

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

**VIABILIDADE DA APLICAÇÃO DE REDES
NEURAIS PARA A CORREÇÃO DE ERROS DE
SEGMENTAÇÃO EM VASOS SANGUÍNEOS**

JAYME SAKAE DOS REIS FURUYAMA

ORIENTADOR: PROF. DR. CESAR HENRIQUE COMIN

São Carlos – SP

2025

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

**VIABILIDADE DA APLICAÇÃO DE REDES
NEURAIS PARA A CORREÇÃO DE ERROS DE
SEGMENTAÇÃO EM VASOS SANGUÍNEOS**

JAYME SAKAE DOS REIS FURUYAMA

Trabalho de Conclusão de Curso apresentado ao curso de Ciência da Computação do Centro de Ciências Exatas e de Tecnologia da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Cesar Henrique Comin

São Carlos – SP

2025

Dedico este trabalho à minha família, em especial ao meu pai, Raul, e à minha mãe, Merilandi, pelo apoio e incentivo nos meus aprofundamentos acadêmicos. Às minhas tias, Júlia e Lígia, pelo carinho e incentivo em cada etapa da minha jornada.

Agradeço também aos colegas e amigos dos grupos de pesquisa INAG e ao laboratório LAVIC, que foram fundamentais com o apoio técnico, intelectual e camaradagem.

AGRADECIMENTOS

Agradeço, primeiramente, a Deus, por me guiar e fortalecer em cada momento desta caminhada, dando-me coragem, resiliência e fé para seguir em frente.

Ao meu pai, Raul, à minha madrastra, Alice, à minha irmã, Caroline, à minha mãe, Merilandi, assim como às minhas tias, Lígia e Júlia, por serem exemplos de dedicação, esforço e determinação. Seu apoio incondicional, incentivo e amor sempre foram pilares fundamentais na minha trajetória. Nos momentos mais difíceis, foram suas palavras e atitudes que me motivaram a persistir e a acreditar no meu potencial.

Ao meu orientador, Cesar Henrique Comin, por sua paciência, confiança e dedicação. Suas valiosas orientações, críticas construtivas e incentivo ao pensamento crítico foram fundamentais não apenas para o desenvolvimento deste trabalho, mas também para o meu crescimento acadêmico e profissional. Agradeço pelas inúmeras discussões enriquecedoras, que ampliaram minha visão e aprofundaram meu conhecimento.

Aos meus grandes amigos, Gabriel e Samuel, com quem compartilhei não apenas momentos felizes, mas também desafios, inseguranças e superações. A amizade de vocês foi um alicerce fundamental durante essa etapa, tornando as dificuldades mais leves e os momentos de celebração ainda mais especiais. Nossa jornada juntos fortaleceu laços que levarei para toda a vida.

Aos amigos que fiz durante essa caminhada, cada um contribuindo de forma única para essa experiência. Compartilhamos risadas, aprendizados, noites de estudo e momentos de incerteza, mas, acima de tudo, crescemos juntos. Agradeço por cada conversa, apoio e parceria ao longo dessa jornada.

Aos professores que, direta ou indiretamente, contribuíram para minha formação, transmitindo conhecimento, inspirando curiosidade e incentivando o pensamento crítico. Cada ensinamento deixou uma marca significativa no meu desenvolvimento.

À minha família e a todos aqueles que, de alguma forma, me apoiaram, seja com palavras de incentivo, gestos de carinho ou simplesmente acreditando em mim. O suporte de vocês foi essencial para que eu chegasse até aqui.

Por fim, agradeço a você, leitor, por dedicar seu tempo à leitura deste trabalho. Espero que este estudo possa contribuir de alguma forma para seu conhecimento e reflexão.

A todos vocês, minha eterna gratidão.

A ciência consiste em substituir o saber que parecia seguro por uma teoria, ou seja, por algo problemático.
(José Ortega y Gasset)

RESUMO

O avanço das redes neurais convolucionais (CNN) representou um marco significativo para diversas áreas da visão computacional, sendo amplamente utilizado em tarefas como classificação de imagens, segmentação e detecção de objetos, tanto em pesquisas quanto na indústria. Em especial, essas redes promoveram importantes progressos no campo biomédico, onde são aplicadas à segmentação de imagens de microscopia e raio-x, auxiliando na identificação de possíveis anomalias em vasos sanguíneos. No entanto, a precisão diagnóstica pode ser comprometida pela presença de erros ou descontinuidades nos segmentos vasculares extraídos por uma CNN. Diante desse desafio, este estudo busca avaliar a viabilidade de utilizar uma segunda rede neural para corrigir falhas nos resultados de uma CNN. Para isso, foram geradas imagens artificiais que representam a geometria dos vasos sanguíneos por meio de curvas de Bézier, permitindo testar a capacidade da CNN em restaurar segmentos incompletos. Com base nessa abordagem, foram treinados modelos CNN em diferentes conjuntos de dados, variando a quantidade de segmentos e o grau de descontinuidade dos ramos. Também foi definida uma nova métrica de performance que permite avaliar a segmentação de curvas. Os resultados demonstraram que a rede foi eficaz na correção das segmentações, alcançando altos valores em métricas de qualidade. Entretanto, para grandes descontinuidades, o modelo apresentou dificuldades em identificar padrões claros para a reconstrução, principalmente por conta da complexidade da curva, assim prejudicando a continuidade dos segmentos. Buscando aprimorar os resultados, também foram testadas duas funções de perda: entropia cruzada e Tversky. Embora a diferença geral na qualidade das segmentações tenha sido pequena, a função Tversky apresentou resultados promissores em casos de pequenas remoções, demonstrando melhor recuperação dos segmentos perdidos.

Palavras-chave: redes neurais convolucionais, segmentação de imagens, continuidade de segmentos, vasos sanguíneos, reconstrução de segmentação .

ABSTRACT

Convolutional Neural Networks (CNNs) represented a significant milestone for several areas of computer vision, being widely used in tasks such as image classification, segmentation and object detection, both in research and in industry. In particular, these networks promoted important advances in the biomedical field, where they are applied to the segmentation of microscopy and X-ray images, helping to identify possible anomalies in blood vessels. However, diagnostic accuracy can be compromised by the presence of errors or discontinuities in the vascular segments extracted by a CNN. Faced with this challenge, this study seeks to evaluate the feasibility of using a second neural network to correct errors in the results of a CNN. To this end, artificial images representing the geometry of blood vessels using Bézier curves were generated, allowing the CNN's ability to restore incomplete segments to be tested. Based on this approach, CNNs were trained on different datasets, varying the number of segments and the degree of discontinuity of the branches. The results demonstrated that the network was effective in correcting the segmentations, achieving high values in quality metrics. However, as the removal of parts of vessels increased, the model had difficulty identifying clear patterns for reconstruction, mainly due to the complexity of the curve, thus compromising the continuity of the segments. In order to improve the results, two loss functions were also tested: cross-entropy and Tversky. Although the overall difference in segmentation quality was small, the Tversky function showed promising results in cases of small removals, demonstrating better recovery of lost segments.

Keywords: convolutional neural networks, image segmentation, segment continuity, blood vessels, segmentation reconstruction.

LISTA DE FIGURAS

Figura 1 – Demonstração de uma imagem de vasos sanguíneos obtida por microscopia e sua respectiva imagem de segmentação binária.	16
Figura 2 – Visualização das camadas da arquitetura U-Net.	20
Figura 3 – Ilustração da criação de curvas de Bézier.	24
Figura 4 – Demonstração de amostras dos conjuntos de dados com uma curva e com cinco curvas.	25
Figura 5 – Ilustração da remoção de uma região (mostrada em verde) de uma curva utilizando um disco.	25
Figura 6 – Demonstração da métrica IDS. Na figura à esquerda, são mostrados em branco os pixels de interseção entre o padrão ouro e a predição do modelo, em azul os pixels do padrão ouro que não foram reconhecidos pelo modelo (falsos negativos) e em vermelho os pixels preditos pelo modelo mas não presentes no padrão ouro (falsos positivos). Para o cálculo do IDS, o padrão ouro é dilatado, como mostrado na figura à direita. Nesse caso, não há mais falsos positivos e os pixels dilatados não são considerados falsos negativos.	27
Figura 7 – Amostras com remoções e raios utilizados para a remoção (linha superior) e seus respectivos padrões-ouro (linha inferior).	29
Figura 8 – Mapa de calor dos resultados de Precisão e Revocação para os experimentos usando Entropia Cruzada.	29
Figura 9 – Mapa de calor dos resultados de Precisão e Revocação para os experimentos utilizando a função de perda Tversky.	30
Figura 10 – Gráficos de barras mostrando os melhores valores das métricas de qualidade obtidos durante o treinamento dos modelos.	31
Figura 11 – Painel com amostras da reconstrução dos modelos para curvas com 1 segmento	32
Figura 12 – Painel com amostras da reconstrução dos modelos para curvas com 5 segmentos	32

LISTA DE SIGLAS

ANN	Rede Neural Artificial
PLN	Processamento de Linguagem Natural
CNN	Rede Neural Convolucional
IoU	Intersecção sobre a união
CIDice	Centerline Dice
NN	Redes Neurais
IDS	Intersecção da dilatação sobre o esqueleto

SUMÁRIO

CAPÍTULO 1–INTRODUÇÃO	12
1.1 Contexto e motivação	12
1.2 Objetivos	13
1.2.1 Geral	13
1.2.2 Específicos	13
1.3 Organização	13
CAPÍTULO 2–FUNDAMENTAÇÃO TEÓRICA	15
2.1 Segmentação de imagens	15
2.2 Aprendizado profundo	15
2.2.1 Retropropagação	17
2.3 Métricas de avaliação de qualidade de segmentação	17
2.3.1 Precisão	17
2.3.2 Revocação	17
2.3.3 Intersecção sobre a união	18
2.3.4 Esqueletização	18
2.3.5 CIDice	19
2.4 Funções de perda	19
2.4.1 Entropia cruzada	19
2.4.2 Função Tversky	19
2.5 Rede de segmentação U-Net	20
CAPÍTULO 3–TRABALHOS RELACIONADOS	21
3.1 Segmentação de objetos	21
3.2 Pós-processamento da segmentação	21
CAPÍTULO 4–METODOLOGIA E DESENVOLVIMENTO	23
4.1 Ferramentas utilizadas	23
4.2 Geração de segmentos de vasos	24
4.3 Experimentos com um segmento na imagem	26
4.4 Experimentos com vários segmentos na imagem	26
4.5 Intersecção da dilatação sobre o esqueleto - IDS	26
CAPÍTULO 5–RESULTADOS	28
5.1 Treinamento da rede	28
5.2 Análise das métricas de qualidade	29

5.3	Interpretação dos resultados	30
CAPÍTULO 6–CONCLUSÃO	33
6.1	Trabalhos futuros	34
REFERÊNCIAS	35

Capítulo 1

INTRODUÇÃO

1.1 Contexto e motivação

Os vasos sanguíneos têm um papel de extrema importância para o corpo de seres vertebrados, neles são distribuídos os nutrientes essenciais para o bom funcionamento de órgãos e tecidos de seres vivos (PITTMAN, 2011). Considerando a importância, há diversos estudos dessas estruturas publicados na literatura, tais como o aprimoramento de diagnósticos para detecção de doenças através de exames de imagem, assim como exames de fundo de olho (KLUFAS et al., 2015), nos quais são destacados os vasos sanguíneos oculares para a detecção de doenças como retinopatia diabética.

Visto que há uma grande importância na análise de vasos sanguíneos, surge a necessidade de desenvolvimento e aprimoramento dos métodos atuais para a captura e análise dessas estruturas por meio de técnicas de imagem. Em conjunto com métodos de processamento de imagens, é possível separar os tecidos de forma que as estruturas de interesse sejam facilmente visualizadas. Com o grande desenvolvimento do aprendizado profundo, hoje é possível que as imagens sejam utilizadas para o treinamento de modelos de aprendizado de máquina que se especializam em fazer tais detecções de forma automática e com pouca revisão humana.

Atualmente, o desenvolvimento das Redes Neurais Convolucionais (CNNs) tem se destacado no contexto de medicina computacional, principalmente por serem eficientes e com uma boa taxa de acerto (SOOMRO et al., 2019). Essas redes têm a capacidade de destacar os objetos de interesse nas imagens para as quais foram treinadas a detectar, neste contexto, com os vasos sanguíneos segmentados a partir da imagem original.

Existem algumas arquiteturas de modelos de aprendizado de máquina que se destacam para tarefas de segmentação, como a U-Net (RONNEBERGER et al., 2015), que tem embasamento sobre as Redes Neurais Convolucionais com o objetivo de codificar uma imagem em atributos relevantes e decodificá-los segundo um padrão desejado, segmentando o objeto para o qual foi treinada.

No contexto de vasos sanguíneos, para algumas imagens os modelos podem ter problemas

com a completude dos vasos, de forma que alguns segmentos podem acabar se desconectando do ramo ao qual pertencem. Considerando este problema, um estudo recente indica que o treinamento de modelos utilizando formas (GEIRHOS et al., 2022) acaba sendo benéfico para o aprendizado pois o modelo se torna invariante à variações de textura nas imagens.

Desta forma, o presente trabalho tem como foco principal verificar a viabilidade da aplicação de CNNs para a correção de estruturas segmentadas. Inicialmente, para validação deste estudo, foram utilizadas curvas de Bézier para a simulação dos ramos de uma imagem de vasos sanguíneos, assim, representando a complexidade e a variabilidade dos diferentes padrões que podem ocorrer em um segmento real.

1.2 Objetivos

1.2.1 Geral

Verificar se é possível treinar com sucesso uma CNN para a tarefa de corrigir imagens binárias contendo erros de segmentação de vasos sanguíneos. Em particular, os erros de segmentação a serem corrigidos são descontinuidades em segmentos de vasos.

1.2.2 Específicos

Os objetivos específicos do trabalho são:

1. Desenvolver um conjunto de dados que permita simular a geometria de vasos sanguíneos;
2. Desenvolver métodos para quantificar a qualidade da segmentação de esqueletos de vasos preditos pelos modelos;
3. Identificar parâmetros de treinamento e funções de perda que possibilitem treinar com sucesso um modelo para correção de erros de segmentação.

1.3 Organização

Este trabalho está organizado da seguinte forma:

- Capítulo 2: são apresentados a fundamentação teórica e conceitos necessários para a compreensão dos assuntos mencionados no desenvolvimento do estudo, incluindo (i) Segmentação de imagens, (ii) Aprendizado profundo, (iii) Métricas de avaliação de qualidade de segmentação, (iv) Funções de perda utilizadas e (v) Rede de segmentação U-Net;
- Capítulo 3: neste capítulo, é realizada uma revisão bibliográfica de trabalhos relacionados ao tema;

-
- Capítulo 4: o capítulo descreve e estabelece a metodologia e técnicas utilizadas no trabalho para concluir os objetivos propostos;
 - Capítulo 5: são descritos os experimentos, as visualizações das métricas estabelecidas e das reconstruções feitas a partir dos experimentos;
 - Capítulo 6 apresenta as considerações finais e propostas para futuros trabalhos.

Capítulo 2

FUNDAMENTAÇÃO TEÓRICA

2.1 Segmentação de imagens

A segmentação de imagens é uma tarefa importante em diversos contextos, como por exemplo em reconhecimento de objetos (CHEN et al., 2019), diagnósticos biomédicos (LEE et al., 2024) e carros autônomos (PAN et al., 2020). A segmentação de imagens consiste em destacar os pixels de objetos de interesse nas imagens, de forma a atribuir um rótulo a cada pixel de objeto.

No contexto de imagens médicas comumente são utilizadas as segmentações binárias, onde são definidos como 0 os pixels de fundo da imagem e como 1 os pixels de objeto. Essa abordagem simplifica a análise, facilitando a aplicação de técnicas computacionais para extração de características, quantificação de áreas e auxílio em diagnósticos. Além disso, a segmentação binária é frequentemente utilizada em combinação com métodos de Aprendizado de Máquina e Redes Neurais Convolucionais para aprimorar a precisão de detecção de estruturas em exames como microscopia (LEE et al., 2024) ou fundo de olho (KLUFAS et al., 2015). Um exemplo de segmentação binária é mostrado na figura 1.

As segmentações são utilizadas em diversos contextos, incluindo na área médica para o auxílio a diagnóstico de doenças que podem ser detectadas por imagem de algum órgão ou região específica do corpo. Os modelos de Aprendizado de Máquina se destacam neste contexto, pois possibilitam que um modelo seja treinado em imagens anotadas por especialistas e aplicado para o diagnóstico em novas imagens. Na seção seguinte, são explicadas as utilizações no contexto de Redes Neurais (NN).

2.2 Aprendizado profundo

A área de Aprendizado de Máquina tem se popularizado nas últimas décadas, principalmente para a resolução de diversos problemas que antes seriam feitos por um humano. Hoje, modelos de Aprendizado de Máquina aprendem a realizar ações de forma precisa e com alta

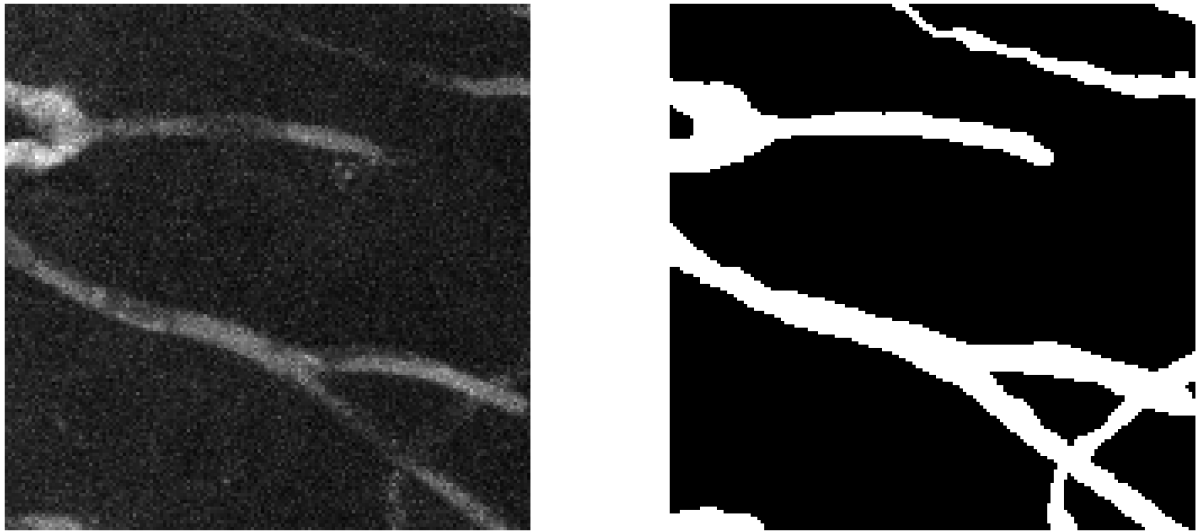


Figura 1 – Demonstração de uma imagem de vasos sanguíneos obtida por microscopia e sua respectiva imagem de segmentação binária.

Fonte: VessMAP (SILVA et al., 2024).

escalabilidade em diferentes contextos. Dentro de Aprendizado de Máquina existem diversos tópicos para diferentes contextos, como por exemplo a Visão Computacional, Reconhecimento de Voz, Processamento de Linguagem Natural, Mecanismos de Recomendação entre outros.

O Aprendizado Profundo, também conhecido pelo termo em inglês Deep Learning, pode ser entendido como o aprendizado de representação, considerando que com poucas operações de pré-processamento, o modelo é capaz de realizar determinada ação que ele foi designado a fazer, podendo ser Classificação, Regressão, Detecção de Objetos, Segmentação de Objetos, entre outras.

A arquitetura que popularizou a área de Deep Learning foi a Rede Neural Artificial (Artificial Neural Network, ou ANN), que foi inspirada no sistema nervoso dos animais, de forma que cada unidade de processamento interna do modelo é chamada de neurônio. Conjuntos de neurônios podem ser utilizados para a extração de características e, por isso, são chamados de codificadores. Os Neurônios utilizados na camada de saída são chamados de cabeça (head) do modelo.

A partir do desenvolvimento das ANNs, foram criadas adaptações dessas arquiteturas para suprir diferentes necessidades, como as Redes Neurais Convolucionais (Convolutional Neural Networks, ou CNNs) aplicadas em Visão Computacional. Nelas, cada uma das camadas internas realiza a tarefa de extração de características, que são passadas para as camadas seguintes e assim por diante até a camada de saída do modelo.

2.2.1 Retropropagação

A Retropropagação (ou Backpropagation) (RUMELHART et al., 1986) foi um marco importante para o desenvolvimento de modelos de Aprendizado de Máquina. Antes da sua criação, havia uma grande limitação na capacidade de abstração dos modelos, o que resultava em uma baixa taxa de acerto. Para contornar essa limitação, foi criada a retropropagação, que visa calcular o gradiente de erro em relação aos pesos da rede de maneira eficiente, penalizando cada camada do modelo caso haja um erro na comparação entre o padrão ouro em relação à predição do modelo.

2.3 Métricas de avaliação de qualidade de segmentação

As métricas para a avaliação dos resultados constituem uma etapa muito importante no desenvolvimento de novos modelos. No contexto de segmentação, a partir de uma imagem anotada manualmente por um humano conseguimos quantificar a qualidade da máscara de segmentação gerada por um modelo. Porém, uma dificuldade que é encontrada quando nos deparamos com o contexto de segmentação é de que a maioria das métricas são análises de pixel a pixel, ou seja, se um pixel da imagem de saída do modelo está incorreto em relação à anotação humana, o valor de acerto será penalizado, mesmo que não seja um pixel relevante para a análise.

A partir disto, podemos considerar os valores de verdadeiros positivos (VP) são os pixels que pertencem à estrutura de interesse tanto na segmentação predita quanto na segmentação do padrão ouro. Os falsos negativos (FN) correspondem aos pixels que pertencem à segmentação do padrão ouro, mas que não foram identificados pelo modelo. Já os falsos positivos (FP) são os pixels que estão presentes na segmentação de saída do modelo, mas não na segmentação do padrão ouro. A seguir, serão descritas algumas métricas que serão utilizadas neste trabalho.

2.3.1 Precisão

A precisão (do inglês Precision) é uma medida de qualidade que visa calcular a quantidade de acertos de uma análise preditiva. Neste contexto, a precisão pode ser entendida como a quantidade de pixels da intersecção entre o padrão ouro e a segmentação predita pelo modelo, dividida pela quantidade de valores positivos preditos pelo modelo. Ela pode ser calculada utilizando a Equação 2.1.

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (2.1)$$

2.3.2 Revocação

A revocação (do inglês Recall) é uma medida de qualidade que consiste em mensurar a quantidade de dados que são relevantes para a análise. Neste contexto, a revocação pode ser

interpretada como a quantidade de pixels positivos presentes no padrão ouro que foram de fato detectados pelo modelo, de forma que pode ser calculada utilizando a Equação 2.2.

$$\text{Revocação} = \frac{VP}{VP + FN} \quad (2.2)$$

2.3.3 Intersecção sobre a união

Uma das métricas de qualidade mais populares hoje quando se trata de Visão Computacional é a medida de Intersecção sobre União (Intersection Over Union, ou IoU) sendo criada inicialmente para Detecção de Objetos, com o intuito de conseguir quantificar se as predições do modelo estavam nos pixels exatos dos padrão ouro criado para aquela imagem. Com o tempo, foi adaptada para tarefas de segmentação de imagens, com o foco de quantificar pixel a pixel a qualidade das segmentações geradas pelos modelos.

A fórmula do IoU consiste na divisão entre a intersecção do padrão ouro e a saída do modelo, sobre a união entre ambos, sendo assim, definido assim como na Equação 2.3.

$$IOU = \frac{GT \cap Label}{GT \cup Label} \quad (2.3)$$

A IoU também pode ser calculada através da Equação 2.4.

$$IOU = \frac{VP}{VP + FP + FN} \quad (2.4)$$

2.3.4 Esqueletização

Os algoritmos de esqueletização, também conhecidos como algoritmos de eixo medial ou afinamento (LEE et al., 1994), são técnicas fundamentais para a extração do eixo central de estruturas tubulares, especialmente na segmentação de vasos sanguíneos. Esses métodos reduzem formas complexas a uma representação unidimensional que preserva sua topologia, permitindo uma visualização mais compacta e eficiente. Além disso, garantem a conectividade da estrutura resultante e preservam a geometria do objeto original, minimizando distorções durante o processo de afinamento.

Segundo Lee et al. (LEE et al., 1994), os algoritmos de esqueletização podem ser implementados em duas ou três dimensões, utilizando abordagens iterativas para a remoção de pontos da borda enquanto mantêm a continuidade da estrutura. O método proposto no artigo enfatiza a importância da preservação do eixo medial por meio de regras específicas de remoção de pixels, garantindo que o esqueleto gerado represente com fidelidade a forma original.

2.3.5 CIDice

O CIDice (ou Centerline Dice) (SHIT et al., 2020) é uma métrica de avaliação da qualidade de segmentação que se concentra na precisão da linha central de objetos tubulares, como em vasos sanguíneos ou neurônios. Essencialmente, o CIDice é uma variação do coeficiente Dice. O CIDice é calculado como descrito na equação 2.7 (SHIT et al., 2020), utilizando a precisão e a revocação do esqueleto, conforme as fórmulas 2.5 e 2.6, respectivamente, onde V_L representa o padrão ouro, V_P a segmentação predita, S_P seu esqueleto correspondente e S_L o esqueleto do padrão ouro.

$$T_{prec}(S_P, V_L) = \frac{|S_P \cap V_L|}{|S_P|} \quad (2.5)$$

$$T_{sens}(S_L, V_P) = \frac{|S_L \cap V_P|}{|S_L|} \quad (2.6)$$

$$CIDice(V_P, V_L) = 2 \cdot \frac{T_{prec}(S_P, V_L) \cdot T_{sens}(S_L, V_P)}{T_{prec}(S_P, V_L) + T_{sens}(S_L, V_P)} \quad (2.7)$$

2.4 Funções de perda

Para a penalização do modelo usando as ideias apresentadas na subseção de retropropagação 2.2.1, é necessária a utilização de funções de perda que melhor representam o escopo de desenvolvimento. Nesta seção serão apresentadas a função de perda Entropia Cruzada e a função Tversky que serão utilizadas nos experimentos deste trabalho.

2.4.1 Entropia cruzada

A Entropia Cruzada (ou Cross Entropy Loss) é uma função de perda muito utilizada em contextos de segmentação e classificação, especialmente quando o objetivo está relacionado a atribuir diferentes valores de probabilidade entre as classes ou rótulos obtidos pelo modelo. A Equação 2.8 destaca a aplicação da entropia cruzada em um conjunto de n pixels, onde \mathbf{p}_i representa o vetor de probabilidades retornado pelo modelo para o i -ésimo pixel e \mathbf{y}_i representa a classe, ou rótulo, do pixel i . O termo $p_i(y_i)$ representa a probabilidade retornada pelo modelo para a classe esperada do pixel i .

$$E(y) = - \sum_i^n \log(p_i(y_i)) \quad (2.8)$$

2.4.2 Função Tversky

A função de perda Tversky (SALEHI et al., 2017) foi criada com o intuito de mensurar os acertos do modelo no caso de segmentação binária, considerando o balanceamento entre as

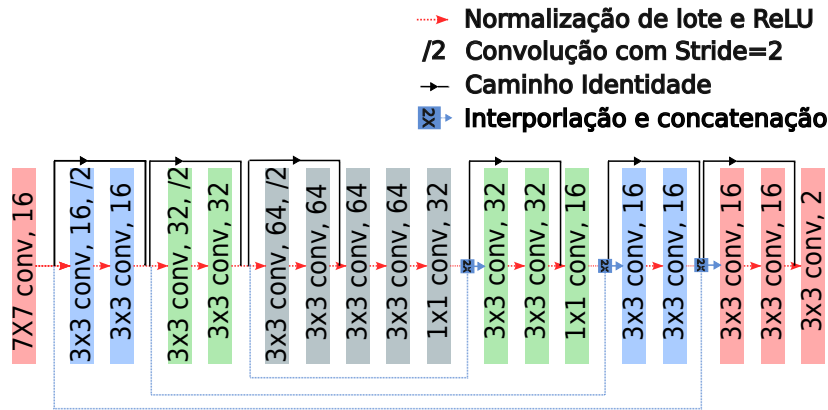


Figura 2 – Visualização das camadas da arquitetura U-Net.

classes na imagem, ou seja, se há mais pixels de fundo ou de interesse. Dessa forma, a função consegue penalizar o modelo caso haja uma quantidade significativa de pixels classificados de maneira errônea. A função é mostrada na Equação 2.9 (SALEHI et al., 2017), onde p_{0i} é a probabilidade retornada pelo modelo de um pixel i ser um pixel de interesse (objeto), p_{1i} a probabilidade de que não seja um pixel de interesse (fundo) e g_{0i} e g_{1i} são os valores dos pixels referentes à imagem padrão ouro. Os parâmetros α e β podem ser ajustados para ponderar a relevância de falsos positivos (FP) e falsos negativos (FN) no valor final da perda.

Quando $\alpha = 0.5$ e $\beta = 0.5$, a função Tversky se torna a função de perda Dice, que é muito utilizada na literatura. Os autores da função Tversky utilizaram os valores $\alpha = 0.3$ e $\beta = 0.7$ no artigo (SALEHI et al., 2017). Esses valores foram utilizados nos nossos experimentos.

$$T(\alpha, \beta) = \frac{\sum_{i=1}^N p_{0i}g_{0i}}{\sum_{i=1}^N p_{0i}g_{0i} + \alpha \sum_{i=1}^N p_{0i}g_{1i} + \beta \sum_{i=1}^N p_{1i}g_{0i}} \quad (2.9)$$

2.5 Rede de segmentação U-Net

A U-Net (RONNEBERGER et al., 2015) é uma arquitetura de Redes Neurais Convolucionais adaptada para a tarefa de segmentação de imagens, com o foco principal de ser utilizada para imagens Biomédicas. O grande destaque desta arquitetura é que ela utiliza como base a arquitetura ResNet e faz as adaptações de acordo com a quantidade de camadas de abstração da arquitetura, de forma que após a extração de características obtidas, existe uma reconstrução da imagem utilizando aumento da resolução (Upscaling). A figura 2 demonstra de maneira resumida como é formada a arquitetura U-Net utilizando a Resnet 18.

Capítulo 3

TRABALHOS RELACIONADOS

Nesta seção serão explorados alguns estudos da literatura relacionados ao tema desenvolvido neste trabalho.

3.1 Segmentação de objetos

A segmentação de objetos é uma área em constante evolução, com diversas aplicações, como a identificação de pessoas e animais, além da análise de imagens médicas, percepção robótica, realidade aumentada e compreensão de imagens (MINAEE et al., 2022). Em contextos críticos, como aplicações médicas e veículos autônomos, a segmentação deve ser eficiente, executável em tempo real e precisa, além disso, os modelos devem ser altamente precisos em suas previsões.

No contexto da segmentação de vasos sanguíneos, diversos modelos são avaliados em diferentes conjuntos de dados para medir seu desempenho na retina (MOOKIAH et al., 2021). Em sua maioria, os modelos destacados mantêm uma alta taxa de acerto, entretanto, as segmentações ainda precisam de revisão humana para validação, pois podem apresentar erros ou incompletudes em determinados ramos.

3.2 Pós-processamento da segmentação

Em determinados contextos, a continuidade ou a separação exata de pixels sobre a imagem de segmentação é de extrema importância, dessa forma, foram desenvolvidas metodologias que atendem a essa necessidade. Por exemplo, SegRefiner (WANG et al., 2023), no qual é desenvolvido um modelo de difusão para o refinamento da segmentação dos objetos da imagem, ajustando a predição conforme o objeto de interesse. Em outro estudo (QIU et al., 2023), descreve-se a utilização de duas redes de segmentação, uma para a segmentação do objeto e outra para o refinamento da saída utilizando a U-Net como arquitetura base.

Uma técnica tradicional para corrigir falhas de segmentação consiste em utilizar os

chamados *Fully Connected Conditional Random Fields* (CRFs completamente conectados) (KRÄHENBÜHL; KOLTUN, 2012), que são modelos probabilísticos discriminativos usados para rotular e segmentar dados estruturados, como sequências de texto ou imagens. Nesse estudo, focado em imagens, é proposta uma abordagem que utiliza máscaras gaussianas para filtragem espacial, aprimorando a representação das características e melhorando a segmentação de regiões com pequenas variações entre o objeto de interesse e o fundo.

As CNNs apresentam limitações na segmentação, especialmente na precisão das bordas, o que pode levar à perda de detalhes finos. Para corrigir esse problema, foram desenvolvidos modelos que utilizam diferentes padrões de kernels para as convoluções, combinando informações de cor e posição a fim de melhorar a segmentação e preservar melhor os contornos dos objetos (CHEN et al., 2016). Além disso, estudos recentes destacam a importância das texturas no aprendizado e na predição de modelos baseados em CNNs (GEIRHOS et al., 2022), de forma que, independentemente das bordas ou tonalidade de cores, se os detalhes da imagem forem diferentes dos padrões do mundo real, os modelos têm grandes dificuldades para realizar a tarefa com êxito.

Recentemente foi desenvolvido um modelo de segmentação que utiliza uma camada de abstração após a predição do modelo que possibilita mesclar a tarefa de refinamento e completude dos segmentos juntamente com ativações intermediárias do modelo (SONG et al., 2024).

Capítulo 4

METODOLOGIA E DESENVOLVIMENTO

Esta seção detalha os métodos e procedimentos adotados neste trabalho. Primeiramente, são apresentadas as ferramentas utilizadas no desenvolvimento. Em seguida, descreve-se a geração de curvas de Bézier para simular estruturas tubulares, considerando tanto um único segmento quanto cinco segmentos por imagem. Por fim, são explicados os passos para a realização dos experimentos, abordando conjuntos de dados com diferentes configurações. A seguir, são listados os tópicos abordados.

1. Ferramentas utilizadas;
2. Geração de segmentos de Vasos;
3. Experimentos com um segmento na imagem;
4. Experimentos com vários segmentos na imagem.

4.1 Ferramentas utilizadas

As ferramentas e bibliotecas utilizadas para o treinamento e desenvolvimento dos modelos deste estudo foram:

- Linguagem de programação: Python 3.10.13
- Framework para desenvolvimento das CNNs: Pytorch
- Utilitários para treinamento das CNNs: TorchTrainer
- API para processamento dos tensores: CUDA
- Computação numérica: NumPy
- Demonstração de gráficos e imagens: Seaborn e Matplotlib

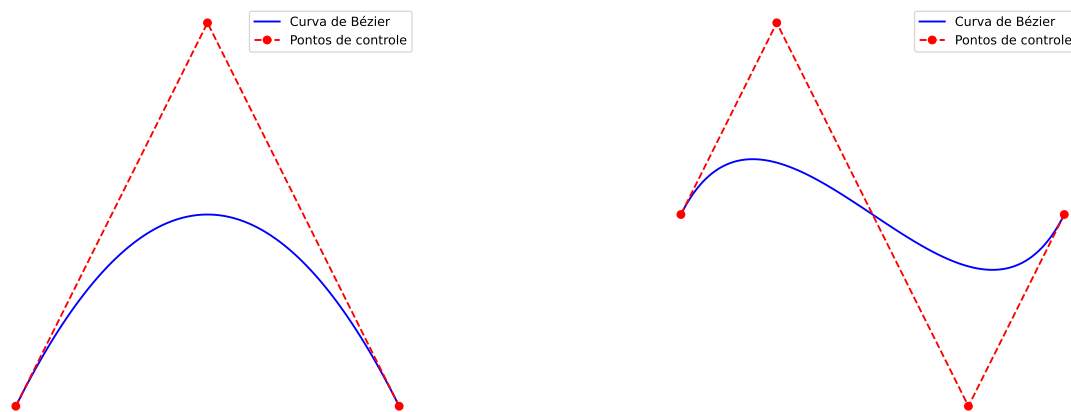
Considerando as ferramentas utilizadas e a necessidade de diversos cálculos com alto custo computacional, o computador utilizado foi disponibilizado pelo laboratório de Visão Computacional (LAVIC) do Departamento de Computação (DC) da Universidade Federal de São Carlos (UFSCar) com as seguintes especificações:

- Processador: AMD Ryzen 7 2700
- Placa Gráfica: Nvidia GeForce RTX 2080 Ti 11GB
- Memória RAM: 16 GB

4.2 Geração de segmentos de vasos

Para a criação do conjunto de dados foram utilizadas curvas de Bézier quadráticas, que possibilitam a criação de segmentos com diferentes tamanhos e diferentes curvaturas apenas com algumas modificações de parâmetros. A Equação 4.1 é utilizada para a criação das curvas, sendo B_0 e B_2 o ponto inicial e final, e B_1 o ponto de controle do arqueamento, podendo ser criados mais pontos intermediários para diferentes curvaturas ao longo do segmento, assim como demonstrado na figura 3.

$$B(t) = (1 - t)^2 b_0 + 2t(1 - t)b_1 + t^2 b_2, \quad t \in [0, 1] \quad (4.1)$$

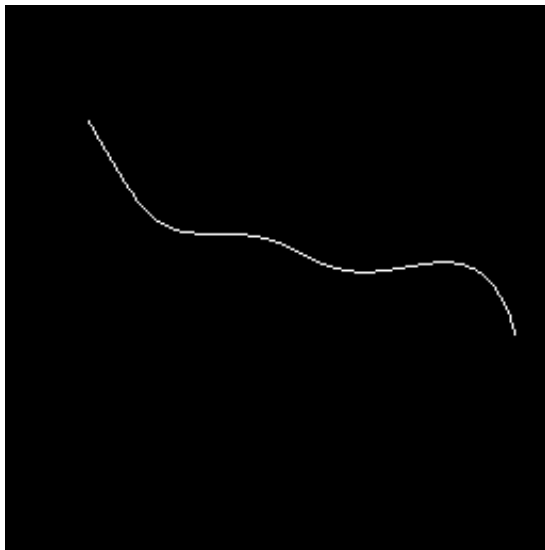


(a) Curva com um ponto de controle.

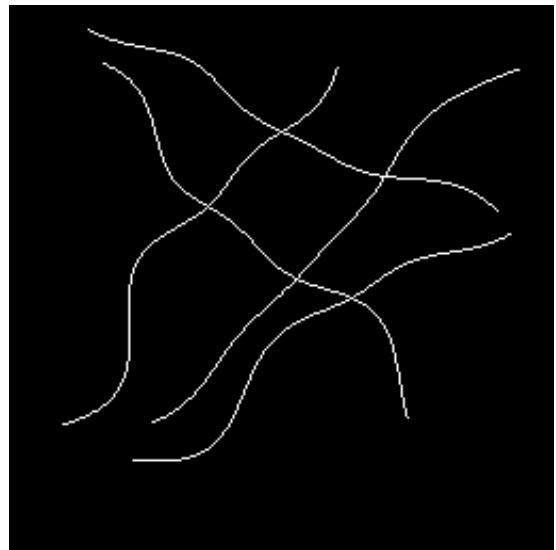
(b) Curva com dois pontos de controle.

Figura 3 – Ilustração da criação de curvas de Bézier.

Os conjuntos criados neste trabalho utilizaram dez pontos de controle definidos aleatoriamente, assim como os pontos de início e fim de cada curva, que foram selecionados dentro de um intervalo próximo às bordas da imagem. Cada amostra gerada possui dimensão [256, 256]. A figura 4 apresenta exemplos dos diferentes conjuntos utilizados nos treinamentos.



(a) Amostra do conjunto de uma curva.



(b) Amostra do conjunto com cinco curvas.

Figura 4 – Demonstração de amostras dos conjuntos de dados com uma curva e com cinco curvas.

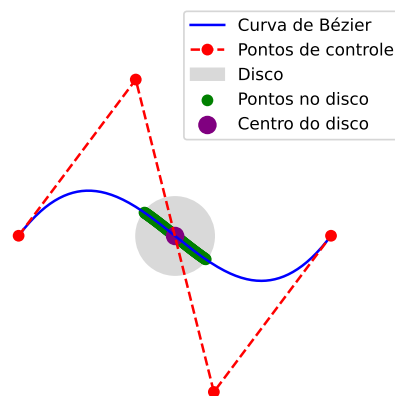


Figura 5 – Ilustração da remoção de uma região (mostrada em verde) de uma curva utilizando um disco.

Para um conjunto de dados que condiz com nossos objetivos propostos, foram feitas adaptações para remoção de alguns pixels da curva. Nesse contexto, chamaremos essas remoções de descontinuidade. O processo consistiu em selecionar aleatoriamente um pixel ao longo do segmento para criar uma oclusão na região, utilizando um disco que remove todos os pixels dentro de seu diâmetro, conforme ilustrado na Figura 5.

4.3 Experimentos com um segmento na imagem

Para a verificação da viabilidade de treinar uma rede neural que possibilite corrigir descontinuidades, inicialmente foram feitos experimentos com imagens contendo apenas um segmento. A entrada do modelo consiste em imagens contendo segmentos descontínuos, e as respectivas imagens de rótulos possuem os segmentos sem as descontinuidades. O modelo deve aprender a preencher as descontinuidades de forma a transformar as imagens de entrada nas imagens de rótulos.

Foram utilizados três conjuntos de dados com diferentes tamanhos de raio para a remoção de pixels, sendo eles de 20, 40 e 60 pixels. Cada conjunto contém 3000 imagens. O objetivo de utilizar diferentes raios é analisar qual é o limite que a rede U-Net de segmentação consegue completar os ramos.

Para o treinamento, foi utilizada a rede U-Net definida na seção 2.5 com a arquitetura Resnet 18 treinada em 500 épocas, com o otimizador Adam. Foi utilizada uma taxa de aprendizado inicial de 0.01, que foi reduzida ao longo do treinamento com um agendador de taxa de aprendizado Polinomial. O tamanho de lote por época foi de 32 imagens.

Além das características destacadas, foram feitos experimentos com a função de perda da Entropia Cruzada e com a função Tversky, Com o intuito de verificar se há ganhos significativos ao utilizar uma penalização recentemente proposta.

4.4 Experimentos com vários segmentos na imagem

No experimento com múltiplos segmentos, foram utilizados conjuntos de dados contendo cinco curvas por imagem, com a remoção de pixels aplicada apenas ao segmento correspondente. Dessa forma, cada imagem apresenta regiões específicas onde partes das curvas foram eliminadas. Foram testados três conjuntos, com remoções de 10, 20 e 40 pixels de raio.

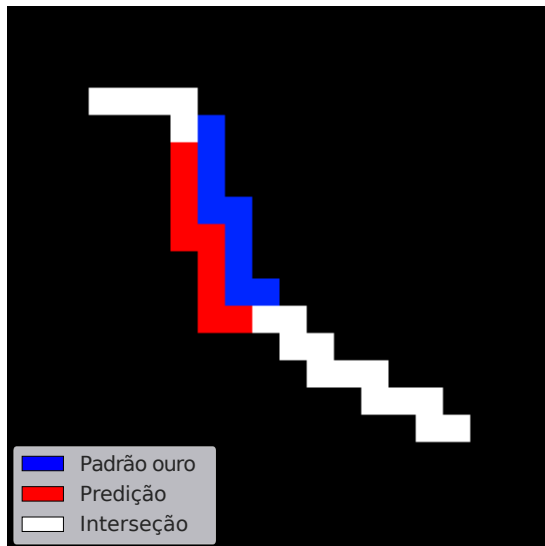
Como características dos experimentos, foram utilizadas as mesmas especificações descritas na seção 4.3, com o intuito de avaliar a capacidade de abstração e reconstrução dos segmentos quando há variações de contextos.

4.5 Intersecção da dilatação sobre o esqueleto - IDS

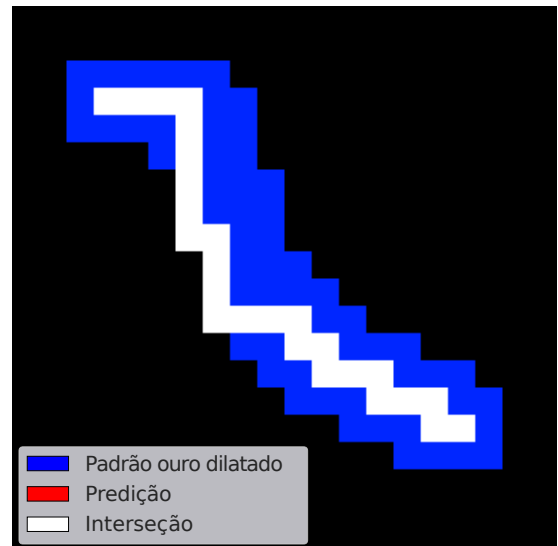
Uma limitação sobre métricas de qualidade de segmentação consiste em normalmente serem calculadas pixel a pixel, que desconsidera pixels de deslocação pequena, ou seja, pixels que estão a distancia de 1 pixel entre o verdadeiro e o predito. Considerando isso, neste estudo foi desenvolvido uma abordagem que supre tal necessidade, chamada de *Intersection of the Dilation over the Skeleton*, ou IDS. A métrica possibilita mensurar a qualidade dos resultados do modelo considerando uma margem de deslocação de um pixel em relação ao resultado esperado.

O IDS é descrito pela fórmula 4.2, onde $pred$ é a predição do modelo, gt é o padrão ouro da imagem e gt_dil é a transformada distância Euclidiana com limitação de 1 pixel de distância. Na Figura 6 é demonstrado o IDS aplicado em uma imagem.

$$IDS(pred, GT) = \frac{soma(gt_dil * pred)}{max(gt, pred)} \quad (4.2)$$



(a) Demonstração dos pixels do padrão ouro em comparação com a predição.



(b) Demonstração do IDS e a dilatação do padrão ouro.

Figura 6 – Demonstração da métrica IDS. Na figura à esquerda, são mostrados em branco os pixels de interseção entre o padrão ouro e a predição do modelo, em azul os pixels do padrão ouro que não foram reconhecidos pelo modelo (falsos negativos) e em vermelho os pixels preditos pelo modelo mas não presentes no padrão ouro (falsos positivos). Para o cálculo do IDS, o padrão ouro é dilatado, como mostrado na figura à direita. Nesse caso, não há mais falsos positivos e os pixels dilatados não são considerados falsos negativos.

Capítulo 5

RESULTADOS

Neste capítulo são apresentados os resultados obtidos a partir dos experimentos descritos no capítulo anterior. Descrevemos os resultados dos treinamentos da rede nos diferentes conjuntos de teste e analisamos as reconstruções dos segmentos.

5.1 Treinamento da rede

O treinamento seguiu as seguintes abordagens para garantir a padronização e reprodutibilidade dos resultados. Portanto, os critérios considerados para o treinamento dos modelos foram os seguintes:

1. Divisão do conjunto de dados: 80% das imagens para treino e 20% para validação;
2. Não foi utilizado aumento de dados, pois as imagens foram geradas artificialmente e os conjuntos de treinamento possuem tamanho adequado;
3. Imagens de entrada com as dimensões de 256 X 256 geradas pelo algoritmo de criação dos segmentos;
4. O modelo com maior valor de IoU encontrado durante o treinamento é utilizado para geração de visualizações;
5. Todos os valores de métricas de qualidade foram calculados utilizando o conjunto de validação.

Além das características destacadas anteriormente, as curvas de Bézier foram geradas seguindo as especificações de remoção de raios nos segmentos. Para o experimento envolvendo uma única curva, foram consideradas remoções de raios de 20, 40 e 60. Já para o conjunto de dados contendo múltiplas curvas, foram criadas cinco curvas com remoções de 10, 20 e 40. A Figura 7 apresenta amostras de cada uma dessas remoções, que serão utilizadas como entrada para o modelo, juntamente com seus respectivos padrões ouro.

Em relação ao tempo de execução, cada época de treinamento durou aproximadamente 34 segundos, enquanto que cada época de validação levou cerca de 9 segundos nos experimentos utilizando a função de perda Entropia Cruzada. Já para a função de perda Tversky, o tempo médio por época foi de 30 segundos para o treinamento e 6 segundos para a validação.

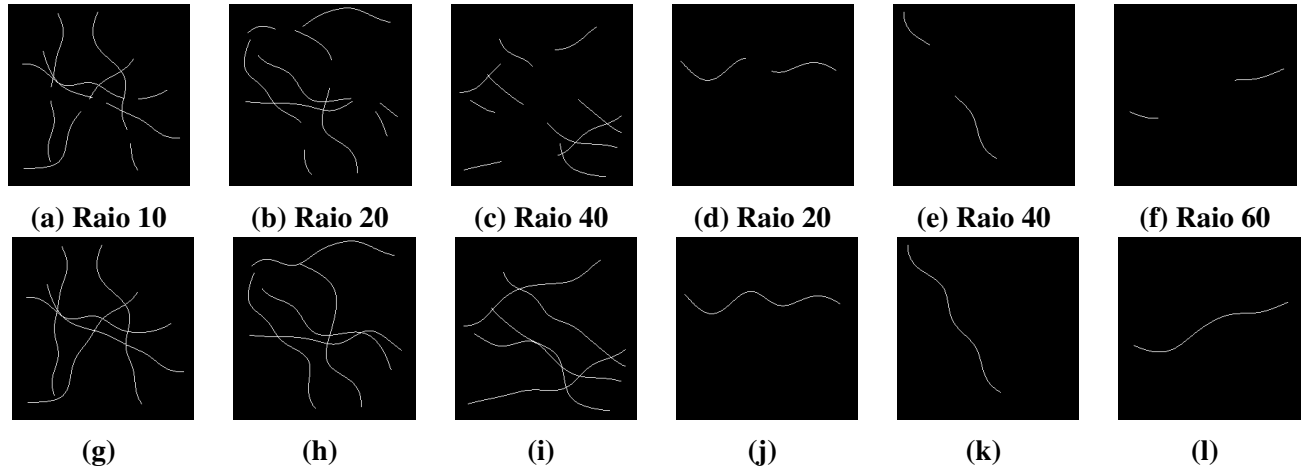
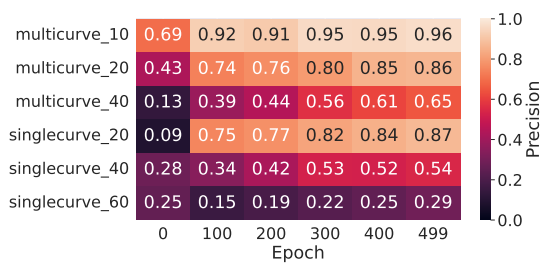


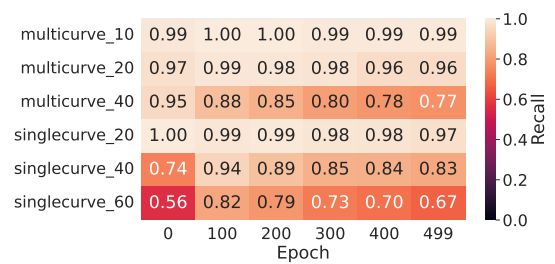
Figura 7 – Amostras com remoções e raios utilizados para a remoção (linha superior) e seus respectivos padrões-ouro (linha inferior).

5.2 Análise das métricas de qualidade

Nos termos definidos neste estudo, as métricas de qualidade foram salvas em diferentes épocas de treinamento, considerando o conjunto de validação utilizado. Desta forma, as métricas de Precisão e Revocação são apresentadas na Figura 8 para a função de perda Entropia Cruzada e na Figura 9 para a função Tversky.



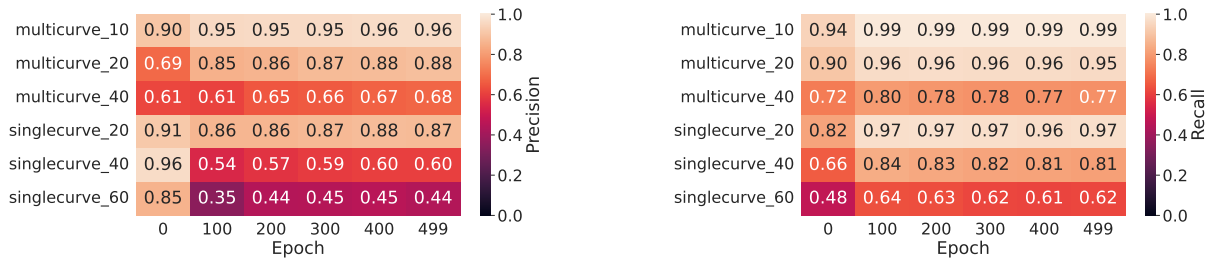
(a) Precisão do conjunto de validação de cada conjunto de dados ao longo das épocas de treinamento.



(b) Revocação do conjunto de validação de cada conjunto de dados ao longo das épocas de treinamento.

Figura 8 – Mapa de calor dos resultados de Precisão e Revocação para os experimentos usando Entropia Cruzada.

Observando os gráficos das execuções é notável de que as métricas mantém um padrão de que ao longo das épocas de treino, o modelo cresce sua taxa de acerto dos pixels, porém, ao longo do crescimento das remoções, o valor se estabiliza ou até diminui. As execuções com a



(a) Precisão do conjunto de validação de cada conjunto de dados ao longo das épocas de treinamento.

(b) Revocação do conjunto de validação de cada conjunto de dados ao longo das épocas de treinamento.

Figura 9 – Mapa de calor dos resultados de Precisão e Revocação para os experimentos utilizando a função de perda Tversky.

função de perda Tversky resultaram em uma melhor taxa de acerto sobre as imagens do conjunto de validação.

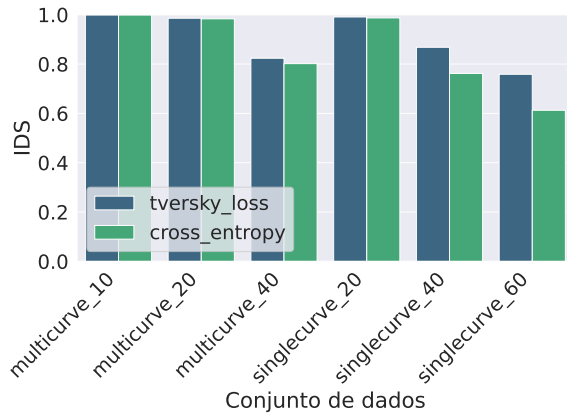
Considerando a revocação, os modelos conseguem reconstruir boa parte dos pixels que foram removidos com facilidade. Nas primeiras épocas, os pixels já começam a ser reconstruídos; porém, existem alguns pixels que foram segmentados como objeto de interesse de maneira errônea, de forma que a Precisão mantém um valor baixo para maiores valores de remoção dos pixels, em outras palavras, o modelo consegue reconstruir, mas há vários pixels falsos positivos que reduzem o valor da Precisão.

Para mensurar os valores de Precisão e Revocação de maneira precisa sobre o contexto deste estudo, foram utilizados as métricas de Intersecção sobre a União (IoU), Center line DICE (CIDice) e Intersection of the Dilation over the Skeleton (IDS), assim como demonstrado na Figura 10.

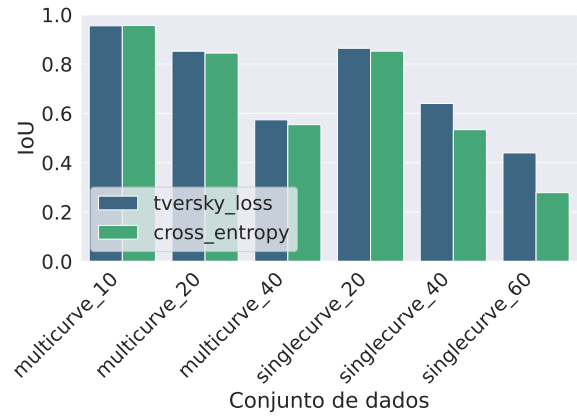
Os gráficos de barras comparam os conjuntos de dados e as funções de perda utilizadas. A qualidade da reconstrução tende a cair com descontinuidades maiores, pois o problema se torna mais difícil. Comparando o IDS e o IoU podemos notar que os diferentes conjuntos conseguem ser reconstruído pelo modelo, porém, com um pequeno deslocamento dos pixels. Observando os valores de CIDice, é possível interpretar que os modelos mantêm uma continuidade boa para as diferentes execuções, com ênfase na função de perda Tversky, que tem a maioria dos melhores resultados.

5.3 Interpretação dos resultados

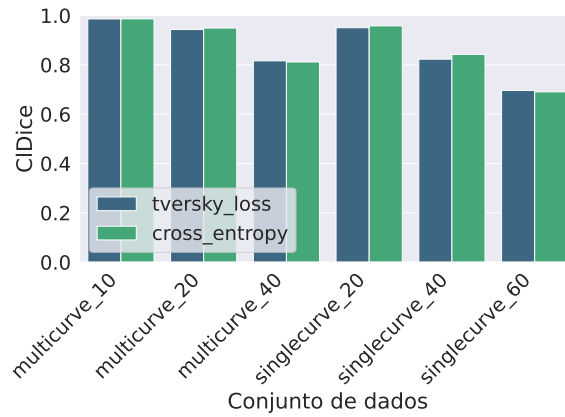
Na Figura 11 são demonstradas amostras dos resultados dos modelos para cada um dos conjuntos de dados obtidos na última época de treinamento. Para pequenos raios de remoção, o modelo reconstrói facilmente com ambas as funções de perda. No entanto, com o aumento da descontinuidade, a complexidade da curvatura impede a reconstrução adequada, comprometendo a continuidade e a precisão em relação ao padrão ouro.



(a) IDS



(b) IoU



(c) CIDice

Figura 10 – Gráficos de barras mostrando os melhores valores das métricas de qualidade obtidos durante o treinamento dos modelos.

No experimento com múltiplas curvas, onde as amostras de demonstração estão na Figura 12, acontece o mesmo caso mencionado no experimento com uma curva. Quando a curvatura é mais complexa, como nos casos com maior remoção, a taxa de acerto diminui, resultando em métricas mais baixas. Em casos simples, com menos remoções, o modelo reconstrói com maior precisão.

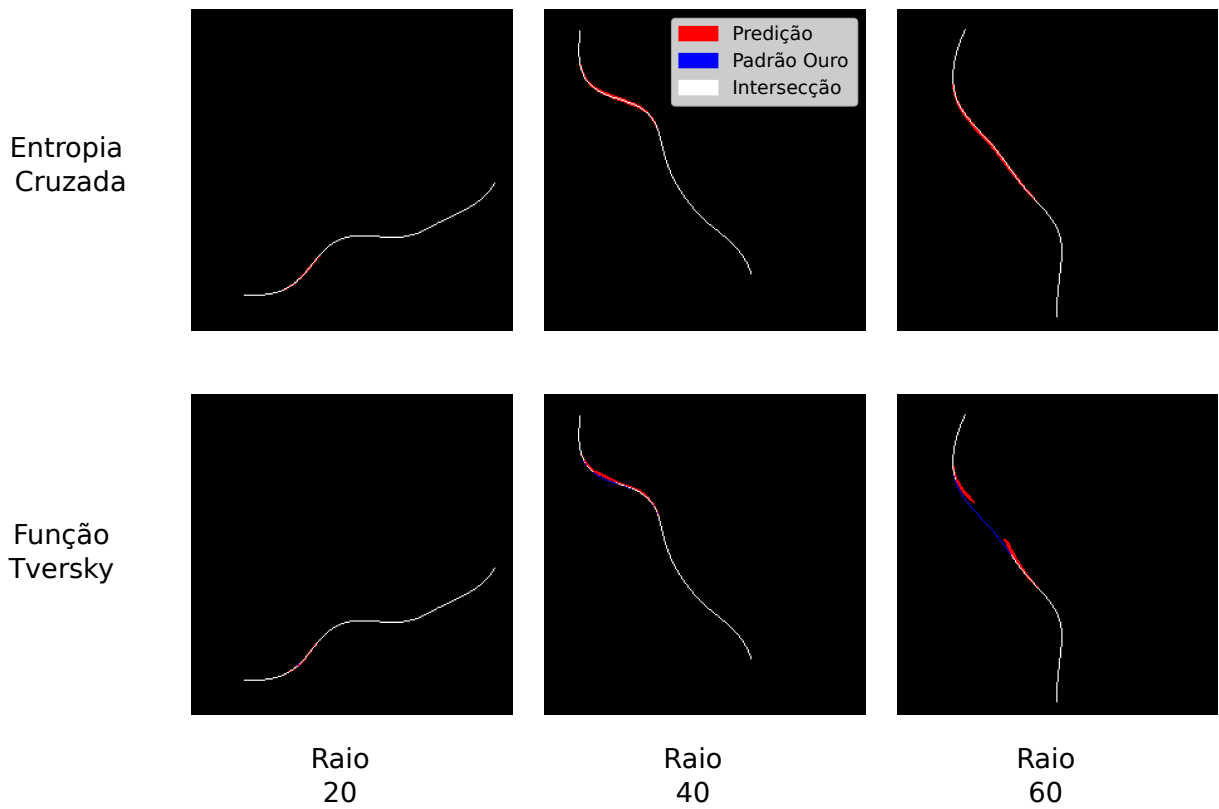


Figura 11 – Painel com amostras da reconstrução dos modelos para curvas com 1 segmento

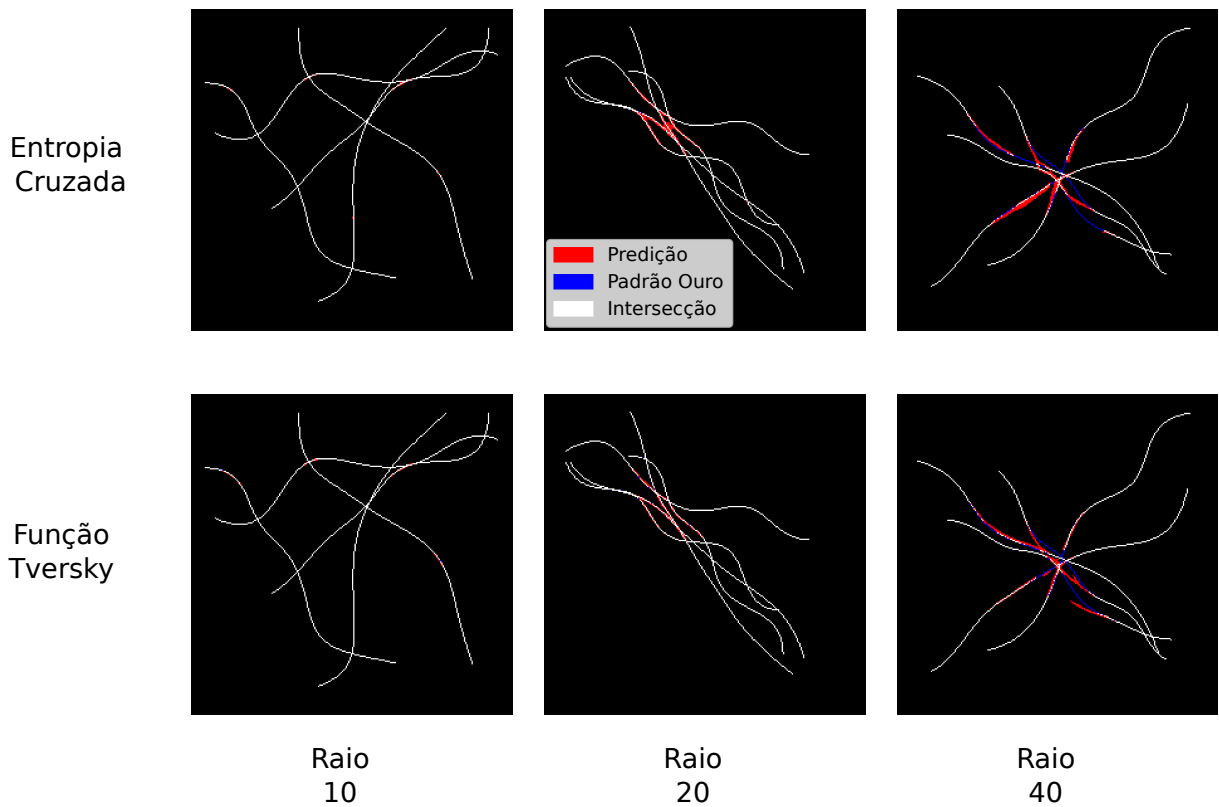


Figura 12 – Painel com amostras da reconstrução dos modelos para curvas com 5 segmentos

Capítulo 6

CONCLUSÃO

A correção de imagens de segmentação é uma tarefa de pós-processamento essencial, especialmente em contextos relacionados à saúde ou a temas sensíveis, onde a margem de erros é mínima e a coerência dos resultados é fundamental. Para aumentar a precisão das predições geralmente se recorre à revisão humana, garantindo que possam ser utilizadas no auxílio ao diagnóstico de doenças.

Diante desse contexto, exploramos um modelo de segmentação para reconstruir segmentos de uma ou múltiplas curvas com um pixel de largura. Além disso, apresentamos métricas predominantemente baseadas em análises pixel a pixel para avaliar o desempenho do modelo. No entanto, foi necessário desenvolver uma métrica complementar, denominada IDS, que permite considerar a segmentação predita em relação ao padrão ouro com uma tolerância de deslocamento entre as curvas, tornando possível avaliar a capacidade do modelo de reconstruir segmentos mesmo que haja um pequeno desvio na posição exata dos pixels.

Na fase de treinamento, utilizamos duas funções de perda: entropia cruzada e Tversky. O objetivo foi permitir que a retropropagação priorizasse os pixels mais relevantes para a análise.

Na avaliação das métricas de qualidade, observamos que ambas as funções de perda produzem resultados similares, mas a função Tversky se destaca nas métricas IDS e IoU. Por outro lado, ao analisarmos a métrica CIDice, ambas funções mantêm valores semelhantes.

A análise das reconstruções preditas pelo modelo indica que, para remoções com raios maiores, o modelo apresenta dificuldades em reconstruir segmentos com precisão. No entanto, para valores baixos de remoção, a reconstrução se aproxima do padrão ouro de maneira quase perfeita.

Retomando os objetivos específicos deste estudo, temos o seguinte:

1. Desenvolver um conjunto de dados que permita simular a geometria de vasos sanguíneos: foram desenvolvidas curvas de Bézier que pudessem representar a geometria e complexidade dos vasos;
2. Desenvolver métodos para quantificar a qualidade da segmentação de esqueletos de vasos

preditos pelos modelos: foi apresentada uma métrica de qualidade chamada IDS que abrange uma margem de 1 pixel de dilatação sobre o padrão ouro;

3. Identificar parâmetros de treinamento e funções de perda que possibilitem treinar com sucesso um modelo para correção de erros de segmentação: foi analisado o desempenho da função de perda Tversky nas segmentações, observando-se resultados promissores quando comparados aos obtidos com a Entropia Cruzada.

6.1 Trabalhos futuros

Para expandir este trabalho, propomos aplicações em diferentes contextos a fim de explorar sua capacidade em dados reais.

1. Aplicar a proposta deste estudo em um conjunto de dados real de segmentação de vasos sanguíneos;
2. Estudar métodos e modelos de segmentação mais eficientes no caso de segmentos possuindo alta curvatura;
3. Investigar a aplicabilidade em outras tarefas envolvendo a identificação de curvas, como rachaduras em materiais, ruas em imagens de satélite e trajetórias em tarefas de rastreamento de objetos.

REFERÊNCIAS

CHEN, L.-C.; PAPANDREOU, G.; KOKKINOS, I.; MURPHY, K.; YUILLE, A. L. *Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs*. 2016. Disponível em: <<https://arxiv.org/abs/1412.7062>>. Citado na página 22.

CHEN, X.; GIRSHICK, R.; HE, K.; DOLLÁR, P. Tensormask: A foundation for dense object segmentation. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2019. p. 2061–2069. Citado na página 15.

GEIRHOS, R.; RUBISCH, P.; MICHAELIS, C.; BETHGE, M.; WICHMANN, F. A.; BRENDEL, W. *ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness*. 2022. Disponível em: <<https://arxiv.org/abs/1811.12231>>. Citado 2 vezes nas páginas 13 e 22.

KLUFAS, M. A.; PATEL, S. N.; RYAN, M. C.; GUPTA, M. P.; JONAS, K. E.; OSTMO, S.; MARTINEZ-CASTELLANOS, M. A.; BERROCAL, A. M.; CHIANG, M. F.; CHAN, R. P. Influence of fluorescein angiography on the diagnosis and management of retinopathy of prematurity. *Ophthalmology*, Elsevier, v. 122, n. 8, p. 1601–1608, 2015. Citado 2 vezes nas páginas 12 e 15.

KRÄHENBÜHL, P.; KOLTUN, V. *Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials*. 2012. Disponível em: <<https://arxiv.org/abs/1210.5644>>. Citado na página 22.

LEE, H. H.; GU, Y.; ZHAO, T.; XU, Y.; YANG, J.; USUYAMA, N.; WONG, C.; WEI, M.; LANDMAN, B. A.; HUO, Y. et al. Foundation models for biomedical image segmentation: A survey. *arXiv preprint arXiv:2401.07654*, 2024. Citado na página 15.

LEE, T.-C.; KASHYAP, R. L.; CHU, C.-N. Building skeleton models via 3-d medial surface axis thinning algorithms. *CVGIP: Graphical Models and Image Processing*, Elsevier, v. 56, n. 6, p. 462–478, 1994. Citado na página 18.

MINAEE, S.; BOYKOV, Y.; PORIKLI, F.; PLAZA, A.; KEHTARNAVAZ, N.; TERZOPOULOS, D. Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 44, n. 7, p. 3523–3542, July 2022. ISSN 1939-3539. Citado na página 21.

MOOKIAH, M. R. K.; HOGG, S.; MACGILLIVRAY, T. J.; PRATHIBA, V.; PRADEEPA, R.; MOHAN, V.; ANJANA, R. M.; DONEY, A. S.; PALMER, C. N.; TRUCCO, E. A review of machine learning methods for retinal blood vessel segmentation and artery/vein classification. *Medical Image Analysis*, v. 68, p. 101905, 2021. ISSN 1361-8415. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1361841520302693>>. Citado na página 21.

- PAN, Z.; EMARU, T.; RAVANKAR, A.; KOBAYASHI, Y. *Applying Semantic Segmentation to Autonomous Cars in the Snowy Environment*. 2020. Disponível em: <<https://arxiv.org/abs/2007.12869>>. Citado na página 15.
- PITTMAN, R. N. The circulatory system and oxygen transport. In: *Regulation of Tissue Oxygenation*. [S.l.]: Morgan & Claypool Life Sciences, 2011. Citado na página 12.
- QIU, Y.; LI, Z.; WANG, Y.; DONG, P.; WU, D.; YANG, X.; HONG, Q. Corsegrec: A topology-preserving scheme for extracting fully-connected coronary arteries from ct angiography. In: _____. [S.l.: s.n.], 2023. p. 670–680. ISBN 978-3-031-43897-4. Citado na página 21.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. Disponível em: <<https://arxiv.org/abs/1505.04597>>. Citado 2 vezes nas páginas 12 e 20.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *nature*, Nature Publishing Group UK London, v. 323, n. 6088, p. 533–536, 1986. Citado na página 17.
- SALEHI, S. S. M.; ERDOGMUS, D.; GHOLIPOUR, A. Tversky loss function for image segmentation using 3d fully convolutional deep networks. In: SPRINGER. *International workshop on machine learning in medical imaging*. [S.l.], 2017. p. 379–387. Citado 2 vezes nas páginas 19 e 20.
- SHIT, S.; PAETZOLD, J. C.; SEKUBOYINA, A.; ZHYLKA, A.; EZHOV, I.; UNGER, A.; PLUIM, J. P. W.; TETTEH, G.; MENZE, B. H. cldice - a topology-preserving loss function for tubular structure segmentation. *CoRR*, abs/2003.07311, 2020. Disponível em: <<https://arxiv.org/abs/2003.07311>>. Citado na página 19.
- SILVA, M. V. da; SANTOS, N. de C.; OUELLETTE, J.; LACOSTE, B.; COMIN, C. H. *A new dataset for measuring the performance of blood vessel segmentation methods under distribution shifts*. 2024. Disponível em: <<https://arxiv.org/abs/2301.04517>>. Citado na página 16.
- SONG, D.; HUANG, W.; LIU, J.; ISLAM, M. J.; YANG, H.; WANG, S. *Optimized Vessel Segmentation: A Structure-Agnostic Approach with Small Vessel Enhancement and Morphological Correction*. 2024. Disponível em: <<https://arxiv.org/abs/2411.15251>>. Citado na página 22.
- SOOMRO, T. A.; AFIFI, A. J.; ZHENG, L.; SOOMRO, S.; GAO, J.; HELLWICH, O.; PAUL, M. Deep learning models for retinal blood vessels segmentation: a review. *IEEE Access*, IEEE, v. 7, p. 71696–71717, 2019. Citado na página 12.
- WANG, M.; DING, H.; LIEW, J. H.; LIU, J.; ZHAO, Y.; WEI, Y. *SegRefiner: Towards Model-Agnostic Segmentation Refinement with Discrete Diffusion Process*. 2023. Disponível em: <<https://arxiv.org/abs/2312.12425>>. Citado na página 21.