

Estudo Comparativo da Função de Perda Combo Loss para Segmentação de Endometriose Profunda em Ressonância Magnética Utilizando DeepLabv3

Marcos V. Farias¹, Wesley K. R. Figueredo¹, Aristófanés C. Silva¹
Anselmo C. de Paiva¹, Alice C. C. B. Salomão², Marco A. P. de Oliveira³

¹Núcleo de Computação Aplicada – Universidade Federal do Maranhão (UFMA)
Av. dos Portugueses, 1966 - Vila Bacanga, São Luís - MA, 65080-805

²Clínica Fonte de Imagem - Rio de Janeiro, RJ - Brasil

³Universidade do Estado do Rio de Janeiro (UERJ) - Rio de Janeiro, RJ - Brasil

{marvinfar852, wesley.kelson, ari, paiva}@nca.ufma.br,

{brandaosalomao, endometriose}@gmail.com

Abstract. *Endometriosis is an inflammatory disease that affects 10% of the female population globally, causing symptoms ranging from bleeding and pain to infertility. Diagnosing endometriosis is a time-consuming and costly process for clinicians. Diagnosis by imaging tests is becoming the standard, with magnetic resonance imaging (MRI) standing out in this field. Using image processing techniques combined with machine learning, it is possible to build algorithms to aid in the location of lesions, functioning as tools to assist the physician. This work seeks to conduct a study on different strategies for segmenting these lesions in MRI images.*

Resumo. *A endometriose é uma doença inflamatória que afeta 10% da população do sexo feminino globalmente, podendo causar desde sangramentos e dor até infertilidade. O diagnóstico da endometriose é um processo demorado e custoso para o clínico. O diagnóstico por exames de imagem vem se tornando o padrão, âmbito onde a ressonância magnética (RM) se destaca. Utilizando técnicas de processamento de imagens aliadas ao aprendizado de máquina é possível construir algoritmos para auxílio na localização das lesões, funcionando como ferramentas de auxílio ao médico. Este trabalho busca realizar um estudo em diferentes estratégias para segmentar essas lesões em imagens de RM.*

1. Introdução

Caracterizada pela presença de tecido endometrial fora do útero, a endometriose trata-se de uma condição inflamatória comum, muitas vezes crônica e prevalente em torno de 5 a 10% da população feminina [Leyland et al. 2010, Allaire et al. 2023]. A patologia pode se apresentar de maneira assintomática ou sintomática, sendo esta última a mais recorrente [Nácul and Spritzer 2010]. Os principais sintomas são dismenorrea, dor pélvica crônica, infertilidade, dispareunia de profundidade, sintomas intestinais e urinários cíclicos, como dor ou sangramento ao evacuar/urinar durante o período menstrual [Bellelis et al. 2010]. Embora não exista uma teoria comprovada acerca da etiopatonegia da endometriose, a

explicação mais aceita postula que as células endometriais alcançam a cavidade peritoneal através da menstruação retrógrada (um processo fisiológico que ocorre em 90% das mulheres) [Allaire et al. 2023]. Assim, a patogênese da doença decorre do refluxo menstrual via tubas uterinas, onde, havendo a incapacidade de eliminação dessas células pelo sistema imune, ocorra sua fixação e crescimento progressivo.

O diagnóstico de endometriose enfrenta múltiplos fatores adversos, impondo atraso entre os primeiros sintomas e um diagnóstico definitivo [De Corte et al. 2025]. O fenômeno de atraso se demonstra extremamente comum, com estudos consistentemente reportando intervalos prolongados de 7-10 anos entre o início dos sintomas e a confirmação de um diagnóstico [De Corte et al. 2025]. Não obstante, a endometriose é facilmente confundida com outras enfermidades por não apresentar sintomas característicos, prejudicando ainda mais a detecção da patologia. Nesse sentido, a confirmação visual por videolaparoscopia em conjunto a confirmação histopatológica há muito tempo são considerados o padrão ouro referente ao diagnóstico clínico de endometriose. No entanto, diretivas médicas recentes advogam por diagnósticos não-cirúrgicos baseados nos sintomas apresentados e exames por imagem [Allaire et al. 2023]. Ultrassom e Ressonância Magnética (RM) são recomendados como métodos primários não-invasivos para o diagnóstico de endometriose, sendo que a RM pode alcançar acima de 90% em sensibilidade e especificidade na maioria dos casos [Liang et al. 2025].

Métodos computacionais para análise de imagem se tornam cada vez mais presentes no fornecimento de subsídios para a tomada de decisão em ambiente clínico. Tais abordagens se destacam por serem capazes de detectar detalhes finos nas imagens, como fibrose, dano inflamatório e depósitos de ferro associados com a endometriose que são difíceis para o olho humano distinguir [Moassefi et al. 2025]. Estudos recentes demonstraram a efetividade de métodos baseados em aprendizado profundo para segmentação de órgãos pélvicos em imagens de RM [Liang et al. 2025]. O trabalho de [Liang et al. 2025], além de introduzir uma nova base de dados de RM, propõe um protocolo automático para segmentação do ovário, um órgão central para o diagnóstico de endometriose. Seu método se baseia no uso de uma arquitetura ResNet [He et al. 2015] para encontrar as fatias dentro do volume que possuem o ovário, seguindo então da utilização de uma Attention U-net [Oktay et al. 2018] para segmentação.

Nesse âmbito, a endometriose é uma patologia complexa que possui considerável impacto na vida daqueles afetados por ela e que não possui cura [Allaire et al. 2023]. Paralelamente, métodos de auxílio ao diagnóstico são cada vez mais utilizados para práticas clínicas, aumentando a capacidade dos médicos de tomar decisões assertivas e oferecer tratamentos mais eficientemente. Assim, este trabalho busca apresentar um estudo acerca da utilização de técnicas de aprendizado profundo para segmentação de lesões de endometriose em imagens de RM. O estudo busca comparar diferentes funções de *loss* e *backbones* para a arquitetura *DeepLab V3*, visando o aprendizado de características relevantes para lesões de diferentes escalas e posições.

Este trabalho está estruturado do seguinte modo: o método proposto é apresentado na Seção 2; os experimentos realizados e resultados obtidos estão presentes na Seção 3; por fim, a Seção 4 contém conclusões e trabalhos futuros.

2. Método

Esta seção elucida a metodologia proposta, assim como o fluxo realizado nos experimentos conduzidos. A Figura 1 apresenta o método de maneira gráfica, indicando as diferentes etapas percorridas pela abordagem.

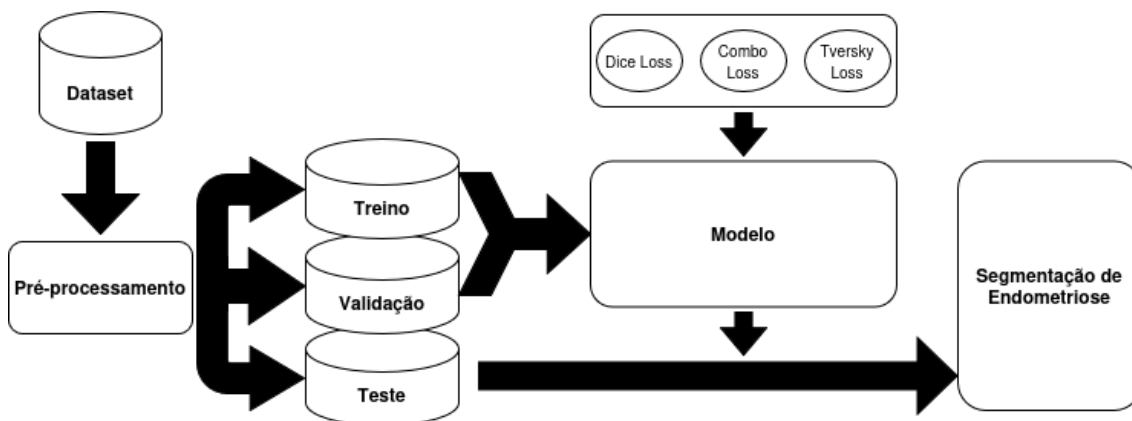


Figura 1. Fluxograma dos experimentos realizados.

2.1. Dataset

O conjunto de dados utilizado neste estudo foi adquirido via RM, gerando volumes do abdômen inferior dos pacientes com resolução de 512×512 pixels no plano sagital. Como critério de seleção para a tarefa de segmentação, foram incluídas apenas fatias que apresentassem evidência de endometriose, descartando-se imagens sem patologia. A base final compreende exames de 105 pacientes, totalizando 451 cortes com lesões superiores a 2 cm de diâmetro, variando de 1 a 9 imagens por paciente. Exemplos de fatias originais e suas respectivas Regiões de Interesse (ROI) extraídas estão ilustrados na Figura 2.

2.2. Pré-processamento

A delimitação de uma ROI que envolva as lesões de maneira precisa é fundamental, pois otimiza o aprendizado do modelo ao descartar informações irrelevantes inerentes aos exames de RM e reduz o custo computacional por meio da compressão dos dados. Neste estudo, as lesões situam-se no reto-sigmóide (porção final do intestino grosso). Para extrair essa região, adotou-se a metodologia de [Figueredo et al. 2023], iniciando com a técnica de Template Matching para isolar uma área comum que englobe a anatomia do paciente. A imagem utilizada como template é um recorte de um exame selecionado por apresentar uma melhor visualização do reto-sigmóide. Após a binarização do resultado do *Template Matching*, identifica-se a parede abdominal via análise pixel a pixel: partindo do primeiro pixel não branco no canto superior esquerdo (exterior do abdômen), percorre-se o eixo x através da zona muscular (pixels pretos) até atingir a região visceral (pixels brancos). O corte final, detalhado na Figura 3, utiliza a altura definida pelo template matching e a coordenada x inicial dos órgãos, resultando em uma imagem de 256×256 pixels.

2.3. Arquitetura DeepLab v3

Arquiteturas convolucionais representam a escolha intuitiva para a realização de diversas tarefas referentes a imagens. A segmentação, que consiste na delimitação precisa de determinado objeto em uma imagem através de uma máscara, representa um desafio para

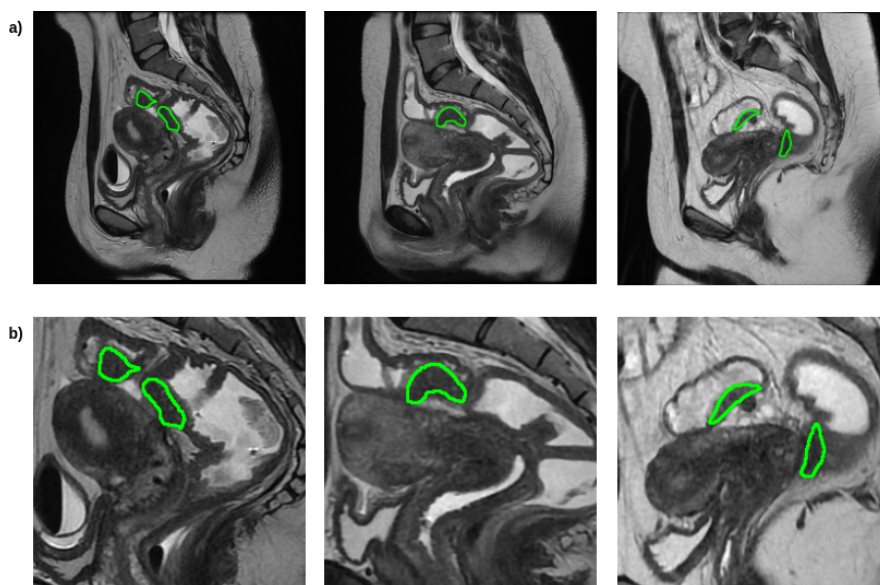


Figura 2. Exemplos de imagens presentes na base de imagens: a) apresenta as imagens antes da extração da região de interesse correspondente ao reto-sigmóide e b) apresenta os exames após essa etapa.

imagens de diversos domínios mesmo nos dias de hoje. Diversos modelos foram propostos para segmentar imagens, dentre eles, a arquitetura U-net [Ronneberger et al. 2015], que destaca-se por apresentar uma lógica de codificação-decodificação. Seu principal foco é realizar a extração de características das imagens de entrada através diversos filtros convolucionais, reduzindo a dimensionalidade da informação para, logo em seguida, reverter esse processo, superamostrando o mapa de características das camadas mais profundas do codificador para a geração de máscaras. No entanto, para a captura de detalhes em objetos de múltipla escala, outras arquiteturas podem se mostrar mais apropriadas. A arquitetura DeepLab v3 [Chen et al. 2017], propõe-se a explorar convoluções dilatadas como um módulo de contexto para aumento do campo receptivo. A arquitetura utiliza um módulo de extração de características dilatado em forma de pirâmide (ASPP), o qual é mais efetivo em compreender características de pequenos objetos [Si et al. 2020]. O módulo ASPP permite que a rede processe a imagem em múltiplas escalas de maneira simultânea, potencializando a capacidade do modelo em identificar objetos de diferentes tamanhos. A Figura 4 apresenta o panorama geral do funcionamento do algoritmo, o qual, similar a Unet, é composto de um *encoder* e um *decoder*.

De maneira a auxiliar a DeepLab v3 na tarefa de extração de características, foram empregados *backbones* convolucionais de ampla adoção na literatura. Estes modelos substituem a camada de codificação (*encoder*) da arquitetura original, servindo como método para o processamento e geração de características robustas para a decodificação posterior. Os *backbones* utilizados neste trabalho foram a ResNet [He et al. 2015] (versão 101 e 50) e a MobileNet v3 [Howard et al. 2019], arquiteturas que aplicam diferentes estratégias durante seu treinamento como modo de extrair características finais de maior qualidade. A ResNet caracteriza-se por possuir blocos residuais em sua estrutura, os quais previnem que ocorra degradação durante o treinamento através da passagem de características de blocos anteriores para camadas mais profundas da rede. Essa estratégia

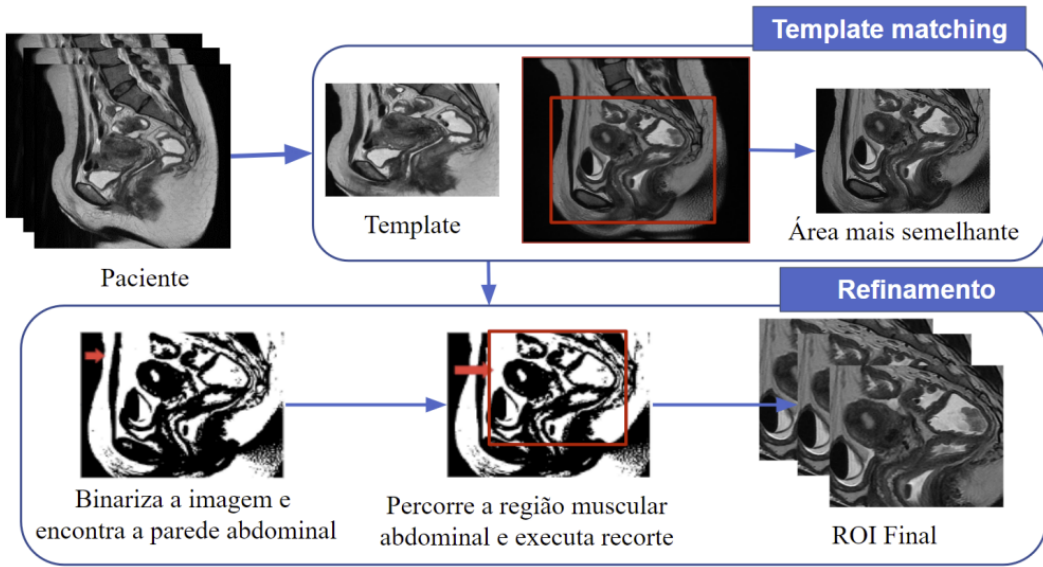


Figura 3. Diagrama extraído de [Figueredo et al. 2023] ilustrando a extração da ROI.

permite que a rede continue aprendendo características de mais alto nível sem perder referência dos padrões visuais da imagem. Em complemento, a MobileNet apresenta uma estrutura que busca aliar eficiência e performance, fazendo uso de *bottlenecks* invertidos e convoluções ponto a ponto para reduzir o número de parâmetros da rede mantendo um desempenho satisfatório.

2.4. Combo Loss

A ComboLoss é uma função que busca aproveitar as propriedades da função Dice, que lida com o desbalanceamento das classes de entrada, ou seja, segmentar um objeto pequeno em uma grande área de fundo, enquanto ao mesmo tempo controla o balanceamento entre falsos positivos e falsos negativos e força uma curva suave de treinamento utilizando a função *Cross Entropy* [Asgari Taghanaki et al. 2018]. A Equação 1 apresenta a função matemática resultante da combinação das funções de perda Dice e entropia cruzada, ponderadas pelo parâmetro α . Adicionalmente, a equação permite ainda controlar a intensidade da penalização para a geração de falsos positivos ou falsos negativos através do parâmetro β na parte esquerda, a qual representa a função de entropia cruzada.

$$L = \alpha \left(-\frac{1}{N} \sum_{i=1}^N \beta (t_i - \ln p_i) + (1 - \beta) [(1 - t_i) \ln (1 - p_i)] \right) - (1 - \alpha) \sum_{i=1}^K \left(\frac{2 \sum_{i=1}^N p_i t_i + S}{\sum_{i=1}^N p_i + \sum_{i=1}^N i_i = S} \right) \quad (1)$$

2.5. Métricas de Avaliação

A validação da segmentação produzida pelos modelos fundamenta-se na comparação direta com o padrão-ouro (ground truth) estabelecido pelo especialista. Essa análise quantitativa baseia-se na contagem de pixels classificados como: Verdadeiros Positivos (VP),

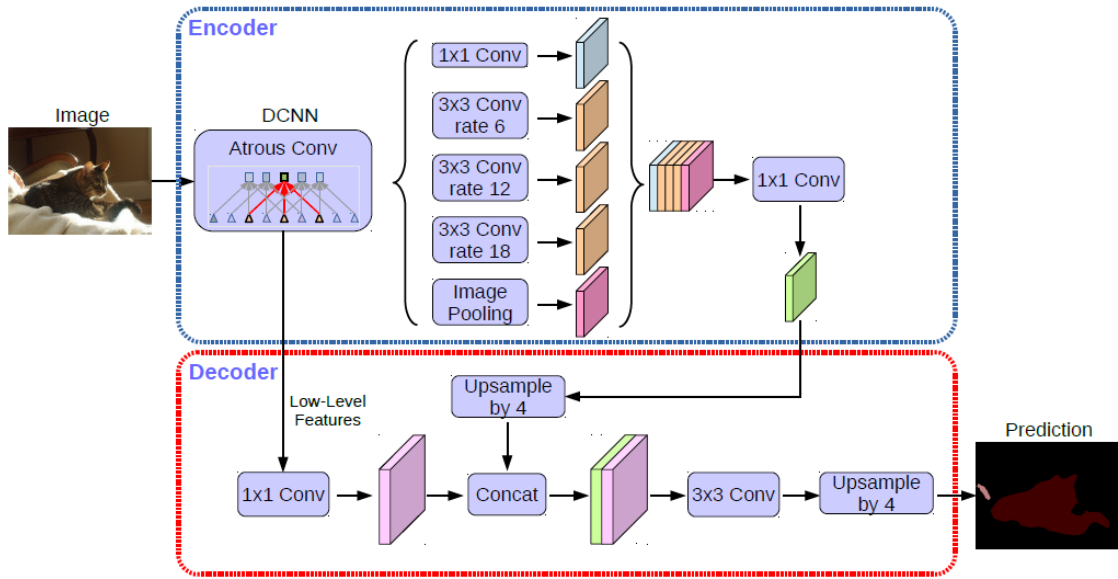


Figura 4. Diagrama explicando o funcionamento da DeepLab v3. [Tsang]

que representam a identificação correta de focos de endometriose; Falsos Positivos (FP), referentes a pixels erroneamente atribuídos à patologia; e Falsos Negativos (FN), que indicam pixels de endometriose não detectados pelo modelo.

A eficácia do método proposto é avaliada por meio das seguintes métricas estatísticas: Dice Similarity Coefficient (Dice) [Eq. 2], Precisão [Eq. 3], Sensibilidade (Recall) [Eq. 4] e Intersecção sobre a União (IoU) [Eq. 5].

$$Dice = \frac{2 * VP}{VP + FP + FN} \quad (2)$$

$$Precisão = \frac{VP}{VP + FP} \quad (3)$$

$$Sensibilidade = \frac{VP}{VP + FN} \quad (4)$$

$$IoU = \frac{VP}{VP + FP + FN} \quad (5)$$

3. Experimentos

3.1. Configurações experimentais

O conjunto de dados foi particionado nos subconjuntos de treinamento, validação e teste, contendo 274, 95 e 82 imagens, respectivamente, todas redimensionadas para a resolução de 128×128 pixels. A avaliação seguiu o protocolo de hold-out, sendo os resultados reportados a partir da média e do desvio padrão obtidos após três execuções independentes de cada experimento.

Para garantir a reprodutibilidade e a consistência entre os diferentes backbones avaliados, os hiperparâmetros foram padronizados com um tamanho de lote (batch size) de 16, otimizador AdamW com taxa de aprendizado inicial de 10^{-4} e um ciclo de treinamento de 200 épocas. A investigação abrangeu três funções de perda distintas: Dice Loss, Tversky Loss e Combo Loss. A Dice Loss foi selecionada devido à sua ampla adoção em tarefas de segmentação, operando sobre o nível de sobreposição entre o mapa de segmentação predito e a máscara de referência para calcular o erro a ser retropropagado. Contudo, em cenários de acentuado desbalanceamento entre as classes (lesão versus fundo), a Dice Loss pode ser insuficiente. Para mitigar essa limitação, utilizou-se a Tversky Loss, uma generalização do coeficiente de Dice que introduz constantes para o controle de pesos sobre os erros. Assim como a Combo Loss, a função Tversky permite penalizar de forma assimétrica os Falsos Positivos (FP) e Falsos Negativos (FN). No diagnóstico por imagem, a minimização de FN é crítica para evitar a não identificação de lesões presentes no exame. Consequentemente, os parâmetros foram ajustados para priorizar a sensibilidade do modelo: na Combo Loss, adotou-se $\alpha = 0.5$ e $\beta = 0.7$; para a Tversky Loss, utilizou-se $\alpha = 0.7$, conferindo maior peso à penalização de omissões de detecção em ambas abordagens. Estes valores foram obtidos empiricamente através de testes com diferentes valores variando de 0.5 a 0.9.

Os experimentos apresentados nesta seção foram conduzidos em uma GPU NVIDIA GeForce RTX 3060 utilizando o *framework* PyTorch. É importante ressaltar que as arquiteturas empregadas como *backbones* utilizam pesos pré-treinados na base **ImageNet1k**.

3.2. Resultados

As métricas obtidas nos experimentos realizados podem ser visualizadas na Tabela 1.

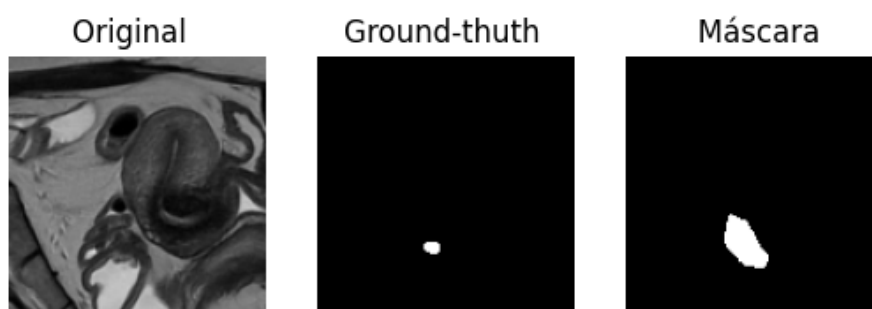


Figura 5. Exemplo de predição realizada pelo *backbone* ResNet-101 com a função *Combo Loss*

Analisando a seção tabela relativa à *Combo Loss*, podemos observar que a alta ponderação dos FN impactou as métricas, em especial a sensibilidade, que é extremamente sensível a variações na quantidade de FN. O modelo codificado pela ResNet-101 obteve a maior média nesta métrica, alcançando 73,54%, resultado bem superior se comparado às demais funções de perda. No entanto, é possível observar que o modelo demonstra comportamentos enviesados nesse sentido, o que é indicado pela baixa precisão. Este fator é um claro indicativo de modelos que superestimam a região de lesão presente na imagem, predizendo máscaras que englobam corretamente a região de lesão mas que

Tabela 1. Resultados obtidos

Backbone	Dice	Precisão	Sensibilidade	IoU
Combo Loss				
ResNet-101	37,56% ±15	31,51% ±14,7	73,54% ± 5,7	27,04% ± 11,7
ResNet-50	45,75% ± 2,5	41,99% ± 5,2	66,66% ± 7,1	33,38% ± 2,6
MobileNet v3	33,63% ±0,7	29,79% ± 2,2	55,86% ± 3,5	22,87% ± 0,8
Tversky Loss				
ResNet-101	46,09% ± 1,9	52,57% ± 1,9	47,61% ± 3,1	33,97% ± 1,7
ResNet-50	48,57% ± 0,6	50,45% ± 0,5	54,71% ± 1,4	37,21% ± 0,8
MobileNet v3	35,52% ± 1,3	40,43% ± 0,5	39,77% ± 2,2	25,70% ± 1,1
Dice Loss				
ResNet-101	49,88% ± 0,5	51,01% ± 2,1	56,93% ± 1,7	38,16% ± 0,6
ResNet-50	52,41% ± 0,4	52,41% ± 1,5	57,12% ± 0,9	38,39% ± 0,2
MobileNet v3	36,61% ± 0,3	40,82% ± 1,0	41,51% ± 0,8	26,34% ± 0,4

também compreendem outras áreas além dos limites da classe correta. A imagem 5 representa uma imagem que exemplifica esse caso. Além disso, os resultados dessa abordagem apresentam alto desvio padrão, indicando que a utilização dessa configuração experimental pode levar a uma maior instabilidade durante o treinamento.

A Figura 5 apresenta um cenário onde nenhuma função de perda foi efetiva em guiar o modelo para a realização de uma predição correta, demonstrando assim que algumas lesões pequenas ainda demonstram-se extremamente difíceis. O mesmo comportamento se apresenta em outros *backbones*, o que demonstra que as informações consideradas pelo modelo não são significativas o suficiente para que este consiga segmentar lesões de pequena dimensão. Nesse ponto, a *Tversky Loss*, apesar de não identificar a lesão corretamente, não realizou a predição de outra região incorretamente.

Paralelamente, os modelos conseguiram atingir boas métricas para diversas imagens, tal qual exemplifica a Figura 7. Aqui temos casos claros onde o modelo obteve uma boa compreensão da localização e morfologia das lesões, assim como de sua extensão. Esses cenários demonstram o potencial desse tipo de abordagem para a identificação dessas lesões e auxílio clínico, onde, os modelos apresentaram métricas acima de 80% para Dice. Mesmo modelos mais simples como a MobileNet conseguiram gerar boas segmentações apesar de suas métricas globais inferiores e quantidade menor de parâmetros. As segmentações foram realizadas com as funções *Combo* e *Tversky Loss* respectivamente, demonstrando situações onde a ponderação realizada pode ser benéfica e ajudar as arquiteturas na correta identificação das regiões de lesão.

Durante os experimentos realizados, evidenciou-se uma acurácia reduzida em

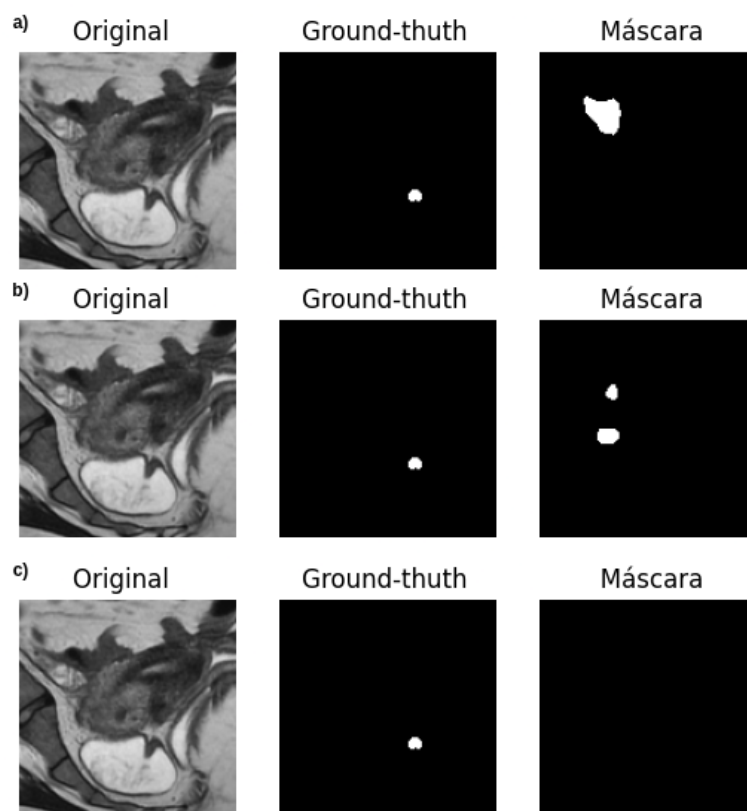


Figura 6. Exemplo de erro na base com a rede ResNet-101 onde: a) *Combo Loss*; b) *Dice Loss*; c) *Tversky Loss*.

lesões periféricas, fenômeno possivelmente atribuído a fatores como a delimitação inadequada da ROI, baixo contraste nas bordas do campo de visão ou ao desbalanceamento de amostras com essa característica no conjunto de treinamento. Essas ocorrências devem ser mais profundamente investigadas e tratadas de maneira condizente, seja por etapas extras de pré-processamento ou por mecanismos internos da rede que possam mitigar estes percauços. Apesar do panorama geral, o modelo ResNet-50 apresentou resultados consistentes em conjunto com a *Dice Loss*, alcançando as melhores métricas tanto em Dice quanto em IoU, com desvios padrão menores que 0,4. Esse fato demonstra que a *Dice Loss* possui propriedades interessantes para guiar o treinamento do modelo, sendo melhor que a *Combo Loss* nesses aspectos. Por fim, a MobileNet V3, apesar de apresentar as menores métricas, demonstra consistência em seus resultados e atinge um bom nível de desempenho considerando seu número de parâmetros é 4 vezes menor que a ResNet-50 e 10 vezes menor que a ResNet-101.

Os modelos de segmentação apresentados desempenham de maneira satisfatória para boa parte do conjunto de dados, possuindo a capacidade de ao menos identificar a localidade das lesões. Para o ambiente clínico, essa capacidade já demonstra-se valiosa para o radiologista, visto o nível de complexidade apresentado pelas imagens. Analisando isoladamente a segunda execução do modelo ResNet101 + ComboLoss, o qual obteve um Dice de 50,52% e Recall de 75,21%, é possível constatar um bom desempenho na identificação de pelo menos uma fatia corretamente por paciente. Considerando um limiar de Dice de 50%, verifica-se que 77,7% dos pacientes do teste (14 de 18 pacientes)

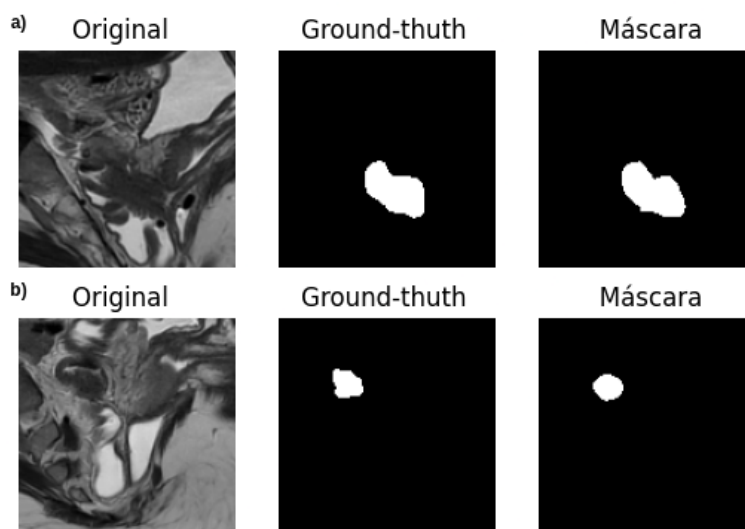


Figura 7. Exemplos de imagens segmentadas corretamente. a) Modelo ResNet-101 utilizando a *Combo Loss*; b) Modelo MobileNetV3 utilizando a *Tversky Loss*.

possuem ao menos 1 fatia identificada corretamente. Assim, devido a natureza sequencial das fatias de RM, a correta marcação de uma fatia já indicaria para o especialista a localização da lesão na fatia analisada e em suas vizinhas.

4. Conclusão

O presente trabalho buscou apresentar um estudo comparativo sobre o impacto de diferentes funções de perda e arquiteturas codificadoras sobre o modelo DeepLab v3. Por fim, faz-se necessário ressaltar as limitações provenientes da abordagem apresenta neste trabalho. Foi possível verificar que a função *Combo Loss* adiciona certo nível de instabilidade no treinamento, o que impacta principalmente o alto desvio padrão decorrente da diferença de resultado entre experimentos. Sua característica de ponderamento de falsos negativos é positiva para a métrica sensibilidade, no entanto, acaba enviesando o modelo a superestimar a área de lesão, contribuindo para uma menor precisão e degradando assim as outras métricas avaliadas de uma maneira geral. As demais arquiteturas apresentaram resultados mais estáveis, demonstrando que existe um limiar intermediário onde as redes, mesmo com uma quantidade de parâmetros menor, consegue extrair resultados similares ao de redes mais profundas.

A adoção de uma estratégia bidimensional para a segmentação, onde a rede não possui informações acerca das fatias vizinhas de um mesmo paciente, impõe dificuldades adicionais à tarefa. Esse argumento é fortalecido considerando o alto grau de desbalanceamento em algumas fatias (lesões muito pequenas), o que poderia ser mitigado observando a morfologia dessa lesão ao longo de múltiplas fatias em uma abordagem tridimensional. É necessário ainda, realizar um estudo acerca dos pré-processamentos que podem ser aplicados para refinar a região de interesse e o contraste das lesões em relação a demais tecidos. Apesar das dificuldades encontradas, a proposta apresenta resultados relevantes para a área do conhecimento, visto que essa é extremamente escassa acerca de trabalhos que contemplem o escopo abordado.

Os autores do presente trabalho agradecem o apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, assim como da Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão - FAPEMA.

Referências

- Allaire, C., Bedaiwy, M. A., and Yong, P. J. (2023). Diagnosis and management of endometriosis. *Cmaj*, 195(10):E363–E371.
- Asgari Taghanaki, S., Zheng, Y., Zhou, S. K., Georgescu, B., Sharma, P., Xu, D., Comaniciu, D., and Hamarneh, G. (2018). Combo loss: Handling input and output imbalance in multi-organ segmentation. *arXiv e-prints*, pages arXiv–1805.
- Bellelis, P., Dias Jr, J. A., Podgaec, S., Gonzales, M., Baracat, E. C., and Abrão, M. S. (2010). Aspectos epidemiológicos e clínicos da endometriose pélvica: uma série de casos. *Revista da associação médica brasileira*, 56:467–471.
- Chen, L.-C., Papandreou, G., Schroff, F., and Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation.
- De Corte, P., Klinghardt, M., von Stockum, S., and Heinemann, K. (2025). Time to diagnose endometriosis: Current status, challenges and regional characteristics—a systematic literature review. *BJOG: An International Journal of Obstetrics & Gynaecology*, 132(2):118–130.
- Figueredo, W. K., da Silva, I. F., Diniz, J. O., Silva, A. C., de Paiva, A. C., Salomão, A. C. B., and de Oliveira, M. A. (2023). Abordagem computacional baseada em deep learning para o diagnóstico de endometriose profunda através de imagens de ressonância magnética. In *Anais do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 138–149. SBC.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition.
- Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., and Adam, H. (2019). Searching for mobilenetv3.
- Leyland, N., Casper, R., Laberge, P., Singh, S. S., Allen, L., Arendas, K., Leyland, N., Allaire, C., Awadalla, A., Best, C., et al. (2010). Endometriosis: diagnosis and management. *Journal of Endometriosis*, 2(3):107–134.
- Liang, X., Alpuing Radilla, L. A., Khalaj, K., Dawoodally, H., Mokashi, C., Guan, X., Roberts, K. E., Sheth, S. A., Tammisetti, V. S., and Giancardo, L. (2025). A multi-modal pelvic mri dataset for deep learning-based pelvic organ segmentation in endometriosis. *Scientific Data*, 12(1):1292.
- Moassefi, M., Faghani, S., Colak, C., Sheedy, S. P., Andrieu, P. L. C., Wang, S. S., McPhedran, R. L., Flicek, K. T., Suman, G., Takahashi, H., et al. (2025). Advancing endometriosis detection in daily practice: a deep learning-enhanced multi-sequence mri analytical model. *Abdominal Radiology*, pages 1–12.
- Nácul, A. P. and Spritzer, P. M. (2010). Aspectos atuais do diagnóstico e tratamento da endometriose. *Revista Brasileira de ginecologia e obstetrícia*, 32:298–307.

- Oktaç, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., Glocker, B., and Rueckert, D. (2018). Attention u-net: Learning where to look for the pancreas.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer.
- Si, H., Shi, Z., Hu, X., Wang, Y., and Yang, C. (2020). Image semantic segmentation based on improved deeplab v3 model. *International Journal of Modelling, Identification and Control*, 36(2):116–125.
- Tsang, S.-H. Review: DeepLabv3-Atrous Separable Convolution (Semantic Segmentation) — sh-tsang.medium.com. [Accessed 13-01-2026].