

Desenvolvimento de Abordagens Baseadas em Redes Neurais Profundas para Detecção e Segmentação de Instância de Lesões Retinianas

Carlos Santos¹, Marilton Sanchotene de Aguiar^{2,3}, Daniel Welfer³

¹Instituto Federal de Educação, Ciência e Tecnologia Farroupilha (IFFar)
Alegrete – RS, Brasil

²Programa de Pós-Graduação em Computação, Universidade Federal de Pelotas (UFPEL)
Pelotas – RS, Brasil

³Departamento de Computação Aplicada – Universidade Federal de Santa Maria (UFSM)
Santa Maria – RS, Brasil

carlos.santos@iffarroupilha.edu.br, marilton@inf.ufpel.edu.br

daniel.welfer@ufsm.br

Abstract. *Diabetic Retinopathy (DR) is one of the leading causes of vision loss and presents fundus lesions in its initial stages, such as microaneurysms, hemorrhages, hard exudates, and soft exudates. Computational models capable of detecting these lesions can support the early diagnosis of the disease and prevent the manifestation of more severe forms of lesions, helping in the screening process and definition of the best form of treatment. However, the detection of microlesions using computational systems is a challenge due to several factors, such as the size and shape of these lesions, the presence of noise and poor contrast in the images, the small number of labeled examples in public DR datasets, and the difficulty of deep learning algorithms in detecting tiny objects due to gradient dissipation during training. Thus, to overcome these problems, this work proposes two new approaches based on image processing techniques, data augmentation, transfer learning, and deep neural networks to support the medical diagnosis of fundus lesions. We trained, adjusted, and evaluated the proposed approaches using different public Diabetic Retinopathy datasets. We partitioned the datasets into sets of training (50%), validation (20%), and test (30%) to carry out the experiments. We used a validation step to fine-tune the hyperparameters and a test step to assess the generalization capacity of the models. The approach to detecting fundus lesions achieved mAP of 0.2630 for the limit of IoU of 0.5 in the validation step using the DDR dataset and Adam optimizer. The approach for segmenting instances of fundus lesions reached mAP of 0.2903 for the limit of IoU of 0.5 in the validation stage using the DDR dataset and Adam optimizer, thus being 10.38% more accurate than the proposed detection approach. The results obtained in the experiments demonstrate that the new approaches presented promising results in detecting fundus lesions associated with DR.*

Resumo. *A Retinopatia Diabética (RD) é uma das principais causas de perda de visão e apresenta em suas fases iniciais lesões de fundo, como microaneu-*

rismas, hemorragias e exsudatos duros e algodinosos. Modelos computacionais capazes de detectar essas lesões podem auxiliar no diagnóstico precoce da doença e prevenir a manifestação de formas mais graves de lesões, auxiliando também no processo de triagem e definição da melhor forma de tratamento. Entretanto, a detecção de microlesões por meio de sistemas computacionais é um desafio por inúmeros fatores, como o tamanho e formato destas lesões, a presença de ruído e contraste ruim das imagens, a pequena quantidade de exemplos rotulados nos conjuntos de dados públicos de RD, e a dificuldade de algoritmos de aprendizado profundo em detectar objetos muito pequenos em função da dissipação de gradiente durante o treinamento. Assim, para contornar estes problemas, este trabalho propõe duas novas abordagens baseadas em técnicas de processamento de imagens, aumento de dados, transferência de aprendizado e redes neurais profundas, com o propósito de auxiliar no diagnóstico médico de lesões de fundo. As abordagens propostas foram treinadas, ajustadas e avaliadas usando diferentes conjuntos de dados públicos de Retinopatia Diabética. Para a realização dos experimentos os datasets foram particionados em conjunto de treinamento (50%), validação (20%) e teste (30%). Utilizou-se uma etapa de validação para realizar o ajuste fino de hiperparâmetros, e uma etapa de teste para aferir a capacidade de generalização dos modelos. A abordagem para detecção das lesões de fundo alcançou mAP de 0,2630 para o limite de IoU de 0,5 na etapa de validação utilizando o conjunto de dados DDR e otimizador Adam. Já a abordagem para segmentação de instância das lesões de fundo alcançou mAP de 0,2903 para o limite de IoU de 0,5 na etapa de validação utilizando o conjunto de dados DDR e otimizador Adam, sendo, portanto, 10,38% mais preciso que a abordagem proposta para detecção. Os resultados obtidos nos experimentos demonstram que as novas abordagens apresentaram resultados promissores na detecção de lesões de fundo associadas à RD.

1. Caracterização do Problema e Motivação da Pesquisa

A Retinopatia Diabética (RD), uma doença diretamente relacionada ao Diabetes e que afeta os olhos, é uma das principais causas de perda de visão em adultos em idade produtiva. Aproximadamente 1/3 (34,6%) das pessoas com Diabetes nos Estados Unidos da América, Europa e Ásia têm RD. A prevalência do Diabetes está aumentando mundialmente e nas últimas duas décadas a perda parcial ou total da visão cresceu devido às complicações causadas pelo aumento do número de pessoas com Diabetes [Vocaturu and Zumpano 2020]. A Federação Internacional de Diabetes (em inglês, *International Diabetes Federation – IDF*)¹ relata que, em 2000, o número global estimado de adultos com Diabetes era de 151 milhões. Em 2009 havia crescido 88%, indo para 285 milhões. A IDF estima que haverá 600 milhões de pessoas com Diabetes em 2035 e 700 milhões em 2045.

A RD tem origem na lesão dos vasos sanguíneos do tecido sensível à luz da retina e trata-se da principal causa de perda de visão em pessoas com idade entre os 20 e 74 anos. No trabalho apresentado por [Ting et al. 2016] foi demonstrado que a perda de visão resultante de RD pode ser evitada quando tratada precocemente. No entanto, a triagem

¹<https://idf.org/>

realizada para identificação precoce da RD permanece um desafio, pois diabéticos são geralmente tratados em departamentos de endocrinologia de hospitais, nos quais costumam-se realizar exames de fundo do olho e este processo geralmente é demorado, pois apenas um número limitado de exames pode ser processado a cada dia. Além disso, o número de oftalmologistas não consegue atender às crescentes demandas ao redor do mundo, particularmente em regiões em desenvolvimento [Chakrabarti et al. 2012].

A RD geralmente é identificada por meio de exames oftalmológicos que visam identificar lesões na retina (lesões de fundo) incluindo Exsudatos Duros (EX), Exsudatos Algodonosos (SE), Microaneurismas (MA) e Hemorragias (HE) [Nayak et al. 2008], conforme ilustrado na Figura 1.

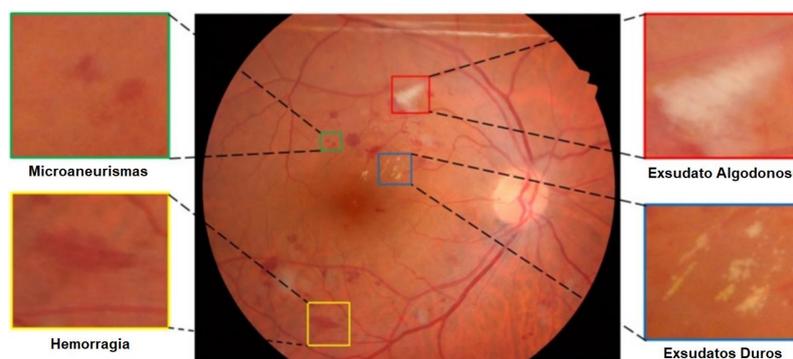


Figura 1. Exemplo de imagem de fundo com as lesões anotadas: Microaneurismas, Hemorragias, Exsudatos Algodonosos e Exsudatos Duros

Considerando o número de pacientes afetados por Diabetes no mundo, é importante destacar que para a realização de uma triagem efetiva a fim de verificar as pessoas afetadas pela RD envolve a mobilização de grande quantidade de recursos. Embora o método principal para avaliar a RD envolva a oftalmoscopia² [Wong et al. 2017], os grandes volumes de dados e o Aprendizado Profundo (em inglês, *Deep Learning* – DL) podem proporcionar soluções de baixo custo e eficazes para apoiar a realização de diagnósticos precoces, auxiliando o trabalho de profissionais de saúde e permitindo amplo acesso à população a exames para identificação desta doença.

Dessa forma, o desenvolvimento de ferramentas computacionais capazes de analisar e processar imagens de fundo pode auxiliar nesse processo de triagem e identificação precoce da doença. Devido ao significado da detecção de lesões em imagem de fundo e à complexidade associada à realização desta tarefa sem o auxílio de um sistema informatizado, muitos métodos automatizados baseados em *Deep Learning* para detectar estas lesões foram desenvolvidos, tais como as soluções apresentadas por [Li et al. 2019, Mateen et al. 2020, Wang et al. 2020, Porwal et al. 2020, Dai et al. 2021, Alyoubi et al. 2021, Santos et al. 2021, Shenavarmasouleh et al. 2021].

²Oftalmoscopia é o exame que visa observar a região posterior do globo ocular, que compreende a retina, o disco óptico e os vasos sanguíneos.

Entretanto, mesmo que redes neurais profundas tenham sido utilizadas para a detecção de lesões em imagens de fundo, ainda assim há limitações nos resultados obtidos, principalmente na detecção de microlesões retinianas, provenientes da baixa representatividade dos atributos extraídos das imagens utilizadas para o treinamento dos modelos e da complexidade associada às características como formato, tamanho e incidência destas lesões.

Em função destas limitações, da gravidade da RD e, também, do impacto desta doença na saúde e qualidade de vida das pessoas, é oportuno propor a criação e/ou otimização de procedimentos, abordagens e instrumentos computacionais capazes de prover um suporte rápido e preciso ao diagnóstico médico, com o mínimo de intervenção humana. Todos estes aspectos motivam o problema de pesquisa considerado nesta Tese de Doutorado.

2. Objetivos e Contribuições do Trabalho

O objetivo geral desta Tese é propor novas abordagens baseadas em técnicas de processamento digital de imagens (PDI) e aprendizado profundo para detectar e segmentar instâncias das lesões de fundo associadas à Retinopatia Diabética que possa alcançar um desempenho preditivo superior a alternativas similares apresentadas recentemente na literatura. Para tanto, no decorrer deste trabalho, foi realizado um estudo aprofundado sobre técnicas de PDI e aprendizado profundo. Com base nesse estudo, foram desenvolvidas duas abordagens principais para detectar e segmentar Exsudatos Duros (EX), Exsudatos Algodonosos (SE), Microaneurismas (MA) e Hemorragias (HE), sendo:

Abordagem para Detecção. Como primeira contribuição desta Tese, apresenta-se uma nova abordagem para a detecção de lesões de fundo. Foi implementado um bloco de pré-processamento para reduzir *outliers* e melhorar o realce, a fim de prover uma extração de características mais eficiente das lesões retinianas. Este bloco conta com um método para *Cropping* parcial do fundo preto das imagens de fundo, para minimizar a geração de falsos positivos, e também um método de *Tiling*, para aumentar o campo receptivo em torno das lesões e minimizar a perda de informação causada pela redução de resolução das imagens originais na camada de entrada do *Backbone* da arquitetura. Além disso, foi apresentada uma estrutura de rede neural convolucional baseada em uma arquitetura YOLO de última geração para melhorar a detecção das lesões de fundo e realizar inferências em tempo real sobre GPUs de baixo custo. A solução para a detecção das lesões de fundo foi treinada e avaliada utilizando dois conjuntos de dados públicos de Retinopatia Diabética: DDR e IDRiD. Particionou-se estes conjuntos de dados em conjunto de treinamento, validação e teste em uma proporção de 50:20:30, respectivamente.

Abordagem para Segmentação de Instância. Como segunda contribuição desta Tese, apresentou-se uma nova abordagem para segmentação de instância de lesões de fundo. Implementou-se uma estrutura de rede neural convolucional implementada com base em uma arquitetura Mask R-CNN para melhorar a precisão na detecção das lesões de fundo, principalmente na identificação das microlesões. A abordagem para segmentação de instância também contém um bloco de pré-processamento composto pelos métodos de *Cropping* parcial do fundo preto das imagens de fundo e *Tiling* para a criação de sub-imagens a partir das imagens originais dos conjuntos de dados públicos de RD para treinamento da rede neural. A abordagem para segmentação de instância também foi treinada

e avaliada utilizando os conjuntos de dados DDR e IDRiD, e particionados em conjunto de treinamento, validação e teste na proporção de 50:20:30, respectivamente.

Em ambas abordagens foi utilizado o método de particionamento dos conjuntos de dados em treinamento, validação e teste para que os modelos utilizados fossem ajustados no conjunto de validação e testados no conjunto de teste. Assim, os modelos não tiveram conhecimento *a priori* dos exemplos do conjunto de teste para que não foi produzido sobreajuste e vieses nos resultados obtidos. Não foi utilizado o método de validação cruzada denominado *k-fold* em função do custo computacional envolvido para treinamento dos modelos e o *hardware* disponível para realização dos experimentos. Por fim, foram utilizados os conjuntos de dados DDR e IDRiD em razão destes disponibilizarem as anotações (em nível de *pixel*) das lesões investigadas neste trabalho.

3. Resultados Obtidos e Discussão

Ambas as abordagens desenvolvidas foram comparadas com trabalhos similares no estado-da-arte encontrados na literatura. Os resultados obtidos pela abordagem proposta para detecção apresentam uma precisão superior a trabalhos de mesmo propósito apresentados na literatura. Além disso, a análise dos resultados permitiu identificar as lesões que apresentaram menores taxas de detecção, o que foi determinante para o desenvolvimento das soluções utilizadas para aumentar a precisão na detecção destas lesões. Os melhores resultados alcançados por esta abordagem no conjunto de dados DDR foram obtidos utilizando o otimizador Adam e o método de *Tilling*, alcançando para o limite de *IoU* de 0,5 na etapa de validação *mAP* de 0,2630, e na etapa de teste *mAP* de 0,1540.

Já os resultados alcançados pela abordagem para realizar a segmentação de instância apresenta uma precisão superior à abordagem proposta para detecção. Também apresenta resultados superiores na identificação de microaneurismas em relação a trabalhos de propósito similar relatados na literatura. Os melhores resultados alcançados por esta abordagem no conjunto de dados DDR foram obtidos utilizando o otimizador Adam e o método de *Tilling*, atingindo para o limite de *IoU* de 0,5 na etapa de validação *mAP* de 0,2903, e na etapa de teste *mAP* de 0,1670.

Embora a abordagem para segmentação de instância seja mais precisa, a abordagem proposta para detecção é mais rápida na realização das predições. A abordagem para detecção consegue manter um bom compromisso entre precisão e velocidade, ao passo que a abordagem proposta para segmentação de instância, embora seja relativamente mais precisa, tem um tempo de inferência elevado se comparado à abordagem para detecção. Contudo, a segmentação de instância apresenta a vantagem de realizar a detecção e segmentação das lesões que pode ser mais útil para o diagnóstico médico das lesões retinianas.

As avaliações demonstraram que as abordagens propostas apresentaram resultados promissores na detecção das lesões retinianas associadas à Retinopatia Diabética investigadas. Entretanto, há um vasto espaço que ainda pode ser explorado para melhorar a precisão na identificação destas lesões em imagens de fundo, inclusive motivando investigações futuras que podem servir de apoio para o desenvolvimento de diferentes sistemas para auxílio de diagnóstico médico, incluindo soluções que combinem diferentes contextos, como no aprendizado multimodal a fim de obter resultados mais precisos e auxiliar no diagnóstico precoce da RD.

Referências

- Alyoubi, W. L., Abulkhair, M. F., and Shalash, W. M. (2021). Diabetic retinopathy fundus image classification and lesions localization system using deep learning. *Sensors*, 21(11).
- Chakrabarti, R., Harper, C. A., and Keeffe, J. E. (2012). Diabetic retinopathy management guidelines. *Expert Review of Ophthalmology*, 7(5):417–439.
- Dai, L., Wu, L., Li, H., and Cai, C. (2021). A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nature Communications*, 12(1).
- Li, T., Gao, Y., Wang, K., Guo, S., Liu, H., and Kang, H. (2019). Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. *Information Sciences*, 501:511–522.
- Mateen, M., Wen, J., Nasrullah, N., Sun, S., and Hayat, S. (2020). Exudate Detection for Diabetic Retinopathy Using Pretrained Convolutional Neural Networks. *Complexity*, 2020.
- Nayak, J., Bhat, P. S., Acharya U, R., Lim, C. M., and Kagathi, M. (2008). Automated identification of diabetic retinopathy stages using digital fundus images. *Journal of Medical Systems*, 32(2):107–115.
- Porwal, P., Pachade, S., Kokare, M., Deshmukh, G., Son, J., and Bae, W. (2020). IDRiD: Diabetic Retinopathy – Segmentation and Grading Challenge. *Medical Image Analysis*, 59.
- Santos, C., De Aguiar, M. S., Welfer, D., and Belloni, B. (2021). Deep neural network model based on one-stage detector for identifying fundus lesions. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, Shenzhen, China, 18–22 July 2021. IEEE.
- Shenavarmasouleh, F., Mohammadi, F. G., Amini, M. H., Taha, T., Rasheed, K., and Arabnia, H. R. (2021). Drdrv3: Complete lesion detection in fundus images using mask r-cnn, transfer learning, and lstm.
- Ting, D. S. W., Cheung, G. C. M., and Wong, T. Y. (2016). Diabetic retinopathy: global prevalence, major risk factors, screening practices and public health challenges: a review. *Clinical and Experimental Ophthalmology*, 44(4):260–277.
- Vocaturro, E. and Zumpano, E. (2020). The contribution of AI in the detection of the Diabetic Retinopathy. In *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2020*, pages 1516–1519, Seoul, Korea, 16–19 December 2020. IEEE.
- Wang, H., Yuan, G., and Zhao, X. (2020). Hard exudate detection based on deep model learned information and multi-feature joint representation for diabetic retinopathy screening. *Computer Methods and Programs in Biomedicine*, 191:105398.
- Wong, T. Y., Aiello, L. P., Ferris, F., Gupta, N., Kawasaki, R., Lansingh, V., Maia, M., Mathenge, W., Moreker, S., Mugit, M., Resnikoff, S., Ruamviboonsuk, P., Sun, J., Taylor, H., Verdaguer, J., and Zhao, P. (2017). ICO guidelines for diabetic eye care. Technical report, International Council of Ophthalmology, Brussels, Belgium.