

# Deep Guidance Network for Biomedical Image Segmentation: O que é preciso falar

**Abstract—** A segmentação de imagens 2D é um problema fundamental para a análise de imagens biométricas. Arquitetura mais utilizada para segmentação de imagens biomédicas é a U-Net. No entanto para algumas tarefas tal como a segmentação de vasos da retina, depois de fazer diversas operações de convolução informações espaciais da imagem são perdidas e não podem ser totalmente recuperados. Nesse trabalho é apresentada uma rede de aprendizado profunda para segmentar imagens biométricas utilizando um filtro de imagens para restaurar a informação dessas estruturas.

## I. INTRODUÇÃO

A performance de redes neurais profundas especialmente do tipo convolucionais é maior do que outros tipos de inteligências computacionais em determinadas tarefas. E o sucesso das CNNs é normalmente por causa do seu maior nível de abstração. Para imagens de segmentação biomédicas a arquitetura mais popular é a U-NET. Ela apresenta grande performance para segmentação de imagens biométricas tais como para o disco óptico e segmentação de copo óptico e segmentação de vasos sanguíneos, segmentação de tumores.

Nesse paper o foco é a segmentação de disco ópticos e vasos usando imagens de retinografia. Essas imagens consistem em imagens de vasos sanguíneos e imagens do disco óptico e do copo ocular. Os vasos da retina da retinografia ajudam médicos a diagnosticar diversas doenças na retina tal como retinopatia diabética, retinopatia hipertensiva e retinopatia solar. O formato do disco ocular e do copo ocular podem diagnosticar glaucoma que é uma das causas de cegueira. A detecção prévia e tratamento do glaucoma normalmente protegem os olhos contra perda de visão.

No entanto, existem diversos desafios para a segmentação de imagens de retinografia devido à baixa qualidade de imagens em função de outras doenças no globo ocular. Assim filtros tais como filtros bilaterais e o filtros guiados são importantes para o processamento de imagens médicas. Esses filtros podem resultar em imagens mais estruturadas e menos borradas. Isso ajuda a melhorar a performance para a segmentação de disco óptico copo ocular e segmentação de veias.

Nesse trabalho foi criado um filtro para restaurar a perda de informações parciais devido às operações de diminuição da resolução, também conhecido como downsampling, e introduz mais informação sobre a estrutura dentro da rede neural. Além disso, o filtro reduz a quantidade de ruído na imagem.

## II. TRABALHOS RELACIONADOS

Métodos de segmentação de copo ocular e disco ocular já foram pesquisados de forma extensiva nas últimas décadas. No entanto, esses modelos não tem poder suficiente para discriminar condições de imageamento e patologias mais complexas, o que afeta a sua performance. Técnicas de aprendizado profundo se tornaram o método popular para análise de imagens médicas, em

particular redes convolucionais mostram uma grande performance para tarefas como segmentação e classificação.

Mais recentemente, métodos de aprendizado profundo foram adotados para realizar segmentações de vaso com resultados muito bons. Comparados com outros métodos, o modelo desse trabalho introduz um filtro que preserva mais informações da imagem inicial.

## III. METODOLOGIA

Nesse trabalho foi criado um filtro de imagens que preserva mais informações da imagem inicial e reduz o ruído de cada imagem foi também adotado uma imagem parcial para cada caminho de da U-Net.

**Network Architecture:** arquitetura da rede consiste em uma rede de formato de U assim como o filtro e saídas individuais. O input é uma imagem RGB.

**U-shape net:** arquitetura da U-NET original é uma rede neural convolucional muito eficiente para segmentação de imagens biomédicas. Cada caminho executa convoluções com uma função de ativação RELU que produz feature maps.

**Multi scale guided filter module:** a operação de diminuição de resolução, o downsample, na U-NET resulta na perda de informação espacial especialmente para vasos muito pequenos na retinografia. Essa informação no fim não é restaurada por métodos comuns. Para resolver esse problema foi criado um filtro que preserva essa informação de imagens originais em tons de cinza. O filtro captura uma imagem guia como input e como output ele gera uma imagem que contém a informação de estrutura da imagem original guia.

Esse filtro atua em toda operação de downsample e upsample de forma a restaurar a perda da informação espacial causada pelas pela operação de downsample além disso novamente o filtro reduz o efeito do ruído.

Saídas Laterais: cada operação downsample gera um resultado para a rede. A função loss para cada camada de saída é dado como equação 1. Onde  $y_i$  é a probabilidade de valor prevista para a classe  $i$  e  $y$  é a probabilidade real para esta classe. Neste trabalho são usados quatro saídas laterais e uma saída resultante das médias das quatro saídas anteriores para combinação. A função perda final é a soma dessas cinco funções perdas individuais. O classificador final trata cada segmentação como uma classificação por Pixel para produzir um mapa de probabilidade para cada pixel.

O input para o filtro é uma imagem original  $F$  e uma imagem de guia  $G$ . Essas imagens são passadas então por um filtro de média, um filtro de modelo linear, novamente por um filtro de média, e então finalmente linearizados em um vetor  $O$ . Então o vetor  $O$  é concatenado com  $F$  original para construir uma representação de características mais poderosa. Então, essas

características são processadas por uma camada de convolução para fundir as características e mudanças dos canais da imagem original. Novamente este filtro restaura a informação perdida causada pela operação de mudança de resolução.

#### IV. EXPERIMENTOS

O modelo foi avaliado em duas tarefas a segmentação de disco óptico e copo ocular e a segmentação de veias na retina.

Os experimentos foram feitos com quatro datasets: Drive, ChaseDB1, ORIGA e REFUGE cada um com as suas determinadas características de imagem e tamanho.

Para a segmentação de retina a rede foi treinada para 150 épocas com a taxa de aprendizado de 0.001 com decaimento de 10 vezes a cada 50 épocas. Toda a rede neural foi treinada do zero não usando nenhum outro dado de treino adicional.

##### OD e CS

**Método de avaliação:** para avaliar o método de segmentação de disco óptico e copo óptico foi adotado um erro de overlapping.

**Performance do experimento:** foi comparado a rede deste trabalho com outros modelos no estado da arte. A performance para o caso de disco óptico e copo óptico tem performance bem maior que a U-NET original. A figura 4 mostra o resultado do modelo deste trabalho comparado com a U-NET. O modelo pode gerar resultados bons até mesmo para imagens de baixa qualidade tal como mostrado nas duas últimas linhas da figura 4. A tabela 3 mostra a performance usando dataset. O método também tem uma performance bem melhor do que U-Net original.

##### Segmentação de veias da retina

**Critério de avaliação:** o processo de segmentação de veias é uma classificação baseada em pixels em que cada Pixel é classificado como veia ou como tecido ao redor. Foram usados os parâmetros **especificidade (SPE)**, **sensibilidade (SEN)**, **acurácia (ACC)**, e a **área sobre ROC (AUC)**.

**SEN** - Sensibilidade é a habilidade para detectar pixels de vasos sanguíneos.

**SPE** – Especificidade é a habilidade para detectar pixels de tecidos ao redor dos vasos.

**ACC** - Acurácia é medida pela razão do número de pixels classificados corretamente em comparação com o número de pixels da imagem total.

Nas equações mostradas TP representa os verdadeiros positivos onde um pixel é identificado tanto na imagem segmentado quanto no ground truth, TN representa um verdadeiro positivo onde um tecido ao redor é corretamente classificado na imagem de segmentação.

A curva ROC mostra a fração de pixels corretamente classificados versus a fração de pixels de tecidos ao redor falsamente classificados. Quanto mais rápido a curva ROC alcança o topo melhor é a performance do sistema. O valor da área da curva é usado como performance extraída da curva ROC onde o valor 1 seria um sistema ótimo.

**Performance do experimento:** foi comparado o resultado desta rede com outros modelos do estado da arte. Para os datasets DRIVE e ChaseDB1 essa rede teve resultados melhores que métodos estado da arte e obteve o melhor valor de ACC AOC e SPE. Este experimento mostra que esse método pode efetivamente identificar pixels de veias. Além disso ele tem uma performance superior ao U-NET original.

A tabela 8 mostra a mudança de performance ao adicionar cada módulo a rede. Além disso nota-se que o filtro aumenta o parâmetro AUC de uma grande margem.

#### V. CONCLUSÃO

Neste trabalho é proposto uma rede neural profunda com filtro para segmentação de veias e disco óptico e copos oculares essa rede incorpora um filtro que introduz mais informações para a imagem de guia para a rede neural e restaura informação espacial perdida nas operações de downsample.