

UMA ABORDAGEM BASEADA EM REGIÕES DE INTERESSE PARA A DETECÇÃO E SEGMENTAÇÃO DE LESÕES RETINIANAS ASSOCIADAS À RETINOPATIA DIABÉTICA

MARCELO DA SILVA DIAS¹, MARILTON SANCHOTENE DE AGUIAR²

¹Universidade Federal de Pelotas – marcelo.sdias@inf.ufpel.edu.br

²Universidade Federal de Pelotas – marilton@inf.ufpel.edu.br

1. INTRODUÇÃO

O diabetes decorre da produção insuficiente ou da inadequada absorção de insulina, hormônio essencial para a conversão da glicose em energia (MARTINS; COELHO; BAFFA, 2021). A Retinopatia Diabética (RD) configura-se como uma complicação microvascular do diabetes e representa a principal causa de perda visual em adultos (OPHTHALMOLOGY, 2017). Embora não seja passível de cura, a cegueira pode ser prevenida mediante a detecção precoce de lesões em imagens da retina obtidas por meio de câmeras de fundo de olho (KUMAR; RAMASWAMY KARTHIKEYAN, 2021). Dessa forma, o diagnóstico e o tratamento antecipados tornam-se essenciais para mitigar o risco de complicações e de perda irreversível da visão.

A classificação dos estágios da RD fundamenta-se na interpretação de distintos padrões identificados em exames de fundo de olho. Contudo, a detecção de anomalias pode ser comprometida pela presença de ruídos, por falhas na captura das imagens e pelo fato de as lesões apresentarem formatos irregulares e dimensões bastante reduzidas (MARTINS; COELHO; BAFFA, 2021). Entre as principais lesões empregadas na avaliação do estágio da RD, destacam-se os exsudatos duros (*hard exudates* – EX), as hemorragias (*hemorrhages* – HE), os exsudatos algodonosos (*soft exudates* – SE) e os microaneurismas (*microaneurysms* – MA).

Como o diagnóstico da RD é realizado por meio da análise de imagens, torna-se viável empregar modelos computacionais para auxiliar na detecção de lesões e na classificação dos estágios da doença. De acordo com MATEEN et al. (2022), o avanço de abordagens automáticas pode possibilitar a identificação precoce de lesões retinianas, favorecendo o controle inicial da RD.

Estudos como os de OMETTO et al. (2017), MUNUERA-GIFRE et al. (2020) e ROMERO-AROCA et al. (2024) evidenciam que determinadas lesões, como MAs e EXs, não se distribuem aleatoriamente na retina, mas ocorrem com maior frequência na região temporal e ao redor da mácula. Dessa forma, direcionar o treinamento de modelos para explorar preferencialmente essas áreas pode ampliar a exposição a regiões relevantes, reduzir o tempo dedicado a áreas sem lesões e potencializar a eficácia do aprendizado.

Nesse contexto, este trabalho propõe o desenvolvimento de um pipeline para o treinamento de modelos computacionais, com o objetivo de aprimorar a detecção e a segmentação de lesões retinianas associadas à RD, tais como EX, HE, SE e, sobretudo, MA. A abordagem integra uma etapa adicional de pré-processamento das imagens, destinada à investigação de regiões de interesse, com ênfase na região temporal e nas áreas adjacentes à mácula, onde a probabilidade de ocorrência de EX e MAs é mais elevada. O aperfeiçoamento da detecção dessas lesões pode contribuir de maneira significativa para o

diagnóstico precoce da RD, possibilitando intervenções mais eficazes e mitigando o risco de complicações graves.

2. METODOLOGIA

Este trabalho adotou uma metodologia estruturada nas seguintes etapas: (1) levantamento do referencial teórico; (2) seleção do conjunto de dados; (3) extração da região de interesse; (4) treinamento do modelo; e (5) análise dos resultados obtidos.

Inicialmente, foi conduzida uma revisão da literatura, contemplando artigos que fornecem embasamento ao estudo. Os tópicos investigados abrangem pesquisas relacionadas à RD, redes neurais convolucionais, aprendizado profundo, pré-processamento de imagens e técnicas de detecção de objetos de pequenas dimensões.

Na sequência, procedeu-se à busca por conjuntos de dados públicos relacionados à RD. Para a tarefa de detecção e segmentação de lesões retinianas, foi selecionado o *Dataset for Diabetic Retinopathy* (DDR), empregado no treinamento e na avaliação do modelo. Para a detecção do disco óptico e da mácula, utilizou-se o *Indian Diabetic Retinopathy Image Dataset* (IDRID).

Para o treinamento do modelo na tarefa de detecção do disco óptico e da mácula, foi empregado o seguinte pré-processamento das imagens: (1) remoção parcial do fundo preto; (2) extração do canal verde; (3) aplicação do filtro da mediana para suavização; (4) equalização de histogramas adaptativa com limite de contraste (*Contrast Limited Adaptive Histogram Equalization* – CLAHE); e (5) aplicação do filtro bilateral para mitigar possíveis ruídos introduzidos pelo CLAHE.

O modelo YOLOv11 foi treinado no conjunto de dados IDRID com a seguinte configuração: *batch size* de 32, até 1000 épocas, com *early stopping* de 100 épocas. As imagens foram redimensionadas para 320×320 pixels, e o treinamento utilizou o otimizador AdamW com taxa de aprendizado de 0,001 e pesos pré-treinados do YOLOv11x. Para mitigar o *overfitting*, aplicou-se *dropout* de 10%, técnica que consiste em desativar parte dos neurônios durante o treinamento.

Com o modelo de detecção da mácula concluído, este foi aplicado ao conjunto de dados DDR para a extração das regiões de interesse, posteriormente utilizadas no treinamento do modelo de detecção e segmentação de lesões retinianas. Nessa etapa, também foi empregado o YOLOv11, configurado com 500 épocas, *patience* de 100 épocas, *batch size* de 4 e imagens redimensionadas para 800×800 pixels, além da aplicação de *dropout* com probabilidade de 10%.

Para ampliar a capacidade de generalização do modelo, foram empregadas técnicas de aumento de dados, incluindo *flips* verticais e horizontais, rotações, translações e recortes aleatórios. Além disso, aplicaram-se estratégias mais sofisticadas, como *mosaic*, que combina quatro imagens em uma única composição; *copy-paste*, que insere objetos de uma imagem em outra; e *mixup*, que mescla duas imagens e seus respectivos rótulos de forma interpolada.

Por fim, para a avaliação do modelo, não se mostrou viável utilizar apenas a região de interesse, uma vez que nem todas as lesões estão presentes nessa área da imagem. Dessa forma, recorreu-se à técnica de *tiling*, que consiste em dividir a imagem em blocos de 2×2. Essa abordagem possibilita a análise integral da imagem, ao mesmo tempo em que amplia a capacidade do modelo de detectar objetos de pequenas dimensões, concentrando a atenção em regiões locais em vez de considerar apenas a imagem global.

3. RESULTADOS E DISCUSSÃO

O modelo proposto foi confrontado com diferentes arquiteturas amplamente utilizadas nas tarefas de detecção e segmentação de lesões retinianas. Nos experimentos, foram avaliados os seguintes modelos: (1) YOLOv8; (2) YOLOv9; (3) YOLOv10; e (4) YOLOv11. Ademais, os resultados foram também comparados com os trabalhos de SANTOS (2023), que empregou a Detectron2, e de PEREIRA et al. (2023), que utilizou a YOLOR-CSP.

A Tabela 1 apresenta os resultados obtidos na detecção e segmentação das lesões EX, HE, SE e MA, considerando as métricas AP e mAP@50, avaliadas sobre o conjunto de testes do DDR. Os valores destacados em negrito indicam os melhores desempenhos em cada métrica.

Tabela 1. Resultados obtidos nas tarefas de detecção e segmentação de lesões retinianas no conjunto de testes do DDR.

Modelo		AP				mAP
		EX	HE	SE	MA	
YOLOv8	Bbox	0,2880	0,2890	0,2320	0,0851	0,2240
	Mask	0,2020	0,1860	0,2180	0,0435	0,1640
YOLOv9	Bbox	0,2860	0,3120	0,2620	0,098	0,2400
	Mask	0,2170	0,2230	0,2260	0,0469	0,1780
YOLOv10	Bbox	0,2890	0,3060	0,2270	0,0956	0,2290
	Mask	-	-	-	-	-
YOLOv11	Bbox	0,3080	0,3090	0,2420	0,1000	0,2400
	Mask	0,2110	0,1980	0,2160	0,0314	0,1640
Detectron2	Bbox	0,2515	0,1577	0,1548	0,1042	0,1670
	Mask	0,2687	0,1388	0,1589	0,1274	0,1736
YOLOR-CSP	Bbox	0,328	0,2190	0,2030	0,1400	0,2225
	Mask	-	-	-	-	-
Trabalho proposto	Bbox	0,4090	0,3410	0,1760	0,1830	0,2770
	Mask	0,4230	0,3300	0,1780	0,2010	0,2830

Verifica-se que o método proposto apresentou resultados expressivos, alcançando valores de AP de 0,4090 para EX, 0,3410 para HE, 0,1760 para SE e 0,1830 para MA. Nas tarefas de segmentação, os valores correspondentes foram 0,4230, 0,3300, 0,1780 e 0,2010, respectivamente. Destacam-se as lesões EX, HE e MA, cujos valores de AP superaram os modelos de referência tanto em detecção quanto em segmentação. Ademais, o método obteve os melhores resultados de mAP@50 em ambas as tarefas, com 0,2770 em detecção e 0,2830 em segmentação.

Apesar dos resultados promissores, a detecção e a segmentação de MAs permanecem como um desafio significativo, em virtude do tamanho extremamente reduzido dessas lesões e, em certos casos, da dificuldade de distingui-las em relação ao fundo da retina. Outro aspecto a ser ressaltado refere-se ao desempenho da lesão SE, para a qual o método proposto foi superado por outros modelos. Esse resultado pode ser atribuído a dois fatores principais: (1) o número reduzido de exemplos dessa lesão no conjunto de dados e (2) a baixa incidência dessa lesão na região macular.

4. CONCLUSÕES

Os resultados sugerem que abordagens que direcionam o treinamento para regiões de interesse podem contribuir de maneira significativa para o desenvolvimento de modelos computacionais mais eficazes na detecção e segmentação de lesões retinianas. Apesar dos avanços alcançados, ainda há espaço para aprimoramentos, sobretudo na detecção de MAs, que continuam a representar o maior desafio. Ademais, o desempenho reduzido observado em SE pode ser explicado, em grande parte, pela limitada quantidade de exemplos dessa lesão.

Como perspectivas futuras, propõe-se a aplicação de técnicas de pré-processamento mais sofisticadas nas imagens de treinamento, a exploração de diferentes arquiteturas, como U-Nets, bem como a adoção de estratégias de balanceamento de classes e a validação em conjuntos de dados adicionais.

5. AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- KUMAR, N. S.; RAMASWAMY KARTHIKEYAN, B. Diabetic Retinopathy Detection using CNN, Transformer and MLP based Architectures. In: **INTERNATIONAL SYMPOSIUM ON INTELLIGENT SIGNAL PROCESSING AND COMMUNICATION SYSTEMS (ISPACS)**, 2021., 2021, Hualien, Taiwan. Proceedings... IEEE, 2021. p.1-2.
- MARTINS, J., COELHO, A., BAFFA, M. Detecção dos Estágios Iniciais da Retinopatia Diabética Utilizando Redes Neurais Profundas. In: **VIII ESCOLA REGIONAL DE COMPUTAÇÃO APLICADA À SAÚDE**, 2021, Porto Alegre, RS, Brasil. Anais... SBC, 2021. p.42-45.
- MATEEN, M. et al. Deep Learning Approach for Automatic Microaneurysms Detection. **Sensors**, Basel, Switzerland, v.22, n.2, 2022.
- MUNUERA-GIFRE, E. et al. Analysis of the location of retinal lesions in central retinographies of patients with Type 2 diabetes. **Acta Ophthalmologica**, Oxford, UK, v.98, n.1, p.e13-e21, 2020.
- OMETTO, G. et al. Spatial distribution of early red lesions is a risk factor for development of vision-threatening diabetic retinopathy. **Diabetologia**, UK, v.60, n.12, p.2361-2367, Dec. 2017.
- OPHTHALMOLOGY, I. C. of. Updated 2017 ICO Guidelines for Diabetic Eye Care. **ICO Guidelines for Diabetic Eye Care**, Brussels, Belgium, p.1–33, 2017.
- RADOSAVOVIC, I., KOSARAJU, R. P., GIRSHICK, R. B., HE, K., e DOLLAR, P. (2020). Designing network design spaces. **CoRR**, abs/2003.13678.
- ROMERO-AROCA, P. et al. Distribution of Microaneurysms and Hemorrhages in Accordance with the Grading of Diabetic Retinopathy in Type Diabetes Patients. **Diagnostics**, Basel, Switzerland, v.14, n.14, 2024.