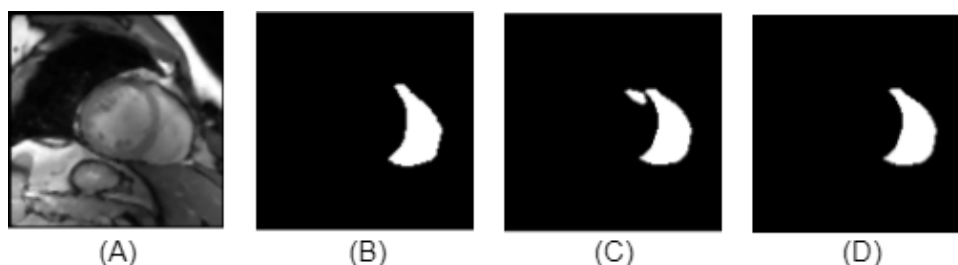
Fonte: Elaborada pelo autor.

E, por fim, esse maior objeto é submetido ao processo de correção de descontinuidades no espaço polar. A transformação para o espaço polar é um passo inspirado por

Tan et al. (2017). Contudo, os autores usam essa técnica como um pré-processamento anterior à segmentação. No presente trabalho, entretanto, usa-se como pós-processamento. A técnica consiste em, dado um centro e um raio, mapear cada pixel da máscara para uma nova imagem que os represente no espaço das coordenadas polares. O centro é obtido via transformada de Hough e o raio definido como a metade da largura da imagem ($160/2 = 80$ pixels). Após a obtenção da imagem polar, calcula-se o histograma das projeções horizontal para encontrar quais linhas não possuem pixel de valor 1, o que indica a descontinuidade. Também são identificadas as linhas anteriores e posteriores à descontinuidade. Em seguida, são encontrados os pontos extremos dessas linhas para então estabelecer a conexão entre elas. Esse processo é mostrado pela Figura 23. Por fim, a imagem é transformada do espaço polar para o cartesiano, finalizando o pós-processamento da máscara Mio.

Por fim, no caso da segmentação do VD, os problemas observados que persistiram após a reconstrução são menos complexos em relação às máscaras Mio. Assim, o pós-processamento visa eliminar pequenas regiões de falsos positivos e lacunas através do uso de (1) filtro de área e (2) fechamento morfológico. O primeiro recupera o maior objeto (VD) e o segundo é utilizado para preencher possíveis lacunas, como mostra a Figura 24.

Figura 24 – Pós-processamento de máscaras VD: exemplo de fatia recortada (A), seu *ground-truth* (B), máscara antes (C) e depois (D) o pós-processamento.



Fonte: Elaborada pelo autor.

E como passo final da etapa de refinamento, após serem aplicados os módulos de reconstrução e pós-processamento, é feita união entre as máscaras do Mio, CV e VD para produzir o mapa final de segmentação, sendo essa a saída gerada pelo método SCR.

4.4 Experimentos e Resultados

Nesta seção, serão apresentados os detalhes da base de imagens utilizada durante o desenvolvimento do método SCR, assim como os resultados alcançados durante os

experimentos. Destaca-se que o método foi implementado em linguagem Python 3.6¹, utilizando-se as bibliotecas Keras (CHOLLET et al., 2015) e Tensorflow². Os experimentos foram executados com as seguintes especificações de hardware: CPU Intel Core i5-6700HQ, 8GB de RAM, placa gráfica NVIDIA GeForce GTX 1050 e sistema operacional Windows.

Para avaliar a similaridade entre as máscaras de segmentação geradas pelos métodos propostos e o *ground-truth*, são utilizados o Coeficiente Dice, Índice de Jaccard (IoU), Sensibilidade (SEN) e Precisão (PRC) (explanadas na Seção 3.5).

4.4.1 Base de Imagens ACDC

Uma das bases de imagens utilizadas para validar o método SCR é a disponibilizada pelo desafio *Automated Cardiac Disease Diagnosis Challenge (ACDC)* (BERNARD et al., 2018). O conjunto possui 150 exames, divididos em 5 subgrupos, sendo um composto por pacientes saudáveis e os outros 4 por exemplares com as seguintes patologias: antecedentes de infarto do miocárdio, cardiomiopatia dilatada, cardiomiopatia hipertrófica e anormalidades no ventrículo direito.

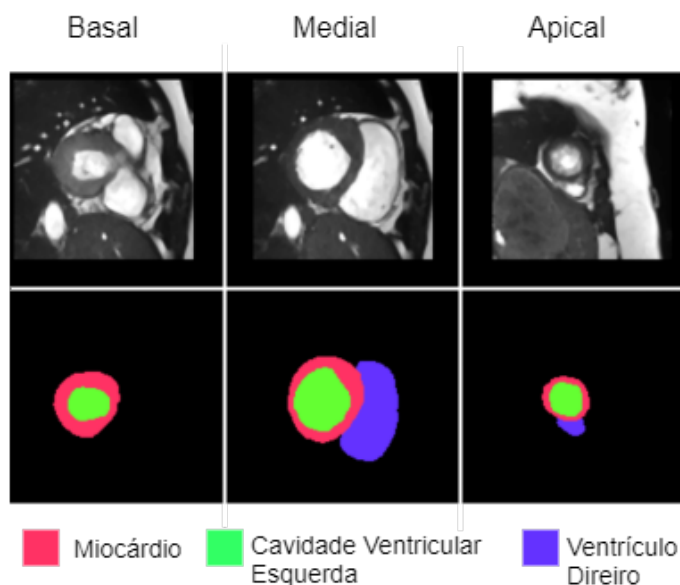
Os volumes de cine-RM cardíaco foram capturados no plano de eixo curto, em momentos nos quais os pacientes precisam segurar a respiração. Outros parâmetros da aquisição são: espessura de 5 ou 8 mm, distância inter-fatias de 5 ou 8 mm, e resolução espacial de 1,37 a 1,68 mm^2 /pixel. A Figura 25 mostra exemplos de fatias e suas anotações.

A base ACDC está dividida em dois conjuntos, um para treinamento e outro para teste. O conjunto de treinamento é composto por exames de 100 pacientes junto com as anotações manuais (*ground-truth*) feitas por especialistas durante as fases finais da sístole (ES - *end systole*) e da diástole (ES - *end diastole*). As anotações delimitam o ventrículo direito (VD), o miocárdio (Mio) e a cavidade ventricular esquerda (CV) por todo o volume do exame, caracterizado por três tipos de fatias: apicais, mediais e basais. Por fim, o conjunto de teste é composto por exames de 50 pacientes, e não possui *ground-truth*. Logo, para avaliar o método usando esse conjunto, foi preciso submeter as máscaras geradas à plataforma online do desafio ACDC.

¹ <https://www.python.org/>

² <https://www.tensorflow.org>

Figura 25 – Exemplos do dataset ACDC com suas respectivas marcações.



Fonte: Elaborada pelo autor.

4.4.2 Resultados com a Base ACDC

Como dito anteriormente, a porção da base ACDC usada nos experimentos possui exames e respectivas anotações (*ground-truth*) de 100 pacientes. Essa base foi separada aleatoriamente em três conjuntos: treino, validação e teste. O primeiro contém 60 pacientes, e os outros 20 pacientes cada. Dessa forma, é garantida uma maior quantidade de exemplos para o treinamento dos modelos de rede, favorecendo a generalização. Ressalta-se também que esses conjuntos, uma vez definidos, se mantiveram em todos os experimentos realizados, os quais são detalhados a seguir.

4.4.2.1 Extração da ROI

Conforme mencionado na Seção 4.1, a extração da ROI é baseada na segmentação do miocárdio (Mio). Nos experimentos relacionados a essa etapa, foram utilizadas duas U-Net (1 e 2). A U-Net 1 objetiva segmentar o Mio somente em fatias mediais e a U-Net 2 visa executar essa tarefa em todos os tipos de fatias. Dessa forma, é possível verificar a necessidade de um treinamento especializado para a geração das segmentações de referência a serem usadas no processo de extração da ROI.

Para o treinamento da U-Net 1, foram selecionadas somente as fatias mediais de cada exame. No entanto, esse processo reduz bastante a quantidade de exemplos de treinamento (apenas 120 fatias). Portanto, foi necessário o uso de *Data Augmentation*

para ampliar o conjunto de treino por meio das operações de rotação $[-5^\circ, 5^\circ]$, translações verticais e horizontais e *flip horizontal*, gerando 1080 fatias no total.

A U-Net 1 foi treinada por 300 épocas usando a função *Jaccard Loss* e o otimizador Adam (KINGMA; BA, 2014) com taxa de aprendizado inicial de 10^{-3} , fator de caimento de 0,1 e *Early Stopping* (PRECHELT, 1998). Em relação à U-Net 2, foram usados os mesmos hiperparâmetros da U-Net 1. Porém, o treinamento dessa rede foi realizado com a base completa, isto é, com todos os tipos de fatia, totalizando 1154 fatias. Ambas as redes possuem um total de 18,634,467 parâmetros treináveis.

O conjunto de teste usado para a avaliação contém apenas fatias mediais, totalizando 100 casos. A Tabela 2 apresenta os resultados obtidos para a segmentação do Mio como parte do processo de extração da ROI.

Tabela 2 – Comparação entre os experimentos realizados para a Extração da ROI

	Dice	IoU
U-Net 1	0,9368	0,8820
U-Net 2	0,9041	0,8261

A U-Net 1, que representa o treinamento especializado em fatias mediais, gerou melhores segmentações para o conjunto de teste, obtendo 0,9368 de Dice e 0,8820 de IoU. Já a U-Net 2 apresentou resultados menos favoráveis em virtude de ter sido treinada com uma base mais variada.

Outra análise foi realizada através das métricas Erro Médio Absoluto (MAE) (SAMMUT; WEBB, 2010) e Mean Average Precision (mAP) (LIU; ÖZSU, 2009), que são usadas para medir a similaridade entre as *bounding boxes* extraídas manualmente a partir do *ground-truth*, e aquelas geradas automaticamente com base nas segmentações de referência geradas pelas U-Nets 1 e 2. Os resultados podem ser vistos na Tabela 3. Ressalta-se que, nesta avaliação, o cálculo do mAP considera o limiar de $IoU \geq 0,80$. Dessa forma, as *bounding boxes* geradas são classificadas como verdadeiros positivos somente se o IoU em relação ao *ground-truth* for igual ou superior o limiar definido.

Tabela 3 – Resultados de MAE e mAP para a etapa de extração da ROI.

	Análise do MAE				
	mAP	MAE	Desv. Padrão	Max.	Min.
U-Net 1	0,90	2,93	4,27	18,25	0,00
U-Net 2	0,60	6,93	6,18	24,00	0,75

A abordagem de extração da ROI baseada nas segmentações geradas pela U-Net 1 obteve mAP de 0,90, e MAE de 2,93. Isso indica que as ROIs extraídas são bastante similares ao *ground-truth*. A análise do MAE também mostra um baixo desvio padrão, indicando resultados consistentes. Além disso, há casos em que o MAE é zero, ou seja, as ROIs extraídas automaticamente pela U-Net 1 são idênticas a aquelas extraídas com base no *ground-truth*.

É importante ressaltar que as fatias de entradas da etapa de extração da ROI são redimensionadas para 160×160 , e isso pode acarretar a alteração de algumas características e a perda de informação relevante. Contudo, os resultados mostram que esse redimensionamento não causou efeitos negativos na geração das *bounding boxes*.

A U-Net 2, treinada para segmentar o Mio em todas as fatias, obteve os resultados menos expressivos. Portanto, é possível concluir que o treinamento especializado em fatias mediais (U-Net 1) é a melhor abordagem para compor a etapa de extração da ROI. A alta similaridade com o *ground-truth* indica que todas as estruturas de interesse (Mio, CV, e VD) são preservadas dentro das ROIs, as quais servirão como entrada para as próximas etapas.

4.4.2.2 Segmentação Inicial

Para o processo de treinamento da EAIS-Net, foi utilizada uma versão da base de treino contendo somente as ROIs extraídas usando o *ground-truth* como referência. *Data augmentation* foi utilizado no conjunto de treino para ampliar a quantidade de exemplos de modo a evitar *overfitting* e favorecer a generalização. As operações aplicadas são as mesmas utilizadas para ampliar a base no experimento da etapa de extração da ROI. Como resultado, foram geradas 10.386 fatias para o conjunto de treinamento.

Conforme dito na Seção 4.2, foram treinados três modelos da EAIS-Net para gerar as segmentações do Mio, CV e VD separadamente. Cada treinamento ocorreu por 300 épocas usando a função Jaccard Loss e o otimizador Adam com taxa de aprendizado de 10^{-3} e fator de caimento de 0,1 acionado sempre que a *loss* se mantivesse estabilizada por mais de 10 épocas. O número de parâmetros treináveis de cada modelo da EAIS-Net é igual a 53.138.869.

Para o teste, são utilizados dois conjuntos: um com ROIs extraídas manualmente e outro resultante da etapa de extração da ROI, simulando o *pipeline* do método SCR.

Ambos são compostos pelos mesmos pacientes, totalizando 382 fatias.

A Tabela 4 apresenta os resultados alcançados pela EAIS-Net e mostra uma análise comparativa entre esta e outras quatro abordagens: a U-Net tradicional (RONNEBERGER et al., 2015) e quatro FCN cada uma combinando as EfficientNets B0, B3, B5 e B7 com um caminho de expansão baseado em convoluções padrão. A U-Net foi escolhida como método base para a comparação haja vista o seu amplo uso em segmentação de imagens médicas. Além disso, o método SCR utiliza essa rede para extração da ROI, o que motiva o seu teste também na presente etapa. Quanto aos experimentos envolvendo as FCN com convoluções padrão, estes têm o intuito de avaliar o uso isolado de cada EfficientNet a fim de verificar como seus desempenhos impactam o processo de segmentação.

Esses resultados se referem ao conjunto de teste com as ROIs obtidas automaticamente pela etapa de extração da ROI. O teste 3 mostra que o uso da EfficientNetB3 para a extração de características no caminho de contração melhorou significativamente os resultados de segmentação em comparação com a U-Net e com as demais EfficientNets. Isso pode ser visto para as três estruturas, dentre as quais a mais notável foi o ventrículo direito, em que o Dice e IoU aumentaram aproximadamente em 0,07 e 0,09 respectivamente. Esses resultados, portanto, justificam a sua escolha para integrar a EAIS-Net, que, no teste 6, no caso do VD, supera a U-Net em cerca de 0,12 e 0,14 de Dice e IoU respectivamente.

Em geral, a EAIS-Net apresenta os melhores resultados. É possível verificar um aumento considerável dos valores obtidos para a segmentação das três estruturas. Para a CV, a segmentação inicial via EAIS-Net obtém 0,9283 de Dice, 0,8866 de IoU, 0,9374 de Sensibilidade e 0,9299 de Precisão; representando um aumento médio de aproximadamente 3,8% em comparação com a U-Net.

No caso do Mio, a EAIS-Net obtém 0,8474 de Dice, 0,7672 de IoU, 0,8355 de Sensibilidade e 0,8657 de Precisão. Isso mostra um aumento médio de 4,8%. E para o VD, o resultado alcançou a EAIS-Net obtém 0,8306 de Dice, 0,7730 de IoU, 0,8515 de Sensibilidade e 0,8256 de Precisão (aumento médio de 12,8% em relação à U-Net).

Dessa forma, observa-se que, além do uso da EfficientNetB3, a combinação proposta entre os blocos Inception, Attention e SAE é mais promissora em relação aos blocos de convolução padrão usados pelas outras abordagens no caminho de expansão, como visto nos testes 1 e 2.

Foi realizada também uma análise comparativa entre o uso da EAIS-Net de forma

Tabela 4 – Segmentação das estruturas cardíacas: cavidade ventricular (CV), miocárdio (Mio) e ventrículo direito (VD). Comparação entre os resultados obtidos pela EAIS-Net e por outros métodos de segmentação.

Testes	Métodos		Dice	IoU	SEN	PRC
1	U-Net	CV	0,8903	0,8400	0,8867	0,9129
2	FCN (EfficientNetB0 + Expansão padrão)		0,0431	0,0224	0,0224	0,8289
3	FCN (EfficientNetB3 + Expansão padrão)		0,9209	0,8807	0,9318	0,9166
4	FCN (EfficientNetB5 + Expansão padrão)		0,8995	0,8558	0,9197	0,8958
5	FCN (EfficientNetB7 + Expansão padrão)		0,8998	0,8564	0,9080	0,8991
6	EAIS-Net		0,9283	0,8866	0,9374	0,9299
1	U-Net	Mio	0,8016	0,7002	0,8141	0,8074
2	FCN (EfficientNetB0 + Expansão padrão)		0,0114	0,0088	0,0089	0,3526
3	FCN (EfficientNetB3 + Expansão padrão)		0,8443	0,7608	0,8344	0,8711
4	FCN (EfficientNetB5 + Expansão padrão)		0,8262	0,7299	0,7779	0,8936
5	FCN (EfficientNetB7 + Expansão padrão)		0,8332	0,7435	0,8112	0,8659
6	EAIS-Net		0,8474	0,7672	0,8355	0,8657
1	U-Net	VD	0,7025	0,6249	0,7905	0,6695
2	FCN (EfficientNetB0 + Expansão padrão)		0,0571	0,0298	0,0298	0,9554
3	FCN (EfficientNetB3 + Expansão padrão)		0,7779	0,7175	0,8343	0,7586
4	FCN (EfficientNetB5 + Expansão padrão)		0,6331	0,5576	0,6687	0,6388
5	FCN (EfficientNetB7 + Expansão padrão)		0,6316	0,5621	0,6961	0,6151
6	EAIS-Net		0,8306	0,7730	0,8515	0,8256

individualizada, isto é, para segmentar separadamente cada estrutura de interesse, e outra abordagem em que essa rede visa a segmentação simultânea das estruturas (multiclasse). Para isso, na saída da rede, a função *Sigmoid* foi substituída pela *Softmax*, por ser mais adequada a problemas multiclasse. Essa análise é apresentada na Tabela 5. E conforme o experimento, a abordagem de treinamento especializado apresenta resultados superiores, visto que o aprendizado de cada rede é direcionado para uma estrutura específica. No caso da abordagem de segmentação multiclasse, percebe-se um desempenho piorado principalmente no caso do ventrículo direito. Com base nisso, optou-se pelo uso da

abordagem individualizada para compor a etapa de segmentação inicial do método SCR. Sendo assim, as fatias são passadas como entrada para três EAIS-Nets para que cada rede segmente sua respectiva estrutura.

Tabela 5 – Análise comparativa entre abordagens de uso da EAIS-Net.

Testes	EAIS-Net		Dice	IoU	SEN	PRC
1	Segmentação individualizada	CV	0,9283	0,8866	0,9374	0,9299
2	Segmentação multiclasse		0,9154	0,8742	0,9271	0,9120
1	Segmentação individualizada	Mio	0,8474	0,7672	0,8355	0,8657
2	Segmentação multiclasse		0,8454	0,7578	0,8553	0,8516
1	Segmentação individualizada	VD	0,8306	0,7730	0,8515	0,8256
2	Segmentação multiclasse		0,7649	0,7013	0,8399	0,7451

Conforme dito anteriormente, para verificar o impacto do processo de extração da ROI, foi efetuado outro experimento utilizando o conjunto de teste contendo os mesmos pacientes, porém com a ROI extraída de forma manual. Esses resultados podem ser vistos na Tabela 6. No geral, observa-se que o impacto da extração automática no processo de segmentação é mínimo, visto que a EAIS-Net alcança resultados similares, o que mostra uma boa interoperabilidade entre as etapas 1 e 2 do método SCR.

Tabela 6 – Experimentos com a EAIS-Net em dois conjuntos de teste: (C1) com as ROIs extraídas manualmente e (C2) ROIs extraídas pela etapa 1 do método SCR.

Métodos		Dice	IoU	SEN	PRC
C1	CV	0,9216	0,8810	0,9276	0,9215
C2		0,9283	0,8866	0,9374	0,9299
C1	Mio	0,8470	0,7648	0,8378	0,8643
C2		0,8474	0,7672	0,8355	0,8657
C1	VD	0,8324	0,7764	0,8477	0,8347
C2		0,8306	0,7730	0,8515	0,8256

Testes de hipóteses (Seção 3.6) foram realizados para avaliar a diferença estatística entre os resultados desse experimento. Os parâmetros utilizados foram o tamanho da amostra (382 fatias), e o nível de significância (α) de 0,05 (recomendado). A hipótese nula é de que há diferença significativa entre as abordagens. Portanto, usando o método de teste Z para as métricas Dice, IoU, SEN, e PRC, foram encontrados os valores p mostrados na Tabela 7. E, conforme pode ser visto, nenhum dos valores p são menores que o nível α . Logo, não há diferença significativa.

A rejeição da hipótese nula mostra que a inclusão de uma etapa automática para a extração das ROIs causa um baixo impacto aos resultados, uma vez que não há uma

Tabela 7 – Resultados dos testes de hipóteses: valores p encontrados para as métricas Dice, IoU, SEN, e PRC obtidos pelas abordagens semi-supervisionada e completamente automática. Nível $\alpha = 0,05$.

	Valores p			
	Dice	IoU	SEN	PRC
CV	0,7253	0,8092	0,5893	0,6580
Mio	0,9877	0,9376	0,9315	0,9548
VD	0,9470	0,9104	0,8832	0,7377

diferença significativa em comparação aos obtidos com a abordagem semi-supervisionada. Isso, portanto, corrobora a viabilidade da proposta de segmentação automática no cenário de aplicação, em que se busca uma solução computacional que minimize os esforços dos especialistas.

Por fim, é feita uma análise comparativa entre três cenários: (A) o método SCR (apenas com as etapas 1 e 2), (B) a EAIS-Net e (C) U-Net, os dois últimos para segmentar as fatias inteiras, dispensando a etapa 1 do referido método. Essa comparação é apresentada pela Tabela 8.

Tabela 8 – Análise comparativa entre o método SCR (A) que faz segmentação dividida em etapas e métodos de segmentação em etapa única via EAIS-Net (B) e U-Net (C).

Métodos		Dice	IoU	SEN	PRC
A		0,9283	0,8866	0,9374	0,9299
B	CV	0,8540	0,7784	0,7915	0,9476
C		0,8592	0,7837	0,7970	0,9532
A		0,8474	0,7672	0,8355	0,8657
B	Mio	0,4399	0,3177	0,3278	0,8943
C		0,4457	0,3213	0,3277	0,9097
A		0,8306	0,7730	0,8515	0,8256
B	VD	0,6309	0,5353	0,5607	0,7864
C		0,5975	0,4995	0,5201	0,7626

Nos cenários B e C, as dimensões das fatias são reduzidas para 160×160 devido às limitações de hardware. Os resultados alcançados são consideravelmente inferiores em relação ao método SCR (A), pois, de forma diferente deste, as outras redes processam a fatia inteira e, por isso, estão sujeitas a lidar com um grande desbalanceamento entre as classes. Esse desbalanceamento é ampliado com a redução das dimensões das fatias, pois esse processo diminui consideravelmente o tamanho das estruturas de interesse principalmente nas regiões apical e basal do exame onde elas são naturalmente pequenas.

Com isso, verifica-se que a redução de escopo provida pela etapa 1 auxilia na geração de melhores resultados de segmentação.

4.4.2.3 Refinamento: Reconstrução e Pós-Processamentos

A terceira etapa do método SCR é responsável pelo refinamento das segmentações iniciais através de dois módulos: reconstrução e pós-processamento. Portanto, esta seção apresenta os resultados obtidos após a aplicação dessas técnicas, sendo esses os resultados finais alcançados pelo método.

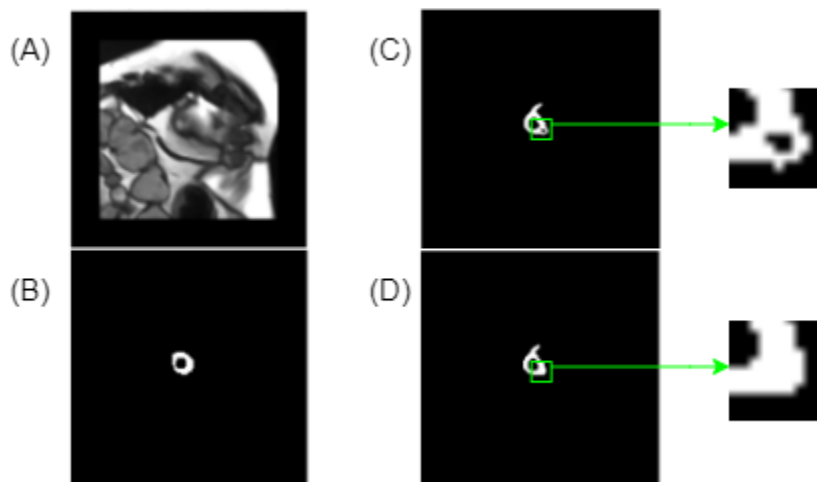
Como dito na Seção 4.3, as entradas para a U-Net usada no módulo de reconstrução são a ROI que delimita as estruturas de interesse e a sua máscara de segmentação inicial; e como saída, obtém-se uma máscara melhorada. Entradas e saídas possuem dimensões 160×160 . A rede de reconstrução foi treinada com os mesmos hiperparâmetros definidos para os treinamentos anteriores, ou seja, quantidade de épocas, otimizador e taxa de aprendizado. Essa arquitetura U-Net é similar à usada na etapa de extração da ROI, logo, o número de parâmetros treináveis é o mesmo. Mudou-se apenas a função de perda para a combinação Binary Crossentropy (BCE) + Jaccard Loss. A incorporação da BCE objetiva reforçar a penalização pela predição de falsos positivos e falsos negativos (JADON, 2020). Os resultados do módulo de reconstrução proposto são apresentados pela Tabela 9. Essa tabela também mostra outros dois experimentos baseados na abordagem proposta por Souza et al. (2019).

Tabela 9 – Resultados dos experimentos realizados com diferentes métodos para o módulo de reconstrução.

Exp.	Métodos		Dice	IoU	SEN	PRC
1	Souza et al. (2019)		0,8222	0,7477	0,7734	0,9025
2	U-Net + BCE	CV	0,9057	0,8473	0,8640	0,9647
3	U-Net + BCE + Jaccard Loss		0,9230	0,8812	0,9328	0,9238
1	Souza et al. (2019)		0,4216	0,2813	0,3195	0,6410
2	U-Net + BCE	Mio	0,8145	0,7144	0,7328	0,9271
3	U-Net + BCE + Jaccard Loss		0,8528	0,7727	0,8420	0,8700
1	Souza et al. (2019)		0,4388	0,3192	0,3364	0,6943
2	U-Net + BCE	VD	0,7962	0,7164	0,7486	0,8706
3	U-Net + BCE + Jaccard Loss		0,8304	0,7728	0,8523	0,8248

A rede proposta por Souza et al. (2019) é uma FCN baseada na ResNet (HE

Figura 26 – Exemplo de reconstrução do miocárdio: (A) fatia de entrada, (B) *ground-truth*, (C) segmentação inicial com Dice de 0,7368 e (D) segmentação após o módulo de reconstrução, com Dice de 0,7528.



Fonte: Elaborada pelo autor.

et al., 2016) e tem no final uma camada completamente conectada com quantidade de neurônios igual às dimensões da entrada. A função de perda é composta por dois termos BCE. Um é usado para analisar o fundo e o outro analisa a região predita em relação ao *ground-truth*. Por limitações de hardware, as fatias de entrada foram redimensionadas para 96×96 neste experimento; e as saídas produzidas nessas mesmas dimensões retornam depois ao tamanho original. Por fim, essa rede não obteve resultados satisfatórios para a reconstrução das estruturas cardíacas, visto que apresentou um desempenho inferior ao alcançado pela segmentação inicial sozinha. O segundo experimento consistiu em utilizar a U-Net com a mesma função de perda usada no experimento anterior. A U-Net apresentou resultados melhores em comparação à FCN do experimento 1, porém ainda insatisfatórios.

Ao trocar a função de perda pela composição BCE + Jaccard Loss, obteve-se no experimento 3 o melhor resultado geral para a reconstrução das estruturas. Observa-se que os melhoramentos ocorreram para as máscaras do Mio e do VD. Em relação à reconstrução de máscaras CV, todas as métricas sofreram queda e isso levou a descartar o uso desse processo de melhoramento para a referida estrutura. No caso do Mio, todas as métricas apresentaram elevações. E para o VD, as métricas Dice e Iou sofreram uma queda ínfima de 0,0002. Em contrapartida, houve o aumento da sensibilidade, indicando a predisposição do modelo a classificar pixels como sendo da região de interesse. A partir dessa análise, optou-se por usar o módulo de reconstrução somente para as estruturas Mio e VD. A Figura 26 mostra um exemplo do melhoramento provido pelo módulo de reconstrução.

Por fim, são aplicados os pós-processamentos específicos para cada estrutura cardíaca. A Tabela 10 mostra o resultado geral obtido pelo método SCR com o conjunto de teste passando por todas as três etapas. E a Tabela 11 mostra a avaliação de Dice e IoU da aplicação do método sobre o mesmo conjunto de teste, porém com as fatias divididas entre dois subconjuntos: fim da diástole (ED) e fim de sístole (ES). Esse modo de análise é similar ao feito pelo desafio ACDC.

Tabela 10 – Resultados alcançados pelo método SCR em dois cenários: execução das etapas 1 e 2; e a execução completa.

Método SCR		Dice	IoU	SEN	PRC
Etapas 1 e 2	CV	0,9283	0,8866	0,9374	0,9299
Etapas 1, 2 e 3		0,9303	0,8886	0,9416	0,9285
Etapas 1 e 2	Mio	0,8474	0,7672	0,8355	0,8657
Etapas 1, 2 e 3		0,8552	0,7747	0,8449	0,8730
Etapas 1 e 2	VD	0,8306	0,7730	0,8515	0,8256
Etapas 1, 2 e 3		0,8312	0,7740	0,8512	0,8269

Tabela 11 – Avaliação do método SCR a partir das métricas Dice e IoU para as fatias nas fases de fim da diástole (ED) e fim da sístole (ES).

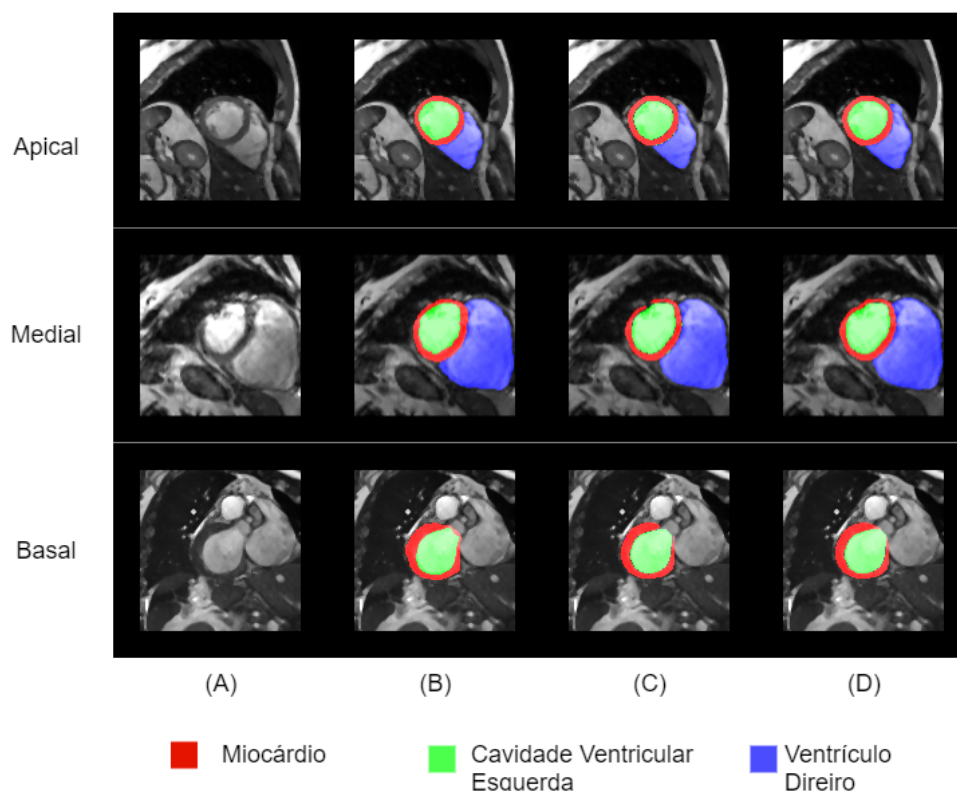
Método SCR			Dice	IoU
Etapas 1 e 2	CV	ED	0,9602	0,9248
Etapas 1, 2 e 3		ED	0,9602	0,9248
Etapas 1 e 2	Mio	ES	0,9246	0,8708
Etapas 1, 2 e 3		ES	0,9337	0,8795
Etapas 1 e 2	VD	ED	0,8870	0,7986
Etapas 1, 2 e 3		ED	0,8871	0,7986
Etapas 1 e 2	VD	ES	0,8875	0,8006
Etapas 1, 2 e 3		ES	0,8902	0,8041
Etapas 1 e 2	VD	ED	0,9262	0,8653
Etapas 1, 2 e 3		ED	0,9257	0,8646
Etapas 1 e 2	VD	ES	0,8517	0,7477
Etapas 1, 2 e 3		ES	0,8543	0,7520

As tabelas mostram os resultados obtidos com a execução das etapas 1 e 2 do método SCR em comparação com a sua aplicação completa. Verifica-se que, na avaliação geral, o uso do módulo de reconstrução combinado com os pós-processamentos impactou positivamente na geração de uma segmentação final melhorada. Isso se confirma pelo incremento significativo que pode ser observado para as métricas de avaliação.

Na análise de fatias agrupadas por fases cardíacas, observa-se que o método apresenta resultados consistentes para o grupo fim de diástole (ED), e se destaca no melhoramento das segmentações em fim de sístole (ES). Isso ocorre porque, nessa fase, as estruturas estão mais retraídas e menos evidentes, o que ocasiona as principais falhas corrigidas pela etapa de refinamento. Dessa forma, o melhoramento observado quantitativamente indica a importante contribuição dessa etapa para o método.

A Figura 27 apresenta os resultados qualitativos da segmentação final para os três tipos de fatias. Nessa figura é possível observar a ação da etapa de refinamento sobre a segmentação inicial de modo a melhorá-la, com destaque para a correção de descontinuidades nas máscaras do miocárdio mostrada nos casos medial e basal. O exemplo apical, por sua vez, mostra resultados bons apesar de um pequeno crescimento da região predita para o VD. Nesse caso, verifica-se uma semelhança entre a textura dessa região com outras adjacências, visto que elas estão mais próximas por causa do movimento de sístole. Mas, no exemplo basal, percebe-se que o método SCR considera não só aspectos texturais, mas também a forma das regiões, visto que, embora haja nessas fatias uma estrutura com textura similar ao VD, mas com formato diferente, não houve nesses casos a geração de falsos positivos, corroborando assim a sua capacidade de generalização.

Figura 27 – Resultados qualitativos do método SCR, sendo (A) a entrada, (B) o *ground-truth*, (C) a segmentação inicial e (D) a segmentação final.



Fonte: Elaborada pelo autor.

4.4.3 Base de Imagens M&Ms

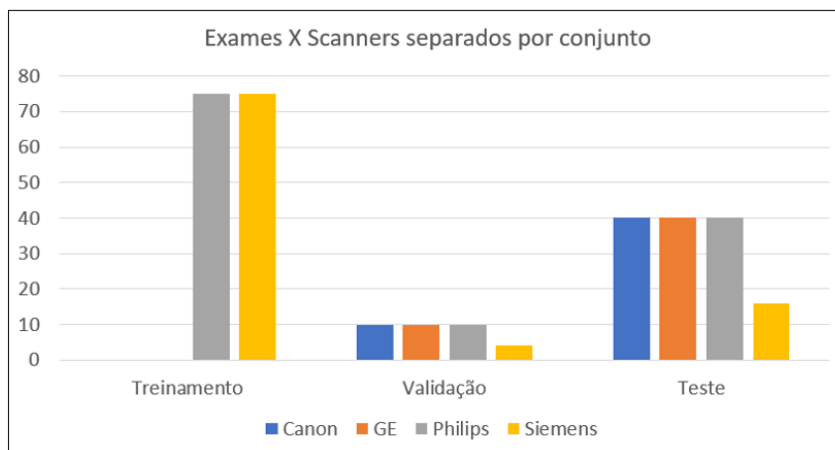
Outra base usada no processo de validação do método SCR é a disponibilizada pelo desafio *Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Image Segmentation Challenge* (M&Ms) (CAMPELLO et al., 2021b). A base M&Ms é composta por exames de cine-RM de eixo curto de 345 pacientes, sendo estes saudáveis ou acometidos pelas seguintes patologias: cardiomiopatias hipertrófica, dilatada e isquêmica; doença coronariana, anormalidades no ventrículo direito, e miocardite. Todos foram escaneados em centros clínicos de três países diferentes (Espanha, Alemanha e Canadá) usando quatro fornecedores distintos de scanners de ressonância magnética, a saber: Siemens, General Electric (GE), Philips e Canon. Isso, portanto, caracteriza uma maior diversidade em relação à aquisição das imagens.

Assim como no caso da base ACDC, a M&Ms contém exames extraídos nas fases cardíacas de fim de sístole e diástole. Do total disponibilizado, apenas 320 contêm as anotações das estruturas CV, Mio, e VD. A base com anotações então é dividida em três conjuntos: 150 pacientes para o treinamento; 34 para a validação; e 136 para o teste. As

fatias possuem dimensões de largura e altura variando, respectivamente, entre 196×240 e 512×512 .

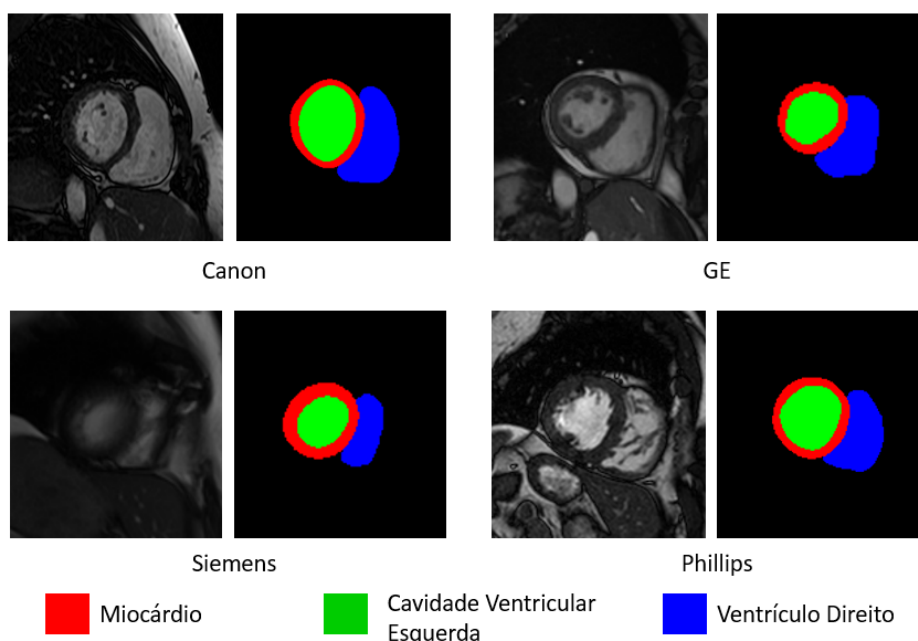
A Figura 28 mostra um gráfico que relaciona o número de exames por tipo de scanner agrupados por cada um dos conjuntos supracitados. Por fim, na Figura 29, são apresentados exemplos da base com suas anotações, cada qual referente a um scanner.

Figura 28 – Quantidade de exames por tipo de scanner contida nos conjuntos de treinamento, validação e teste da base M&Ms.



Fonte: Elaborada pelo autor.

Figura 29 – Base M&Ms: exemplos de fatias mediais obtidas pelos diferentes scanners utilizados na elaboração da base.



Fonte: Elaborada pelo autor.

4.4.4 Resultados com a Base M&Ms

Nesta seção, serão apresentados os resultados obtidos pelo método SCR nos testes com a base M&Ms. Ressalta-se que esses experimentos foram realizados após a consolidação do referido método. Portanto, essa seção enfatiza somente o resultado alcançado com a aplicação das técnicas escolhidas para compôr as etapas do método SCR.

Foram realizados dois experimentos. No primeiro, chamado de SCR_ACDC, utiliza-se o método SCR em uma abordagem de *transfer learning*, em que os modelos de rede são treinados com a base ACDC, e o teste é feito com a base M&Ms. No segundo, chamado de SCR_MMS, a base M&Ms é usada tanto no treinamento dos modelos, quanto no teste. É importante ressaltar que, nesse experimento, as redes foram treinadas com os mesmos hiperparâmetros apresentados na Seção 4.4.2. O mesmo *Data augmentation* foi aplicado ao conjunto de treinamento, que, de 3286 fatias, passou a totalizar 24616 fatias. Os conjuntos de validação e teste possuem, respectivamente, 806 e 3218 fatias. A Tabela 12 mostra o comparativo dos resultados obtidos.

Tabela 12 – Método SCR: resultados dos experimentos realizados com a base M&Ms.

Experimento	Estrutura	Dice	IoU	SEN	PRC
SCR_ACDC	CV	0,8137	0,7217	0,7958	0,7611
	Mio	0,6928	0,5483	0,5993	0,7932
	VD	0,7250	0,6050	0,6314	0,6776
SCR_MMS	CV	0,8453	0,7662	0,8652	0,8457
	Mio	0,7834	0,6709	0,7701	0,7899
	VD	0,7434	0,6320	0,7602	0,8266

Observa-se, de modo geral, que o desempenho do método SCR foi melhor no experimento SCR_MMS, em virtude do treinamento dos modelos com a própria base M&Ms. Essa base, conforme dito anteriormente, possui mais exames que a base ACDC, bem como uma maior variabilidade de características devido os diferentes *scanners* utilizados em sua construção. Logo, os modelos de rede do experimento SCR_ACDC, por terem sido treinados com a base ACDC, não contemplaram determinados aspectos dos exames do conjunto de teste, formado por exames da base M&Ms. Consequentemente, esse experimento obteve um resultado inferior.

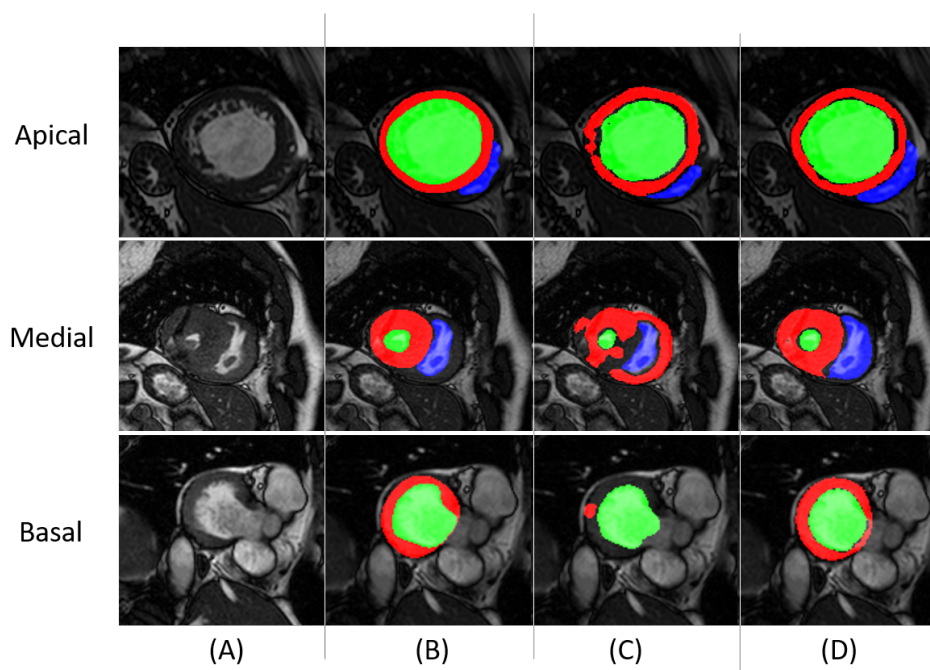
Na Tabela 13, é mostrada uma comparação entre os resultados de Dice e IoU separados pelas fases cardíacas ED e ES. Com a análise das métricas, é possível observar que o método SCR, em ambos os experimentos, apresenta um melhor desempenho na

segmentação da CV e do VD no caso das fatias em ED, quando essas estruturas estão mais aparentes nos exames devido à dilatação do coração. No caso do Mio, o método mostra resultados melhores nas fatias em ES, pois, dada a contração do coração, surgem os principais problemas mitigados pela etapa de refinamento, como as discontinuidades. Na Figura 30, são apresentados os resultados qualitativos para alguns exemplos de teste.

Tabela 13 – Resultados da aplicação do método SCR na base M&Ms separados por fase cardíaca.

Experimento	Estrutura	Dice ED	Dice ES	IoU ED	IoU ES
SCR_ACDC	CV	0,8552	0,7721	0,7812	0,6622
	Mio	0,6601	0,7245	0,6923	0,5084
	VD	0,7818	0,6683	0,7250	0,6761
SCR_MMS	CV	0,8704	0,8202	0,8052	0,7272
	Mio	0,7639	0,8022	0,6443	0,6966
	VD	0,7648	0,6963	0,6641	0,5733

Figura 30 – Resultados qualitativos do método SCR nos experimentos com a base M&Ms, sendo (A) a entrada, (B) o *ground-truth*, (C) o resultado do experimento SCR_ACDC e (D) o resultado do experimento SCR_MMS.



Fonte: Elaborada pelo autor.

No exemplo da fatia apical, o experimento SCR_ACDC alcança Dices de 0,9550 para a CV, 0,8072 para o Mio, e 0,6077 para o VD. Por sua vez, o experimento SCR_MMS obteve valores superiores para essa métrica, sendo eles 0,9624 (CV), 0,8969 (Mio), e 0,7689 (VD). Neste caso, observa-se que, no primeiro experimento, o método não segmentou

partes das estruturas de interesse, principalmente Mio e VD. Esse efeito é reduzido no caso do segundo experimento, pois, além da redução desses falsos negativos, as máscaras apresentam um formato com menos irregularidades.

O exemplo da fatia medial mostra um caso em que, no experimento SCR_ACDC, a etapa de refinamento não foi capaz de corrigir as falhas da segmentação inicial relacionadas ao Mio. Os Dices obtidos foram de 0,5147 (CV), 0,5076 (Mio), e 0,9344 (VD). No experimento SCR_MMS, os resultados de Dice da CV e Mio apresentaram melhorias, sendo, respectivamente, de 0,9123 e 0,8666. Entretanto, houve uma queda em relação ao VD, que apresentou Dice de 0,7397. Por se tratar de uma fatia em fase ES, o VD se apresenta muito contraído, e, neste exemplo, a segmentação gerada envolveu pixels ao entorno dessa estrutura.

Por fim, no exemplo basal, em ambos os experimentos, não foram geradas máscaras para o VD. Este exemplo mostra novamente um melhor desempenho do método no experimento SCR_MMS. Em comparação com o experimento SCR_ACDC, os Dices evoluíram de 0,9465 para 0,9484 (CV), e de 0,1224 para 0,8060 (Mio).

Testes de hipóteses também foram realizados para avaliar a diferença estatística entre os resultados dos experimentos. Os valores p calculados podem ser vistos na Tabela 14. Nos casos da CV e do Mio, os valores p são inferiores ao nível de significância ($\alpha = 0,05$), indicando, dessa forma, que há diferença significativa. Em relação ao VD, essa diferença não se mostra significativa quanto à métrica Dice, porém, isso não é observado no que toca às demais métricas para essa estrutura. Isso, portanto, confirma que o treinamento dos modelos com a base M&Ms acarretou a obtenção de melhores resultados.

Tabela 14 – Valores p encontrados para as métricas de avaliação do método SCR nos experimentos SCR_ACDC e SCR_MMS.

	Valores p			
	Dice	IoU	SEN	PRC
CV	0,0007	0,0000	0,0000	0,0000
Mio	0,0000	0,0000	0,0000	0,7436
VD	0,0935	0,0252	0,0000	0,0000

Apesar das diferenças observadas entre os resultados apresentados, nota-se que o método se mostrou capaz de produzir segmentações de modo satisfatório no contexto da base M&Ms, que tem um grau maior de heterogeneidade que aquele representado pela base ACDC.

4.5 Considerações Finais

O método de segmentação em cascata com reconstrução obteve resultados promissores, conforme mostram os experimentos com as bases ACDC e M&Ms. O processo de avaliação mostra que a ação integrada das técnicas desenvolvidas para cada etapa acarretaram a geração de máscaras com expressiva semelhança ao *ground-truth*.

Nos testes realizados com a base ACDC, nota-se que a etapa de extração da ROI impacta positivamente o método, pois a sua inclusão impõe melhoramentos expressivos aos resultados, ao passo que reduz problemas como o desbalanceamento de classes e o custo de processamento, além de ser uma alternativa ao redimensionamento das entradas, cuja aplicação pode comprometer o aprendizado de acordo com os resultados.

Além disso, os experimentos relativos à etapa de segmentação inicial mostram que a combinação dos mecanismos convolucionais que resultou na EAIS-Net é capaz de produzir resultados melhores que as convoluções padrão usadas na U-Net, que é um método comumente bem avaliado na segmentação de imagens médicas. Por fim, o módulo de reconstrução e os pós-processamentos usados também mostram a sua importância para se corrigir falhas na segmentação inicial, melhorando, assim, o resultado final do método SCR.

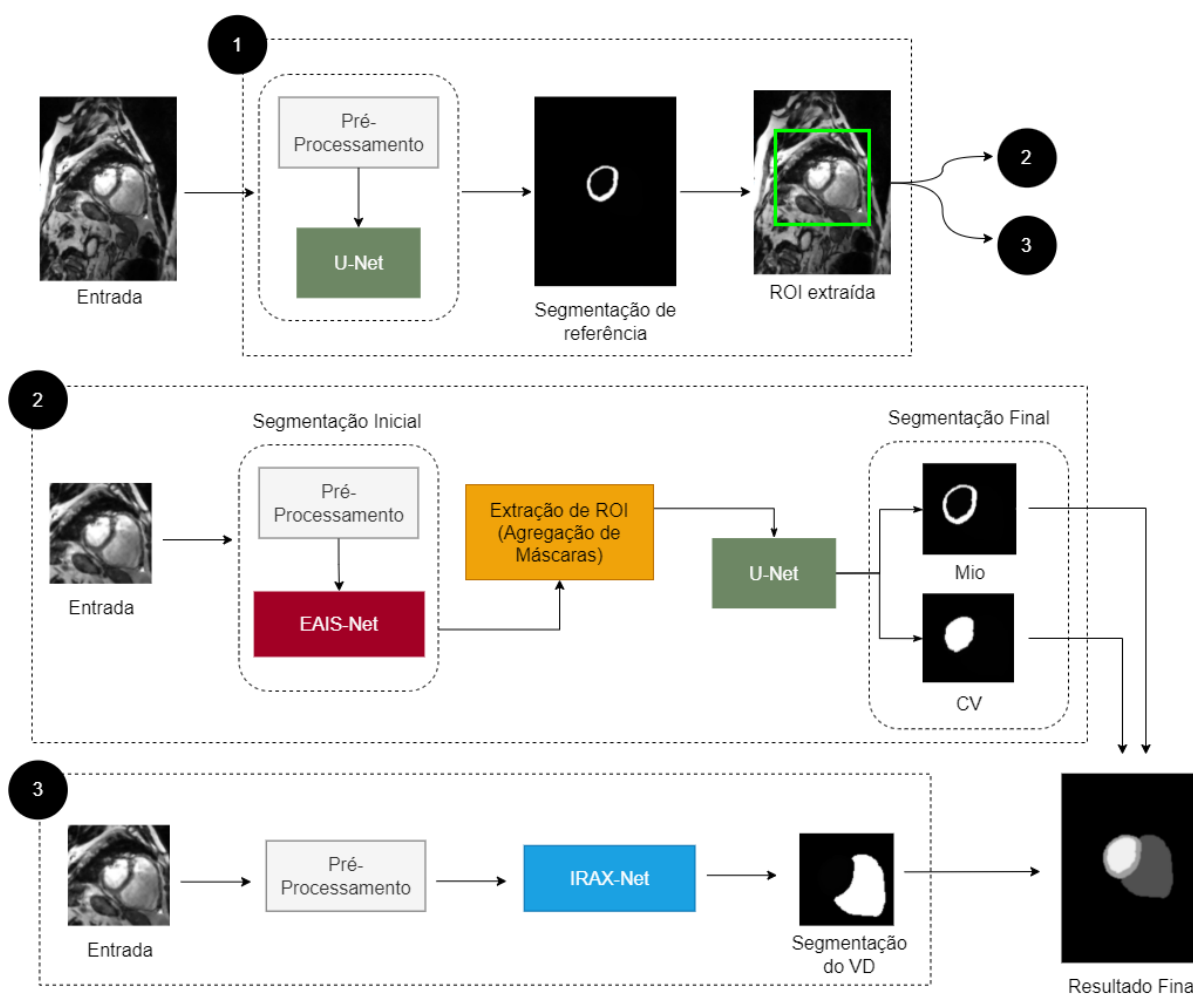
Por fim, com a submissão do método já consolidado aos testes com a base M&Ms, observa-se que o método também alcança um desempenho promissor, principalmente quando os modelos de rede são treinados com os exemplares dessa base, que tem um contexto de maior quantidade e heterogeneidade de exames de cine-RM de eixo curto.

5 MÉTODO DE SEGMENTAÇÃO ESPECIALIZADA POR CONTEXTO

O presente capítulo descreve o segundo método proposto por este trabalho. A abordagem apresentada por esse método é denominada de Segmentação Especializada por Contexto (SEC). Ela compreende no uso de um conjunto específico de técnicas para a realização da segmentação em cada contexto definido, que, neste caso, são dois: um relacionado ao Mio e à CV, e o outro se refere ao VD.

O método SEC também é composto por três etapas principais, conforme mostra a Figura 31. A primeira objetiva a extração de uma ROI para reduzir o escopo da imagem. Essa ROI é submetida à segunda etapa, que visa a segmentação do Mio e da CV; e para a terceira, que é responsável pela segmentação do VD.

Figura 31 – Visão geral das etapas do método SEC: (1) Extração da ROI, (2) Segmentação do Mio e CV, e (3) Segmentação do VD.



Fonte: Elaborada pelo autor.

O desenvolvimento do método SEC é pautado em observações feitas sobre o

método SCR, descrito no Capítulo 4. São elas:

- O método SCR utiliza seis redes convolucionais ao longo de suas etapas: uma U-Net para a extração da ROI; três modelos da mesma rede (EAIS-Net) na etapa de segmentação inicial; e duas U-Nets na fase de Reconstrução. O método SEC busca reduzir esse total, mantendo resultados satisfatórios.
- O método SCR conta com uma etapa de extração de ROI anterior à segmentação inicial. O método SEC, por seu turno, apresenta como subetapa um processo para remover mais regiões de fundo após a segmentação inicial, na tentativa de melhorar os resultados.
- O método SCR utiliza a EAIS-Net para segmentar as três estruturas de interesse individualmente. O método SEC, por sua vez, usa estratégias específicas para cada estrutura, incluindo a proposição de outra FCN.

É importante ressaltar que a etapa de extração da ROI usada no método SEC é a mesma desenvolvida para o método SCR (Seção 4.1). Isso se manteve devido aos resultados expressivos obtidos nos experimentos realizados para essa etapa. Assim sendo, as seções seguintes apresentam mais detalhes sobre a etapa de segmentação especializada aos contextos do Mio e da CV; e do VD.

5.1 Segmentação do Mio e CV

Esta etapa recebe como entrada as ROIs extraídas pela etapa anterior e é responsável pela geração das máscaras do Mio e da CV utilizando uma abordagem dividida em três subetapas: segmentação inicial, extração de uma nova ROI e segmentação final. Cada etapa será explanada detalhadamente a seguir.

5.1.1 Segmentação Inicial

A subetapa de segmentação inicial se inicia com a aplicação de pré-processamento das ROIs de entrada de modo similar ao do método SCR. Como resultado, as entradas pré-processadas terão dimensões 160×160 com três canais e serão submetidas à EAIS-Net. Os pré-processamentos, bem como a arquitetura da EAIS-Net, podem ser consultada na Seção 4.2.

No âmbito do método SEC, essa rede sofre uma alteração na camada de saída com a aplicação de uma função de ativação Softmax para realizar a segmentação simultânea do Mio e da CV, em contrapartida ao que é feito no método SCR, no qual as referidas estruturas de interesse são segmentadas individualmente.

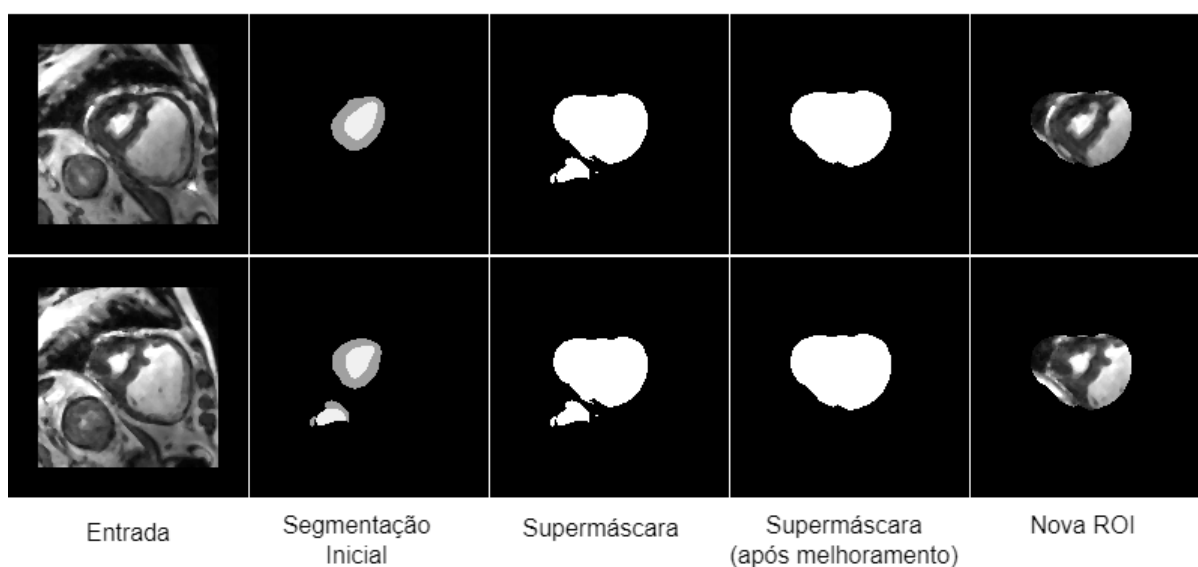
Entretanto, foi verificado por meio dos experimentos que, a despeito dessa alteração, o resultado pode apresentar algumas falhas. Surgiu, portanto, uma hipótese de que os resultados poderiam ser melhorados com a inclusão de uma segunda subetapa de segmentação aplicada sob ROIs mais reduzidas. Esse processo será descrito na sequência.

5.1.2 Agregação de Máscaras para a Extração de uma Nova ROI

Esta subetapa objetiva utilizar as máscaras geradas pela segmentação inicial para delimitar uma nova ROI. Assim, busca-se a eliminação de mais regiões de fundo, e, com isso, reduzir o desbalanceamento entre classes de pixels.

O processo de agregação de máscaras é realizado seguindo alguns passos, como mostra a Figura 32. Para cada volume, primeiro é feita a união das máscaras do Mio e da CV geradas pela EAIS-Net, o que resulta em uma nova máscara, chamada de supermáscara. Essa supermáscara compreende todas as regiões do volume onde certamente as estruturas de interesse são encontradas.

Figura 32 – Processo de agregação de máscaras para a extração de uma nova ROI.



Fonte: Elaborada pelo autor.

No entanto, a supermáscara herda as falhas das máscaras que a compõem,

principalmente regiões de fundo preditas erroneamente (falsos positivos). Então, para reduzi-los, é aplicada uma abertura morfológica de elemento estruturante elíptico de tamanho 5×5 . Com isso, algumas regiões são desconectadas da supermáscara. Depois, é aplicado um filtro de região para manter somente o objeto de maior área mais próximo do centro da imagem, que sempre será a supermáscara. Ao final, é feita uma dilatação morfológica na supermáscara usando os mesmos parâmetros da abertura.

Por fim, a supermáscara resultante desse processo é utilizada para extrair uma nova ROI de todas as fatias do volume que passaram por pré-processamento antes da submissão à EAIS-Net. Então, essa nova ROI será submetida à segmentação final.

5.1.3 Segmentação Final do Mio e da CV

As novas ROIs extraídas via agregação de máscaras são passadas para a subetapa de segmentação final, na qual a U-Net é usada para a segmentação do Mio e da CV. A escolha dessa rede se deu por meio dos experimentos realizados. A U-Net utilizada possui a mesma arquitetura daquela descrita na Seção 4.1 que é baseada na VGG16, e contém os mecanismos tradicionais de convolução. Portanto, a rede recebe as novas ROIs geradas na subetapa anterior e obtém como resultado as máscaras da CV e Mio. Essas máscaras serão combinadas com aquelas geradas pela etapa de segmentação do VD, cujos detalhes serão apresentados na sequência.

5.2 Segmentação do VD

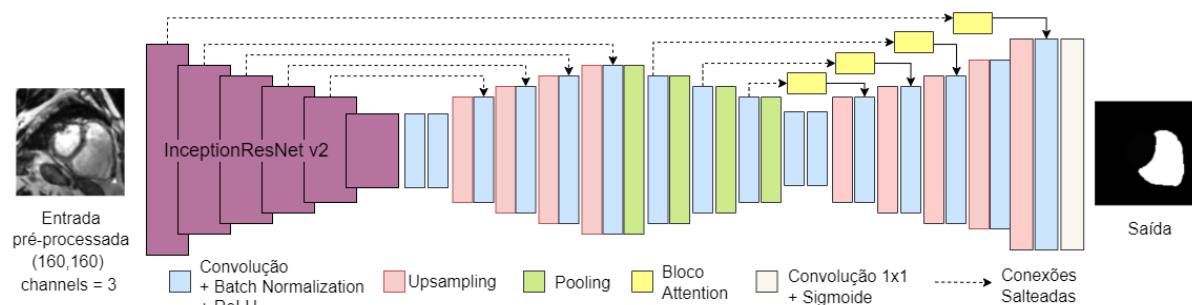
A segmentação do VD é abordada em uma etapa separada, pois, entre as estruturas de interesse, é a que apresentou um comportamento mais complexo. O VD é uma estrutura com tamanho e forma variável não só ao longo do volume, como também entre os pacientes.

Para realizar a segmentação do VD, o método SEC utiliza uma outra FCN, chamada de IRAX-Net, cuja arquitetura proposta se baseia na X-Net (BULLOCK et al., 2019). Como entrada, essa rede recebe as ROIs extraídas pela etapa de extração da ROI, que passam pelo mesmo pré-processamento descrito na Seção 5.1.1.

A Figura 33 mostra a arquitetura proposta da IRAX-Net. A rede possui dois estágios de extração de características. O primeiro deles usa em seu caminho de contração a InceptionResnet-v2 (IRV2) (SZEGEDY et al., 2017). Essa rede é composta por blocos Inception que incorporam conexões residuais substituindo o estágio final desses blocos,

que é a concatenação de filtros. Dessa forma, os blocos residuais Inception realizam a extração de características em múltiplas escalas de filtros e com custo computacional mais reduzido.

Figura 33 – Arquitetura da IRAX-Net proposta para a segmentação do VD no método SEC.



Fonte: Elaborada pelo autor.

No caminho de contração do segundo estágio, são utilizados apenas blocos convolucionais (kernel 3×3) e Max Pooling (kernel 2×2). Operadores Upsampling (kernel 2×2) são usados nos dois caminhos de expansão. Em todos os estágios, ocorre o compartilhamento de mapas de característica entre as camadas através das conexões salteadas. No segundo estágio, especificamente, utiliza-se blocos Attention (similares aos da EAIS-Net) para atribuir pesos às regiões desses mapas que são mais relevantes ao processo de aprendizado. A última camada da IRAX-Net usa convoluções 1×1 e função de ativação sigmoide para a classificação pixel a pixel entre os valores 0 (fundo) e 1 (ventrículo direito). Por fim, como passo final do método SEC, é feita a combinação das máscaras geradas do Mio, CV, e VD para produzir o resultado final.

5.3 Experimentos e Resultados

Nesta seção, são apresentados os resultados obtidos pelo método SEC, bem como a análise destes por meio de estudos de caso e testes de hipóteses. Isso, e os detalhes da condução dos experimentos com as bases ACDC e M&Ms são apresentados a seguir.

5.3.1 Resultados com a base ACDC

Nos experimentos realizados com o método SEC, os quais são relatados a seguir, foi utilizada a base ACDC (descrita na Seção 4.4.1). Além disso, conforme dito anteriormente, o método SEC conta com a mesma etapa de extração da ROI usada pelo método SCR, cujos resultados foram apresentados na Seção 4.4.2.1. Portanto, nesta seção, serão detalhados

os resultados alcançados no âmbito do método SEC para as etapas de segmentação do Mio e da Cv, e segmentação do VD.

5.3.1.1 Segmentação do Mio e da CV

Iniciando-se pela subetapa de segmentação inicial, esse processo é realizado pela EAIS-Net com saída Softmax. O treinamento dessa rede foi realizado com o uso das mesmas operações de *Data augmentation*, e hiperparâmetros do método SCR, como detalhado na Seção 4.4.2.2. A principal diferença é que, no método SEC, a EAIS-Net visa segmentar simultaneamente a CV e o Mio.

Para fins comparativos, também foram realizados experimentos com a U-Net tradicional e uma variação desta rede contendo os Decoder Blocks propostos neste trabalho. Ambas foram treinadas com os mesmos hiperparâmetros definidos para a EAIS-Net, tais quais número de épocas, otimizador, taxa de aprendizado e fator de decaimento. A Tabela 15 mostra os resultados alcançados. O conjunto de testes contém 382 fatias e é composto pelas ROIs produzidas pela etapa de extração da ROI, simulando o fluxo de execução do método SEC.

Tabela 15 – Comparação entre os resultados obtidos nos experimentos para a segmentação inicial do Mio e da CV.

Estrutura	Métodos	Dice	IoU	SEN	PRC
CV	U-Net	0,8903	0,8400	0,8867	0,9129
	U-Net + Decoder Blocks	0,8876	0,8432	0,9111	0,8776
	EAIS-Net	0,9154	0,8742	0,9271	0,9120
Mio	U-Net	0,8016	0,7002	0,7905	0,6665
	U-Net + Decoder Blocks	0,8141	0,7183	0,8252	0,8187
	EAIS-Net	0,8454	0,7578	0,8553	0,8516

A U-Net com Decoder Blocks no caminho de expansão produziu segmentações melhores para o Mio que a U-Net tradicional. Quanto à segmentação da CV, houve uma redução no Dice de 0,0027, e um pequeno aumento de 0,0032 no valor de IoU. Com esse resultado, é possível validar o uso dos Decoder Blocks propostos como uma estrutura que pode prover melhoramentos ao resultado.

A EAIS-Net, por sua vez, apresenta resultados com uma diferença expressiva em comparação aos valores de métricas obtidos pelas outras redes. Essa rede é composta por Decoder Blocks no caminho de expansão e tem seu caminho de contração baseado na EfficientNet B3. Os resultados alcançados pela EAIS-Net para a segmentação da CV são:

Dice de 0,9154, e IoU de 0,8742, o que representa um aumento médio de 2,9% em relação à U-Net tradicional. Para a segmentação do Mio, os valores de Dice e IoU obtidos pela EAIS-Net são, respectivamente, 0,8454 e 0,7578. Isso mostra um aumento médio de 5% em relação ao resultado da U-Net.

Os resultados mostram que, no âmbito do método SEC, a EAIS-Net supera os resultados da U-Net tradicional. Entretanto, como dito na Seção 5.1, surgiu uma hipótese de que os resultados poderiam ser melhorados com a inclusão de uma segunda etapa de segmentação aplicada sob ROIs mais reduzidas. Portanto, para validar essa hipótese, foram realizados experimentos com as mesmas redes do experimento anterior. Entretanto, neste caso, elas recebem como entrada as novas ROIs extraídas pelo processo de agregação de máscaras relatado na Seção 5.1.2.

As redes foram treinadas utilizando um conjunto de treino composto por ROIs extraídas pelo processo de agregação de máscara, usando como referência o *ground-truth*. As ROIs do conjunto de teste, por outro lado, são extraídas conforme o fluxo de execução do método SEC. Logo, a agregação de máscaras usa como referência o resultado da segmentação inicial via EAIS-Net, que foi a rede que apresentou o melhor desempenho no experimento anterior.

Para o treinamento das redes, foram utilizados os mesmos hiperparâmetros (épocas, otimizador, taxa de aprendizado) daquelas experimentadas na etapa de segmentação inicial. A Tabela 16 mostra os resultados alcançados. Observa-se, por meio das métricas, que a U-Net supera as demais redes na segmentação de ambas as estruturas.

Tabela 16 – Comparação entre os resultados obtidos nos experimentos para a segmentação final do Mio e da CV.

Estrutura	Métodos	Dice	IoU	SEN	PRC
CV	U-Net	0,9236	0,8851	0,9312	0,9244
	U-Net + Decoder Blocks	0,9203	0,8782	0,9302	0,9180
	EAIS-Net	0,9047	0,8640	0,9132	0,9102
Mio	U-Net	0,8547	0,7708	0,8464	0,8715
	U-Net + Decoder Blocks	0,8393	0,7491	0,8173	0,8700
	EAIS-Net	0,8356	0,7498	0,8173	0,8677

As ROIs de entrada da segmentação final contém menos regiões de fundo se comparadas às entradas da etapa de segmentação inicial. Com isso, as estruturas de interesse ficam mais destacadas, e o desbalanceamento de classes é reduzido ainda mais.

A EAIS-Net foi elaborada para a segmentação em um contexto com mais infor-

mação. Por isso, é mais profunda, e possui blocos convolucionais mais complexos que os tradicionais que compõem a U-Net. A U-Net com Decoder Blocks é uma rede menos robusta que a EAIS-Net, e obteve resultados melhores que ela. No entanto, a U-Net tradicional superou as demais. Ela é uma rede com arquitetura mais simples, e por isso apresentou melhor desempenho quando aplicada sob as novas ROIs extraídas, visto que elas possuem uma quantidade mais reduzida de informação.

Por fim, é possível concluir que a hipótese que inspirou o desenvolvimento da subetapa de segmentação final foi validada. E os resultados justificam a escolha das redes EAIS-Net e U-Net, respectivamente, para compôr as subetapas inicial e final do processo de segmentação do Mio e da CV.

5.3.1.2 Segmentação do VD

Os conjuntos de treinamento, validação e teste usados nesta etapa são compostos pelas mesmas ROIs usadas no experimento da subetapa inicial da segmentação do Mio e VD. Entretanto, o *ground-truth* de ambos contém somente anotações do ventrículo direito.

O método SEC usa uma outra FCN proposta, chamada de IRAX-Net, para realizar a segmentação do VD. O processo de treinamento dessa rede utiliza a função de perda Jaccard Loss, otimizador Adam e taxa de aprendizado de 10^{-3} e fator de decaimento de 0,1, com duração de 300 épocas. A IRAX-Net tem 73.334.662 parâmetros treináveis.

Nos experimentos, outras redes, além da IRAX-Net, foram testadas com o intuito de se fazer uma análise comparativa. São elas a U-net tradicional, a EAIS-Net, a X-Net em sua arquitetura padrão, e variações contendo os *backbones* EfficientNet B3 e InceptionResnet-v2 (IRv2). Os resultados obtidos podem ser vistos na Tabela 17.

Tabela 17 – Análise comparativa entre as redes usadas no processo de segmentação do VD.

Método	Dice	IoU	SEN	PRC
U-Net	0,7025	0,6249	0,7905	0,6695
EAIS-Net	0,7649	0,7013	0,8399	0,7451
X-Net	0,6864	0,6150	0,7006	0,7149
X-Net (EfficientNet B3)	0,5835	0,4826	0,6800	0,7116
X-Net (IRv2)	0,7208	0,6595	0,6856	0,7050
IRAX-Net (IRv2+ Blocos Attention)	0,8213	0,7623	0,8198	0,8383

Os experimentos com a U-Net e a EAIS-Net foram conduzidos inicialmente para segmentar as três estruturas. Contudo, essas redes apresentaram resultados promissores

somente para a segmentação do Mio e da CV. No caso do VD, foram obtidos resultados menos expressivos, o que motivou a execução de testes envolvendo outras redes.

Foram realizados testes com a X-Net, que obteve baixo desempenho. Então, outras adaptações foram feitas nessa arquitetura, tais quais o uso das redes EfficientNet B3 e IRv2 como backbone, sendo que essa última apresentou resultados melhores. Dado isso, a última alteração feita foi a inclusão dos blocos Attention no segundo caminho de expansão da X-Net com backbone IRv2, desenvolvendo, assim, a IRAX-Net. Essa rede alcançou os melhores resultados para a segmentação do VD, obtendo Dice de 0,8213, e IoU de 0,7623, que, em média, representam um aumento de 5,8% em relação à EAIS-Net.

Observa-se, portanto, que o processo de extração de característica feito pela IRv2 foi mais efetivo quando comparado à arquitetura tradicional da X-Net, e ao uso da EfficientNet B3 como backbone. Além disso, a inclusão de blocos Attention mostrou-se importante para a obtenção de resultados mais expressivos. Em virtude disso, a IRAX-Net foi a rede escolhida para integrar a etapa de segmentação do VD.

5.3.1.3 Resultado Final e Estudos de Caso

O resultado final é definido pela junção das saídas produzidas pelas etapas que compõem o método SEC. Para a segmentação do Mio e da CV, é utilizada a EAIS-Net em sua subetapa inicial, e a U-Net na subetapa final. Na segmentação do VD, é aplicada a IRAX-Net. A Tabela 18 apresenta os resultados gerais de Dice e IoU. Além disso, essa tabela também mostra a avaliação do método aplicado sobre o mesmo conjunto de teste, cujas fatias estão divididas em dois grupos referentes às fases cardíacas de fim de diástole (ED) e fim de sístole (ES). Esse modo de análise é similar ao feito pelo desafio ACDC.

Tabela 18 – Abordagens escolhidas para compor o método SEC e seus respectivos resultados de segmentação da CV, Mio, e VD.

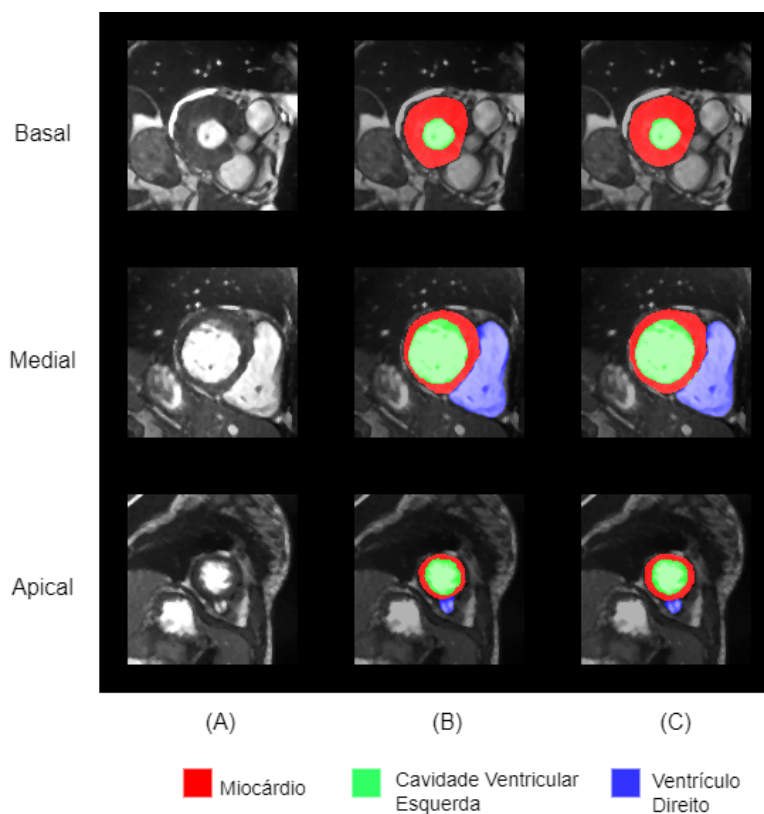
Etapas		Dice	Dice	Dice	IoU	IoU	IoU
		Geral	ED	ES	Geral	ED	ES
Segmentação do Mio e da CV (EAIS-Net + U-Net)	CV	0,9236	0,9473	0,9000	0,8851	0,9145	0,8557
	Mio	0,8547	0,8616	0,8479	0,7708	0,7703	0,7713
Segmentação do VD (IRAX-Net)	VD	0,8213	0,8546	0,7880	0,7623	0,8003	0,7244

Através dos resultados, é possível observar que o método SEC obtém melhor desempenho no caso das fatias em ED. Durante essa fase cardíaca, o coração está mais dilatado, e, devido a isso, as estruturas de interesse se tornam mais visíveis no exame.

No caso das fatias em ES, os resultados apresentam uma diminuição, principalmente aqueles relacionados ao VD. Durante essa fase, as estruturas ficam mais comprimidas, e as alterações de tamanho e forma contribuem para os erros na segmentação.

A Figura 34 mostra os resultados qualitativos do método SEC aplicado em um caso basal, medial e apical, cujos valores de Dice obtidos foram, respectivamente, 0,9604, 0,9838, e 0,9672 (CV); 0,9479, 0,8904, e 0,8194 (Mio); 1,00, 0,9723, e 0,8540 (VD). Nestes exemplos, é possível verificar as mudanças na forma e no tamanho das estruturas de interesse ao longo do volume. Apesar disso, o método gerou bons resultados de segmentação, com destaque aos valores de Dice acima dos 95% no caso basal. No exemplo medial, as estruturas estão mais bem definidas, por conseguinte, os resultados melhoram, exceto para o Mio, cujo Dice diminuiu. Neste caso, o método produziu uma segmentação um pouco maior que o *ground-truth* dada a semelhança de textura com regiões muito próximas. O mesmo se pode observar no exemplo apical, inclusive para o VD. No entanto, apesar de essa estrutura ter um tamanho bastante reduzido, o método foi capaz de apresentar um bom resultado.

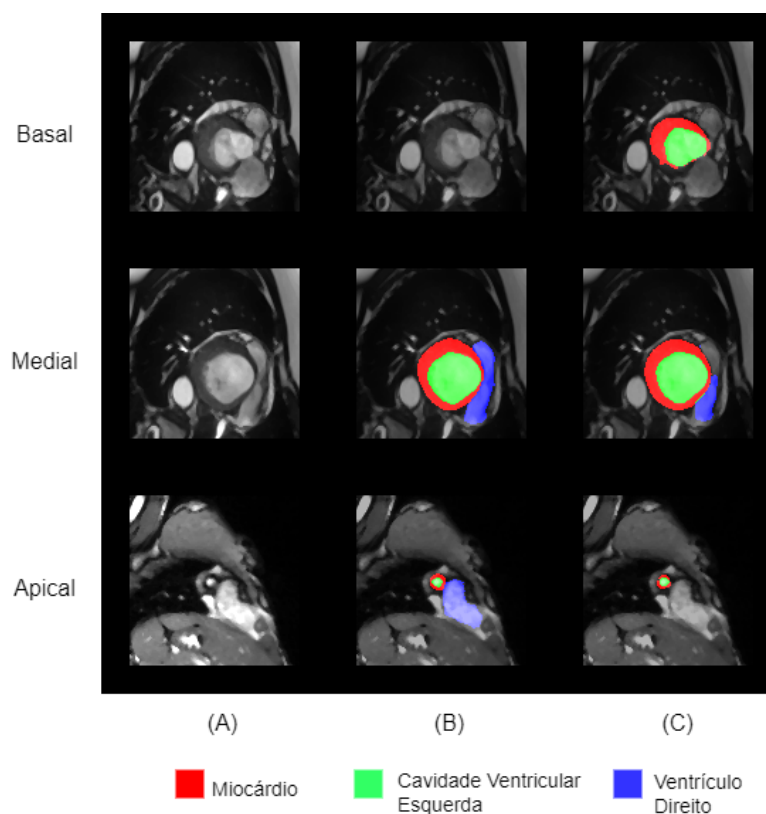
Figura 34 – Resultados qualitativos do método SEC aplicado em fatias basais, mediais e apicais, sendo (A) a entrada, (B) o *ground-truth*, e (C) o resultado final.



Fonte: Elaborada pelo autor.

A Figura 35 apresenta casos em que o método é aplicado sob fatias durante a fase de ES, que, conforme dito anteriormente, são as que apresentam erros de segmentação com mais frequência. Os valores obtidos de Dice para o exemplo basal, medial, e apical, foram, respectivamente, 0,00, 0,9723, e 0,9166 (CV); 0,00, 0,9288, e 0,6727 (Mio); 1,00, 0,6624, e 0,00 (VD). No caso basal, não há a presença das estruturas de interesse, porém o método gerou falsos positivos devido à semelhança entre regiões da fatia com o Mio e a CV. O caso medial exemplifica a presença das três estruturas. A segmentação do Mio e da CV obteve resultados bastante expressivos. Contudo, como o VD apresenta uma forma mais contraída e esticada, o método só identificou parte dos pixels que compõe essa estrutura. Por fim, no exemplo apical, CV e Mio possuem um tamanho muito reduzido, mas mantêm o suas características de forma. Por outro lado, o VD assume uma forma diferente e uma textura que se assemelha às regiões de fundo, levando o método a produzir falsos negativos.

Figura 35 – Exemplos de casos de erros de segmentação produzidos pelo método SEC em fatias na fase de ES, sendo (A) a entrada, (B) o *ground-truth*, e (C) o resultado final.



Fonte: Elaborada pelo autor.

5.3.1.4 Análise de Impacto na Segmentação

Alguns experimentos foram conduzidos para avaliar o impacto causado no resultado da segmentação de estruturas cardíacas por mudanças feitas no método SEC. O primeiro experimento consiste em avaliar o desempenho individual das redes utilizadas (U-Net, EAIS-Net, e IRAX-Net). Neste caso, essas redes representam o processo de segmentação de única etapa. Esse resultado, portanto, é comparado com aquele obtido pelo método SEC, que se trata de um processo dividido em mais etapas. Essa análise comparativa pode ser vista na Tabela 19.

Tabela 19 – Comparativo entre os resultados obtidos pelo método SEC e outras abordagens de segmentação de etapa única.

Estrutura	Métodos	Dice	IoU
CV	U-Net	0,8592	0,7837
	EAIS-Net	0,8840	0,7784
	IRAX-Net	0,8000	0,7147
	Método SEC	0,9236	0,8851
Mio	U-Net	0,4457	0,3213
	EAIS-Net	0,4399	0,3177
	IRAX-Net	0,6169	0,4715
	Método SEC	0,8547	0,7708
VD	U-Net	0,5975	0,4995
	EAIS-Net	0,6309	0,5353
	IRAX-Net	0,5307	0,4187
	Método SEC	0,8213	0,7623

Nos testes com a U-Net, EAIS-Net, e IRAX-Net, as dimensões das fatias de entrada foram reduzidas para 160x160 devido às limitações de hardware. No geral, os resultados alcançados por essas abordagens são bastante inferiores ao método SEC. Uma das causas observadas é que o redimensionamento das fatias reduz o tamanho das estruturas de interesse. Isso dificulta o processo de extração de características, principalmente nas regiões apicais do exame, onde as estruturas são naturalmente pequenas. Por outro lado, o método SEC divide o processo de segmentação em etapas, o que permite lidar melhor com o desbalanceamento de classes, bem como manter inalterado o tamanho das estruturas cardíacas.

O outro experimento realizado visa analisar o impacto da etapa de extração da ROI no resultado final da segmentação. Esse experimento consiste em substituir as ROIs extraídas automaticamente por um conjunto de teste contendo os mesmos pacientes, porém

composto por ROIs extraídas de forma semi-supervisionada, com base no *ground-truth*. Os resultados são apresentados na Tabela 20.

Tabela 20 – Experimentos realizados com o método SEC seguindo duas abordagens: a extração semi-supervisionada da ROI, e a proposta completamente automatizada.

Estrutura	Método SEC	Dice	IoU
CV	Extração semi-supervisionada da ROI	0,9360	0,8975
	Completamente automático	0,9236	0,8851
Mio	Extração semi-supervisionada da ROI	0,8655	0,7825
	Completamente automático	0,8547	0,7708
VD	Extração semi-supervisionada da ROI	0,8105	0,7524
	Completamente automático	0,8213	0,7623

Através desse experimento, é possível notar que o processo automático acarreta em algumas perdas para a CV e Mio. Por outro lado, os resultados relativos ao VD obtém uma pequena vantagem comparado ao processo semi-supervisionado. Sendo assim, para avaliar se a diferença entre os resultados é estatisticamente significativa, foram realizados testes de hipótese (Seção 3.6). Os parâmetros utilizados foram o tamanho da amostra (382 fatias), e o nível de significância (α) de 0,05. Novamente, a hipótese nula é de que há diferença significativa entre os resultados. O resultado dos testes podem ser vistos na Tabela 21.

Tabela 21 – Valores p encontrados para as métricas Dice e IoU resultantes das abordagens semi-supervisionada e automática de extração da ROI no método SEC.

	Valores p	
	Dice	IoU
CV	0,5024	0,5819
Mio	0,6670	0,6978
VD	0,7001	0,7422

Os valores p encontrados para as métricas Dice e IoU são todos maiores que α . Assim sendo, rejeita-se a hipótese nula, e conclui-se que não há diferença significativa entre os resultados. Portanto, assim como no outro método proposto (método SCR), o uso da etapa de extração de ROI apresenta um baixo impacto, corroborando a viabilidade da proposta de segmentação automática no cenário de aplicação.

5.3.2 Resultados com a Base M&Ms

Nesta seção, serão apresentados os resultados obtidos pelo método SEC nos testes com a base M&Ms. Assim como no caso do método SCR, esses testes foram feitos após a

consolidação do método, portanto, os resultados mostrados na sequência foram alcançados com o uso das técnicas escolhidas para compôr as subetapas do método SEC.

Também foram realizados dois experimentos. Naquele chamado de SEC_ACDC, o método SEC utiliza os modelos de rede previamente treinados com a base ACDC e é testado com a base M&Ms. No outro, denominado SEC_MMS, os modelos são treinados e testados com a referida base. As redes foram treinadas com os mesmos hiperparâmetros relatados nas Seções 5.3.1.1 e 5.3.1.2; e o mesmo processo de *Data augmentation* também foi utilizado.

A Tabela 22 mostra os resultados obtidos nos dois experimentos. As métricas de avaliação alcançaram valores mais expressivos no experimento SEC_MMS que no SEC_ACDC. De maneira análoga ao que foi observado nos experimentos com o método SCR, é possível verificar que o treinamento dos modelos com exemplares da base M&MS acarretou melhores resultados em comparação ao uso de modelos treinados previamente com a base ACDC. O treinamento com a base M&Ms, portanto, consegue apresentar às redes alguns aspectos presentes nessa base que não são contemplados entre os exemplos da base ACDC, dado que a primeira possui maior diversidade entre os exames devido aos diferentes modelos de scanner usados na aquisição.

Tabela 22 – Método SEC: resultados dos experimentos realizados com a base M&Ms.

Experimento	Estrutura	Dice	IoU	SEN	PRC
SEC_ACDC	CV	0,7761	0,6851	0,7445	0,7139
	Mio	0,6757	0,5529	0,6529	0,6105
	VD	0,6945	0,5902	0,7712	0,7432
SEC_MMS	CV	0,8339	0,7553	0,8620	0,8215
	Mio	0,7464	0,6263	0,7546	0,7679
	VD	0,7083	0,6067	0,7890	0,7675

No caso das estruturas CV e Mio, todas as métricas apresentaram um aumento notório. Tomando como exemplo o coeficiente Dice, no experimento SEC_ACDC, o método SEC obteve 0,7761 e 0,6757, respectivamente, para CV e Mio. No experimento SEC_MMS, esses valores se elevaram para 0,8339 e 0,7464. No que se refere ao VD, também se observa um melhor desempenho do método no experimento SEC_MMS. Entretanto, os valores das métricas não são expressivamente maiores. Por exemplo, Dice e IoU foram, respectivamente, de 0,6945 e 0,5902 para 0,7083 e 0,6067, representando um acréscimo de cerca de 0,1. O aumento mais expressivo pode ser visto para a precisão (de

0,7432 para 0,7675), significando que, no experimento SEC_MMS, o método produziu menos falsos positivos na segmentação do VD.

A Tabela 23 mostra os resultados divididos entre as fases cardíacas de ED e ES. No experimento SEC_ACDC, verifica-se que o método alcança melhores resultados nas fatias em ED, fase em que as estruturas estão mais aparentes. Esse comportamento se mantém nos casos da CV e VD no experimento SEC_MMS. De maneira geral, neste experimento, todos os resultados são melhorados. Entretanto, no caso do Mio, essa melhora é mais acentuada para as fatias em ES. Isso é influenciado pela quantidade maior de exemplos que possui a base M&Ms, permitindo que as redes aprendam mais características sobre esse tipo de fatia.

Tabela 23 – Resultados da aplicação do método SEC na base M&Ms separados por fase cardíaca.

Experimento	Estrutura	Dice ED	Dice ES	IoU ED	IoU ES
SEC_ACDC	CV	0,8385	0,7136	0,7707	0,5994
	Mio	0,6806	0,6707	0,5524	0,5535
	VD	0,7252	0,6639	0,6299	0,5505
SEC_MMS	CV	0,8618	0,8059	0,7987	0,7119
	Mio	0,7330	0,7598	0,6076	0,6451
	VD	0,7218	0,6949	0,6294	0,5839

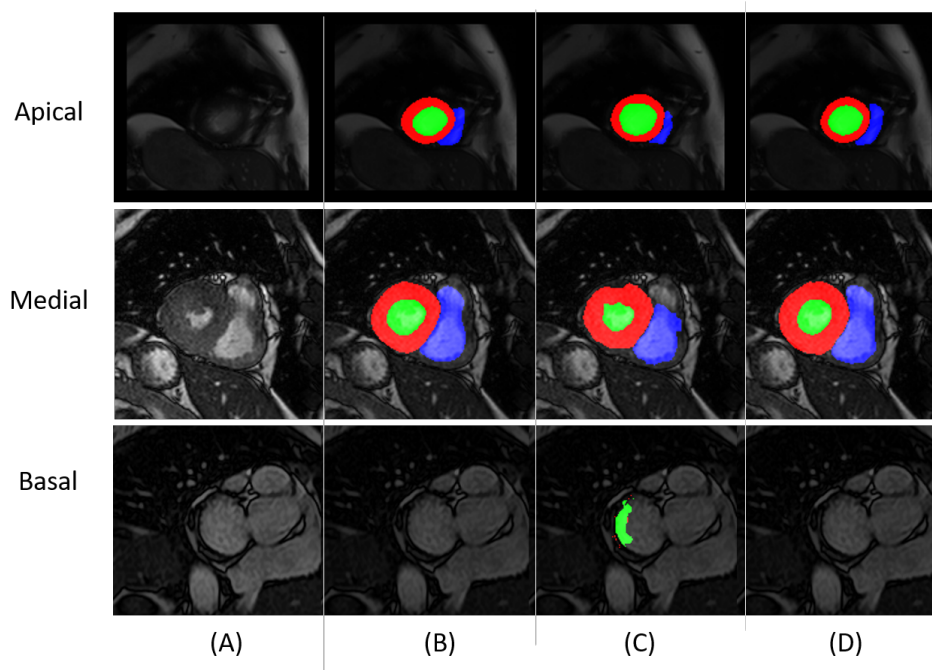
Na Figura 36, pode ser visto o resultado da segmentação produzida pelo método SEC para cada experimento realizado. No exemplo apical, os valores de Dice obtidos foram de 0,8726 (CV), 0,7498 (Mio), e 0,8023 (VD) no experimento SEC_ACDC, para 0,9478 (CV), 0,8579 (Mio), e 0,8472 (VD) no experimento SEC_MMS.

O caso medial ilustra uma das principais falhas observadas no experimento SEC_ACDC, que é a segmentação parcial das estruturas, indicando a geração de falsos negativos. Essas falhas ocorrem com menos frequência no experimento SEC_MMS, reiterando a importância do treinamento das redes com exemplos da base M&Ms. No exemplo da Figura 36, os valores de Dice foram de 0,7830 (CV), 0,9352 (Mio) e 0,8351 (VD) para 0,9237 (CV), 0,8047 (Mio), e 0,9048 (VD).

No caso basal, é ilustrada outra falha vista no experimento SEC_ACDC, que é a geração de falsos positivos. As fatias basais não possuem VD, e, em sua maioria, têm CV e Mio. Contudo, algumas fatias podem não conter nenhuma das estruturas de interesse, mas, apresentam regiões similares quanto à forma e textura, o que acarreta erros de segmentação. Esse efeito é reduzido no experimento SEC_MMS, pois a base

M&Ms possui mais fatias com essas características que a base ACDC.

Figura 36 – Resultados qualitativos do método SEC nos experimentos com a base M&Ms, sendo (A) a entrada, (B) o *ground-truth*, (C) o resultado do experimento SEC_ACDC e (D) o resultado do experimento SEC_MMS.



Fonte: Elaborada pelo autor.

Por fim, testes de hipóteses foram realizados para avaliar a diferença estatística entre os resultados encontrados para cada experimento. Nesses testes, assume-se que a hipótese nula é a de que há diferença significativa. O nível de significância (α) usado é de 0,05, de modo que, se o valor p for menor que α , a hipótese nula é aceita. Na Tabela 24, são apresentados os valores p obtidos nos testes.

Tabela 24 – Valores p encontrados para as métricas de avaliação do método SEC nos experimentos SEC_ACDC e SEC_MMS.

	Valores p			
	Dice	IoU	SEN	PRC
CV	0,0000	0,0000	0,0000	0,0000
Mio	0,0000	0,0000	0,0000	0,0000
VD	0,2247	0,1754	0,0836	0,0229

A partir dos resultados, identifica-se que há diferença significativa em relação a todas as métricas que avaliam a segmentação do Mio e da CV. Isso corrobora a análise apresentada de que o método SEC também mostra um desempenho promissor quando aplicado na base M&Ms, principalmente no experimento em que os modelos de rede são

treinados com essa base. No caso do VD, a diferença é significativa somente nos resultados de precisão, reforçando a constatação de que, no experimento SEC_MMS, o método foi capaz de reduzir os falsos positivos.

5.4 Considerações Finais

Analisando os experimentos, é possível verificar que as hipóteses levantadas para a elaboração do método SEC foram atestadas. Esse método utiliza uma quantidade menor de redes que o método SCR, e, ainda assim, alcança resultados promissores. Novamente, verifica-se que a divisão do processo de segmentação em etapas é importante para a obtenção de resultados melhores, como visto na comparação com abordagens de passo único.

No contexto da segmentação do Mio e da CV, a extração da nova ROI via agregação de máscaras se mostrou relevante para o aperfeiçoamento dos resultados inicialmente produzidos pela EAIS-Net. Assim, a U-Net aplicada sobre ROIs menores, e com menos desbalanceamento entre fundo e estruturas de interesse, pôde alcançar melhoramentos na segmentação final. No contexto do VD, os resultados obtidos pela IRAX-Net foram superiores aos de outras arquiteturas compostas por mecanismos tradicionais de convolução, e também à EAIS-Net, elaborada para outro contexto.

Nos testes com as bases ACDC e M&Ms, o método SEC apresentou resultados significativos, principalmente para as fatias na fase cardíaca de ED, na qual as estruturas de interesse estão mais aparentes. Portanto, é possível considerar que a definição de contextos e o uso de técnicas específicas para cada um deles também é aplicável no cenário da segmentação em cine-RM.

6 ANÁLISE COMPARATIVA DOS MÉTODOS PROPOSTOS

Este capítulo apresenta uma análise comparativa entre os métodos propostos (SCR e SEC), tendo como base os experimentos relatados anteriormente com as bases ACDC e M&Ms.

6.1 Base ACDC

Nesta seção, são abordados os resultados obtidos pelos métodos propostos aplicados sobre o conjunto de testes local quanto aqueles avaliados via submissão à plataforma online do desafio ACDC.

6.1.1 Conjunto Local de Testes

Os experimentos locais foram efetuados para ambos os métodos com o mesmo conjunto de teste, composto por 20 pacientes da base ACDC com total de 382 fatias. Na sequência, portanto, serão feitas comparações relativas aos resultados finais da segmentação e aos tempos de execução.

A Tabela 25 apresenta uma comparação entre os resultados finais dos métodos SCR e SEC, considerando a segmentação do conjunto de teste de um modo geral, isto é, sem dividi-las entre as fases de ED e ES. O método SCR supera o método SEC na maioria das métricas, exceto a precisão no caso da segmentação do VD.

Tabela 25 – Base ACDC: resultados finais obtidos pelos métodos propostos.

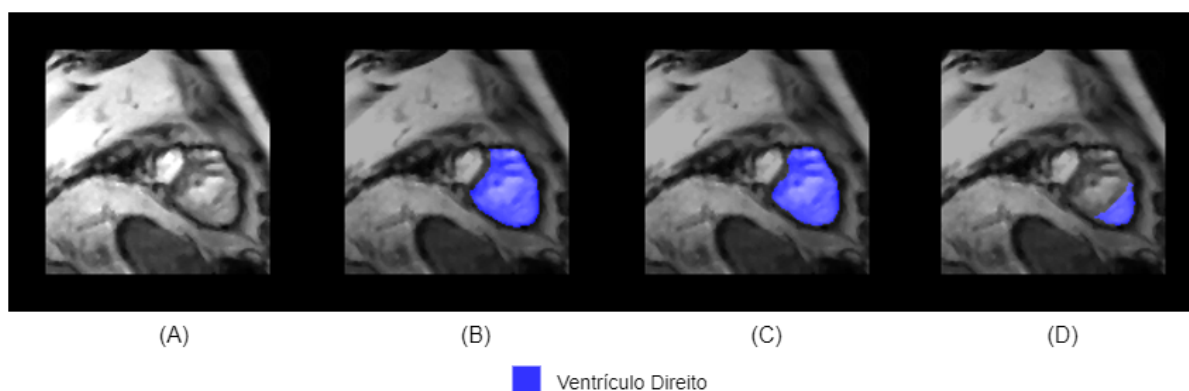
Abordagens	Estruturas	Dice	IoU	SEN	PRC
Método SCR	CV	0,9303	0,8886	0,9416	0,9285
	Mio	0,8552	0,7747	0,8449	0,8730
	VD	0,8312	0,7740	0,8512	0,8269
Método SEC	CV	0,9236	0,8851	0,9312	0,9244
	Mio	0,8547	0,7708	0,8464	0,8715
	VD	0,8213	0,7623	0,8198	0,8383

A precisão é inversamente proporcional aos falsos positivos, logo, no caso do VD, infere-se que o método SEC os produz em menor quantidade. Entretanto, analisando a sensibilidade, vê-se que método SCR é superior, o que indica uma menor geração de falsos negativos, isto é, menos pixels pertencentes ao VD são perdidos. Isso mostra a atuação do

módulo de reconstrução, que busca minimizar essa perda, impactando, por conseguinte, Dice e IoU de forma positiva.

A Figura 37 mostra um exemplo que corrobora essa observação. Neste, o resultado do método SCR (Figura 37C) obtém as seguintes métricas: 0,9380 (Dice), 0,8833 (IoU), 0,9948 (SEN), e 0,8836 (PRC). Por sua vez, o método SEC (Figura 37D) obtém 0,3442 (Dice), 0,2078 (IoU), 0,2078 (SEN), e 1,00 (PRC).

Figura 37 – Exemplo comparativo entre os métodos SCR e SEC no caso da segmentação do VD, sendo (A) a entrada, (B) o ground-truth, (C) o resultado gerado pelo método SCR, e (D) o resultado do método SEC.



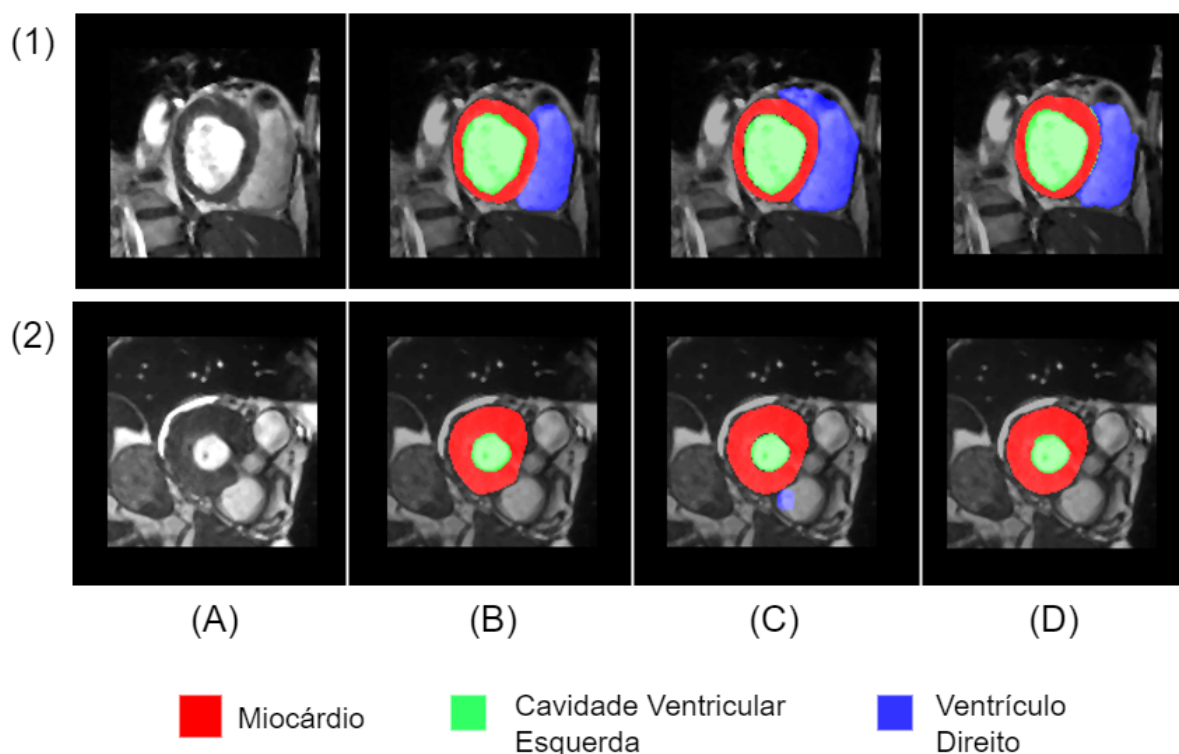
Fonte: Elaborada pelo autor.

A segmentação do VD produzida pelo método SEC possui precisão de 1,00, pois não houve a geração de falsos positivos. Contudo, grande parte da máscara do VD não foi predita, explicando, assim, o motivo para Dice, IoU e sensibilidade obterem valores mais baixos. Diferentemente, o método SCR possui uma precisão menor. Mas, gerou uma máscara que cobre mais regiões do VD, obtendo, portanto, maiores valores de Dice, IoU, e sensibilidade.

Sobre a sensibilidade do método SCR em relação à segmentação do VD, esta é maior que a do método SEC na avaliação geral. Isso indica uma maior taxa de acerto na classificação dos pixels pertencentes à referida estrutura. Entretanto, como consequência, há casos, como visto na Figura 38, em que o método SCR produz falsos positivos em regiões próximas ao VD, estrapolando os limites da estrutura de interesse (Figura 38-1); e também em fatias basais que não possuem marcação para o VD (Figura 38-2). Nestes exemplos, o método SEC apresenta um desempenho melhor, dado o treinamento especializado da IRAX-Net para o contexto do VD.

No exemplo da Figura 39, os dois métodos obtêm resultados de Dice aproximados

Figura 38 – Casos em que o método SCR apresenta mais falhas que o método SEC, sendo (A) a entrada, (B) o ground-truth, (C) o resultado gerado pelo método SCR, e (D) o resultado do método SEC.

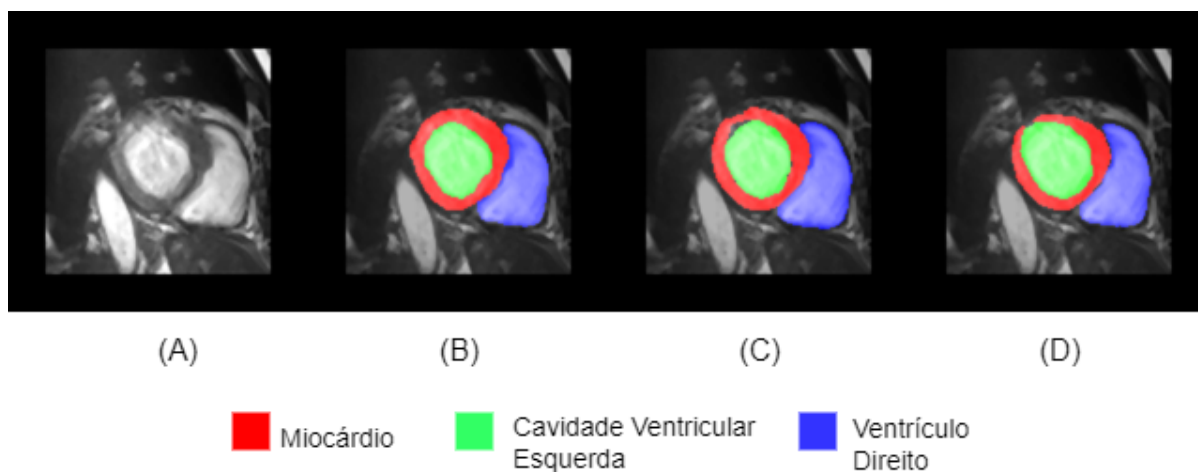


Fonte: Elaborada pelo autor.

na segmentação do VD, sendo de 0,9585 e 0,9618 respectivamente para os métodos SCR e SEC, com este último superando o primeiro. Quanto às estruturas Mio e CV, o método SCR obtém Dices de 0,8992 (Mio) e 0,9600 (CV), ultrapassando, dessa forma, o método SEC que, por sua vez, obtém 0,8148 (Mio) e 0,9293 (CV). Os principais motivos para essa diferença nos resultados, a qual desfavorece o método SEC, são as descontinuidades presentes na segmentação do Mio em virtude de uma máscara mais ampliada da CV produzida por esse método. Por outro lado, o método SCR conseguiu diferenciar melhor os pixels da CV e do Mio, acarretando melhores resultados.

Essa máscara ampliada da CV é um efeito observado também no exemplo da Figura 40. Entretanto, neste caso, o método SEC apresenta uma segmentação melhor para a CV em comparação ao método SCR, visto que este último gerou uma segmentação menor à CV, indicando mais falsos negativos. Os resultados de Dice são 0,7657 (SCR) e 0,9333 (SEC). A extração da nova ROI, que é subetapa do método SEC, contribuiu para esse resultado, pois, gera uma região menor, e, portanto, reduz o desbalanceamento entre os pixels da CV e do fundo, permitindo que o método identifique melhor as características

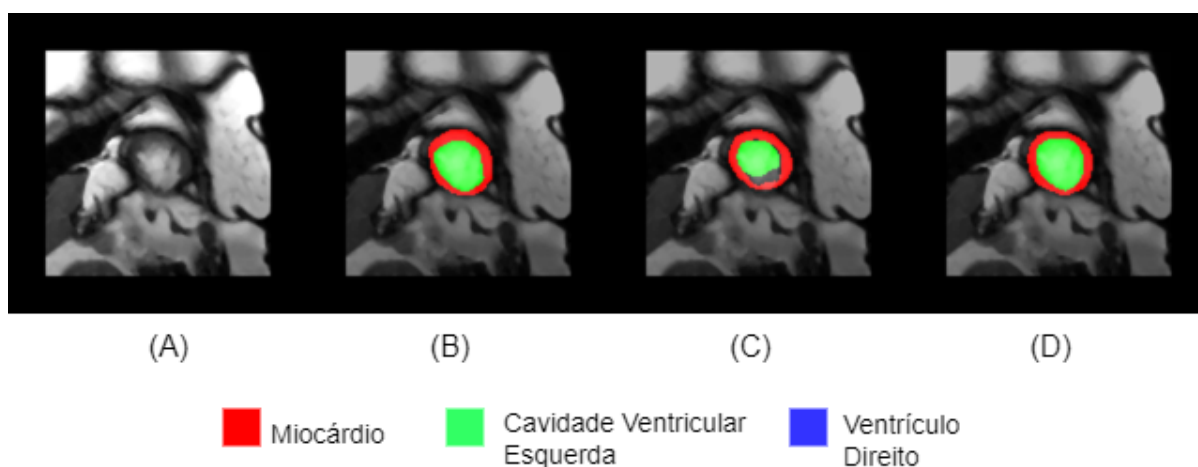
Figura 39 – Exemplo comparativo entre os métodos SCR e SEC, com a segmentação das três estruturas de interesse. Neste caso, (A) é a entrada, (B) é o ground-truth, (C) é o resultado gerado pelo método SCR, e (D) é o resultado do método SEC.



Fonte: Elaborada pelo autor.

das estruturas.

Figura 40 – Exemplo comparativo entre os métodos SCR e SEC, com foco na segmentação do Mio e da CV. Neste caso, (A) é a entrada, (B) é o ground-truth, (C) é o resultado gerado pelo método SCR, e (D) é o resultado do método SEC.



Fonte: Elaborada pelo autor.

Outra comparação entre os métodos pode ser vista na Tabela 26. Neste caso, as métricas foram calculadas separadamente para as fases cardíacas ED e ES, numa abordagem semelhante à utilizada pelo desafio ACDC conforme mencionado anteriormente. Novamente, nota-se que o método SCR supera os resultados do método SEC de modo geral. Verifica-se também que ambos os métodos têm resultados menos expressivos na segmentação de fatias na fase ES do que na fase ED, principalmente no caso do VD.

Apesar de o método SCR ultrapassar o método SEC na maioria das métricas

Tabela 26 – Base ACDC: resultados finais separados por fase cardíaca (ED e ES).

Abordagens	Estruturas	Dice	Dice	IoU	IoU
		ED	ES	ED	ES
Método SCR	CV	0,9602	0,9337	0,9248	0,8795
	Mio	0,8871	0,8902	0,7986	0,8041
	VD	0,9257	0,8543	0,8646	0,7520
Método SEC	CV	0,9413	0,9000	0,9145	0,8557
	Mio	0,8616	0,8479	0,7703	0,7713
	VD	0,8546	0,7880	0,8003	0,7244

de modo geral, testes de hipóteses foram realizados para verificar se há uma diferença significativa entre esses resultados. A análise foi feita de acordo com o elucidado na Seção 3.6 e teve o nível de significância $\alpha = 0,05$. Os valores p são calculados a partir da diferença entre os valores de Dice, Iou, SEN e PRC obtidos por ambos os métodos. A hipótese nula é de que há diferença significativa entre esses valores.

Os resultados dos testes de hipóteses podem ser vistos na Tabela 27. Em nenhuma métrica, o valor p encontrado foi menor que o nível α , o que indica a rejeição da hipótese nula. Logo, não há diferença significativa entre os resultados produzidos pelos métodos SCR e SEC.

Tabela 27 – Resultados dos testes de hipóteses: valores p encontrados para as métricas Dice, IoU, SEN, e PRC produzidos pelos métodos SCR e SEC para cada estrutura de interesse. Nível $\alpha = 0,05$.

	Valores p			
	Dice	IoU	SEN	PRC
CV	0,7220	0,8786	0,5559	0,8282
Mio	0,9843	0,8977	0,9542	0,9505
VD	0,7180	0,7016	0,2418	0,6730

Por fim, é feita uma comparação com respeito ao tempo médio de execução dos métodos ao efetuar a segmentação por exame. Dessa forma, busca-se avaliar a complexidade computacional. O método SCR executa todo o seu fluxo em tempo médio de 4 segundos por exame. Entre as três etapas deste método, a que consome mais tempo é a segmentação inicial, com uma média de 2 segundos por exame. Isso ocorre porque, na referida etapa, é feito o uso de três redes para segmentar especificamente cada estrutura de interesse.

Quanto ao método SEC, o tempo médio de execução deste é de 3,3 segundos por exame. Dentre as suas etapas, a de segmentação do Mio e da CV é a que apresenta maior tempo médio de execução, sendo este de 1,69 segundos por exame. Como essa etapa é dividida em três subetapas, e utiliza duas redes (EAIS-Net e U-Net), conseqüentemente

há um maior tempo de processamento. Ainda assim, esse método tem tempo de execução menor que o método SCR em virtude da utilização de menos FCNs. Enfim, apesar das peculiaridades de cada abordagem proposta, observa-se que ambas alcançam resultados promissores em tempo considerado satisfatório.

6.1.2 Avaliação na Plataforma Online e Comparação com Trabalhos Relacionados

O desafio ACDC disponibiliza uma base separada com exames de 50 pacientes (sem *ground-truth*) exclusivamente para a avaliação de métodos. As métricas aplicadas pelo desafio são o coeficiente Dice e a distância de Hausdoff (HD) para avaliar o acerto da segmentação nas fases cardíacas de fim da diástole (ED) e fim da sístole (ES). E para uma avaliação clínica, são calculadas a correlação e *bias* para medir os volumes das estruturas em ED e ES, fração de ejeção (EF), e massa do miocárdio. As Tabelas 28, 29 e 30 apresentam os resultados obtidos pelos métodos propostos no conjunto de teste do desafio ACDC assim como a comparação com demais trabalhos classificados neste desafio.

Tabela 28 – Resultados dos métodos propostos para a segmentação da cavidade ventricular (CV) na base de teste e comparação com os melhores métodos elencados no desafio ACDC.

Método	Dice ED	Dice ES	HD ED	HD ES	EF corr.	EF bias	Vol. ED corr.	Vol. ED bias
Método SCR	0,963	0,912	8,062	10,432	0,975	1,030	0,994	0,110
Método SEC	0,960	0,904	14,976	17,488	0,983	0,310	0,995	-0,710
Simantiris e Tziritas (2020)	0,967	0,928	6,366	7,573	0,993	-0,360	0,998	2,032
Isensee et al. (2017)	0,967	0,928	5,476	6,921	0,991	0,490	0,997	1,530
Zotti et al. (2018)	0,964	0,912	6,180	8,386	0,990	-0,476	0,997	3,746
Painchaud et al. (2019)	0,961	0,911	6,152	8,278	0,990	-0,480	0,997	3,824
Khened et al. (2019)	0,964	0,917	8,129	8,968	0,989	-0,548	0,997	0,576
Baumgartner et al. (2017)	0,963	0,911	6,526	9,170	0,988	0,568	0,995	1,436
Calisto e Lai-Yuen (2020)	0,958	0,903	5,592	8,644	0,981	0,494	0,997	3,072
Wolterink et al. (2017)	0,961	0,918	7,515	9,603	0,988	-0,494	0,993	3,046
Rohé et al. (2017)	0,957	0,900	7,483	10,747	0,989	-0,094	0,993	4,182
Ammar et al. (2021)	0,968	0,911	7,993	10,528	0,982	-0,390	0,997	0,650

O métodos propostos apresentam desempenhos promissores. Entre eles, os resultados obtidos pelo método SCR estão entre as melhores colocações para a segmentação da CV, Mio e VD. No cenário da segmentação da CV, este método obtém Dice ED similar ao proposto por Baumgartner et al. (2017), superando este por 0,001 de Dice ES. Em relação à segmentação do Mio, o método SCR obtém um dos três melhores resultados de Dice ED e ES, superando o método de Simantiris e Tziritas (2020) por 0,002 (ED) e 0,001 (ES). Destaca-se o valor alcançado para a métrica HD ED, que, comparado com os

Tabela 29 – Resultados dos métodos propostos para a segmentação do miocárdio (Mio) na base de teste e comparação com os melhores métodos elencados no desafio ACDC.

Método	Dice ED	Dice ES	HD ED	HD ES	Vol. ES corr.	Vol. ES bias	Massa ED corr.	Massa ED bias
Método SCR	0,894	0,905	7,906	9,912	0,980	-1,100	0,988	-1,820
Método SEC	0,880	0,892	13,440	12,590	0,972	3,210	0,971	-1,870
Isensee et al. (2017)	0,904	0,923	7,014	7,328	0,988	-1,984	0,987	-2,547
Simantiris e Tziritas (2020)	0,891	0,904	8,264	9,575	0,983	-2,134	0,992	-2,904
Calisto e Lai-Yuen (2020)	0,873	0,895	8,197	8,318	0,988	-1,79	0,989	-2,100
Zotti et al. (2018)	0,886	0,902	9,586	9,291	0,980	1,160	0,986	-1,827
Painchaud et al. (2019)	0,881	0,897	8,651	9,598	0,979	0,296	0,987	-2,906
Khened et al. (2019)	0,889	0,898	9,841	12,582	0,979	-2,572	0,990	-2,873
Patravali et al. (2017)	0,882	0,897	9,757	11,256	0,986	-4,464	0,989	-11,586
Baumgartner et al. (2017)	0,892	0,901	8,703	10,637	0,983	-9,602	0,982	-6,861
Zotti et al. (2017)	0,884	0,896	8,708	9,264	0,960	-7,804	0,984	-12,405
Wolterink et al. (2017)	0,875	0,894	11,121	10,687	0,971	0,906	0,963	-0,960
Ammar et al. (2021)	0,891	0,901	10,575	13,891	0,934	1,590	0,986	2,977

Tabela 30 – Resultados dos métodos propostos para a segmentação do ventrículo direito (VD) na base de teste e comparação com os melhores métodos elencados no desafio ACDC.

Método	Dice ED	Dice ES	HD ED	HD ES	EF corr.	EF bias	Vol. ED corr.	Vol. ED bias
Método SCR	0,900	0,860	14,660	17,560	0,743	1,810	0,931	7,370
Método SEC	0,910	0,860	16,580	17,300	0,746	2,190	0,944	8,940
Isensee et al. (2017)	0,951	0,904	8,205	11,655	0,910	-3,750	0,992	0,900
Calisto e Lai-Yuen (2020)	0,936	0,884	10,183	12,234	0,899	-2,118	0,989	3,550
Simantiris e Tziritas (2020)	0,936	0,889	13,289	14,367	0,894	-1,292	0,990	0,906
Zotti et al. (2018)	0,934	0,885	11,052	12,650	0,869	-0,872	0,986	2,372
Zotti et al. (2017)	0,941	0,882	10,318	14,053	0,872	-2,228	0,991	-3,722
Painchaud et al. (2019)	0,933	0,884	13,718	13,323	0,865	-0,874	0,986	2,078
Khened et al. (2019)	0,935	0,879	13,994	13,930	0,858	-2,246	0,982	-2,896
Baumgartner et al. (2017)	0,932	0,883	12,670	14,691	0,851	1,218	0,977	-2,290
Wolterink et al. (2017)	0,928	0,872	11,879	13,399	0,852	-4,610	0,980	3,596
Rohé et al. (2017)	0,916	0,845	14,049	15,926	0,781	-0,662	0,983	7,340

demais, é um dos menores, significando que as segmentações geradas para os exames na fase ED são consideravelmente similares ao *ground-truth*.

O método SEC também alcança resultados que se aproximam das abordagens melhor classificadas pelo desafio. Na segmentação da CV, este método supera os resultados de Dice ED e ES obtidos por abordagens mais robustas, como a AdaEn-Net (CALISTO; LAI-YUEN, 2020) e a segmentação multi-atlas proposta por Rohé et al. (2017), esse último sendo também superado quanto ao Dice ES na segmentação do VD.

Em relação à segmentação do Mio, o método SEC apresenta Dice ED superior, novamente, à AdaEn-Net, e ao método de (WOLTERINK et al., 2017). As referidas abordagens utilizam uma CNN para a segmentação simultânea das três estruturas nas fatias em seu tamanho original. O método SEC, assim como o método SCR, é composto

por etapas, sendo que a primeira delas é responsável por extrair uma ROI, eliminando, assim, mais regiões de fundo, e auxiliando na obtenção de resultados melhores.

O método SEC não possui uma etapa específica para pós-processamentos, mas, ainda assim, obtém valores de Dice ED e ES muito próximos aos alcançados por (PAIN-CHAUD et al., 2019), que utilizam um Auto-Encoder Adversário (AAE) para refinar especificamente as máscaras geradas por uma FCN. Essa abordagem é similar ao módulo de reconstrução proposto pelo método SCR. Entretanto, o AAE recebe somente a segmentação inicial para então gerar uma nova máscara. O módulo de reconstrução, por outro lado, recebe a fatia e sua segmentação inicial, portanto, mais informação, para, enfim, gerar a máscara reconstruída. Isso refletiu resultados melhores ao método SCR em termos de Dice ED e ES.

As propostas que superam os métodos propostos utilizam técnicas que exigem muitos recursos computacionais, como memória RAM, processador e placas de vídeo de alto desempenho. Neste caso, há destaque para o trabalho de Isensee et al. (2017) que alcança a primeira colocação no desafio e é baseado no uso de um ensemble de U-Nets 2D e 3D. Ainda assim, a diferença entre este e o método SCR é de aproximadamente 0,004 em relação ao Dice ED na segmentação da CV. Quanto ao método SEC, nesse mesmo cenário, a diferença é de 0,007.

Ressalta-se também que os métodos propostos usam menos recursos de hardware e ainda assim apresentam resultados aproximados ou superiores aos de métodos mais robustos. Porém, entre eles, o método SCR é o que mais se destaca. Por fim, quanto à segmentação do VD, ambos os métodos propostos obtém resultados inferiores aos das outras estruturas (CV e Mio), mas ainda se mantêm aproximados com as demais abordagens.

6.2 Base M&Ms

Nesta seção, é feita a análise dos resultados obtidos pelos métodos SCR e SEC com o conjunto de testes disponibilizado pelo desafio M&Ms, bem como a comparação com os trabalhos relacionados.

6.2.1 Comparação entre os Métodos SCR e SEC

O conjunto de testes é composto por 136 pacientes, totalizando 3218 fatias. Um comparativo dos resultados obtidos pelos métodos SCR e SEC pode ser visto na Tabela 31. Esses resultados se referem aos experimentos em que os modelos de rede de cada método foram treinados com a base M&Ms, conforme descrito anteriormente nas Seções 4.4.4 e 5.3.2.

Tabela 31 – Base M&Ms: resultados finais obtidos pelos métodos propostos.

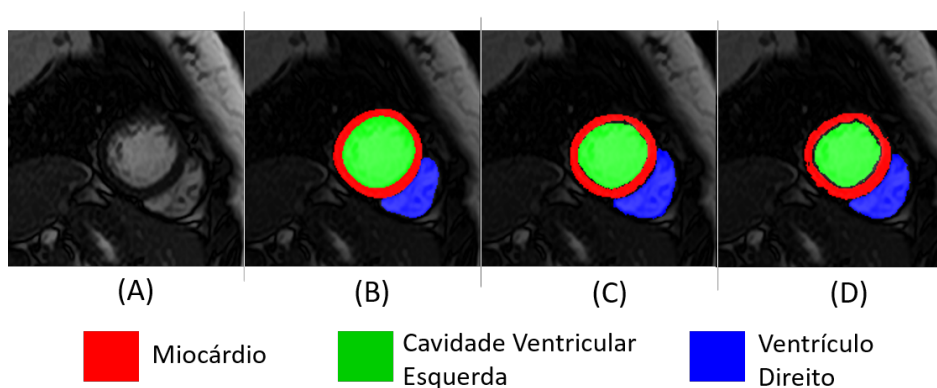
Abordagens	Estruturas	Dice	IoU	SEN	PRC
Método SCR	CV	0,8453	0,7662	0,8652	0,8457
	Mio	0,7834	0,6709	0,7701	0,7899
	VD	0,7434	0,6320	0,7602	0,8266
Método SEC	CV	0,8339	0,7553	0,8620	0,8215
	Mio	0,7464	0,6263	0,7546	0,7679
	VD	0,7083	0,6067	0,7890	0,7675

Assim como visto nos experimentos com a base ACDC, no caso daqueles feitos com a M&Ms também se observa que o método SCR apresenta resultados superiores ao SEC. Ele possui uma quantidade maior de modelos treinados, além de contar com um módulo de refinamento composto pela reconstrução e pós-processamentos específicos para cada estrutura. Diferentemente, o método SEC não contém uma etapa de refinamento, portanto, o seu resultado final pode conter algumas falhas.

No caso da CV, os valores de Dice, IoU e SEN não apresentam uma diferença expressiva, do ponto de vista numérico, na comparação entre os métodos. Por outro lado, isso não se repete no caso do Mio, em que as referidas métricas do método SCR são bastante superiores. Na Figura 41, é possível visualizar um exemplo com essas características. Os resultados de Dice do método SCR (Figura 41C) foram de 0,9363 para a CV, 0,8277 para o Mio, e 0,9720 para o VD. Já para o método SEC (Figura 41D), foram de 0,9304; 0,7624; e 0,9685, respectivamente para a CV, Mio, e VD. Em relação ao Mio, observa-se que o método SEC gerou falhas que são, no método SCR, corrigidas via módulo de refinamento, como pequenos alongamentos e lacunas.

No que diz respeito ao VD, os resultados gerais mostram uma sensibilidade maior no resultado do método SEC que no do SCR, o que indica uma menor perda na predição dos pixels dessa estrutura cardíaca. Isso pode ser visto através do exemplo mostrado na Figura 42, que se trata de uma fatia apical. As segmentações do VD geradas pelos métodos

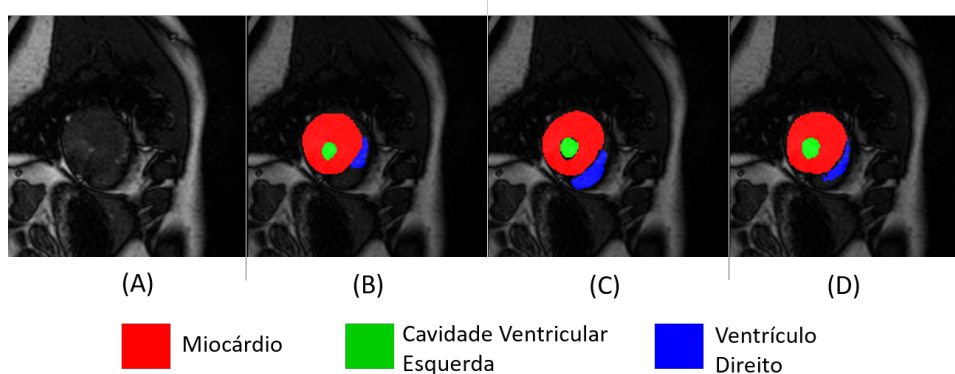
Figura 41 – Exemplo comparativo entre os métodos SCR e SEC no experimento com a base M&Ms, sendo (A) a entrada, (B) o *ground-truth*, (C) o resultado gerado pelo método SCR, e (D) o resultado do método SEC.



Fonte: Elaborada pelo autor.

envolvem uma área maior que o *ground-truth*. Contudo, aquela produzida pelo método SCR perde mais partes da estrutura de interesse que a gerada pelo método SEC. Isso se reflete nos resultados de Dice, que são, para o VD, de 0,4059 e 0,6153 (métodos SCR e SEC, respectivamente). Em relação às outras estruturas, há superioridade do método SCR: Dices para CV e Mio, respectivamente, de 0,8048 e 0,9265 contra 0,7706 e 0,9084 produzidos pelo método SEC.

Figura 42 – Outro exemplo comparativo dos métodos desenvolvidos no experimento com a base M&Ms, sendo (A) a entrada, (B) o *ground-truth*, (C) o resultado gerado pelo método SCR, e (D) o resultado do método SEC.



Fonte: Elaborada pelo autor.

A avaliação separada por fase cardíaca pode ser vista na Tabela 32. Verifica-se, neste caso, que os métodos seguem a tendência analisada nos resultados anteriores com a base ACDC, que é a de apresentar melhor desempenho geral nas fatias em fim de diástole (fase ED). Pontua-se também que os valores de Dice e IoU do miocárdio são maiores para fatias em ES, pois os problemas principais relativos à segmentação dessa estrutura, que

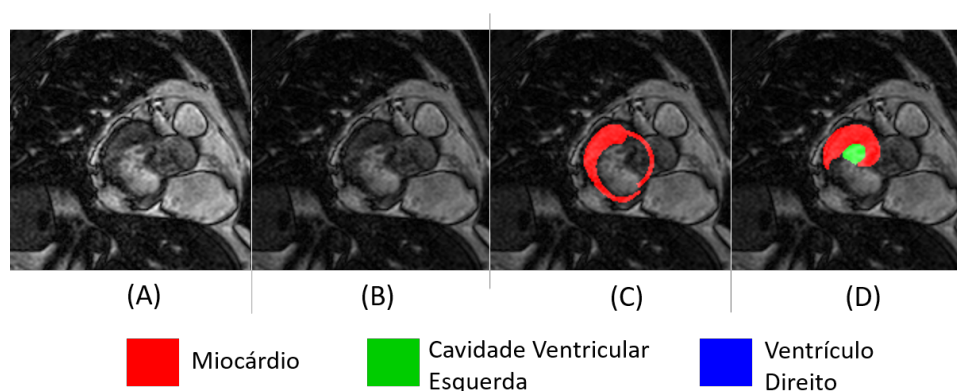
ocorrem durante essa fase cardíaca, são corrigidos pelas etapas de refinamento (método SCR) e segmentação final (método SEC).

Tabela 32 – Base M&Ms: resultados finais dos métodos SCR e SEC separados por fase cardíaca (ED e ES).

Abordagens	Estruturas	Dice	Dice	IoU	IoU
		ED	ES	ED	ES
Método SCR	CV	0,8704	0,8202	0,8052	0,7272
	Mio	0,7639	0,8022	0,6443	0,6966
	VD	0,7648	0,6963	0,6641	0,5733
Método SEC	CV	0,8618	0,8059	0,7987	0,7119
	Mio	0,7330	0,7598	0,6076	0,6451
	VD	0,7218	0,6949	0,6294	0,5839

Os resultados apresentados acima mostram que o desempenho do método SCR supera o SEC no contexto da base M&Ms, que contém mais exames e possui uma diversidade maior de características de aquisição das imagens em relação à base ACDC. Contudo, há casos em que ambos os métodos falham, como pode ser visto na Figura 43, que se trata de uma fatia basal na fase cardíaca de fim de sístole (ES). Esse exemplo não contém marcações para as estruturas de interesse, mas possui regiões semelhantes ao Mio e à CV em termos de forma e textura, ocasionando a geração de máscaras por ambos os métodos.

Figura 43 – Exemplo de falha de ambos os métodos, em que (A) representa a entrada, (B) o *ground-truth*, (C) o resultado gerado pelo método SCR, e (D) o resultado do método SEC.



Fonte: Elaborada pelo autor.

A despeito dos resultados qualitativos explanados anteriormente, testes de hipótese também foram realizados para verificar se há diferença estatística significativa na avaliação dos métodos. Esses testes seguem a mesma configuração usada em experimentos anteriores, ou seja, nível de significância α de 0,05. Os valores p encontrados podem ser vistos na

Tabela 33. Ressalta-se que esses testes foram executados com base no resultado geral das métricas Dice, IoU, SEN e PRC.

Tabela 33 – Resultados dos testes de hipóteses aplicados na comparação entre os métodos SCR e SEC nos experimentos com a base M&Ms.

	Valores p			
	Dice	IoU	SEN	PRC
CV	0,2110	0,3036	0,7074	0,0089
Mio	0,0004	0,0002	0,1426	0,8335
VD	0,0015	0,0359	0,0055	0,0000

Resultados com diferença significativa podem ser observados no caso da segmentação do miocárdio, em termos de Dice e IoU, e na segmentação do VD (todas as métricas), evidenciando que o método SCR, dada a sua robustez, apresenta um desempenho superior ao método SEC, que possui menos modelos de rede, e também não apresenta uma etapa exclusiva para o refinamento. No caso da CV, não há diferença estatística, o que era esperado visto que essa estrutura tem um comportamento mais uniformizado em comparação às demais, não exigindo um processo de reconstrução ou pós-processamentos mais complexos em ambos os métodos.

As diferenças estatísticas significativas entre os métodos não ocorrem no experimento com a base ACDC, mas se evidenciam no contexto da base M&Ms, que retrata um cenário com mais heterogeneidade devido à maior quantidade de exames obtidos por meio de diversos *scanners*. Nesse contexto, portanto, o método SCR mostra um desempenho mais efetivo. Isso, entretanto, não imputa ao método SEC uma avaliação desfavorável, pois, com a sua proposta de reduzir a quantidade de redes e processos, ainda obtém resultados promissores.

Por fim, também foram avaliados os tempos de execução dos métodos SCR e SEC no experimento com a base M&Ms. O método SCR realiza todo o seu fluxo de execução em tempo médio de 7,58 segundos por exame. Assim como observado nos testes com a base ACDC, a etapa de segmentação inicial desse método é aquela que mais consome tempo, com uma média de 6,36 segundos por exame, em virtude da execução de três redes EAIS-Net para segmentar separadamente CV, Mio e VD.

Quanto ao método SEC, o tempo médio de execução é de 3,97 segundos por exame, sendo que a etapa de segmentação do Mio e da CV é a que se realiza de modo mais lento, com tempo médio de 2,44 segundos por exame. Isso ocorre dado ao uso de dois

modelos de rede, além do processo de agregação de máscaras. Em uma análise geral, assim como nos testes com a base ACDC, nestes também se observam resultados considerados satisfatórios.

6.2.2 Comparação com os Trabalhos Relacionados

A Tabela 34 mostra um comparativo entre os resultados obtidos pelos métodos propostos com os de trabalhos que também experimentaram seus métodos com a base M&Ms. A métrica usada para análise é o coeficiente Dice, conforme indicado pelo desafio. Ressalta-se que a maioria dos métodos apresentados informam o Dice separadamente por fase cardíaca (ED e ES), mas há aqueles que usam o Dice generalizado. Portanto, para fins comparativos, os métodos SCR e SEC apresentam, na referida tabela, tanto o Dice geral quanto por fase cardíaca.

Tabela 34 – Resultados de Dice Geral (G) e por fase cardíaca (ED e ES) dos métodos propostos para a segmentação da CV, Mio, e VD na base de teste e comparação com os trabalhos relacionados.

Método	CV			Mio			VD		
	G	ED	ES	G	ED	ES	G	ED	ES
Scannell et al. (2020)	-	0,905	0,848	-	0,772	0,820	-	0,876	0,809
Huang et al. (2020)	-	0,896	0,772	-	0,761	0,721	-	0,820	0,698
Full et al. (2020)	-	0,939	0,886	-	0,839	0,867	-	0,910	0,860
Li et al. (2021)	-	0,930	0,894	-	0,764	0,828	-	0,883	0,822
Lin et al. (2022)	-	0,924	0,889	-	0,827	0,859	-	0,878	0,843
Huang et al. (2022)	0,945	-	-	0,869	-	-	0,912	-	-
Método SCR	0,845	0,870	0,820	0,783	0,763	0,802	0,743	0,764	0,696
Método SEC	0,833	0,861	0,805	0,746	0,733	0,759	0,708	0,721	0,694

O trabalho de Scannell et al. (2020) consiste no uso da U-Net 2D, cujo treinamento se dá por meio de uma técnica chamada de *adversaria back-propagation* para tratar principalmente da variabilidade entre os exames da base por conta dos diferentes *scanners*. Em comparação, dentre os métodos desenvolvidos no âmbito do presente trabalho, o método SCR é o que apresenta Dices mais aproximados, porém sendo superado. Ressalta-se, no entanto, que as fatias usadas no trabalho mencionado foram previamente recortadas nas dimensões 192x192, com as estruturas cardíacas devidamente centralizadas, caracterizando um cenário semi-supervisionado. Por sua vez, os métodos SCR e SEC contam com uma etapa de extração da ROI que mantém as estruturas, mas não garante a sua centralidade, considerando, portanto, além da diferença de características oriunda da

aquisição dos exames, os deslocamentos dentro da ROI, o que simula um cenário totalmente automatizado.

Os métodos SCR e SEC apresentam resultados superiores ao trabalho de Huang et al. (2020) para a CV nas fatias em ES, e para o Mio em todos os cenários. Esse trabalho usa duas redes: uma chamada de ST, voltada para o pré-processamento das fatias, que são redimensionadas para 256×256 ; e uma U-Net 2D, que recebe a saída da rede ST para produzir, enfim, a segmentação das estruturas. Por outro lado, os métodos propostos aplicam no pré-processamento apenas técnicas voltadas para a redução dos ruídos característicos da cine-RM e a normalização entre 0 e 1, e a segmentação é realizada em etapas, de modo que as fatias não são previamente redimensionadas para não comprometer as características. Além disso, esses métodos buscam corrigir falhas observadas na segmentação das estruturas, principalmente na fase ES.

Os trabalhos que superam os métodos propostos são caracterizados pelo uso de técnicas que demandam um custo computacional mais elevado. Full et al. (2020), por exemplo, utilizam a nnU-Net, um *framework* que permite o treinamento automatizado de redes a partir da estimação de hiperparâmetros e busca pelo melhor conjunto de operações de *Data augmentation*. Lin et al. (2022) e Huang et al. (2022) propõem métodos que recebem o volume 3D do exame na entrada, o que exige hardware mais robusto. O primeiro, por exemplo, usa 128 GB de RAM, enquanto os métodos SCR e SEC foram desenvolvidos em ambiente com 8GB de RAM. O segundo usa uma GPU com 10.496 núcleos cuda, enquanto os métodos propostos foram treinados em ambiente com 640 núcleos cuda (uma diferença superior a 90%).

A despeito, os métodos SCR e SEC apresentam resultados competitivos. Esses métodos foram desenvolvidos e aprimorados através de experimentos com a base ACDC, que possui um contexto semelhante à M&Ms, porém, essa última contém mais heterogeneidade em virtude dos diferentes tipos de *scanner* usados na aquisição dos exames. Ainda assim, os métodos propostos obtiveram resultados elevados para as métricas de avaliação, superando alguns trabalhos que foram desenvolvidos considerando, desde sua concepção, o cenário retratado na base M&Ms.

6.3 Discussão

Os métodos propostos por este trabalho apresentam como principal ideia a divisão do processo de segmentação em etapas, o que, de acordo com os experimentos, se mostrou importante para a obtenção de resultados promissores, visto que as técnicas desenvolvidas para cada etapa e a sua atuação conjunta auxiliam a geração de máscaras mais acuradas, resolvendo alguns problemas inerentes ao contexto da aplicação.

Ambos os métodos possuem como a primeira etapa o processo de extração da ROI, que mitiga o desbalanceamento entre os pixels das estruturas de interesse e de fundo. Além disso, reduz o custo de processamento através do recorte da ROI como alternativa ao redimensionamento das entradas, que, conforme visto nos testes, pode modificar suas características a ponto de comprometer o aprendizado.

Sobre o método SCR, especificamente, a sua segunda etapa produz segmentações iniciais a partir da ROI extraída usando uma rede proposta que combina mecanismos bem avaliados pela literatura. A última etapa busca melhorar a segmentação inicial usando o módulo de reconstrução e pós-processamentos elaborados de modo específico para cada estrutura cardíaca. Destaca-se ainda que o módulo de reconstrução acelerou o processo de melhoramento dos resultados, visto que evita a experimentação de diferentes técnicas de pós-processamento que pode ser exaustiva.

O método SEC, que foi desenvolvido a partir de observações levantadas sobre o método SCR, utiliza uma quantidade menor de redes que esse método. As redes propostas (EAI-Net e IRAX-Net) têm arquiteturas compostas pela combinação de outros *backbones* com os *Decoder Blocks* propostos e os resultados obtidos se mostraram promissores em contrapartida a outras arquiteturas compostas por mecanismos tradicionais de convolução. Entretanto, em todos os experimentos realizados, tanto em âmbito local, quanto com os conjuntos de testes disponibilizados pelas plataformas online dos desafios ACDC e M&Ms, o método SEC é superado pelo método SCR.

De modo geral, os métodos propostos alcançam bons resultados para a segmentação das estruturas cardíacas. Isso é confirmado através da validação via plataforma online do desafio ACDC, e com o conjunto de teste da base M&Ms. Em ambos os desafios, observa-se que a maioria dos trabalhos relacionados utilizam variações da U-Net, *frameworks* para otimização de hiperparâmetros, ou *ensembles* e realizam a segmentação em etapa única. Esses métodos foram desenvolvidos em hardware mais robusto, utilizando

GPUs como a Nvidia GTX modelos Titan X ou V, ou da família RTX. Diferentemente, os métodos propostos SCR e SEC dividem o processo em etapas menores, cada uma voltada para parte específica do problema, o que permitiu o seu desenvolvimento em um hardware mais limitado, e treinamento na GPU NVidia GTX 1050 que tem 640 núcleos cuda, cerca de 80 a 90% menos núcleos que as mencionadas anteriormente.

Na avaliação com a base ACDC, é notado que o desempenho dos métodos propostos é melhor para as fatias na fase cardíaca de ED. Através das métricas de avaliação, observa-se que ambos se aproximam das abordagens melhor posicionadas nos casos das segmentações da CV e do Mio, com destaque ao método SCR, cujos resultados para a CV podem ser classificados entre os oito melhores e, no caso do Mio, entre os cinco melhores.

A base M&Ms, que apresenta mais exemplos que a base ACDC, retrata um cenário de maior heterogeneidade dado aos diferentes *scanners* usados para a aquisição dos exames. Na avaliação com essa base, observa-se que os métodos apresentam resultados melhores para as fatias na fase ES somente no caso do Mio. Para as demais estruturas, os resultados são superiores na fase ED. De modo geral, os resultados obtidos podem ser considerados promissores, visto que as métricas de validação mantêm valores próximos aos de trabalhos elaborados com técnicas e recursos de hardware mais robustos.

Outra constatação observada é de que os métodos propostos produziram resultados inferiores para o ventrículo direito em comparação com as demais estruturas. Isso ocorre nos experimentos com ambas as bases. Entende-se que a causa se dá em virtude da variabilidade de formas que o VD pode tomar ao longo do volume do exame. Essa estrutura costuma ser mais bem definida nas fatias mediais, porém, conforme observado nos testes, nas fatias apicais ele pode assumir um tamanho mais reduzido, o que acarreta a geração de falsos negativos, pois os pixels do VD são classificados erroneamente como pertencentes ao fundo. No caso das fatias basais, em geral, o VD é ocluído por estruturas de textura similar, o que ocasiona a produção de falsos positivos. Não obstante esses apontamentos, e outras limitações mencionadas, as análises quantitativas e qualitativas indicam que os métodos propostos obtêm resultados promissores, assim como validam sua aplicabilidade.

Com base nos experimentos realizados, também é possível afirmar que houve a comprovação de quatro das cinco hipóteses levantadas para o presente trabalho (Seção 1.1). A primeira hipótese é comprovada à medida em que a abordagem usada pelos métodos

propostos de divisão do processo de segmentação em etapas alcança resultados superiores em comparação aos cenários em que o processo de segmentação é feito em passo único. A validação da segunda hipótese é realçada através do desempenho promissor obtido pelas FCNs propostas. A EAIS-Net usada nas etapas de segmentação inicial de ambos os métodos é composta por mecanismos de convolução mais robustos, como os da EfficientNet e Decoder blocks. Esses últimos também compõem a IRAX-Net utilizada na segmentação do VD pelo método SEC. Ambas as FCNs propostas apresentam desempenho superior ao da U-Net, cuja arquitetura é constituída por mecanismos de convolução tradicional. A terceira hipótese é validada por meio do melhoramento dos resultados produzido pela integração do módulo de reconstrução à etapa de refinamento do método SCR.

Somente a quarta hipótese foi refutada. Esperava-se que os métodos obtivessem resultados melhores com propostas semi-supervisionadas em comparação às abordagens automáticas por causa da intervenção humana (mesmo que reduzida). Entretanto, nos experimentos, e através dos testes de hipóteses, foi verificado que não há diferença estatística significativa entre ambas.

Por fim, a quinta hipótese é validada através da análise com a base M&Ms. Em geral, os resultados sofrem uma queda quando comparados com aqueles obtidos no teste com a base ACDC. Entretanto, pode se considerar que os métodos mantêm um nível de desempenho satisfatório, visto que a base M&Ms apresenta uma simulação mais próxima ao cenário real da aplicação, com uma maior quantidade de exames adquiridos por *scanners* de diferentes fabricantes.

7 CONCLUSÃO

As doenças cardiovasculares são responsáveis por milhões de óbitos todos os anos. Nesse cenário, exames não invasivos como o de cine-ressonância magnética (cine-RM) têm favorecido a aquisição de melhores informações acerca dessas patologias, auxiliando os especialistas na obtenção de diagnósticos precoces e, por conseguinte, na condução de tratamentos essenciais para melhorar a qualidade de vida dos indivíduos. Através da cine-RM, os especialistas podem analisar de forma mais precisa as estruturas cardíacas para a avaliação de suas funções. Nesse contexto, métodos computacionais podem auxiliá-los nessa tarefa, reduzindo o tempo da análise e identificação de patologias.

Neste trabalho, foram propostos dois métodos para a segmentação de estruturas cardíacas em imagens de cine-RM em eixo curto: método de segmentação em cascata com reconstrução (SCR), e método de segmentação especializada por contextos (SEC). As estruturas de interesse são a cavidade ventricular (CV), o miocárdio (Mio) e o ventrículo direito (VD). A segmentação delas é uma tarefa importante no âmbito do diagnóstico precoce de doenças cardiovasculares bem como do acompanhamento de pacientes com alguma dessas anormalidades. Os métodos propostos têm como característica comum a divisão do processo de segmentação em etapas, de modo que, em cada passo, são usadas técnicas conforme aspectos específicos de cada estrutura cardíaca e demais características dos exames.

Nos experimentos realizados, os resultados obtidos pelos métodos propostos mostraram-se promissores. A divisão em etapas apresentou melhor desempenho em relação à segmentação em etapa única. Em relação ao método SCR, a EAIS-Net utilizada na etapa de segmentação inicial supera a U-Net que é um modelo considerado meritório pela literatura. E a etapa de refinamento, representada pela reconstrução e pós-processamentos, mostra-se essencial para o melhoramento da segmentação.

Quanto ao método SEC, este usa uma quantidade de FCNs menor que o método SCR. Ainda assim, mostrou-se promissor, visto que obteve, nos testes locais, resultados aproximados aos do método SCR, os quais não possuem diferença estatística significativa conforme os testes de hipóteses realizados. Não obstante essa constatação, entre eles, o método SCR produz resultados melhores de acordo com as métricas de avaliação calculadas para ambos.

Por fim, no processo de validação na plataforma online do desafio ACDC, e com o conjunto de testes da base M&Ms, ambos os métodos apresentam resultados expressivos, superando ou mantendo um resultado numericamente próximo ao de abordagens compostas por técnicas mais robustas, que exigem mais custo computacional. Os métodos obtém melhor desempenho para as segmentações do CV e do Mio, com destaque ao método SCR, que alcança resultados aproximados aos dos primeiros colocados. Isso, portanto, corrobora a aplicabilidade dos métodos propostos pelo presente trabalho.

7.1 Trabalhos Futuros

Não obstante os resultados promissores obtidos, algumas tarefas podem ser desempenhadas para contribuir com os métodos propostos e complementar o estudo realizado. Desta feita, a seguir serão apresentadas algumas sugestões de trabalhos futuros.

Primeiro, aponta-se o estudo de outras composições de blocos Attention, como a multi-escala (SINHA; DOLZ, 2020) de modo a aplicá-los nas principais redes utilizadas pelos métodos propostos (EAIS-Net e IRAX-Net), com o intuito de melhorar os resultados principalmente nos casos das fatias na fase de ES.

Propõe-se também acrescentar ao módulo de pós-processamento um procedimento que leve em consideração as características anatômicas do Mio e da CV para o melhoramento da segmentação dessas estruturas. A CV é delimitada pelos contornos internos do Mio. Portanto, uma boa definição desses contornos pode acarretar a detecção de regiões da CV não previstas anteriormente pela segmentação inicial.

Outra sugestão é a aplicação do módulo de reconstrução, que faz parte do método SCR, no método SEC, especialmente na etapa de segmentação do VD, para a qual esse método apresentou mais falhas. E, por fim, verificar a viabilidade do uso de outras redes, como as da família EfficientNet e Inception, para a composição do módulo de reconstrução, que é atualmente baseado na U-Net.

7.2 Produções Científicas

A Tabela 35 apresenta o artigo relacionado aos métodos apresentados neste trabalho para a segmentação de estruturas cardíacas. Além disso, a Tabela 36 lista os artigos submetidos e publicados com relação a outras atividades envolvendo processamento de imagens e visão computacional realizadas desde o início do doutorado.

Tabela 35 – Artigo relacionado à segmentação de estruturas cardíacas.

Tipo	Artigo	Qualis	Status
Periódico	Silva, I.F.S, et al. (2022). A Cascade Approach for Automatic Segmentation of Cardiac Structures in Short-axis Cine-MR Images using Deep Neural Networks. Em Expert Systems with Applications.	A1	Publicado

Tabela 36 – Publicações envolvendo processamento de imagens e visão computacional.

Tipo	Artigo	Qualis	Status
Conferência	Oliveira, J, et al. (2021). Applying Exploratory Testing and Ad-Hoc Usability Inspection to Improve the Ease of Use of a Mobile Power Consumption Registration App: An Experience Report. Em 23rd International Conference on Human-Computer Interaction (HCII).	B1	Publicado
Conferência	Rocha, C, et al. (2021). A Comparative Study of Methods based on Deep Neural Networks for Self-reading of Energy Consumption in a Chatbot Application Context. Em 34th Conference on Graphics, Patterns and Images (SIBGRAPI).	A3	Publicado
Simpósio	Anchieta Junior, R. J. F, et al. (2021). Detecção de nódulos da tireoide em exames de termografia utilizando redes neurais convolucionais em cascata. Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS).	B3	Publicado
Periódico	Rocha, C. V. M, et al. (2021). A Chatbot Solution for Self-reading Energy Consumption via Chatting Applications. Em Journal of Control, Automation and Electrical Systems.	B1	Publicado
Congresso	Cabrejos, L. J. E. R, et al. (2021). Implementando o Gitflow para Gerencia de Configuração em um Projeto de Desenvolvimento de Software Ágil: Um Relato de Experiência. Em 12o <i>Computer on the Beach</i> .	B3	Publicado
Periódico	Serra, A. C, et al. (2020). Reconhecimento de Dígitos em Imagens de Medidores de Energia no Contexto de um Aplicativo de Autoleitura. Em Brazilian Applied Science Review (BASR).	B3	Publicado
Periódico	Mendes, P. R. C, et al. (2020). Reconhecimento de Dígitos de Medidores de Energia por meio da Voz no Contexto de um Aplicativo de Autoleitura. Em Brazilian Applied Science Review (BASR).	B3	Publicado
Periódico	Serra, A. C, et al. (2020). Segmentação Semântica de Medidores de Energia Elétrica e Componentes de Identificação. Em Brazilian Applied Science Review (BASR).	B3	Publicado
Conferência	Rocha, C. V. M, et al. (2020). Self-Reading by Speech: An End-to-end Speech Recognition Method Applied to the Energy Consumption Measurement on Mobile Devices. Em International Conference on Systems, Signals and Image Processing (IWSSIP).	B1	Publicado

REFERÊNCIAS

- ABDELTAWAB, H.; KHALIFA, F.; TAHER, F.; ALGHAMDI, N. S.; GHAZAL, M.; BEACHE, G.; MOHAMED, T.; KEYNTON, R.; EL-BAZ, A. A deep learning-based approach for automatic segmentation and quantification of the left ventricle from cardiac cine mr images. **Computerized Medical Imaging and Graphics**, Elsevier, v. 81, p. 101717, 2020.
- ABRAHAM, N.; KHAN, N. M. A novel focal tversky loss function with improved attention u-net for lesion segmentation. In: **2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)**. [S.l.: s.n.], 2019. p. 683–687.
- AMMAR, A.; BOUATTANE, O.; YOUSSEFI, M. Automatic cardiac cine mri segmentation and heart disease classification. **Computerized Medical Imaging and Graphics**, v. 88, p. 101864, 2021. ISSN 0895-6111.
- AREL, I.; ROSE, D. C.; KARNOWSKI, T. P. Deep machine learning - a new frontier in artificial intelligence research [research frontier]. **IEEE Computational Intelligence Magazine**, v. 5, n. 4, p. 13–18, Nov 2010. ISSN 1556-603X.
- AZIZ, Y. F. A.; FADZLI, F.; AZMAN, R. R.; SANI, F. M.; VIJAYANANTHAN, A.; NAZRI, M. State of the heart: Cmr in coronary artery disease. **Current Medical Imaging Reviews**, Bentham Science Publishers, v. 9, n. 3, p. 201–213, 2013.
- BADRINARAYANAN, V.; KENDALL, A.; CIPOLLA, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 39, n. 12, p. 2481–2495, 2017.
- BAI, W.; SHI, W.; LEDIG, C.; RUECKERT, D. Multi-atlas segmentation with augmented features for cardiac mr images. **Medical image analysis**, Elsevier, v. 19, n. 1, p. 98–109, 2015.
- BAUMGARTNER, C. F.; KOCH, L. M.; POLLEFEYS, M.; KONUKOGLU, E. An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation. In: SPRINGER. **International Workshop on Statistical Atlases and Computational Models of the Heart**. [S.l.], 2017. p. 111–119.
- BAZIRA, P. J. Clinically applied anatomy of the heart. **Surgery (Oxford)**, Elsevier, 2021.
- BERNARD, O.; LALANDE, A.; ZOTTI, C.; CERVENANSKY, F.; YANG, X.; HENG, P.-A.; CETIN, I.; LEKADIR, K.; CAMARA, O.; BALLESTER, M. A. G. et al. Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved? **IEEE transactions on medical imaging**, ieee, v. 37, n. 11, p. 2514–2525, 2018.
- BEZERRA, L.; MARCHIORI, E.; PONTES, P. Avaliação da função cardíaca por ressonância magnética com seqüências em equilíbrio estável: segmentadas × tempo real. **Radiologia Brasileira**, v. 39, 10 2006.
- BLAND, M. **An introduction to medical statistics (4th ed.)**. [S.l.]: Oxford University Press (UK), 2015.

BULLOCK, J.; CUESTA-LÁZARO, C.; QUERA-BOFARULL, A. XNet: a convolutional neural network (CNN) implementation for medical x-ray image segmentation suitable for small datasets. In: GIMI, B.; KROL, A. (Ed.). **Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging**. [S.l.]: SPIE, 2019. v. 10953, p. 453 – 463.

CAELLES, S.; MANINIS, K.-K.; PONT-TUSET, J.; LEAL-TAIXÉ, L.; CREMERS, D.; GOOL, L. V. One-shot video object segmentation. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2017. p. 221–230.

CALISTO, M. B.; LAI-YUEN, S. K. Adaen-net: An ensemble of adaptive 2d–3d fully convolutional networks for medical image segmentation. **Neural Networks**, Elsevier, v. 126, p. 76–94, 2020.

CAMPELLO, V. M.; GKONTRA, P.; IZQUIERDO, C.; MARTIN-ISLA, C.; SOJOUDI, A.; FULL, P. M.; MAIER-HEIN, K.; ZHANG, Y.; HE, Z.; MA, J.; PARREÑO, M.; ALBIOL, A.; KONG, F.; SHADDEN, S. C.; ACERO, J. C.; SUNDARESAN, V.; SABER, M.; ELATTAR, M.; LI, H.; MENZE, B.; KHADER, F.; HAARBURGER, C.; SCANNELL, C. M.; VETA, M.; CARSCADDEN, A.; PUNITHAKUMAR, K.; LIU, X.; TSAFTARIS, S. A.; HUANG, X.; YANG, X.; LI, L.; ZHUANG, X.; VILADÉS, D.; DESCALZO, M. L.; GUALA, A.; MURA, L. L.; FRIEDRICH, M. G.; GARG, R.; LEBEL, J.; HENRIQUES, F.; KARAKAS, M.; ÇAVUŞ, E.; PETERSEN, S. E.; ESCALERA, S.; SEGUÍ, S.; RODRÍGUEZ-PALOMARES, J. F.; LEKADIR, K. Multi-centre, multi-vendor and multi-disease cardiac segmentation: The m amp;ms challenge. **IEEE Transactions on Medical Imaging**, v. 40, n. 12, p. 3543–3554, 2021.

CAMPELLO, V. M.; GKONTRA, P.; IZQUIERDO, C.; MARTIN-ISLA, C.; SOJOUDI, A.; FULL, P. M.; MAIER-HEIN, K.; ZHANG, Y.; HE, Z.; MA, J. et al. Multi-centre, multi-vendor and multi-disease cardiac segmentation: the m&ms challenge. **IEEE Transactions on Medical Imaging**, IEEE, v. 40, n. 12, p. 3543–3554, 2021.

CHOLLET, F. et al. **Keras**. 2015. Disponível em: <<https://keras.io>>. Acesso em: 01/05/2021.

CUI, H.; YUWEN, C.; JIANG, L.; XIA, Y.; ZHANG, Y. Multiscale attention guided u-net architecture for cardiac segmentation in short-axis mri images. **Computer Methods and Programs in Biomedicine**, v. 206, p. 106142, 2021. ISSN 0169-2607.

FILHO, C. F. d. A.; HADLICH, M.; PETRIZ, J. L. F.; MENDONÇA, L. A.; FILHO, J. N. M.; ROCHITTE, C. E. Quantificação da massa infartada do ventrículo esquerdo pela ressonância magnética cardíaca: comparação entre a planimetria e o método de escore visual semi-quantitativo. **Arquivos Brasileiros de Cardiologia**, SciELO Brasil, v. 83, n. 2, p. 111–117, 2004.

FULL, P. M.; ISENSEE, F.; JÄGER, P. F.; MAIER-HEIN, K. Studying robustness of semantic segmentation under domain shift in cardiac mri. In: SPRINGER. **International Workshop on Statistical Atlases and Computational Models of the Heart**. [S.l.], 2020. p. 238–249.

GAO, L.; ZHANG, L.; LIU, C.; WU, S. Handling imbalanced medical image data: A deep-learning-based one-class classification approach. **Artificial Intelligence in Medicine**, Elsevier, v. 108, p. 101935, 2020.

- GONZALEZ, R. C.; WOODS, R. C. **Processamento Digital de Imagens**. [S.l.]: Pearson Prentice Hall, 2010.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016. Disponível em: <<http://www.deeplearningbook.org>>. Acesso em: 14/03/2022.
- GRINIAS, E.; TZIRITAS, G. Fast fully-automatic cardiac segmentation in mri using mrf model optimization, substructures tracking and b-spline smoothing. In: SPRINGER. **International Workshop on Statistical Atlases and Computational Models of the Heart**. [S.l.], 2017. p. 91–100.
- GUO, Y.; LIU, Y.; OERLEMANS, A.; LAO, S.; WU, S.; LEW, M. S. Deep learning for visual understanding: A review. **Neurocomputing**, v. 187, p. 27 – 48, 2016. ISSN 0925-2312. Recent Developments on Deep Big Vision.
- HAZRA, A.; MANDAL, S. K.; GUPTA, A.; MUKHERJEE, A.; MUKHERJEE, A. Heart disease diagnosis and prediction using machine learning and data mining techniques: a review. **Advances in Computational Sciences and Technology**, v. 10, n. 7, p. 2137–2159, 2017.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 770–778.
- HIGGINS, C. B. **Which standard has the gold?** [S.l.]: American College of Cardiology Foundation Washington, DC, 1992.
- HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural Networks**, v. 2, n. 5, p. 359 – 366, 1989. ISSN 0893-6080.
- HOWARD, A. G.; ZHU, M.; CHEN, B.; KALENICHENKO, D.; WANG, W.; WEYAND, T.; ANDREETTO, M.; ADAM, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. **arXiv preprint arXiv:1704.04861**, 2017.
- HU, H.; PAN, N.; WANG, J.; YIN, T.; YE, R. Automatic segmentation of left ventricle from cardiac mri via deep learning and region constrained dynamic programming. **Neurocomputing**, Elsevier, v. 347, p. 139–148, 2019.
- HU, J.; SHEN, L.; SUN, G. Squeeze-and-excitation networks. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2018. p. 7132–7141.
- HUA, K.-L.; HSU, C.-H.; HIDAYATI, S. C.; CHENG, W.-H.; CHEN, Y.-J. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. **OncoTargets and therapy**, Dove Press, v. 8, 2015.
- HUANG, X.; CHEN, W.; LIU, X.; WU, H.; WEN, Z.; SHEN, L. Left and right ventricular segmentation based on 3d region-aware u-net. In: IEEE. **2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS)**. [S.l.], 2022. p. 137–142.

HUANG, X.; CHEN, Z.; YANG, X.; LIU, Z.; ZOU, Y.; LUO, M.; XUE, W.; NI, D. Style-invariant cardiac image segmentation with test-time augmentation. In: SPRINGER. **International Workshop on Statistical Atlases and Computational Models of the Heart**. [S.l.], 2020. p. 305–315.

ISENSEE, F.; JAEGGER, P. F.; FULL, P. M.; WOLF, I.; ENGELHARDT, S.; MAIER-HEIN, K. H. Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features. In: SPRINGER. **International workshop on statistical atlases and computational models of the heart**. [S.l.], 2017. p. 120–129.

JADON, S. A survey of loss functions for semantic segmentation. In: IEEE. **2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)**. [S.l.], 2020. p. 1–7.

KHAN, M. G. **Encyclopedia of heart diseases**. [S.l.]: Elsevier, 2005.

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**, Elsevier, v. 51, p. 21–45, 2019.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. **arXiv preprint arXiv:1412.6980**, 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, Nov 1998. ISSN 0018-9219.

LI, B.; LIU, Y.; OCCLESHAW, C. J.; COWAN, B. R.; YOUNG, A. A. In-line automated tracking for ventricular function with magnetic resonance imaging. **JACC: Cardiovascular Imaging**, American College of Cardiology Foundation Washington, DC, v. 3, n. 8, p. 860–866, 2010.

LI, F.; LI, W.; GAO, X.; LIU, R.; XIAO, B. Dcnet: Diversity convolutional network for ventricle segmentation on short-axis cardiac magnetic resonance images. **Knowledge-Based Systems**, p. 110033, 2022. ISSN 0950-7051. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0950705122011261>>.

LI, S.; ZHANG, Y.; YANG, X. Semi-supervised cardiac mri segmentation based on generative adversarial network and variational auto-encoder. In: **2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)**. [S.l.: s.n.], 2021. p. 1402–1405.

LIN, M.; JIANG, M.; ZHAO, M.; UKWATTA, E.; WHITE, J. A.; CHIU, B. Cascaded triplanar autoencoder m-net for fully automatic segmentation of left ventricle myocardial scar from three-dimensional late gadolinium-enhanced mr images. **IEEE Journal of Biomedical and Health Informatics**, IEEE, v. 26, n. 6, p. 2582–2593, 2022.

LIU, L.; ÖZSU, M. T. **Encyclopedia of Database Systems**. [S.l.]: Springer US, 2009. ISBN: 978-0-387-39940-9.

LIU, W.; WANG, Z.; LIU, X.; ZENG, N.; LIU, Y.; ALSAADI, F. E. A survey of deep neural network architectures and their applications. **Neurocomputing**, v. 234, p. 11 – 26, 2017. ISSN 0925-2312.

- LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2015. p. 3431–3440.
- MURDOCH, T. B.; DETSKY, A. S. The inevitable application of big data to health care. **Jama**, American Medical Association, v. 309, n. 13, p. 1351–1352, 2013.
- MYERSON, S. G.; BELLENGER, N. G.; PENNELL, D. J. Assessment of left ventricular mass by cardiovascular magnetic resonance. **Hypertension**, Am Heart Assoc, v. 39, n. 3, p. 750–755, 2002.
- NACIF, M. S.; JUNIOR, A. C. d. O.; CARVALHO, A. C. P.; ROCHITTE, C. E. Ressonância magnética cardíaca e seus planos anatômicos: como eu faço? **Arquivos Brasileiros de Cardiologia**, SciELO Brasil, v. 95, n. 6, p. 756–763, 2010.
- NAIR, V.; HINTON, G. E. Rectified linear units improve restricted boltzmann machines. In: FÜRKNRANZ, J.; JOACHIMS, T. (Ed.). **Proceedings of the 27th International Conference on Machine Learning (ICML-10)**. [S.l.: s.n.], 2010. p. 807–814.
- NASR-ESFAHANI, M.; MOHREKESH, M.; AKBARI, M.; SOROUSHMEHR, S. R.; NASR-ESFAHANI, E.; KARIMI, N.; SAMAVI, S.; NAJARIAN, K. Left ventricle segmentation in cardiac mr images using fully convolutional network. In: IEEE. **2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)**. [S.l.], 2018. p. 1275–1278.
- NETO, J. B. R.; SILVA, N. D.; SILVA, W. G. Mortalidade por doenças cardiovasculares no estado do maranhão no quadriênio 2016-2019. **Revista da Saúde da AJES**, v. 8, n. 15, 2022.
- NOH, H.; HONG, S.; HAN, B. Learning deconvolution network for semantic segmentation. In: **Computer Vision (ICCV), 2015 IEEE International Conference on**. [S.l.: s.n.], 2015.
- OKTAY, O.; SCHLEMPER, J.; FOLGOC, L. L.; LEE, M.; HEINRICH, M.; MISAWA, K.; MORI, K.; MCDONAGH, S.; HAMMERLA, N. Y.; KAINZ, B. et al. Attention u-net: Learning where to look for the pancreas. **arXiv preprint arXiv:1804.03999**, 2018.
- PAINCHAUD, N.; SKANDARANI, Y.; JUDGE, T.; BERNARD, O.; LALANDE, A.; JODOIN, P.-M. Cardiac mri segmentation with strong anatomical guarantees. In: SPRINGER. **International Conference on Medical Image Computing and Computer-Assisted Intervention**. [S.l.], 2019. p. 632–640.
- PATRAVALI, J.; JAIN, S.; CHILAMKURTHY, S. 2d-3d fully convolutional neural networks for cardiac mr segmentation. In: SPRINGER. **International Workshop on Statistical Atlases and Computational Models of the Heart**. [S.l.], 2017. p. 130–139.
- PEDRINI, H.; SCHWARTZ, W. R. **Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações**. [S.l.]: Editora Thomson Learning, 2007. 528 p. ISBN 978-85-221-0595-3.
- PINHEIRO, T. B.; TIRADENTES, C. A.; JORGE, A. J. L.; RIBEIRO, M. R.; MESQUITA, E. T. Insuficiência cardíaca com fração de ejeção normal em paciente adulto jovem. **Revista Brasileira de Cardiologia**, p. 54–57, 2013.

- PRECHELT, L. Early stopping-but when? In: **Neural Networks: Tricks of the trade**. [S.l.]: Springer, 1998. p. 55–69.
- RADAU, P.; LU, Y.; CONNELLY, K.; PAUL, G.; DICK, A.; WRIGHT, G. Evaluation framework for algorithms segmenting short axis cardiac mri. **The MIDAS Journal-Cardiac MR Left Ventricle Segmentation Challenge**, v. 49, 2009.
- RECHT, B.; ROELOFS, R.; SCHMIDT, L.; SHANKAR, V. Do imagenet classifiers generalize to imagenet? In: PMLR. **International Conference on Machine Learning**. [S.l.], 2019. p. 5389–5400.
- REDMON, J.; DIVVALA, S. K.; GIRSHICK, R. B.; FARHADI, A. You only look once: Unified, real-time object detection. **2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, abs/1506.02640, 2015. Disponível em: <<http://arxiv.org/abs/1506.02640>>.
- ROCHITTE, C. E.; PINTO, I. M. F.; FERNANDES, J. L.; FILHO, C. F. A.; JATENE, A.; CAMARGO, A. C. de; CARVALHO, J. P. R.; RAMIRES, J. A. F.; OLIVEIRA, S. A.; CATTANI, C. A. M. et al. Diretrizes sbc ressonância e tomografia cardiovascular. **Arquivos Brasileiros de Cardiologia**, v. 87, n. 3, 2006.
- ROHÉ, M.-M.; SERMESANT, M.; PENNEC, X. Automatic multi-atlas segmentation of myocardium with svf-net. In: SPRINGER. **International Workshop on Statistical Atlases and Computational Models of the Heart**. [S.l.], 2017. p. 170–177.
- ROMANO, J. P.; LEHMANN, E. **Testing statistical hypotheses** (3rd ed.). [S.l.]: Springer Berlin, 2005. ISBN: 0-387-98864-5.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. **International Conference on Medical Image Computing and Computer-Assisted Intervention**. [S.l.], 2015. p. 234–241.
- SAKUMA, H.; FUJITA, N.; FOO, T.; CAPUTO, G. R.; NELSON, S. J.; HARTIALA, J.; SHIMAKAWA, A.; HIGGINS, C. B. Evaluation of left ventricular volume and mass with breath-hold cine mr imaging. **Radiology**, v. 188, n. 2, p. 377–380, 1993.
- SAMMUT, C.; WEBB, G. I. **Encyclopedia of Machine Learning**. [S.l.]: Springer US, 2010. ISBN: 978-0-387-30164-8.
- SANDLER, M.; HOWARD, A.; ZHU, M.; ZHMOGINOV, A.; CHEN, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2018. p. 4510–4520.
- SANTURKAR, S.; TSIPRAS, D.; ILYAS, A.; MADRY, A. How does batch normalization help optimization? **arXiv preprint arXiv:1805.11604**, 2018.
- SARA, L.; SZARF, G.; TACHIBANA, A.; SHIOZAKI, A. A.; VILLA, A. V.; OLIVEIRA, A. C. d.; ALBUQUERQUE, A. S. d.; ROCHITTE, C. E.; NOMURA, C. H.; AZEVEDO, C. F. et al. Li diretriz de ressonância magnética e tomografia computadorizada cardiovascular da sociedade brasileira de cardiologia e do colégio brasileiro de radiologia. **Arquivos Brasileiros de Cardiologia**, SciELO Brasil, v. 103, n. 6, p. 1–86, 2014.

- SBC. **Sociedade Brasileira de Cardiologia. Cardiômetro**. 2022. <<http://www.cardiometro.com.br/default.asp>>. "Acesso em 01/11/2022".
- SCANNELL, C. M.; CHIRIBIRI, A.; VETA, M. Domain-adversarial learning for multi-centre, multi-vendor, and multi-disease cardiac mr image segmentation. In: SPRINGER. **International Workshop on Statistical Atlases and Computational Models of the Heart**. [S.l.], 2020. p. 228–237.
- SECHTEM, U.; PFLUGFELDER, P. W.; GOULD, R. G.; CASSIDY, M.; HIGGINS, C. B. Measurement of right and left ventricular volumes in healthy individuals with cine mr imaging. **Radiology**, v. 163, n. 3, p. 697–702, 1987.
- SERRA, J. Introduction to mathematical morphology. **Computer Vision, Graphics, and Image Processing**, Academic Press, v. 35, n. 3, p. 283 – 305, 1986. ISSN 0734-189X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0734189X86900022>>.
- SHI, J.; YE, Y.; ZHU, D.; SU, L.; HUANG, Y.; HUANG, J. Automatic segmentation of cardiac magnetic resonance images based on multi-input fusion network. **Computer Methods and Programs in Biomedicine**, Elsevier, v. 209, p. 106323, 2021.
- SILVA, I. F. S. da. **Detecção automática da presença de patologia na visão baseada em imagens do teste de Brückner**. Dissertação (Mestrado) — Universidade Federal do Maranhão, 2019.
- SIMANTIRIS, G.; TZIRITAS, G. Cardiac mri segmentation with a dilated cnn incorporating domain-specific constraints. **IEEE Journal of Selected Topics in Signal Processing**, IEEE, v. 14, n. 6, p. 1235–1243, 2020.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014.
- SINHA, A.; DOLZ, J. Multi-scale self-guided attention for medical image segmentation. **IEEE journal of biomedical and health informatics**, IEEE, 2020.
- SOUZA, J. C.; DINIZ, J. O. B.; FERREIRA, J. L.; SILVA, G. L. F. da; SILVA, A. C.; PAIVA, A. C. de. An automatic method for lung segmentation and reconstruction in chest x-ray using deep neural networks. **Computer methods and programs in biomedicine**, Elsevier, v. 177, p. 285–296, 2019.
- SUINESIAPUTRA, A.; COWAN, B. R.; AL-AGAMY, A. O.; ELATTAR, M. A.; AYACHE, N.; FAHMY, A. S.; KHALIFA, A. M.; MEDRANO-GRACIA, P.; JOLLY, M.-P.; KADISH, A. H. et al. A collaborative resource to build consensus for automated left ventricular segmentation of cardiac mr images. **Medical image analysis**, Elsevier, v. 18, n. 1, p. 50–62, 2014.
- SZEGEDY, C.; IOFFE, S.; VANHOUCHE, V. Inception-v4, inception-resnet and the impact of residual connections on learning. **Computing Research Repository (CoRR)**, 2016. Disponível em: <<http://arxiv.org/abs/1602.07261>>.
- SZEGEDY, C.; IOFFE, S.; VANHOUCHE, V.; ALEMI, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning. In: **Thirty-first AAAI conference on artificial intelligence**. [S.l.: s.n.], 2017.

SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCKE, V.; RABINOVICH, A. Going deeper with convolutions. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2015. p. 1–9.

TAHA, A. A.; HANBURY, A. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. **BMC medical imaging**, BioMed Central, v. 15, n. 1, p. 29, 2015.

TAN, L. K.; LIEW, Y. M.; LIM, E.; MCLAUGHLIN, R. A. Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine mr sequences. **Medical image analysis**, Elsevier, v. 39, p. 78–86, 2017.

TAN, M.; LE, Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks. **arXiv preprint arXiv:1905.11946**, 2019.

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

WANG, J.; LIU, X. Medical image recognition and segmentation of pathological slices of gastric cancer based on deeplab v3+ neural network. **Computer Methods and Programs in Biomedicine**, Elsevier, v. 207, p. 106210, 2021.

WOLTERINK, J. M.; LEINER, T.; VIERGEVER, M. A.; IŠGUM, I. Automatic segmentation and disease classification using cardiac cine mr images. In: SPRINGER. **International Workshop on Statistical Atlases and Computational Models of the Heart**. [S.l.], 2017. p. 101–110.

ZOTTI, C.; LUO, Z.; HUMBERT, O.; LALANDE, A.; JODOIN, P.-M. Gridnet with automatic shape prior registration for automatic mri cardiac segmentation. In: SPRINGER. **International Workshop on Statistical Atlases and Computational Models of the Heart**. [S.l.], 2017. p. 73–81.

ZOTTI, C.; LUO, Z.; LALANDE, A.; JODOIN, P.-M. Convolutional neural network with shape prior applied to cardiac mri segmentation. **IEEE journal of biomedical and health informatics**, IEEE, v. 23, n. 3, p. 1119–1128, 2018.