

## **APLICAÇÃO DE INTERPRETABILIDADE EM REDE NEURAL ARTIFICIAL CONVOLUCIONAL PARA A DETECÇÃO DE GLAUCOMA <sup>1</sup>**

Maísa Fernandes Gomes<sup>2</sup>, Rafael Stubs Parpinelli<sup>3</sup>

<sup>1</sup> Vinculado ao projeto “Reconhecimento de Padrões com Inteligência Computacional e Processamento de Alto-desempenho para Auxílio de Diagnóstico”

<sup>2</sup> Acadêmica do Curso de Bacharelado em Ciência da Computação - CCT – Bolsista PROBIC

<sup>3</sup> Orientador, Departamento de Ciência da Computação – CTT – rafael.parpinelli@udesc.br

O glaucoma é uma doença ocular que atinge a área do nervo óptico. A doença é silenciosa e em seu caso mais avançado causa a total perda de visão. A detecção precoce permite o monitoramento da progressão da doença no paciente. Os avanços das técnicas de aprendizado profundo e os resultados expressivos no campo de reconhecimento de imagens, inspiraram as pesquisas para a utilização de modelos inteligentes na classificação de imagens médicas, como é o caso da oftalmologia que utiliza imagens médicas para o monitoramento e detecção de doenças. Dentre os modelos de aprendizado profundo (Deep Learning -DL), as redes neurais convolucionais (RNC) são as mais utilizadas nas tarefas de reconhecimento de imagens.

Tendo em vista a importância do uso de interpretabilidade em redes neurais convolucionais para o diagnóstico médico, o presente trabalho teve como objetivo implementar uma RNC para o classificação de imagens de fundo de olho em duas classes, glaucoma e normal. Para o modelo da rede neural convolucional, foi escolhido com base na revisão da literatura a arquitetura residual convolutional network (ResNet), e para o treinamento, teste e validação foram utilizadas duas bases públicas, a RIM-ONE e ACRIMA, uma terceira base foi criada e denominada de RIM-ONE+ACRIMA com a junção das imagens da duas bases, algumas imagens da base da ACRIMA possuíam ruídos e foram descartadas.

Para a implementação do modelo foi utilizado a linguagem de programação Python com o auxílio da biblioteca Keras. Todas as imagens passaram por um redimensionamento para o tamanho de 224x224 que é a entrada para a RNC. Os conjuntos de dados foram balanceados para que tivessem a mesma quantidade de imagens para cada classe e separados de forma randômica em treino e teste, sendo 70% das imagens para treino e 30% para validação. Os hiperparâmetros foram ajustados de maneira empírica, sendo: *batch-size* igual a oito, 200 épocas, otimizador Adam. A taxa de aprendizado de 0.0001, epsilon de  $1^{-10}$  e decaimento de 0.001 para os conjuntos de dados RIM-ONE e RIM-ONE+ACRIMA, e taxa de aprendizado 0.0005, epsilon  $1^{-10}$  e decaimento 0.001 para base ACRIMA.

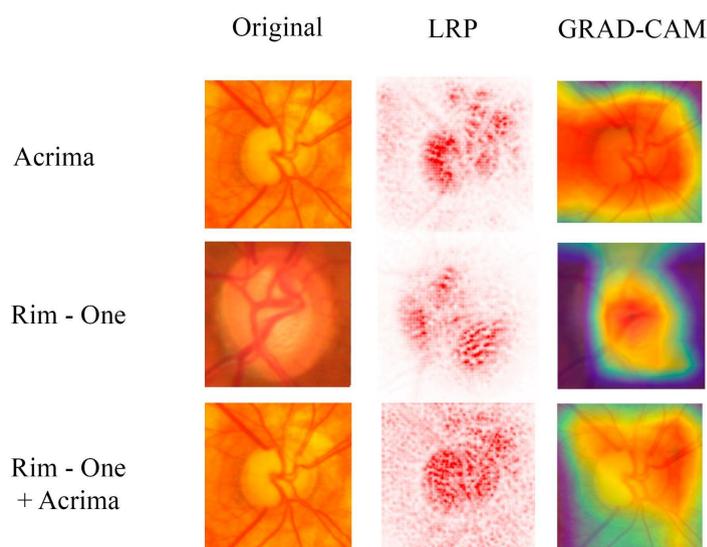
Foi utilizado sistema operacional Linux Ubuntu 20.04 e hardware AMD Ryzen 7 2700X, 3.7GHz/4.35GHz, 16 CPUs, 16GB RAM, 1TB HD, 240GB SSD, GeForce RTX 2060 Super disponível no Laboratório de Pesquisa em Inteligência Computacional (LABICOM), onde a GPU utilizada possibilitou treinamento do modelo com melhor desempenho de tempo de processamento. Foram realizadas 10 execuções do modelo proposto para obtenção de média e desvio-padrão das métricas de acurácia e perda, e são mostrados na Tabela 3. Pode ser observado que a média das 10 execuções para os 3 modelos resultaram em acurácia alta e perda baixa, o que demonstra que os modelos estão tendo resultados promissores, ou seja, estão acertando quase a maioria das imagens de validação, e os erros que obtém são pequenos.

**Tabela 2.** Resultados obtidos do modelo treinado em cada uma das bases de dados.

	ACRIMA	RIM-ONE	RIM-ONE+ACRIMA
Acurácia ( <i>acc</i> )	99.37 ± 7.0	98.75 ± 2.5	99.35 ± 7.0
Perda ( <i>loss</i> )	0.01 ± 0.03	0.06 ± 1.17	0.01 ± 0.05

As duas técnicas de interpretabilidade utilizadas, LRP e Grad-CAM, tem objetivos similares que é demonstrar o quanto um pixel foi importante para a classificação da imagem para a classe que foi classificada, as áreas vermelhas da imagem representam pontos com importância maior. A diferença entre os dois métodos está na forma em que é calculado a importância do pixel. A Figura 1 mostra a comparação entre os resultados obtidos em uma predição para classe glaucoma para cada modelo treinado. Pode ser visualizado que a imagem de saída tanto do LRP quanto do Grad-CAM que as áreas de importância para classificação se assemelham sendo a parte central do disco óptico, a diferença está na forma de visualização, enquanto o LRP é mais pontual o Grad-CAM aponta áreas de maior importância.

**Figura 1 .** A primeira coluna indica a base de dados, e as demais colunas indicam a imagem original, a imagem interpretada com a técnica LRP, e a imagem interpretada com a técnica Grad-CAM.



O uso de técnicas de aprendizado profundo para o reconhecimento de imagens tem apresentando bons resultados e as aplicações de métodos de interpretabilidade trazem um entendimento de como o modelo funciona. Neste trabalho foi alcançado bons resultados utilizando RNC para detecção de glaucoma em imagens de fundo de olho obtidas de bases públicas, e os resultados da interpretabilidade mostraram que os modelos estão considerando as áreas certas para sua predição. Para trabalhos futuros, técnicas de pré-processamento podem ser utilizadas nas imagens de entrada, e as técnicas de interpretabilidade podem ser validadas junto com uma equipe médica.

**Palavras-chave:** Glaucoma. Interpretabilidade. Aprendizado de Máquina.