

“Fábrica de erros”: um arcabouço para simulação de erros na segmentação do ventrículo esquerdo em imagens cardíacas

Bruno F. Raquel¹, Matheus A. O. Ribeiro¹, Fátima L. S. Nunes¹

¹Laboratório de Aplicações de Informática em Saúde,
Escola de Artes, Ciências e Humanidades, Universidade de São Paulo
Rua Arlindo Bettio, 1000, São Paulo – SP, Brazil, 03828-000

{brunofraquel, matheus.alberto.oliveira, fatima.nunes}@usp.br

Abstract. *Left ventricle segmentation in Cardiac Magnetic Resonance exams is important for medical diagnosis. Deep learning methods have excelled in obtaining segmentations similar to those of experts. However, one of the current limitations is the arbitrary production of anatomical errors that can compromise the diagnosis. Given this problem, this work proposes a framework for categorization and controlled automatic simulation of different anatomical errors. The framework favors the development of methods aimed at detecting and correcting these errors. Results indicate that the proposed framework is capable of generating errors of several categories and replicating the same errors produced by networks ($Dice > 0.8$).*

Resumo. *A segmentação do ventrículo esquerdo em exames de Ressonância Magnética Cardíaca é importante para o diagnóstico médico. Métodos de aprendizado profundo têm se destacado ao obter segmentações semelhantes a de especialistas. Entretanto, uma das limitações atuais é a produção arbitrária de erros anatômicos que podem comprometer o diagnóstico. Dado esse problema, esse trabalho propõe um arcabouço para categorização e simulação controlada e automática de diferentes erros anatômicos. O arcabouço favorece o desenvolvimento de métodos voltados à detecção e à correção desses erros. Resultados indicam que o arcabouço proposto é capaz de gerar erros de diversas categorias e replicar os mesmos erros produzidos por redes ($Dice > 0.8$).*

1. Introdução

A correta segmentação do ventrículo esquerdo (VE) é essencial para a extração de biomarcadores cardíacos, sendo exames de Ressonância Magnética Cardíaca (RMC) o padrão-ouro para essa análise devido à sua precisão e reprodutibilidade [Lohr et al. 2024]. A segmentação manual, feita por especialistas, é exaustiva e sujeita a variabilidade inter e intraobservador [Waite et al. 2017], motivando o desenvolvimento de métodos automáticos para maior eficiência na extração de biomarcadores [Wang et al. 2021].

Métodos automáticos baseados em aprendizado profundo (AP) frequentemente alcançam desempenho próximo ao de especialistas [Bernard et al. 2018], mas podem gerar erros anatômicos arbitrários que especialistas não cometeriam [Bernard et al. 2018; Ribeiro and Nunes 2022], impactando a confiabilidade dos biomarcadores e potencialmente levando a diagnósticos incorretos. Apesar da recorrência desses erros, há uma escassez de estudos focados na caracterização e replicação sistemática dessas falhas.

Diante desse cenário, este trabalho propõe um arcabouço para a simulação e análise de erros na segmentação do VE, estruturado como uma “fábrica de erros” capaz de gerar e manipular diferentes falhas de forma calibrável e replicável. As principais contribuições incluem: (i) a categorização dos erros mais comuns em segmentações automáticas, definindo suas características específicas; e (ii) o desenvolvimento de um sistema que reproduz esses erros a partir de segmentações corretas, permitindo um maior controle sobre sua aplicação e intensidade.

2. Trabalhos correlatos

A segmentação de estruturas cardíacas em imagens médicas tem sido amplamente estudada, com métodos que vão desde abordagens clássicas, como modelos deformáveis, até técnicas baseadas em inteligência artificial [Ribeiro and Nunes 2022]. No entanto, poucos trabalhos exploram a caracterização e a replicação de erros anatômicos em segmentações, uma etapa importante para avaliar a qualidade e aprimorar métodos automáticos.

Nesse contexto, Tajbakhsh et al. [2020] propõem uma rede que injeta erros em segmentações de vasos da retina para treinar modelos de predição de falhas. De forma similar, Guan et al. [2023] desenvolvem uma ferramenta parametrizada que gera segmentações com erros, mas sem considerar informações da imagem, como intensidades dos *pixels*, restringindo sua aplicabilidade.

Embora técnicas para aumento de dados, como rotações, translações e deformações, sejam amplamente utilizadas na geração de dados artificiais [Yuan et al. 2018], elas nem sempre refletem falhas reais de segmentação. Para reduzir essa limitação, este trabalho propõe um arcabouço capaz de simular erros em segmentações, gerando exemplos mais próximos a falhas reais observáveis. O método proposto favorece o treinamento de modelos especializados na detecção e correção dessas inconsistências.

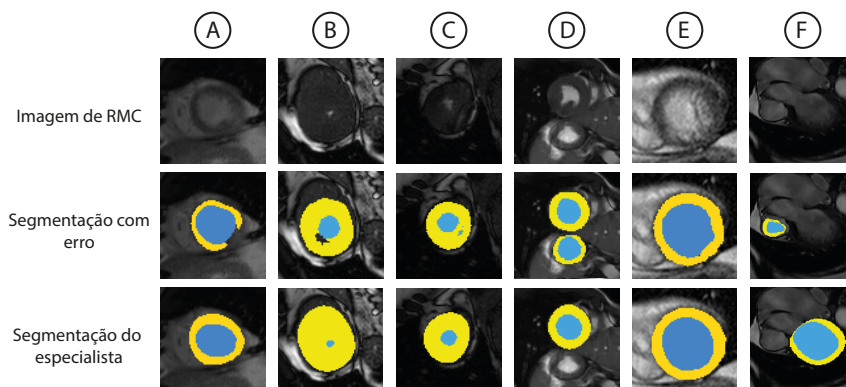
3. Materiais e métodos

A segmentação do VE tem como objetivo identificar duas estruturas principais: o miocárdio e a câmara cardíaca. O miocárdio corresponde ao músculo cardíaco, enquanto a câmara cardíaca indica a região interna do VE, preenchida de sangue. Dessa forma, a segmentação é representada por duas máscaras que delimitam essas regiões, conforme ilustrado na Figura 1. O método proposto para simulação de erros recebe como entrada uma segmentação correta, usada como referência, e a imagem original do exame de RMC. Em seguida, aplica operações selecionadas de forma aleatória ou manual para reproduzir falhas observadas em segmentações automáticas.

3.1. Dataset e visão geral do método

Os exames de RMC foram obtidos a partir de três bases disponibilizadas publicamente: ACDC [Bernard et al. 2018], Sata13 [Landman and Warfield 2013] e Sunny [Radau et al. 2009]. Essas bases são amplamente utilizadas pela literatura e possuem segmentações produzidas por especialistas que são utilizadas como segmentações corretas [Ribeiro and Nunes 2022]. Para obter segmentações com erros anatômicos reais, redes de AP foram treinadas e testadas considerando diferentes pares de bases. O método utilizado possui etapas de pré-processamento, obtenção da região de interesse e segmentação com uma rede U-net [Ribeiro and Nunes 2021]. Um total de 390 casos contendo erros anatômicos

Figura 1. Visão geral dos tipos de erros anatômicos identificados. Para cada erro, a primeira imagem não possui segmentação, a segunda imagem é a segmentação com erro anatômico e a terceira imagem é a segmentação do especialista. As regiões destacadas em azul e amarelo indicam respectivamente, a máscara da câmara cardíaca e do miocárdio: (A) Buraco no miocárdio; (B) Câmara cardíaca com tamanho maior do que o correto; (C) Máscara da câmara cardíaca invadindo o miocárdio; (D) Dois componentes desconexos segmentados; (E) Borda do miocárdio espiculada; (F) Segmentação está com seu centro deslocado para a esquerda.



diversos foram selecionados e utilizados como referência para a catalogação e construção do arcabouço.

3.2. Simulação de imagens com erros

A partir da análise manual das segmentações contendo erros anatômicos produzidos por redes de AP, seis categorias de erros foram definidas, conforme exemplificadas na Figura 1. A identificação de erros considera as seguintes propriedades anatômicas: i) VE possui formato arredondado; ii) miocárdio envolve completamente a câmara cardíaca; iii) não existência de buracos dentro das máscaras de segmentação; iv) regiões de uma mesma máscara sejam conexas. Destaca-se que, a existência de erros não é mutuamente exclusiva, isto é, é possível que um ou mais erros existam simultaneamente em uma mesma segmentação, com diferentes intensidades. O processo para simular cada categoria de erro em segmentações é detalhado a seguir.

Buraco no miocárdio e/ou câmara cardíaca (Figura 1A): as áreas correspondentes ao miocárdio e à câmara cardíaca devem aparecer como regiões únicas e contínuas, portanto, não devem apresentar buracos ou lacunas. Para simular esse erro, o método seleciona um ponto de partida para a geração do buraco. Em seguida, o algoritmo expande o buraco de forma iterativa, removendo os pixels que possuem intensidades mais distantes da média da intensidade da máscara na qual será inserido o buraco.

Tamanho dilatado ou reduzido do miocárdio ou câmara cardíaca (Figura 1B): o erro de dimensionamento em segmentações médicas ocorre quando uma estrutura é identificada com um tamanho incorreto. Para simular esse erro, é aplicada uma transformação geométrica baseada em uma matriz de escala afim, aplicada com coordenadas homogêneas para garantir a alteração proporcional ao redor do centro da estrutura.

Invasão de uma máscara em outra (Figura 1C): este erro ocorre quando uma região anatômica invade indevidamente outra, alterando as fronteiras naturais do VE, cri-

ando componentes segmentados isolados dentro da máscara, ou pela borda, distorcendo o contorno da estrutura. Para simular a invasão interna, o procedimento define qual máscara atuará como fonte e qual como alvo, a partir de um ponto inicial. Em seguida, o método expande progressivamente a invasão. A invasão de uma máscara em outra pela borda é realizada alterando a forma do contorno da segmentação, deslocando pontos da borda de acordo com uma modulação que modifica sua posição original, resultando em uma nova estrutura com a borda deformada.

Presença de múltiplos componentes (Figura 1D): esse erro ocorre quando uma única estrutura anatômica aparece como várias regiões desconexas na segmentação. Para simular essa falha, o método identifica áreas com intensidades semelhantes as dos *pixels* da câmara cardíaca para extrair seus contornos, e, dentre os contornos, seleciona um com base em sua proximidade a um ponto definido. Em seguida, redesenha esse contorno em uma nova máscara, podendo incluir outras modificações como, por exemplo, um novo componente que possui um buraco na câmara cardíaca.

Espiculação de borda (Figura 1E): segmentações adequadas do VE devem apresentar bordas suaves, pois a morfologia do ventrículo é geralmente arredondada. Um erro de borda espiculada ocorre quando a fronteira da segmentação apresenta uma aparência irregular, com artefatos espiculados. Para simular esse erro, primeiramente, extrai-se o contorno exterior do miocárdio e, para cada iteração, é gerado um ruído binário aleatório que é aplicado exclusivamente aos *pixels* dessa borda, removendo-os arbitrariamente. Esse processo causa a formação de "degraus", ou interrupções, na suavidade da linha de contorno.

Segmentação em local errado (Figura 1F): ocorre quando a segmentação é realizada em uma região anatômica fora dos limites esperados para o VE. Isso pode acontecer devido a semelhanças de intensidade com estruturas vizinhas, como o ventrículo direito e a cárdia. Para simular esse erro, o método aplica uma transformação afim de rotação e deslocamento à máscara segmentada, mantendo o centro da estrutura inalterado.

3.3. Experimento de validação

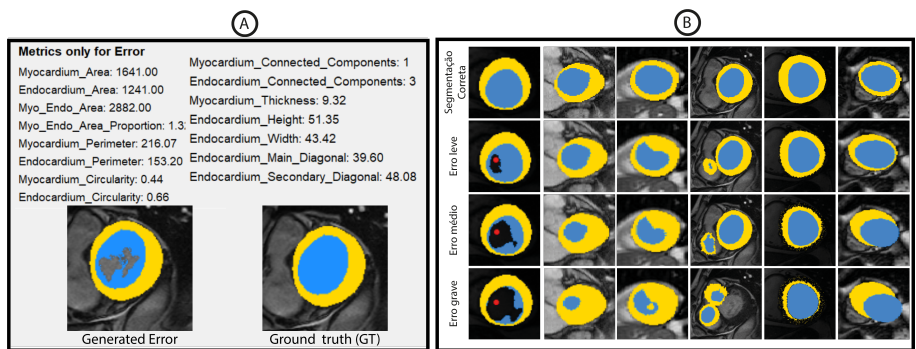
Para demonstrar a capacidade do método proposto em replicar erros reais observados nas segmentações automáticas analisadas, foi conduzido um experimento comparando as segmentações geradas artificialmente com aquelas produzidas por redes de AP. Erros reais pertencentes a cada um dos tipos de falhas previamente definidas (Figura 1) foram selecionados. Em seguida, o arcabouço foi utilizado para degradar a segmentação do especialista com o intuito de replicar a mesma falha. A similaridade das segmentações contendo erros sintéticos e erros reais foi quantificada por meio do coeficiente de *Dice* [Ribeiro and Nunes 2022], métrica utilizada para avaliar a sobreposição entre máscaras.

4. Resultados e discussões

A ferramenta desenvolvida possibilita a inserção de erros de modo automático e manual. A interface gráfica facilita a configuração dos parâmetros associados e a visualização dos erros produzidos por cada método de modificação (Figura 2A). A flexibilidade de combinar múltiplos métodos geradores de erros, cada qual com diferentes valores de parâmetros, confere ao arcabouço alta capacidade de gerar erros diversificados com diferentes intensidades, conforme demonstrado na Figura 2B. Adicionalmente, a ferramenta

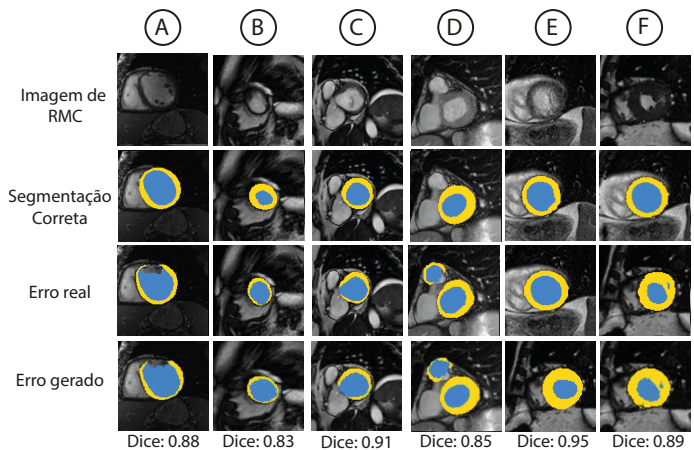
dispõe de uma funcionalidade de monitoramento e visualização de indicadores extraídos da segmentação, facilitando a análise do impacto de cada erro inserido e favorecendo o desenvolvimento de métodos futuros voltados à detecção e correção.

Figura 2. Ferramenta para geração de erros em segmentações. A) interface para visualização de indicadores. B) diferentes tipos de erro introduzidos na segmentação, com variação de parâmetros que gera diferentes intensidades de erro.



A Figura 3 apresenta um conjunto de exemplos nos quais se compara segmentações com erros reais e erros sinteticamente gerados pela ferramenta proposta. Cada coluna (A a F) ilustra um caso diferente. O valor de *Dice* exibido em cada exemplo reflete o grau de sobreposição entre as segmentações, sendo considerado alto (variando de 0,83 a 0,95). Esse resultado indica que o arcabouço é capaz de simular erros anatômicos observáveis em segmentações produzidas por redes de AP, sugerindo que a ferramenta pode ser utilizada para análise de falhas e desenvolvimento de estratégias de detecção e correção das mesmas. Além disso, a possibilidade de parametrização detalhada do arcabouço permite controlar e reproduzir diferentes tipos de erro, o que amplia sua aplicabilidade. Devido à ausência de trabalhos relacionados na literatura, não é possível realizar uma comparação desses resultados.

Figura 3. Reprodução de erros reais a partir dos métodos descritos. (A-F) mesma ordem de erros que a Figura 1. Os valores de *Dice* informados se referem à média das regiões do miocárdio e câmara cardíaca.



5. Conclusão

Nesse trabalho, um arcabouço voltado à categorização e à simulação de erros anatômicos comuns em segmentações automáticas produzidas por redes de AP foi proposto. A ferramenta desenvolvida fornece um ambiente estruturado para a criação de erros sintéticos calibráveis, análise de falhas e extração de indicadores. Os resultados obtidos sugerem que o arcabouço é capaz de replicar erros reais observados em segmentações automáticas. Como trabalhos futuros, pretende-se utilizar o arcabouço no desenvolvimento de novas estratégias voltados à detecção e à correção desses erros.

Agradecimentos

Este trabalho foi apoiado, em parte, pelo Instituto Nacional de Ciência e Tecnologia – Medicina Assistida por Computação Científica (INCT-MACC) [processo 2014/50889-7], pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) [processos 307710/2022-0 e 122803/2024-8] e pela Pró-Reitoria de Pesquisa e Inovação da USP.

Referências

- Bernard, O. et al. (2018). Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved? IEEE Transactions on Medical Imaging, 37(11):2514–2525.
- Guan, S., Samala, R. K., Arab, A., and Chen, W. (2023). Miss-tool: medical image segmentation synthesis tool to emulate segmentation errors. In Iftekharuddin, K. M. and Chen, W., editors, Medical Imaging 2023: Computer-Aided Diagnosis, page 41. SPIE.
- Landman, B. and Warfield, S. (2013). 2013 cardiac atlas project standard challenge - participant project.
- Lohr, D. et al. (2024). Precision imaging of cardiac function and scar size in acute and chronic porcine myocardial infarction using ultrahigh-field mri. Communications Medicine, 4(1).
- Radau, P. et al. (2009). Evaluation framework for algorithms segmenting short axis cardiac MRI. The MIDAS Journal.
- Ribeiro, M. A. O. and Nunes, F. L. S. (2021). Evaluating the pre-processing impact on the generalization of deep learning networks for left ventricle segmentation. In 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), page 3505–3512. IEEE.
- Ribeiro, M. A. O. and Nunes, F. L. S. (2022). Left ventricle segmentation in cardiac MR: A systematic mapping of the past decade. ACM Comput. Surv., 54(11s).
- Tajbakhsh, N., Lai, B., Ananth, S. P., and Ding, X. (2020). Errornet: Learning error representations from limited data to improve vascular segmentation. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), page 1364–1368. IEEE.
- Waite, S. et al. (2017). Tired in the reading room: The influence of fatigue in radiology. Journal of the American College of Radiology, 14(2):191–197.
- Wang, Y. et al. (2021). Deep learning based fully automatic segmentation of the left ventricular endocardium and epicardium from cardiac cine MRI. Quantitative Imaging in Medicine and Surgery, 11(4):1600–1612.
- Yuan, T. et al. (2018). Fully automatic segmentation of the left ventricle using multi-scale fusion learning. In 2018 24th International Conference on Pattern Recognition (ICPR), page 3838–3843. IEEE.