

# Segmentação de Miocárdio por meio de Redes Neurais Convolucionais

Filipe Augusto de Castro Oliveira - NUSP: 13796886  
PTC5892 - Processamento de Imagens Médicas

## I. INTRODUÇÃO

O cálculo de índices cardíacos por meio da análise de suas estruturas desempenha um papel importante para realização de diagnóstico de patologias, cálculo de riscos e tomada de decisão clínica [1]. Aspectos estruturais e funcionais cardíacos podem ser avaliados por meio da segmentação de estruturas cardíacas em imagens de seções do coração [2].

O padrão-ouro para captação de imagens cardíacas é a Ressonância Magnética Cardíaca (CMR), que é captada em sincronia com o Eletrocardiograma, permitindo captar imagens estáticas, mesmo com o órgão em movimento [3].

Entretanto, a segmentação de imagens cardíacas ainda demanda grandes esforços da área médica para identificar e traçar as estruturas de interesse [4]. A segmentação automática de estruturas cardíacas por meio de algoritmos de inteligência artificial e aprendizado profundo tem se mostrado como uma alternativa mais rápida e de menor custo para realizar o trabalho de segmentação. Esse tipo de ferramenta, contudo, está em processo de aprimoramento, para seus resultados sejam confiáveis e de validade clínica. Um tipo de arquitetura de aprendizado profundo que tem ganhado destaque é a U-Net.

A U-Net é uma arquitetura de rede neural convolucional (CNN) amplamente utilizada em tarefas de segmentação de imagens, como a segmentação de objetos em imagens médicas [5].

A arquitetura da rede apresenta uma estrutura em forma de U, composta por duas partes principais: a via de codificação (*encoder*) e a via de decodificação (*decoder*). A via de codificação é responsável por capturar características e reduzir a resolução da imagem, enquanto a via de decodificação reconstrói uma máscara de segmentação refinada, aumentando gradualmente a resolução. A arquitetura U-Net possui como característica a conexão direta entre as camadas de codificação e decodificação, através da concatenação de mapas de características da primeira à segunda. Essa estratégia permite a união de informações de diferentes etapas, preservando detalhes contextuais.

Além disso, redes do tipo U-Net têm se mostrado capazes de apresentar resultados com métricas mais elevadas do que modelos maiores e utilizando menos imagens para treinamento.

### A. Objetivo

Neste trabalho, é proposta uma rede de aprendizado profundo baseada em U-Net para segmentação de miocárdio em imagens de CMR de pacientes sem cardiopatias. A

segmentação será realizada em todos os cortes do coração disponibilizados para cada paciente.

## II. METODOLOGIA

### A. Base de dados

Os dados de segmentação de miocárdio utilizados neste trabalho foram obtidos da base de dados apresentada no *Automated Cardiac Diagnosis Challenge* - ACDC em 2017 [4]. A base de dados contém exames de CMR de 150 pacientes, estes divididos em cinco grupos de 30, onde cada grupo apresenta uma cardiopatia distinta, sendo um dos grupos de pacientes saudáveis.

Cada exame possui as seguintes anotações para segmentação das seguintes estruturas:

- Cavidade do ventrículo direito
- Cavidade do ventrículo esquerdo
- Miocárdio

Com o intuito de evitar a variabilidade das estruturas causadas por cardiopatias, foi escolhido o grupo de pacientes saudáveis para treinamento e validação do modelo de aprendizado. De 30 pacientes incluídos no grupo, 20 foram escolhidos para treinamento e validação da rede e 10 para teste.

O miocárdio foi selecionado como estrutura de interesse de segmentação, portanto a anotação de miocárdio de cada paciente foi utilizada como *ground truth*. Ao utilizar apenas uma anotação, o modelo de aprendizado deverá ser capaz de diferenciar apenas duas classes: plano de fundo ou a estrutura de interesse (miocárdio).

### B. U-Net

O modelo de U-Net utilizado neste trabalho possui estrutura similar à apresentada por [5], entretanto, é utilizada uma camada a menos. Dessa maneira, a camada de menor resolução espacial possui 512 mapas de características, diferente a U-Net apresentada por [5], que possui 1024 mapas. Uma representação visual da U-Net utilizada pode ser vista na Figura 1.

Cada camada convolucional possui um kernel 3x3 e ativação ReLU em sua saída. Após cada camada convolucional há uma etapa de normalização de lote. Para redução de dimensionalidade, é utilizado o MaxPool com kernel 2x2. A etapa de decodificação utiliza etapas de convolução 2x2.

### C. Métricas

Para avaliar o desempenho da rede utilizada, foram selecionadas três métricas:

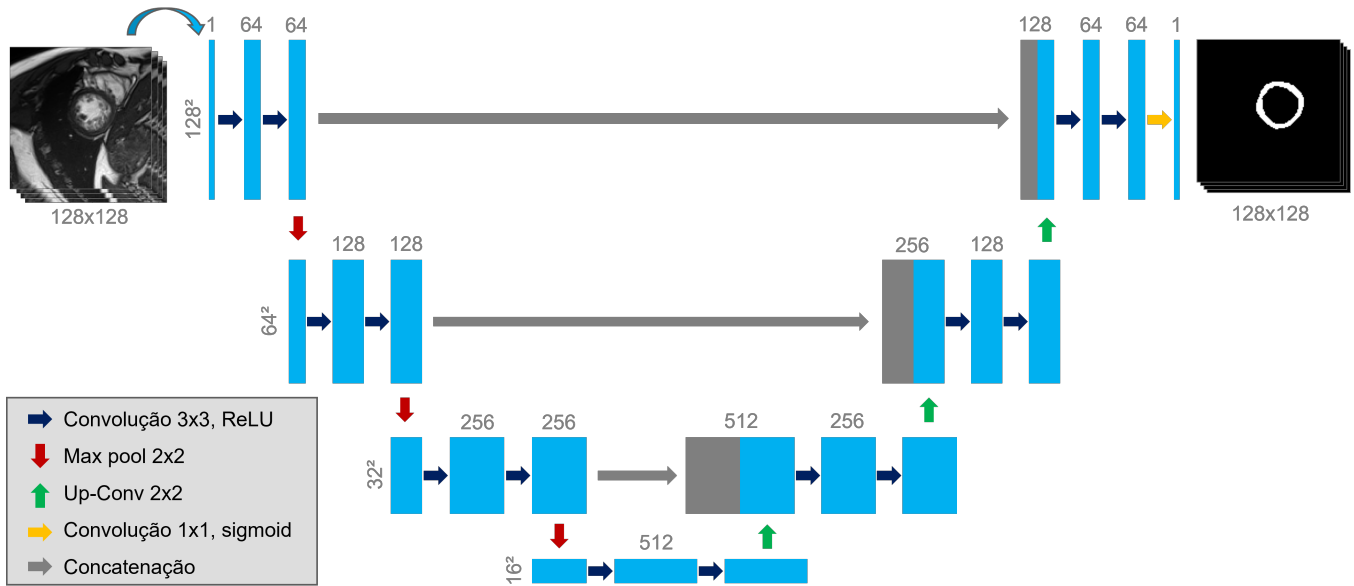


Fig. 1. Estrutura da U-Net utilizada.

- Dice: Conhecido como Índice de Sørensen-Dice, é uma métrica que mede a sobreposição entre duas regiões segmentadas em uma imagem. Varia de 0 (sem sobreposição) a 1 (sobreposição perfeita), representando um indicador da acurácia da segmentação em relação à área de interesse.
- N-RMSE: Representa a raiz do erro médio quadrático normalizado. Neste trabalho, a normalização foi feita por intensidade da imagem de referência.
- SSIM: Representa o Índice de Similaridade Estrutural. Essa métrica mede a semelhança entre duas imagens levando em conta informações de luminância, contraste e estrutura. Valores mais próximos de 1 indicam maior similaridade, enquanto valores mais baixos sugerem maior diferença entre as imagens.

### III. RESULTADOS

Foi realizado treinamento da U-Net utilizando dados de 16 pacientes, e a validação com dados de 4 pacientes. Foram utilizados lotes de 16 imagens durante a etapa de treinamento e função de perda *binary crossentropy*. O treino da rede ocorreu em 25 épocas.

Antes de calcular as métricas, é necessário definir um limiar para binarização da imagem de saída da U-Net. Para defini-lo, foi realizada uma análise da métrica Dice utilizando diferentes limiares. O gráfico exibindo essa variação pode ser visto na Figura 2.

Com base na avaliação de diferentes limiares, o maior valor Dice obtido foi de 0.809 utilizando o limiar de 0.25 no conjunto de validação. Esse valor implica que o valor de cada pixel da imagem de saída da U-Net será comparado a esse limiar. Se seu valor for maior, o pixel será considerado como a estrutura de interesse, e caso contrário, como plano de fundo.

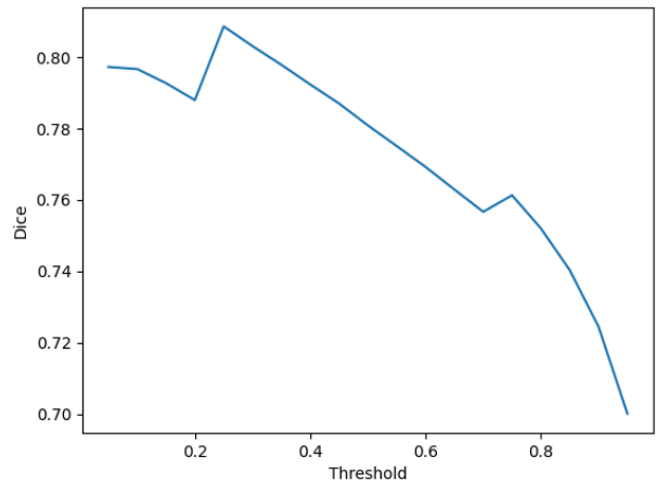


Fig. 2. Valor do Dice para o conjunto de validação com a variação do limiar de binarização.

Para o cálculo das métricas no conjunto de testes, foi utilizado o limiar de 0.25 obtido no conjunto de validação. A Tabela I mostra as métricas obtidas para o conjunto de validação e de teste.

TABLE I  
MÉTRICAS CALCULADAS PARA OS CONJUNTOS DE VALIDAÇÃO E TESTE

Conjunto	Pacientes	Dice	NRMSE	SSIM
Validação	4	0.809 ± 0.243	0.098 ± 1.172	0.952 ± 0.019
Teste	10	0.694 ± 0.292	0.102 ± 1.205	0.949 ± 0.021

A Figura 3 mostra um exemplo de segmentação de duas imagens do conjunto de validação.

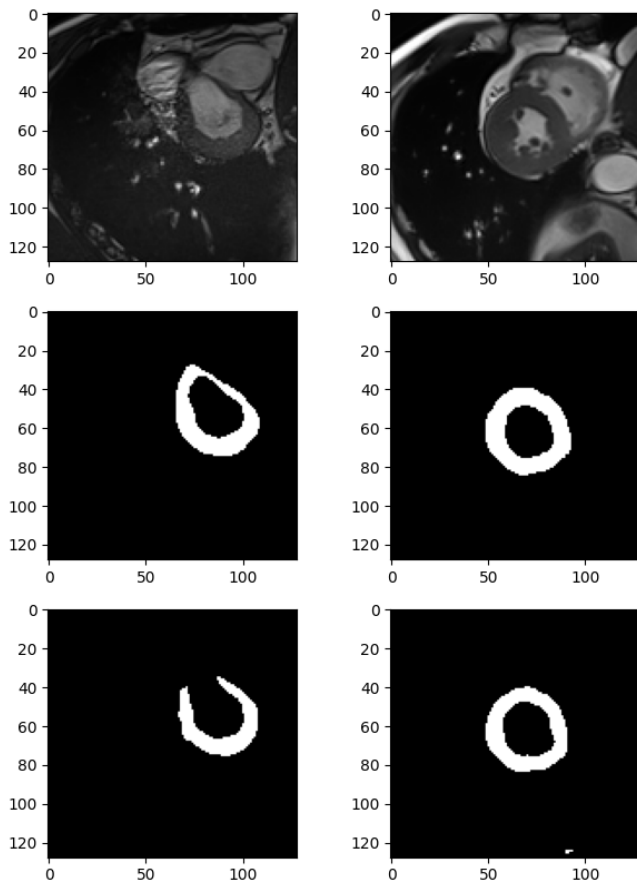


Fig. 3. Segmentação de duas imagens pela U-Net. De cima para baixo: imagens de entrada; imagens de segmentação real; imagens segmentadas pela U-Net.

#### IV. DISCUSSÃO

O modelo de aprendizagem proposto baseado em U-Net foi capaz de segmentar de maneira razoavelmente precisa o miocárdio em imagens de CMR. Esse aspecto pode ser visto por meio de uma análise das métricas apresentadas na Tabela I.

Tanto no conjunto de validação quanto no de teste, as métricas Dice e SSIM mantiveram valores próximos de 1, sendo que o SSIM foi praticamente o mesmo para ambos os conjuntos e o Dice teve uma queda de aproximadamente 0.1 no conjunto de teste. Esses valores indicam que a sobreposição das áreas segmentadas real e estimada pela U-Net foi quase completa em grande parte das imagens segmentadas.

O erro calculado pelo NRMSE teve sua média praticamente inalterada para ambos os conjuntos, entretanto seu desvio padrão manteve um valor que equivale a aproximadamente doze vezes o valor da média em ambos os casos, o que indica que houve falha de segmentação em diversas imagens, seja por partes do miocárdio que não foram segmentadas ou pequenas áreas no plano de fundo consideradas como estruturas do miocárdio. Ambos os casos podem ser vistos na 3.

Diversas estratégias poderiam ser implementadas

para obter melhores resultados utilizando a U-Net para segmentação de miocárdio. A utilização de *data augmentation*, por exemplo, aumentaria a quantidade e a variabilidade de imagens utilizadas no treinamento da rede sem a necessidade de mais pacientes. Essa abordagem poderia melhorar consideravelmente a capacidade de generalização da rede.

Além disso, utilizar também a U-Net completa, como apresentada por [5], que atinge até 1024 mapas de características, poderia aumentar a capacidade da rede de extrair informações de cada imagem e melhorar o processo de segmentação.

#### V. CONCLUSÃO

Este trabalho apresentou uma versão simplificada de uma rede tipo U-Net para segmentação de miocárdio de pacientes sem cardiopatias por meio de imagens de Ressonância Magnética Cardíaca. A rede proposta foi capaz de realizar a segmentação de miocárdio com grande nível de sobreposição em comparação com a segmentação realizadas por especialistas, como vistos pelo Dice e SSIM próximos de 0.7 e 0.95, respectivamente, no conjunto de teste.

Como melhorias futuras, podem ser implementadas estratégias para melhorar o desempenho da rede na segmentação de miocárdio utilizando estratégias como *data augmentation* e implementação da U-Net original.

#### VI. DISPONIBILIZAÇÃO DE CÓDIGO FONTE

O código fonte do algoritmo implementado neste trabalho pode ser visualizado e baixado por meio da plataforma GitHub no endereço: [https://github.com/filipeaocastro/ptc5892\\_img\\_segmentation](https://github.com/filipeaocastro/ptc5892_img_segmentation)

#### REFERÊNCIAS

- [1] H. D. White, R. M. Norris, M. A. Brown, P. W. Brandt, R. M. Whitlock, and C. J. Wild, "Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction," *Circulation*, vol. 76, no. 1, pp. 44–51, 1987.
- [2] N. Kawel-Boehm, S. J. Hetzel, B. Ambale-Venkatesh, G. Captur, C. J. Francois, M. Jerosch-Herold, M. Salerno, S. D. Teague, E. Valsangiacomo-Buechel, R. J. Van Der Geest, and D. A. Bluemke, "Reference ranges ("normal values") for cardiovascular magnetic resonance (CMR) in adults and children: 2020 update," *Journal of Cardiovascular Magnetic Resonance*, vol. 22, no. 1, p. 87, 2020.
- [3] Wen-Yih Isaac Tseng, Mao-Yuan Marine Su, and Yao-Hui Elton Tseng, "Introduction to cardiovascular magnetic resonance: Technical principles and clinical applications," *Acta Cardiologica Sinica*, vol. 32, no. 2, 2016.
- [4] O. Bernard, A. Lalande, C. Zotti, F. Cervenansky, X. Yang, P.-A. Heng, I. Cetin, K. Lekadir, O. Camara, M. A. Gonzalez Ballester, G. Sanroma, S. Napel, S. Petersen, G. Tziritis, E. Grinias, M. Khened, V. A. Kollerathu, G. Krishnamurthi, M.-M. Rohe, X. Pennec, M. Sermesant, F. Isensee, P. Jager, K. H. Maier-Hein, P. M. Full, I. Wolf, S. Engelhardt, C. F. Baumgartner, L. M. Koch, J. M. Wolterink, I. Igum, Y. Jang, Y. Hong, J. Patravali, S. Jain, O. Humbert, and P.-M. Jodoin, "Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved?," *IEEE Transactions on Medical Imaging*, vol. 37, no. 11, pp. 2514–2525, 2018.
- [5] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.