

# APLICANDO REDES ADVERSÁRIAS GENERATIVAS NA TAREFA DE AUMENTO DE DADOS DE IMAGENS DA RETINA E UTILIZAÇÃO EM SISTEMAS DE DETECÇÃO DE LESÕES ASSOCIADAS À RETINOPATIA DIABÉTICA

MARCELO CLASEN RIBEIRO<sup>1</sup>, MARILTON SANCHOTENE DE AGUIAR<sup>2</sup>

<sup>1</sup> Universidade Federal de Pelotas - [mcribeiro@inf.ufpel.edu.br](mailto:mcribeiro@inf.ufpel.edu.br)

<sup>2</sup> Universidade Federal de Pelotas - [marilton@inf.ufpel.edu.br](mailto:marilton@inf.ufpel.edu.br)

## 1. INTRODUÇÃO

De acordo com a Federação Internacional de Diabetes, a diabetes é uma das emergências relacionadas à saúde que mais cresce mundialmente. Uma das complicações causadas pela doença é a retinopatia diabética, problema que afeta a retina e pode levar desde lesões oculares até a perda de visão (International Council of Ophthalmology, 2017).

Caso a retinopatia diabética seja identificada em estágio inicial, é possível evitar casos graves da doença, proporcionando ao paciente um tratamento adequado para impedir que a doença avance. Um dos métodos utilizados para diagnóstico da doença consiste na identificação de sinais oculares que ocorrem no tecido da retina, tais como: hemorragias (HE), microaneurismas (MA), exsudatos duros (EX) e exsudatos algodonosos (SE).

Métodos baseados em redes neurais profundas são comumente utilizados em tarefas de classificação, segmentação e detecção de lesões retinianas. No entanto, esses métodos dependem de uma grande quantidade de dados, incluindo anotações das diferentes classes de lesões associadas à retinopatia diabética para a realização das etapas de treinamento e validação. Embora na literatura sejam disponibilizados *datasets* de imagens de fundo de domínio público, esses conjuntos de dados apresentam problemas relacionados à baixa qualidade das imagens e a indisponibilidade de anotações para os principais tipos de lesões retinianas (LI et al., 2019), o que leva os mesmos a serem frequentemente descartados para utilização no treinamento de modelos de detecção.

Para contornar esse problema, é possível utilizar técnicas de *data augmentation*, que consistem na criação de novos exemplos de treinamento a partir de dados de treinamento já existentes (GOODFELLOW et al., 2020), possibilitando um treinamento mais eficaz e permitindo que o modelo seja capaz de realizar inferências para uma maior variedade de situações. Uma das possíveis técnicas a serem utilizadas nessa tarefa são as Redes Adversárias Generativas (GAN), método baseado em redes neurais profundas capaz de realizar a síntese de dados artificiais.

Este trabalho tem como objetivo apresentar uma abordagem baseada em Redes Adversárias Generativas para realizar a síntese de imagens artificiais da retina para treinamento de modelos de redes neurais profundas visando a detecção de diferentes sinais oculares associados à retinopatia diabética.

## 2. METODOLOGIA

Os conjuntos de dados selecionados estão disponibilizados publicamente, sendo esses o *Dataset for Diabetic Retinopathy* (DDR) (LI et al., 2019) e o *Indian Diabetic Retinopathy image Dataset* (IDRiD) (PORWAL et al., 2020). Ambos conjuntos possuem dois tipos de anotações, sendo um deles relacionado ao grau de retinopatia diabética e o outro das lesões indicativas da doença.

O primeiro passo do desenvolvimento foi o de pré-processamento dos dados. Baseado nos experimentos realizados, surgiu a hipótese de aumentar o nível de contexto semântico fornecido para o modelo generativo, onde, para isso, foram realizadas as segmentações de duas novas estruturas da imagem: a máscara de fundo e os vasos retinianos. Posteriormente, foi aplicado o *cropping*, onde foram removidas as bordas pretas da imagem, diminuindo a quantidade de informações supérfluas, e o redimensionamento, com o objetivo de padronizar o tamanho das imagens para as próximas etapas.

Ainda na etapa de pré-processamento dos dados, foi realizada a montagem dos mapas de *labels*, imagem composta das anotações das lesões retinianas e estruturas segmentadas, que, em conjunto com as imagens de retina, foram utilizadas na etapa de treinamento da GAN. Essa etapa foi necessária devido ao tipo de informação condicional utilizada na GAN selecionada, que é uma imagem composta de informações semânticas do que deve ser gerado.

A próxima etapa da abordagem foi o treinamento da GAN, rede generativa selecionada para etapa de síntese do projeto. Em específico, o modelo escolhido foi o pix2pixHD (WANG et al., 2018). O treinamento da GAN se dá de forma adversarial, onde uma das redes, geradora, tem o trabalho de gerar imagens artificiais com o objetivo de enganar a outra rede, discriminadora, que deve avaliar se o dado faz parte do conjunto original ou é sintético. Com o modelo treinado, podemos gerar novos exemplares de dados a partir de mapas de *labels*, sem utilizar dados reais.

Por fim, foi realizado o treinamento de uma rede neural com arquitetura Mask R-CNN, com Backbone ResNeXt-101-32x8d-FPN, construída utilizando a biblioteca Detectron2 e sendo pré-treinada com o conjunto de dados COCO, de acordo com o realizado em trabalhos anteriores (SANTOS et al., 2023). Essa rede é capaz de gerar caixas delimitadoras e segmentações dos objetos que deseja detectar, e foi utilizada como etapa de validação da eficácia da abordagem proposta.

Com o objetivo de comparar a eficácia da abordagem proposta, foram realizados diversos experimentos. É importante salientar que todos treinamentos de GAN utilizaram, em conjunto com as anotações originais, a máscara de fundo, porém nem todos utilizaram as anotações das veias retinianas.

O primeiro experimento, nomeado **DDR 80%**, utilizou 80% do conjunto de treinamento do dataset DDR para treinamento da GAN e os 20% restantes do conjunto foram utilizados como mapa de *labels* na síntese e substituição no conjunto original após a finalização do treinamento do modelo. O segundo experimento, nomeado **DDR 80% + VE**, incluiu as anotações dos vasos retinianos, visando aumentar o contexto semântico da imagem e diminuir a ocorrência de artefatos na imagem. Os terceiro e quarto experimentos, nomeados respectivamente de **DDR + IDRiD** e **DDR + IDRiD + VE**, seguem a mesma lógica de adição das anotações de vasos retinianos, porém utilizaram todo conjunto de treinamento do DDR para treinamento da GAN e as anotações do conjunto IDRiD como mapa de *labels* e adições ao conjunto original.

A Figura 1 apresenta um exemplar do experimento **DDR + IDRiD + VE**. Nela é possível observar o mapa de *labels* utilizado, a imagem sintética e a imagem original, permitindo analisar a capacidade do modelo de representar as anotações e a similaridade com a imagem original.

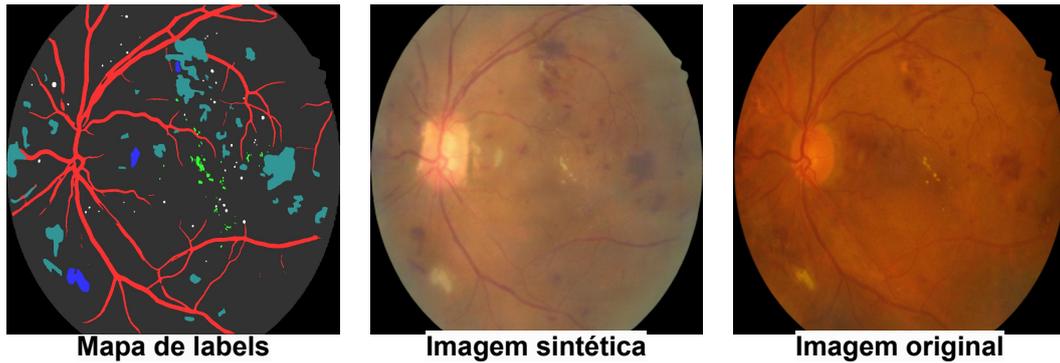


Figura 1: Exemplos do experimento **DDR + IDRiD + VE**. As estruturas anotadas utilizadas foram MA (branco), HE (ciano), SE (azul), EX (verde) vasos retinianos (vermelho) e máscara de fundo (cinza).

### 3. RESULTADOS E DISCUSSÃO

Para avaliar a qualidade das imagens geradas, além do treinamento com os dados dos quatro experimentos citados anteriormente, foi realizado o treinamento com o conjunto **DDR base**, como métrica de comparação.

A Tabela 1 apresenta os resultados encontrados dos experimentos realizados para caixa delimitadora e segmentação no conjunto de teste, sob as métricas *AP* (*Average Precision*) com limite *IoU* (*Intersection over Union*). É possível observar que experimentos utilizando a abordagem proposta foram capazes de superar o conjunto base em alguns casos. Apesar disso, considerando a média de *APs*, o conjunto base se saiu melhor no conjunto de teste.

Tabela 1: Resultados obtidos com as métricas *AP* e *mAP* para o limite de *IoU* de 0,5 no conjunto de teste. O melhor resultado obtido para cada coluna nos respectivos modos está destacado em negrito, para facilitar a visualização.

Modelos	<i>AP</i>				<i>mAP</i>
	EX	HE	SE	MA	
caixa delimitadora					
DDR base	0,19724	0,12111	<b>0,20214</b>	0,05017	<b>0,14267</b>
DDR 80%	0,17916	0,10468	0,18228	0,05242	0,12964
DDR 80% + VE	0,18073	0,11332	0,17064	0,06158	0,13157
DDR + IDRiD	0,19651	0,11566	0,11755	0,04342	0,11828
DDR + IDRiD + VE	<b>0,20404</b>	<b>0,12784</b>	0,16670	<b>0,06313</b>	0,14043
segmentação					
DDR base	0,20284	<b>0,10630</b>	<b>0,20049</b>	0,05416	<b>0,14095</b>
DDR 80%	0,18023	0,08481	0,18206	0,05350	0,12515
DDR 80% + VE	0,18161	0,09804	0,17489	0,04750	0,12551
DDR + IDRiD	<b>0,22344</b>	0,10215	0,12557	<b>0,07339</b>	0,13114
DDR + IDRiD + VE	0,18957	0,10553	0,17565	0,05563	0,13159

## 4. CONCLUSÕES

Os experimentos apresentados demonstram que a técnica proposta teve um desempenho promissor na criação de imagens retinianas sintéticas, alcançando resultados semelhantes na detecção e segmentação de lesões retinianas quando comparada a utilização do conjunto de dados DDR original, inclusive superando-o em alguns casos. Esses resultados validam a ideia de síntese de imagens retinianas a partir de mapas de anotações semânticas. No entanto, são necessárias mais pesquisas para validar a abordagem em diferentes cenários, inclusive utilizando uma taxa de *data augmentation* maior, introduzindo mais exemplares sintéticos ao conjunto de dados original. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

- GOODFELLOW, I.; POUGET-ABADIE, J.; MIRZA, M.; XU, B.; WARDE-FARLEY, D.; OZAI, S.; COURVILLE, A.; BENGIO, Y. Generative adversarial networks. **Communications of the ACM**, v. 63, n. 11, p. 139–144, 2020. ISSN 15577317.
- International Council of Ophthalmology. Updated 2017 ICO Guidelines for Diabetic Eye Care. **ICO Guidelines for Diabetic Eye Care**, Brussels, Belgium, p. 1–33, 2017. Disponível em: <<http://www.icoph.org/downloads/ICOGuidelinesforDiabeticEyeCare.pdf>>.
- LI, T.; GAO, Y.; WANG, K.; GUO, S.; LIU, H.; KANG, H. Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. **Information Sciences**, Elsevier Inc., v. 501, p. 511–522, 2019. ISSN 00200255. Disponível em: <<https://doi.org/10.1016/j.ins.2019.06.011>>.
- PORWAL, P.; PACHADE, S.; KOKARE, M.; DESHMUKH, G.; SON, J.; BAE, W.; LIU, L.; WANG, J.; LIU, X.; GAO, L.; WU, T. B.; XIAO, J.; WANG, F.; YIN, B.; WANG, Y.; DANALA, G.; HE, L.; CHOI, Y. H.; LEE, Y. C.; JUNG, S. H.; LI, Z.; SUI, X.; WU, J.; LI, X.; ZHOU, T.; TOTH, J.; BARAN, A.; KORI, A.; CHENNAMSETTY, S. S.; SAFWAN, M.; ALEX, V.; LYU, X.; CHENG, L.; CHU, Q.; LI, P.; JI, X.; ZHANG, S.; SHEN, Y.; DAI, L.; SAHA, O.; SATHISH, R.; MELO, T.; ARAÚJO, T.; HARANGI, B.; SHENG, B.; FANG, R.; SHEET, D.; HAJDU, A.; ZHENG, Y.; MENDONÇA, A. M.; ZHANG, S.; CAMPILHO, A.; ZHENG, B.; SHEN, D.; GIANCARDIO, L.; QUELLEC, G.; MÉRIAUDEAU, F. IDRiD: Diabetic Retinopathy – Segmentation and Grading Challenge. **Medical Image Analysis**, Elsevier B.V., v. 59, 2020. ISSN 13618423.
- SANTOS, C.; AGUIAR, M.; WELFER, D.; DIAS, M.; PEREIRA, A.; RIBEIRO, M.; BELLONI, B. A new approach for fundus lesions instance segmentation based on mask r-cnn x101-fpn pre-trained architecture. **IEEE Access**, p. 1–1, 2023.
- WANG, T.-C.; LIU, M.-Y.; ZHU, J.-Y.; TAO, A.; KAUTZ, J.; CATANZARO, B. High-resolution image synthesis and semantic manipulation with conditional gans. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2018. p. 8798–8807.