

Detecção de Estruturas em Imagens Fundo de Olho para o Diagnóstico de Patologias da Visão

Lucas Cunha de Carvalho, Geraldo Braz Junior

¹Centro de Ciências Exatas e Tecnológicas - Campus Dom Delgado - UFMA.

²Laboratório de Visão e Processamento de Imagens - Núcleo de Computação Aplicada.

Abstract. *Early diagnosis of pathologies such as diabetic retinopathy and glaucoma is essential for treatment, which must be immediate due to the risk of blindness. The analysis of fundus eye structures is essential in detecting diseases and establishing the stage and severity of the patient. This work aims to develop image segmentation techniques using U-net neural networks, seeking to segment images of the fundus of the eye, the optic disc region and blood veins, favoring the automation and analysis of such pathologies. In this work, the models obtained results of 92% for the segmented optic disc region and 49% for regions of reached blood veins.*

Key Words: *Fundus eye image, U-Net, Deep Learning.*

Resumo. *O diagnóstico precoce de patologias como a retinopatia diabética e glaucoma é fundamental para o tratamento que deve ser imediato por conta do risco de cegueira. A análise de estruturas do fundo de olho são essenciais na detecção de doenças e no estabelecimento do estágio e severidade na qual se encontra o paciente. Este trabalho visa desenvolver técnicas de segmentação de imagens utilizando redes neurais do tipo U-net, buscando segmentar de imagens de fundo de olho, a região do disco óptico e veias sanguíneas, favorecendo a automatização e análises de tais patologias. Neste trabalho, os modelos obtiveram resultados de 92% da região do disco óptico segmentada e 49% para regiões de veias sanguíneas alcançadas.*

Palavras-chave: *Imagem de Fundo de Olho, U-Net, Aprendizagem Profundo.*

1. Introdução

Doenças relacionadas á visão como a retinopatia diabética (DR) e o glaucoma afetam pessoas por todo o mundo. A detecção precoce de tais patologias se faz necessária para prognósticos positivos, entretanto, tal trabalho demanda tempo, esforço intensivo e profissionais capacitados, tais variáveis representam desafios ao diagnóstico. Esta problemática se agrava à medida em que as prevalências destas doenças crescem de forma contínua ao longo dos anos, por consequência, aumentando a demanda por serviços de triagem.

Tradicionalmente, imagens de fundo de olho são utilizadas para detectar e acompanhar a evolução de patologias, para tal, faz-se necessário o reconhecimento de diferentes estruturas afetadas como o disco óptico (DO) e veias sanguíneas da retina, segmentos dos olhos que por vezes possuem lesões e alterações que configuram informações e padrões necessários ao diagnóstico, com isto, a automatização do processo de detecção de tais estruturas utilizando técnicas de deep learning se mostra uma ferramenta viável ao auxílio da detecção precoce de patologias.

No trabalho de [Sevastopolsky 2017], é desenvolvida uma metodologia em que objetiva a segmentação da região do disco óptico, onde é introduzida uma arquitetura de rede neural de convolução do tipo U-net. Vale destacar também, métodos de pré-processamento de imagens utilizado, o método CLAHE, que consiste em uma técnica de melhoramento de contraste local em que é efetuada uma alteração para tornar mais visível estruturas de fundo de olho tornando mais visíveis regiões de interesse [Zuiderveld 1994].

As bases de dados utilizadas pelo artigo em questão foram a Drions-DB, Drishti-GS e Rim-One. O método proposto obteve resultados que demonstraram uma qualidade competitiva à trabalhos anteriores, efetuando a segmentação em um tempo de predição e treinamento inferior além de otimizar parâmetros de forma mais eficiente.

Em [Joshua et al. 2020], os autores propõem uma rede neural convolucional do tipo U-net modificada afim de segmentar veias sanguíneas de imagens de fundo de olho, utilizando também métodos simples de pré-processamento de imagens para melhoramento do contraste de imagens. A metodologia obteve score de estado de arte em comparação a trabalhos relacionados, utilizando uma U-net otimizada de forma simples e mais acessível.

[Shankar et al. 2020] propõe um modelo automático de otimização de parâmetros usando uma rede neural convolucional inception v4 afim de classificar retinopatia diabética em imagens de fundo de olho da base Messidor, além de utilizar o método CLAHE como pré-processamento de imagens [Zuiderveld 1994], a metodologia proposta utiliza métodos bayesianos de otimização de parâmetros no modelo proposto.

Métodos de otimização como a utilização de bibliotecas como a Keras Tuner [O'Malley et al. 2019] são ferramentas importantes na construção de uma rede neural mais eficiente e adaptada ao cenário em que a mesma será utilizada.

Finalmente, [Fu et al. 2018] propõe um sistema multinível de classificação de imagens objetivando a detecção de glaucoma, tal arquitetura envolve uma rede neural U-net de segmentação do disco óptico e redes de diferentes configurações de arquiteturas de redes neurais convolucionais afim de extrair características, tendo como principal objetivo, unir tanto qualidades de resultados extraídos a partir de imagens com alta resolução, o que facilita a extração de características globais, quanto de resultados da imagem de fundo de olho com resolução reduzida, facilitando o manuseamento computacional e preservando características locais. Esta rede, batizada de Disc-Aware Ensemble Network (DE-net), é composta por 4 diferentes fluxos: Fluxo de classificação da imagem como um todo, fluxo de segmentação da região do disco óptico (DO) e dois fluxos de classificação da região do DO.

O método proposto pôde incorporar métodos de deep learning, a partir da união de múltiplos níveis de redes neurais, a arquitetura De-Net produz resultados validados de forma hierárquica, solidificando os resultados finais de classificação do glaucoma em imagens de fundo de olho e performando de forma superior comparado a outros algoritmos de detecção desta patologia, ajudando a demonstrar a importância da análise de estruturas de fundo de olho no diagnóstico.

Nesse contexto, esta pesquisa tem em vista a contribuição na automatização do diagnóstico de doenças como a retinopatia diabética e o glaucoma, por meio da construção de redes neurais de convolução do modelo U-net que realizarão a segmentação do disco

óptico e veias sanguíneas em imagens de fundo de olho, viabilizando assim, a futura classificação e diagnóstico de doenças relacionadas à estas estruturas por meio de métodos de deep learning.

2. Fundamentação Teórica

Com o envelhecimento da população, doenças como o glaucoma e diabetes naturalmente se tornarão mais comuns e evidentes com o passar dos anos. Utilizando fotografias do disco óptico (Figura 1), sistemas de deep learning aparentaram ser capazes de diagnosticar glaucoma e outras doenças com maior acurácia que seres humanos. Eliminando a necessidade de profissionais para classificação de imagens de fundo de olho, consequentemente reduzirá de forma drástica, o custo, esforço e tempo necessário para detecção de tais patologias. Estudos na China e Índia, indicam que a utilização de tais técnicas possuem um ótimo custo-benefício.[Tan Nicholas Y.Q.a; Friedman David S.b 2020]

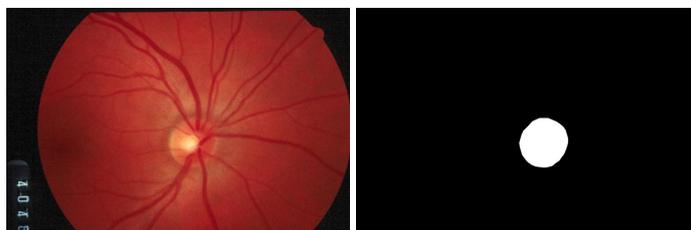


Figura 1. Respectivamente, imagem e máscara do disco óptico da base Drions.

A metodologia para a detecção de doenças como a retinopatia diabética e glaucoma se inicia com a fundamental detecção das estruturas do fundo de olho realizando a segmentação das mesmas, sendo esta podendo ser efetuada de forma supervisionada, ou seja, por meio de máscaras ou anotações das regiões de interesse, ou sem supervisão, além da efetuação de extração de características, para classificação.

Um dos maiores desafios de inteligências artificiais projetadas para aplicações do mundo real é a influência da variação dos mais diversos fatores que podem diferenciar a interpretação de cada dado a ser observado, como exemplo disso, temos imagens noturnas que a depender da iluminação, podem ser tomadas por pixels pretos ou, aproximando do estudo deste trabalho, uma imagem de fundo de olho muito lesionada, obstruindo a visão de estruturas como o disco óptico, tais fatores provocam ruídos e elevam o nível de abstração, dificultando a extração de características.

Para se obter o poderio de reduzir determinados problemas de abstração, pode-se introduzir representações de imagens que são expressadas a partir de versões mais simples das mesmas, ou seja, podemos identificar pessoas em imagens à partir do contornos de seus corpos, assim como podemos detectar estruturas de fundo de olho à partir de suas anotações, configurando um aprendizado supervisionado (Figura 1) [Goodfellow et al. 2016].

3. Metodologia

Esta seção apresenta a metodologia proposta para a realização da segmentação de estruturas de fundo de olho. O objetivo consiste na detecção do disco óptico e veias

sanguíneas afim de facilitar posteriores processos de diagnóstico de patologias como retinopatia diabética e glaucoma. Para tanto, a metodologia desenvolvida explora o uso de redes neurais U-Net realizando a otimização automática de parâmetros da mesma. [O'Malley et al. 2019] A Figura 2 exemplifica a sequência de passos tomados neste estudo, explicado em detalhes nas subseções seguintes.

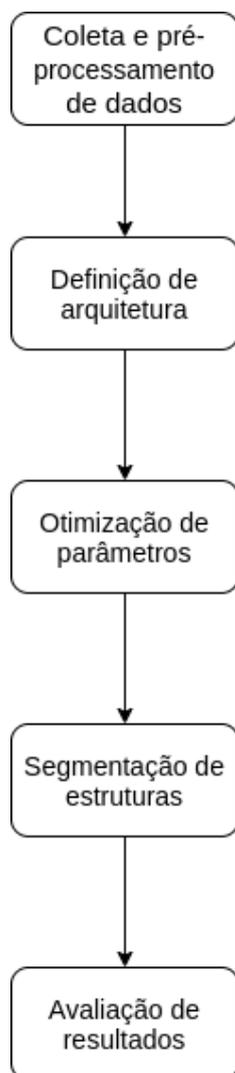


Figura 2. Fluxograma da metodologia proposta.

3.1. Coleta e Pré-Processamento

Primeiramente, efetuou-se uma coleta de dados, sendo estes, imagens de fundo de olho e suas respectivas máscaras do disco óptico (DO) ou veias sanguíneas (VS). Para a segmentação do disco óptico, coletou-se imagens pertencentes às bases de dados Drishti e Drions, totalizando 209 imagens. Para segmentação de veias sanguíneas, foram utilizadas imagens das bases Drive, STARE e HRF (Figura 3), totalizando 85 imagens de fundo de olho.

Para a etapa de segmentação de veias sanguíneas, foi aplicado o método de pré processamento CLAHE, que consiste em uma técnica de melhoramento de contraste local

que busca tornar sobressair estruturas do fundo de olho, tornando mais visíveis regiões de interesse. [Zuiderveld 1994]

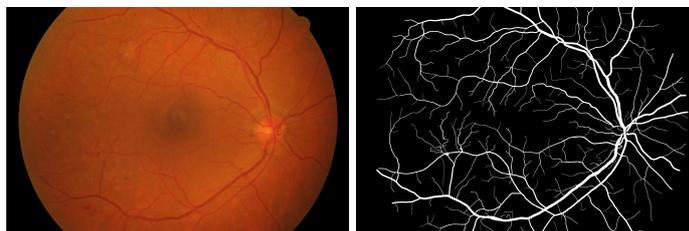


Imagem de fundo de olho. Sua respectiva máscara.

Figura 3. Imagens da base HRF

Para esta e outras etapas posteriores, as bases de dados foram divididas em 70% para treinamento, 10% para validação e 20% para teste, mantendo a proporção previamente existente entre as classes.

3.2. Definição de arquitetura

Neste trabalho, exploramos o uso de redes neurais de arquitetura U-Net para segmentação de estruturas do olho. A rede consiste em uma série de convoluções em bloco, divididas entre encoder, decoder e gargalo, formando um característico formato de "U" (Figura 4).

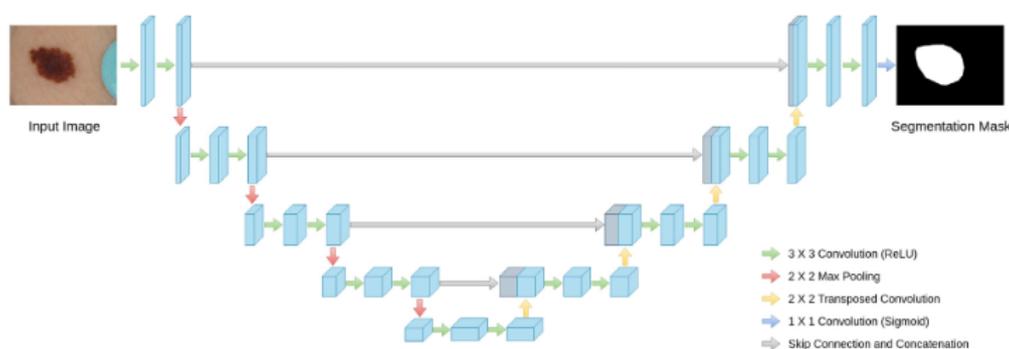


Figura 4. Exemplo de arquitetura de CNN do tipo U-net.

Na região do encoder são efetuadas convoluções de forma a reduzir gradualmente o tamanho da imagem, nos chamados "down-samplings", o input passa por filtros de convoluções e max pooling, até que o mesmo passe pelo processamento até a região do gargalo. O encoder é responsável por extrair informações espaciais do input, após isso, o input é levado ao gargalo, onde a rede compreende informações e as guarda para a construção do mapa de características a ser formado.

Na região do decoder, a imagem passa por uma transformação de forma a recuperar sua resolução gradualmente, construindo um mapa de segmentação a partir das informações extraídas pelo encoder, tal processo é realizado por filtros de convolução que a cada estágio dos "up-samplings" são reduzidos pela metade, além de ser efetuada a concatenação dos outputs de cada estágio do decoder com respectivos outputs da região do encoder, formando um mapa de características a ser propagado por toda a região do decoder até a convolução final onde será utilizada a função de ativação [NabilIbtehaza 2020].

O motivo da utilização da U-net neste estudo, consiste no fato de que essas redes são capazes de generalização mesmo com poucos dados de treinamento e na grande capacidade da mesma de mapeamento de imagens, algo essencial na segmentação.

3.3. Otimização de parâmetros

Afim de encontrar melhores resultados de segmentação, foram efetuadas séries de otimizações buscando parâmetros ideais, nesta fase, diferentes combinações de arquiteturas, otimizadores e funções perda foram testados utilizando keras tuner [O'Malley et al. 2019]. Estima-se que mais de 1500 combinações foram testadas na aprendizagem da segmentação de estruturas de fundo de olho (com opções listadas abaixo).

- Arquiteturas avaliadas: 'vgg16', 'vgg19', 'resnet18', 'resnet34', 'resnet50', 'resnet101', 'resnet152', 'seresnet18', 'seresnet34', 'seresnet50', 'seresnet101', 'seresnet152', 'resnext50', 'resnext101', 'seresnext50', 'seresnext101', 'senet154', 'densenet121', 'densenet169', 'densenet201', 'inceptionv3', 'inceptionresnetv2', 'mobilenet', 'mobilenetv2', 'efficientnetb0', 'efficientnetb1', 'efficientnetb2', 'efficientnetb3', 'efficientnetb4', 'efficientnetb5', 'efficientnetb6' e 'efficientnetb7'.
- Otimizadores: 'SGD', 'Adam', 'Adadelta', 'Adagrad' e 'Adamax'.
- Funções Perda: 'Binary Cross Entropy dice loss', 'Binary crossentropy', 'Tversky', 'Iou' (Jaccard), 'Dice (f1)', 'Categorical Cross Entropy loss', 'Binary Focal loss', 'Categorical Focal Loss', 'Categorical Cross Entropy Dice', 'Categorical Cross Entropy Jaccard', 'Binary Focal Dice', 'Binary Focal Jaccard', 'Categorical Focal Dice Loss' e 'Categorical Focal Jaccard Loss'.

Na avaliação dos resultados da parametrização, foi utilizada a métrica IoU, que consiste em uma divisão entre a área das regiões de interesse retratadas na máscara pela área segmentada, quanto mais próximo de 1, mais precisa a segmentação.

3.4. Segmentação e Avaliação

Efetuada a hiperotimização de parâmetros, partiu-se para a etapa de detecção de estruturas de fundo de olho. Utilizando os modelos U-net otimizados com os parâmetros ideais, realizou-se a segmentação das estruturas do disco óptico e veias sanguíneas em imagens de fundo de olho das bases organizadas anteriormente.

Os testes da segmentação foram realizados usando a base *Ocular Disease Recognition* (ODIR) é composta de 8000 imagens de fundo de olho (Figura 5), dividida em 8 diferentes classes que representam diferentes patologias relacionadas ao fundo de olho (Tabela 1), entre elas, glaucoma, diabetes e pacientes saudáveis.

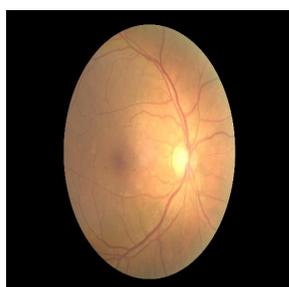
Para este estudo apenas as classes normal, diabetes e glaucoma foram utilizadas para a segmentação.

4. Resultados

Neste trabalho, as principais métricas analisada nos resultados de segmentação foram o IoU score, AUC (Area Under Curve) e o Dice Score nas redes U-net apresentadas. Os testes focaram-se em otimizar ao máximo tais métricas e as imagens segmentadas resultantes, para capacitá-las a serem utilizadas futuramente.

Tabela 1. ODIR - classes de imagens

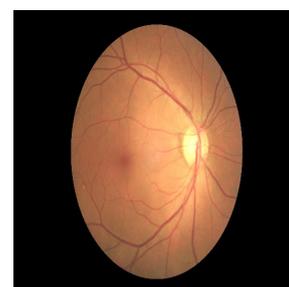
Classe	Número de imagens
Fundo de olho normal	2280
Diabetes	2256
Glaucoma	430
Catarata	424
Degeneração	328
Hipertensão	206
Miopia	348
Outras doenças	1958



Classe diabetes.



Classe glaucoma.



Classe normal.

Figura 5. Exemplos de imagens da base Odir

4.1. Parametrização

Após a etapa inicial de segmentação, realizou-se a etapa de otimização, onde foram encontradas, a partir da análise de diferentes combinações de parâmetros, redes mais capazes de alcançar resultados.

Como frutos desta parametrização temos a rede U-net com arquitetura 'inceptionresnetv2', otimizador 'Adam' e função perda 'Binary Cross Entropy dice loss' encontrada como ideal para a segmentação do disco óptico e a rede U-net com arquitetura 'ResNet152', otimizador 'Adam' e função perda 'Binary focal dice' encontrada para a segmentação de veias sanguíneas.

A resolução das imagens é um dos fatores que influenciam a segmentação final, enquanto resolução menores economizam recursos computacionais, resoluções maiores fornecem mais informações, por este motivo, a resolução das imagens foi aumentada para 256x256, além da adição do método de pré-processamento CLAHE para a segmentação de veias sanguíneas. Tais modelos U-net com parâmetros encontrados à partir da etapa de otimização, os modelos foram treinados por 480 épocas cada.

A Tabela 2 mostra os resultados após o predict efetuado pelas redes otimizadas para segmentar a região do disco óptico e veias sanguíneas.

Pode-se perceber que por conta da otimização de parâmetros, houve um notável avanço na capacidade de aprendizagem dos modelos, além da comprovação da eficiência de um treinamento em épocas reduzidas, o que economiza recursos computacionais.

Como comparação da efetividade da otimização, a tabela 3 nos mostra resulta-

Tabela 2. Segmentação de estruturas de fundo de olho

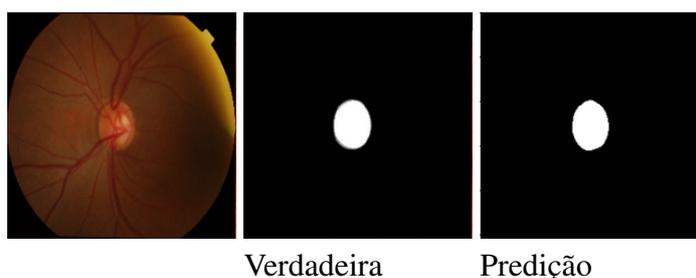
Métrica	Disco óptico	Veias Sanguíneas
IoU (média)	0,92	0,49
AUC (média)	0,98	0,91
Dice (média)	0,89	0,56

dos de treinamentos prévios à etapa de parametrização da rede U-net, efetuados por 160 épocas em dois modelos de rede neural de convolução U-net, formada por camadas de no máximo 128x128 como resolução de input, para este primeiro momento, apenas a métrica IoU score foi utilizada para análise de resultados.

Tabela 3. Resultados da etapa inicial de segmentação de estruturas

	Disco óptico	Veias Sanguíneas
IoU (média)	0,79	0,25

A Figura 6 apresenta a precisão de reconhecimento da região de interesse do disco óptico (DO) das imagens que os modelos adquiriram após a otimização do treinamento.

**Figura 6. Exemplos de resultado na segmentação do disco ótico**

Pode-se perceber que a predição efetuada pelo modelo treinado configura à região verdadeira do disco óptico, tendo poucas diferenças notadas visualmente, o que demonstra que o modelo obteve a capacidade de detectar e segmentar a região do disco óptico.

A Figura 7 pode ser utilizada para demonstrar a precisão adquirida pelo modelo no reconhecimento da região de interesse das veias sanguíneas (VS).

Percebe-se que o modelo foi capaz de segmentar com sucesso de imagens de fundo de olho a grande maioria das características referentes a veias sanguíneas visíveis à olho nu, entretanto, não foi capaz de alcançar a totalidade de veias da anotação, pois muitas delas não se encontram totalmente visíveis.

4.2. Segmentação de imagens da base Odir

Utilizando os modelos treinados na etapa de segmentação anterior, realizou-se a detecção e predição de estruturas de fundo de olho em uma nova base, a ODIR. Foram obtidos diversos resultados onde foi possível separar as determinadas regiões de interesse do disco óptico ou veias sanguíneas da imagem.

A Figura 8 mostra os resultados obtidos após a segmentação da região do disco óptico (DO) e veias sanguíneas (VS) de uma mesma imagem de fundo de olho.

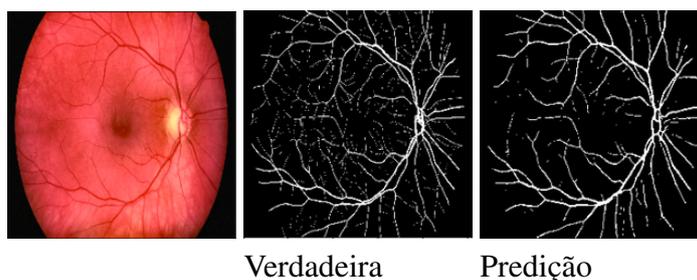


Figura 7. Exemplo de resultado na segmentação dos vasos sanguíneos



Figura 8. Segmentação da região do DO e Veias Sanguíneas em uma imagem de fundo de olho.

Pode-se perceber que os modelos treinados puderam reconhecer e segmentar com precisão as respectivas estruturas do disco óptico e veias sanguíneas, entretanto, em uma minoria de imagens da base ODIR, não foi possível ser efetuada a segmentação com tamanha qualidade, isso se dá devido à imagens com demasiados ruídos, o que é natural em pacientes com o fundo de olho severamente lesionado, o que torna impossível a visualização de estruturas.



Figura 9. Segmentação da região do DO e Veias Sanguíneas em uma imagem de fundo de olho com ruídos.

No caso da figura 9, foi possível extrair a região do DO e algumas regiões de veias sanguíneas, além de outras informações da imagens que foram obtidas pelo modelo por conta da demasiada distorção da imagem de fundo de olho causada pelas lesões do paciente.

5. Conclusão

Analisando os resultados numéricos e as imagens, conclui-se que as redes neurais do tipo U-net utilizadas neste estudo foram capazes de aprender a detectar as regiões do disco óptico e vasos sanguíneos em imagens de fundo de olho, efetuando a segmentação de forma precisa na grande maioria de imagens.

Percebe-se que um dos empecilhos fundamentais para a segmentação de tais estruturas se dá por conta de graus mais elevados de severidade da doença, que podem provocar distorções na imagem devido à um fundo de olho lesionado e por muitas vezes com suas estruturas deterioradas.

A otimização de parâmetros e a etapa de pré-processamento mostraram-se eficientes, aumentando drasticamente o IoU quando tomamos como comparação a etapa inicial, demonstrando a eficiência do modelo em aprender as regiões de interesse.

Os resultados deste trabalho se mostram competitivos à trabalhos anteriores, mesmo propondo uma rede neural mais otimizada em termos de parâmetros, buscando economia computacional, os modelos obtiveram IoU de segmentação do disco óptico superior à [Sevastopolsky 2017] e Auc de segmentação de vasos sanguíneos superior à [Joshua et al. 2020].

De forma eficiente, pôde-se treinar um modelo de rede neural convolucional do tipo U-net capaz de segmentar estruturas de fundo de olho fundamentais no diagnóstico de patologias como a retinopatia diabética e glaucoma, abrindo espaço também para a utilização de tais imagens segmentadas como dados para futuras aplicações ou para análise médica, facilitando o trabalho de profissionais da saúde e reduzindo custos.

Referências

- Fu, H., Cheng, J., Xu, Y., Zhang, C., Wong, D. W. K., Liu, J., and Cao, X. (2018). Disc-aware ensemble network for glaucoma screening from fundus image. *IEEE Transactions on Medical Imaging*, 37(11):2493–2501.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*.
- Joshua, A. O., Nelwamondo, F. V., and Mabuza-Hocquet, G. (2020). Blood vessel segmentation from fundus images using modified u-net convolutional neural network. *Journal of Image and Graphics*, 8(1):21–25.
- Nabilbtehaza, M. (2020). Multiresunet : Rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural Networks*.
- O'Malley, T., Bursztein, E., Long, J., Chollet, F., Jin, H., Invernizzi, L., et al. (2019). Keras tuner. <https://github.com/keras-team/keras-tuner>.
- Sevastopolsky, A. (2017). Optic disc and cup segmentation methods for glaucoma detection with modification of u-net convolutional neural network. *Pattern Recognit. Image Anal.*, 27:618–624.
- Shankar, K., Zhang, Y., Liu, Y., Wu, L., and Chen, C.-H. (2020). Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification. *IEEE Access*, 8:118164–118173.
- Tan Nicholas Y.Q.a; Friedman David S.b, c; Stalmans Ingeborgd; Ahmed Iqbal Ike K.e; Sng, C. C. f. (2020). Glaucoma screening: where are we and where do we need to go? *Current Opinion in Ophthalmology*.
- Zuiderveld, K. (1994). *Contrast Limited Adaptive Histogram Equalization*, page 474–485. Academic Press Professional, Inc., USA.