

# Detecção Automática do Glaucoma em Imagens Retinianas Utilizando Redes Neurais Convolucionais

Jeanderson de Sousa Gomes<sup>1</sup>, Luis Eduardo Silva Brito<sup>2</sup>, Alcilene Dalília<sup>2</sup>,  
Romuere Rodrigues Veloso<sup>123</sup>, Antônio Oseas de Carvalho<sup>123</sup>, Flavio H.D. Araujo<sup>123</sup>

<sup>1</sup> Departamento de Engenharia Elétrica,  
Universidade Federal do Piauí - Teresina, Brasil

<sup>2</sup> Departamento de Sistemas de Informação,  
Universidade Federal do Piauí - Picos, Brasil

<sup>3</sup> Departamento de Ciência da Computação,  
Universidade Federal do Piauí - Teresina, Brasil

{jeandsgomes10,alcilene,romuere,flavio86}@ufpi.edu.br

{duardos36,antoniooseas}@gmail.com

**Abstract.** *Glaucoma is a leading cause of blindness worldwide, affecting the optic nerve and causing progressive damage. Early detection is crucial to prevent vision loss, but manual diagnosis, based on retinal images, is subjective and prone to errors. This work proposes using deep learning with the YOLOv8n model to detect the optic disc, followed by classification with a refined ResNet-50. The method achieves 91,7% precision, 91,1% recall, 91,5% accuracy, and 82,7% kappa coefficient. We conclude that using the full retinal image outperforms using just the optic disc region, as confirmed by Grad-CAM activation analysis.*

**Resumo.** *O glaucoma é uma das principais causas de cegueira no mundo, afetando o nervo óptico e causando danos progressivos. A detecção precoce é essencial para evitar a perda da visão, mas o diagnóstico manual baseado em imagens retinianas pode ser subjetivo e propenso a erros. Propõe-se o uso de deep learning com o modelo YOLOv8n para detectar o disco óptico, seguido da classificação com uma ResNet-50 refinada. O método obteve 91,7% de precisão, 91,1% de recall, 91,5% de acurácia e 82,7% no coeficiente kappa. Conclui-se que o uso da imagem completa da retina teve melhor desempenho do que o uso exclusivo da região do disco óptico, conforme confirmado pela análise de ativação do Grad-CAM.*

## 1. Introdução

Segundo o Ministério da Saúde [da Saúde 2021], o glaucoma é uma das principais causas de cegueira no Brasil, atingindo mais de 900 mil pessoas. Essa doença pode não apresentar sintomas nas fases iniciais e é caracterizada por danos progressivos ao nervo óptico, o que pode resultar na cegueira permanente de pacientes não diagnosticados precocemente [Hayreh 2007]. Nesse contexto, a análise precisa das imagens do fundo do olho surge como uma ferramenta para auxiliar os oftalmologistas na identificação de casos suspeitos. A detecção precoce é crucial para evitar a perda total da visão e garantir um

tratamento eficaz [Silva 2016]. Dessa forma, a análise detalhada dessas imagens torna-se essencial para o diagnóstico precoce e o acompanhamento da doença.

Apesar dos avanços tecnológicos, a interpretação e análise das imagens do fundo do olho ainda é um processo exaustivo e suscetível a erros [Tamim et al. 2021]. Essa prática exige concentração e experiência dos profissionais, além de demandar tempo significativo, especialmente em casos que envolvem grandes volumes de imagens retinianas. Como consequência, esses problemas elevam custos, atrasam tratamentos e comprometem a eficácia das intervenções, afetando a qualidade de vida dos pacientes.

Diante desses desafios, a busca por soluções mais eficientes se torna essencial. As técnicas de inteligência artificial têm demonstrado cada vez mais avanços no campo da oftalmologia. Neste contexto, algoritmos de inteligência artificial e redes neurais convolucionais identificam características nas imagens que podem passar despercebidas pelos oftalmologistas. A detecção automatizada do disco óptico é um passo fundamental para auxiliar no diagnóstico precoce do glaucoma, oferecendo maior precisão na classificação de imagens. Além disso, com o uso dessas técnicas é possível analisar grandes conjuntos de dados com eficiência e robustez. Isso permite uma detecção precoce, e consequentemente melhoramento da qualidade do tratamento e saúde ocular dos pacientes.

Para realizar o diagnóstico do glaucoma, os médicos analisam de forma minuciosa a região do disco óptico. Com base nisso, um dos indicadores da doença é a proporção escavação-disco aumentada, resultando no aumento da escavação óptica em relação ao disco óptico. Esse aumento é conhecido como escavação, na qual resulta na perda de células da retina, refletindo em uma escavação óptica mais pronunciada e de aparência mais clara [Yunitasari et al. 2021]. Portanto, detectar essas alterações é fundamental para o diagnóstico e tratamento precoce da doença.

## **2. Trabalhos relacionados**

Nesta seção, são apresentados trabalhos que aplicam técnicas de classificação de glaucoma em imagens do fundo do olho.

O método apresentado pelo trabalho [Shanmugam et al. 2021] utiliza a proporção CDR (Cup to Disc Ratio) obtida de imagens do fundo do olho para identificar o glaucoma, focando na segmentação do disco óptico e do copo óptico como etapas críticas. A técnica inclui aquisição de imagem com aprimoramento de contraste, extração de características, e avaliação de glaucoma, utilizando um classificador florestal aleatório para classificar as imagens com base nos valores de CDR. Testes comparativos com diferentes versões da técnica U-Net mostraram que o método proposto alcançou uma precisão de classificação de 99% e uma melhoria de 14% na precisão de segmentação em relação ao U-Net Original.

O trabalho [Sulot et al. 2021] avaliou a eficácia das imagens frontais do nervo óptico, obtidas por oftalmoscopia a laser de varredura (SLO) durante tomografias de coerência óptica do segmento posterior, e explorou um método de aprendizado profundo para diferenciar pacientes com e sem glaucoma em um contexto de dados limitados. Utilizando 227 imagens SLO de 227 participantes (105 com glaucoma e 122 controles), categorizados por testes clínicos como medição da pressão intraocular e espessura da camada de fibras nervosas da retina, foi desenvolvida uma nova arquitetura de rede neural

convolucional para a classificação. A avaliação de vários classificadores mostrou que o método de aprendizado profundo proposto teve um alto desempenho de discriminação, com uma precisão balanceada de até 0,962, superando outros métodos testados.

O estudo de [Sanghavi and Kurhekar 2024], apresenta um método capaz de realizar a segmentação precisa do disco óptico e a classificação do glaucoma. Ressaltando a importância dos indicadores como a relação escavação/disco, a atrofia papilar e a diminuição da borda neural da retina (NRR) para a detecção do glaucoma. Diante disso, a proposta utiliza o processamento de histograma para determinar se uma imagem precisa passar por segmentação antes da classificação. Para imagens completas da retina, foi utilizado o algoritmo SLIC (Simple Linear Iterative Clustering) e o corte de gráfico normalizado para segmentar o disco óptico. Já para as imagens já segmentadas, a Rede Neural Convolucional (CNN) do estudo as recebe diretamente para classificação. O desempenho do modelo foi avaliado e comparado com arquiteturas pré-treinadas como VGG19, InceptionV3 e ResNet50V2, utilizando métricas de precisão para validação. Com testes realizados em um conjunto de 3115 imagens de seis bases de dados padrão, a solução proposta obteve uma precisão de 96,33%.

No trabalho de [Elmannai et al. 2024], é proposto um método para segmentar o disco óptico (OD) e detectar o glaucoma, utilizando redes neurais profundas. Neste estudo, é feita a segmentação do disco óptico com a Mask Region-Based Convolutional Neural Network (Mask-RCNN), na qual é utilizada para a classificação do glaucoma. Os pesquisadores, empregaram redes neurais pré-treinadas, como VGG-16, ResNet-18 e Inception-V3, além do classificador de Máquina de Vetores de Suporte (SVM), utilizando características extraídas por Histogram of Oriented Gradients (HOG) e Filtros de Gabor. A metodologia foi validada em três conjuntos de dados públicos, apresentando um desempenho superior a 98% de precisão na classificação da doença.

O artigo [Ruby Elizabeth et al. 2024] propõe o uso de uma versão modificada do modelo de aprendizado profundo AlexNet para detectar e classificar o glaucoma em imagens retinianas. O processo envolve o pré-processamento das imagens, seguido da detecção e segmentação das regiões do disco óptico e do copo óptico usando um filtro circulatório e o algoritmo K-means, respectivamente. As regiões segmentadas são então classificadas pelo modelo AlexNet para determinar se a retina está saudável ou afetada pelo glaucoma. O desempenho do método foi avaliado e comparado com técnicas anteriores usando os conjuntos de dados de imagens retinianas públicas HRF e RIGA. O método alcançou uma alta taxa de detecção de glaucoma (GDR), com 91,6% para casos leves e 100% para casos graves no conjunto HRF, e 97,7% para casos leves e 100% para casos graves no conjunto RIGA, demonstrando uma eficácia significativa na classificação do glaucoma.

O trabalho [Wang et al. 2024] estudo utilizou tecnologias avançadas de aprendizado profundo e processamento de imagem para refinar o diagnóstico de glaucoma, analisando uma variedade de dados, incluindo imagens do fundo da retina e regiões de interesse específicas. Um novo modelo, baseado na linguagem de visão grande (VLLM), foi desenvolvido para ajustar os pesos de quatro tipos diferentes de dados de entrada, alinhando suas probabilidades previstas com as respostas reais. Esse modelo melhorou significativamente a precisão e a sensibilidade na classificação do glaucoma, alcançando aumentos de até 10% na precisão e 29% na sensibilidade. Observando os resultados nos conjuntos de

dados REFUGE e ORIGA, foi destacado o potencial transformador do modelo na prática clínica.

O artigo [Tan et al. 2024] apresenta um modelo híbrido de aprendizado profundo que integra uma rede totalmente conectada com uma rede neural convolucional para analisar mapas e parâmetros obtidos por Tomografia de Coerência Óptica. Este modelo foi desenvolvido para diferenciar olhos saudáveis de olhos afetados por glaucoma perimétrico, com base na observação e análise de 106 indivíduos normais e 164 pacientes com glaucoma. O modelo alcançou uma acurácia de 94,8%, especificidade de 100%, e uma área sob a curva ROC (AROC) de 97,9%, desempenhos que superaram significativamente os modelos baseados em regressão logística. A Tabela 1 representa os trabalhos relacionados a detecção do glaucoma.

O estudo de [Sivakumar and Penkova 2025], foi proposto um sistema multimodal para detecção de glaucoma. Na qual utilizou três modelos híbridos para extração e análise de características: Vision Transformers (ViT) + Redes Neurais Convolucionais (CNN - ResNet), Object-Window-Location Vision Transformer (OWL-ViT) + ResNet, Hierarchical Vision Transformer com Shifted Windows (Swin Transformer) + ResNet. O Swin Transformer com ResNet demonstrou o melhor desempenho, atingindo uma precisão global de 99,4% ao analisar um conjunto de 2874 novos casos, classificando corretamente 2857 amostras. Além disso, os F1-scores foram 0,993 para glaucoma e 0,995 para não-glaucoma.

### 3. Metodologia

Nesta seção, é apresentada a metodologia proposta, dividida em seis fases. Inicialmente, as imagens foram adquiridas e processadas para a extração manual da área de interesse do disco óptico. Em seguida, os recortes gerados foram usados para treinar o modelo de detecção YOLOv8n, que foi posteriormente utilizado para detectar e extrair automaticamente a região do disco óptico. Após a detecção, a arquitetura ResNet-50 foi treinada para classificar as imagens com base nas regiões extraídas do disco óptico. Finalmente, o desempenho do modelo foi avaliado por meio de métricas como loss, acurácia, precisão, recall e coeficiente kappa. Uma ilustração do fluxograma é apresentada na Figura 1.

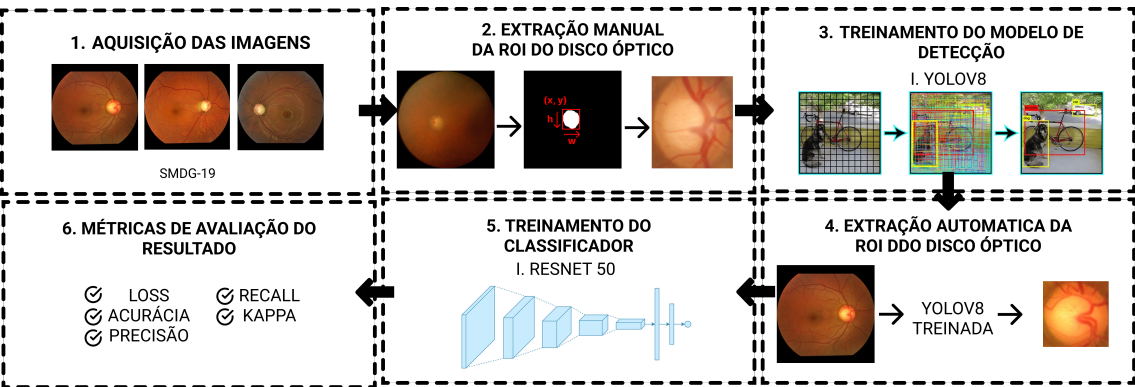


Figura 1. Fluxograma Representando as Etapas do Método Proposto

**Adição:** Com base nisso, este artigo propõe uma abordagem para a detecção automatizada do glaucoma. Para isso, foi utilizada a base de dados SMDG-19, que contém

**Tabela 1. Resumo dos trabalhos encontrados sobre detecção de glaucoma.**

Referência	Método Utilizado	Base de Dados	Resultados
[Sulot et al. 2021]	CNN personalizada com imagens SLO para classificação	227 imagens SLO (105 com glaucoma, 122 controles)	96,2% (Precisão Balanceada)
[Shanmugam et al. 2021]	Segmentação CDR (Cup to Disc Ratio) e Random Forest para classificação	Comparação com diferentes versões da U-Net	99% (Precisão)
[Sanghavi and Kurhekar 2024]	Segmentação do disco óptico com SLIC e CNN para classificação	3115 imagens de 6 bases de dados padrão	96,33% (Precisão)
[Elmannai et al. 2024]	Segmentação com Mask-RCNN e classificação com redes pré-treinadas e SVM	3 conjuntos de dados públicos	Superior a 98% (Precisão)
[Ruby Elizabeth et al. 2024]	AlexNet modificado com segmentação K-means	HRF e RIGA	GDR: 91,6%-100% (HRF), 97,7%-100% (RIGA) (Taxa de Detecção de Glaucoma)
[Wang et al. 2024]	Modelo baseado em VLLM para ajuste probabilístico	REFUGE, G1020 e ORIGA	G1020: Acurácia: 0,9853, Sensibilidade: 0,9831, Especificidade: 0,9583
[Tan et al. 2024]	Modelo híbrido combinando redes totalmente conectadas e CNN	106 indivíduos normais e 164 pacientes com glaucoma	94,8% (Precisão), AROC: 97,9%, especificidade 100%
[Sivakumar and Penkova 2025]	Modelos híbridos ViT+ResNet, OWL-ViT+ResNet e Swin Transformer+ResNet	2874 casos novos	99,4% (Precisão)

um amplo conjunto de imagens retinianas. Inicialmente, as imagens foram adquiridas e processadas para possibilitar a extração manual da região de interesse do disco óptico. Em seguida, o modelo YOLOv8n foi treinado para detectar essa região. Após essa etapa, os recortes do disco óptico foram utilizados para treinar um modelo de classificação baseado na arquitetura ResNet-50. Por fim, foram avaliadas métricas como loss, acurácia, precisão, recall e coeficiente kappa para validar a eficácia do modelo desenvolvido.

### **3.1. Aquisição das imagens**

Nesta etapa, foi feita a aquisição das imagens. Utilizou-se somente a base de dados SMDG-19 [Kiefer et al. 2023]. Esta base de dados é uma mesclagem de 19 conjuntos de dados públicos contendo imagens de fundo de olho de glaucoma. Ela é considerada o maior repositório público de imagens de fundo ocular com glaucoma. Além disso, esse conjunto de dados é dividido em três classes, benigna, maligna, suspeita. Sendo 7549 da classe benigna, 4767 da classe maligna, 133 da classe suspeita.

Inicialmente, a classe “suspeito” foi removida devido ao seu baixo número de amostras, o que poderia prejudicar a capacidade de aprendizado do modelo. Após essa exclusão, a base de dados consistiu em 12.449 imagens distribuídas entre duas classes: maligna e benigna.

### **3.2. Detecção do Disco Óptico**

A detecção do disco óptico é uma etapa fundamental para o diagnóstico de diversas doenças oculares, principalmente o glaucoma. Os oftalmologistas utilizam o disco óptico como um ponto de referência na avaliação da saúde do nervo óptico, pois alterações na sua forma, tamanho e relação escavação/disco (CDR) são indicativos precoces da doença. Diante disso, um dos principais sinais analisados por esses profissionais é o aumento da escavação da cabeça do nervo óptico, que pode indicar a progressão do glaucoma [Mehmood et al. 2023].

Com o objetivo da extração da região do disco óptico (DO), foi inicialmente treinado um modelo de detecção YOLOv8n. Esse modelo foi treinado com 2.298 imagens e validado com 575 imagens, selecionadas aleatoriamente da base de dados, nas quais a região do disco óptico foi marcada manualmente por um especialista.

Com o modelo de detecção devidamente treinado, foi possível automatizar o recorte da área de interesse correspondentes ao disco óptico. Após essa etapa, o conjunto de imagens utilizado no treinamento do modelo de detecção foi descartado. Na próxima seção foram apresentados os resultados da avaliação.

Além disso, com o objetivo de avaliar a influência da área ao redor do disco óptico no desempenho do modelo de classificação, foram aplicadas expansões nas dimensões das bounding boxes detectadas pelo YOLOv8n. As regiões detectadas foram aumentadas em 20%, 100% e 200% para gerar variantes dos recortes utilizados na etapa de classificação.

### **3.3. Classificação**

Nesta etapa, foram utilizadas as regiões de interesse (ROIs), que foram obtidas utilizando o modelo YOLOv8n, permitindo a identificação precisa das áreas relevantes nas imagens analisadas. Com o objetivo de realizar a identificação do Glaucoma, foi utilizada a arquitetura ResNet-50. O modelo foi inicialmente pré-treinado com os pesos da IMAGENET e

implementado por meio da ferramenta PyTorch. O modelo pré-treinado foi refinado para a tarefa específica de classificação de imagens relacionadas ao glaucoma. Como etapa de pré-processamento, todas as imagens foram redimensionadas para  $224 \times 224$  pixels, garantindo compatibilidade com a entrada da rede neural utilizada.

Com o intuito de aprimorar a estimativa de desempenho do modelo e reduzir a variabilidade entre as diferentes amostras de treinamento e teste, foi aplicada a técnica de validação cruzada com 5 *folds*. Nesse procedimento, o conjunto de treinamento foi subdividido em cinco subconjuntos, e, a cada iteração, um dos subconjuntos foi utilizado para validação enquanto os demais foram empregados para o treinamento do modelo. Essa estratégia permitiu uma avaliação mais robusta da capacidade preditiva da rede. Cada variação dos subconjuntos treinou o modelo por 30 épocas, com *learning rate* de 0,0001 e decaimento de 0,1 a cada 7 épocas.

### 3.4. Métricas de Avaliação

Na avaliação de modelos de classificação, foi utilizado métricas como precisão, revocação (recall), acurácia e a função de perda (loss) para medir a eficácia e a eficiência do modelo na identificação correta das classes. Além dessas, o coeficiente kappa de Cohen é empregado para quantificar o grau de concordância entre classificadores [Powers 2020].

O Grad-CAM [Selvaraju et al. 2017] (Gradient-weighted Class Activation Mapping) também foi aplicado aos modelos refinados, gerando mapas de calor sobre as imagens de entrada e destacando as regiões com maior influência na predição da rede. Dessa forma é possível a interpretação dos modelos, permitindo a compreensão das características visuais que contribuíram para cada classificação.

## 4. Resultados e discussão

Esta seção apresenta os resultados dos testes realizados após o treinamento do modelo de detecção de glaucoma. Inicialmente realizou-se a avaliação da detecção da Yolov8n. Em seguida, foi apresentado a avaliação do classificador Resnet50. Por fim, foi utilizado o Grad-CAM para analisar a interpretabilidade do modelo de classificação.

### 4.1. Avaliação da Detecção do DO

A escolha dos modelos aplicados foi fundamentada em sua eficácia em tarefas similares. O modelo YOLOv8n foi selecionado por apresentar bom desempenho para diversas aplicações, incluindo outras aplicações de imagens médicas, conforme demonstrado em [Kubrak 2024] e [Lin et al. 2024]. Já a arquitetura ResNet-50 foi escolhida por seu histórico positivo em tarefas de visão computacional, incluindo classificações médicas, como apresentado em [da Silva Torres and Falcao 2006] e [Pattanaik et al. 2025].

Para avaliar o desempenho da detecção do disco óptico (DO), foram analisadas métricas como precisão, recall e tempo de inferência, que permitem medir sua acurácia e eficiência na identificação do DO. Essas métricas são essenciais para garantir que o modelo possa identificar corretamente a região do DO. Foram utilizadas um total de 575 imagens no teste, e os tempos médios de processamento foram de 1,7 ms para pré-processamento, 4,4 ms para inferência e 2,3 ms para pós-processamento. A Tabela 2 apresenta as principais métricas de avaliação do modelo YOLOv8:

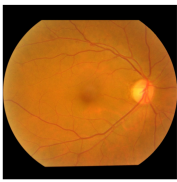




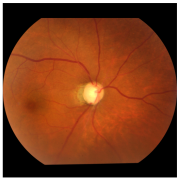
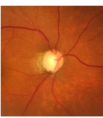



Modelo	Precisão	Recall	mAP@50	mAP@50-95	Inferência (ms)
YOLOv8n	1,000	0,998	0,995	0,890	4,4

**Tabela 2. Métricas de avaliação do modelo YOLOv8n na detecção do disco óptico.**

O modelo YOLOv8 apresentou alta precisão (100%) e recall (99,8%), indicando excelente desempenho na detecção de objetos. O mAP@50 foi 99,5%, enquanto o mAP@50-95 ficou em 89%, o que representa uma boa generalização em diferentes limiares de IoU. O tempo de inferência de 4,4 ms torna o modelo adequado para aplicações em tempo real.

#### 4.2. Avaliação da Classificação

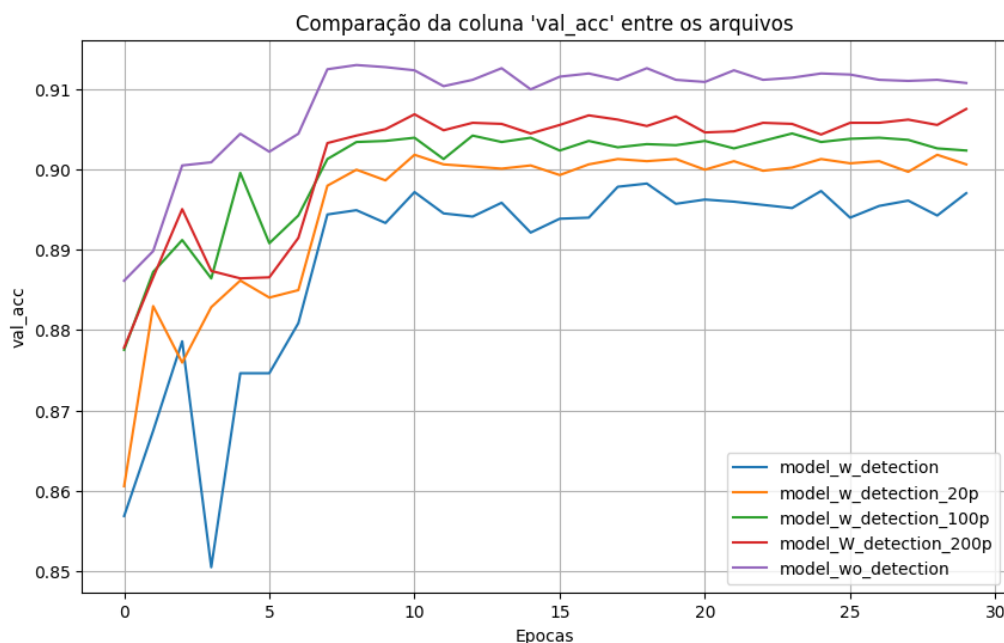
Na Figura 2 são apresentados exemplos de imagens originais e suas respectivas regiões detectadas.

BENIGNA					
MALIGNA					
	IMAGE ORIGINAL	RECORTE COM EXPANSÃO 200% DA BBOX	RECORTE COM EXPANSÃO 100% DA BBOX	RECORTE COM EXPANSÃO 20% DA BBOX	DETECÇÃO SEM EXPANSÃO

**Figura 2. Exemplos de imagens originais e suas respectivas região detectadas**

Na Figura 3, é apresentada a curva das médias da função de acurácia ao longo das épocas de treinamento, mostrando a redução progressiva do erro do modelo. Na legenda do gráfico, a linha identificada como `model_w_detection` representa o treinamento do modelo utilizando a região detectada pelo YOLOv8. As linhas `model_w_detection20p`, `model_w_detection100p` e `model_w_detection200p` correspondem ao mesmo modelo, porém com um aumento de 20%, 100% e 200% no tamanho do bounding box detectado, respectivamente. Já a linha `model_wo_detection` representa o modelo treinado utilizando a imagem original, sem qualquer ajuste na região detectada. A linha contínua representa a média dos valores de perda por época. No entanto, ambos os modelos evoluíram ao longo das épocas, mas o modelo treinado sem detecção obteve melhor desempenho.

A Tabela 3 apresenta as médias dos resultados obtidos pelo cross-validation no modelo ResNet50. Os valores de variância não foram incluídos, pois, em todos os casos, permaneceram abaixo de 0,0001. A análise dos resultados presentes na Tabela 3 e na Figura 3 indicam que, à medida que a região de interesse da imagem recortada se concentra mais no disco óptico, o desempenho do modelo tende a diminuir. Essa observação pode indicar que o modelo não utilizou apenas o DO para realizar a classificação.



**Figura 3. Evolução da Função de Acurácia Durante o Treinamento utilizando o conjunto de validação.**

**Tabela 3. Métricas de desempenho dos modelos na classificação de glaucoma considerando diferentes estratégias de detecção**

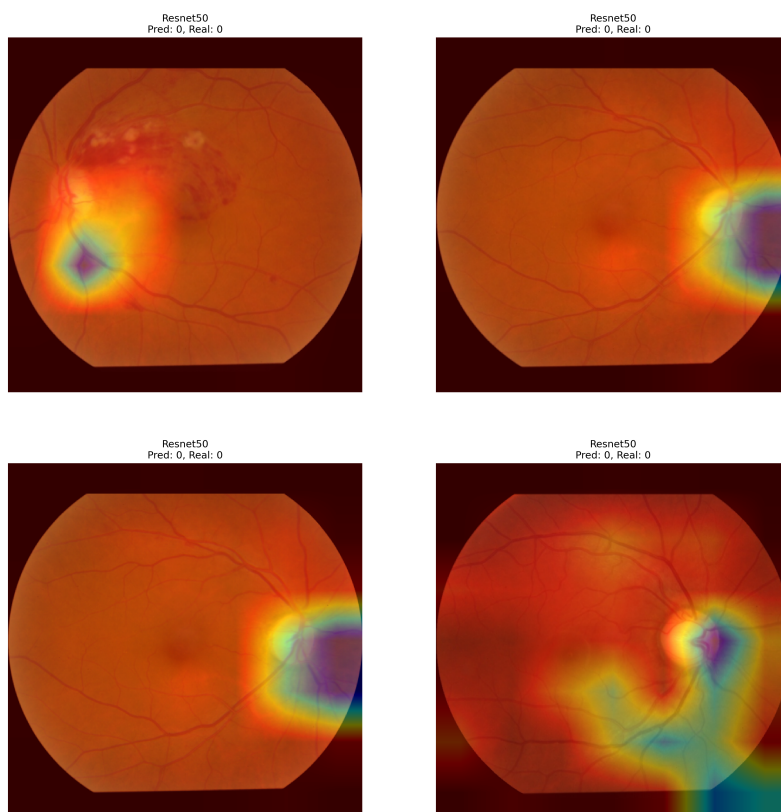
	Precisão	Recall	Acurácia	Kappa
Sem detecção	<b>0,917</b>	<b>0,911</b>	<b>0,915</b>	<b>0,827</b>
Com detecção	0,896	0,896	0,897	0,791
Com detecção 20p	0,902	0,901	0,904	0,804
Com detecção 100p	0,910	0,909	0,911	0,819
Com detecção 200p	0,910	0,909	0,911	0,819

### 4.3. Resultados do Grad-CAM

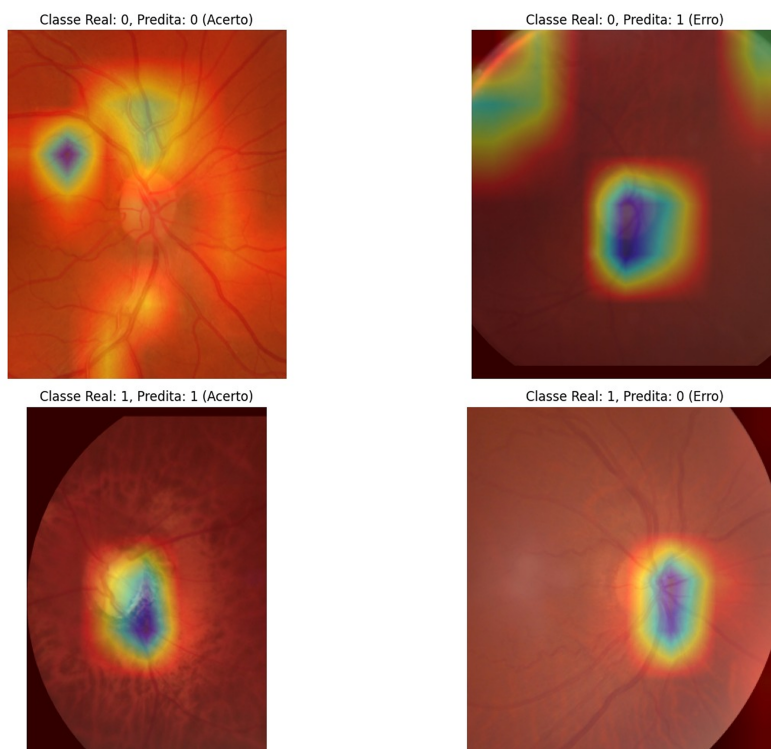
Para analisar a interpretabilidade do modelo proposto, foi aplicada a técnica Grad-CAM (Gradient-weighted Class Activation Mapping) sobre as imagens de entrada. Com isso, foi possível visualizar as regiões de maior influência na predição da rede, destacando as áreas relevantes utilizadas pelo modelo para auxiliar a tomada de decisão. As Figuras 4 e 5 apresentam os mapas de ativação gerados pelo Grad-CAM, evidenciando as características mais relevantes para a classificação.

Com base nos mapas de calor apresentados na Figura 4, nota-se que o modelo não focou exclusivamente nas regiões centrais do disco óptico, mas também em áreas periféricas da retina. Essa distribuição de atenção pode indicar que o modelo está considerando padrões adicionais, diferentes dos presentes no disco óptico, o que pode impactar nos resultados das predições.

Com base nos mapas de calor apresentados na Figura 5, destaca-se que o modelo também leva em consideração regiões externas ao DO, indicando que está aprendendo características consideradas relevantes para a detecção da doença e que são localizadas



**Figura 4. Grad-CAM aplicado sem detecção prévia.**



**Figura 5. Grad-CAM aplicado com detecção aumentada 200%.**

fora da região de interesse.

## 5. Conclusão

Este trabalho abordou o uso de aprendizado profundo para a classificação de imagens do fundo do olho. Para a criação do método foi investigado o uso da YoloV8n para a detecção da região do DO e o impacto dessa região na classificação das imagens em Glaucomatosas ou não. Os resultados obtidos foram promissores e são comparáveis com os trabalhos presentes na literatura. A principal conclusão obtida foi que a utilização da imagem completa produziu resultados ligeiramente superiores ao uso da região detectada do DO. Outra conclusão obtida foi que os resultados melhoram a medida que se aumenta o tamanho da região ao redor do DO utilizada para a classificação. Essa informação foi confirmada pela análise do Grad-CAM, que revelou que o modelo não se limitou a extração de características da região do DO, mas também considerou regiões periféricas da retina, sugerindo a influência de padrões fora da região do DO.

Para futuras pesquisas, recomenda-se a exploração de novas arquiteturas de CNNs. Além disso, sugere-se a implementação de novas abordagens de pré-processamento para melhorar a precisão da classificação como o aumento de dados para o balanceamento das classes.

## Referências

- da Saúde, M. (2021). Glaucoma: diagnóstico precoce e tratamento evitam perda da visão. <https://www.gov.br/saude/pt-br/assuntos/noticias/2021/maio/glaucoma-diagnostico-precoce-e-tratamento-evitam-perda-da-visao>. Acesso em: 26 novembro 2024.
- da Silva Torres, R. and Falcao, A. X. (2006). Content-based image retrieval: theory and applications. *RITA*, 13(2):161–185.
- Elmannai, H., Alhussan, A. A., Ayadi, M., Ksibi, A., et al. (2024). An improved deep learning framework for automated optic disc localization and glaucoma detection. *CMES-Computer Modeling in Engineering & Sciences*, 140(2).
- Hayreh, S. S. (2007). Neovascular glaucoma. *Progress in retinal and eye research*, 26(5):470–485.
- Kiefer, R., Abid, M., Ardali, M. R., Steen, J., and Amjadian, E. (2023). A catalog of public glaucoma datasets for machine learning applications: A detailed description and analysis of public glaucoma datasets available to machine learning engineers tackling glaucoma-related problems using retinal fundus images and oct images. In *Proceedings of the 2023 7th International Conference on Information System and Data Mining*.
- Kubrak, T. (2024). Automated detection of glaucoma and diagnostic features for justraigs challenge. In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1–3. IEEE.
- Lin, H., Apostolidis, C., and Katsaggelos, A. K. (2024). Brighteye: Glaucoma screening with color fundus photographs based on vision transformer. In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1–4. IEEE.

- Mehmood, M., Naveed, K., Aurangzeb, K., Khan, H. A., Alhussein, M., and Naqvi, S. S. (2023). Eddense-net: Fully dense encoder decoder network for joint segmentation of optic cup and disc. *arXiv preprint arXiv:2308.10192*.
- Pattanaik, S., Behera, S., Majhi, S. K., Pradhan, R., and Dwibedy, P. (2025). An ensemble stacked bi-lstm with resnet50 method for glaucoma classification in iot framework: An ensemble method for glaucoma classification iot framework. *Journal of Scientific & Industrial Research (JSIR)*, 84(1):24–35.
- Powers, D. M. (2020). Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.
- Ruby Elizabeth, J., Kesavaraja, D., and Ebenezer Juliet, S. (2024). A deep learning model based glaucoma detection using retinal images. *Journal of Intelligent & Fuzzy Systems*, (Preprint):1–12.
- Sanghavi, J. and Kurhekar, M. (2024). An efficient framework for optic disk segmentation and classification of glaucoma on fundus images. *Biomedical Signal Processing and Control*, 89:105770.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pages 618–626.
- Shanmugam, P., Raja, J., and Pitchai, R. (2021). An automatic recognition of glaucoma in fundus images using deep learning and random forest classifier. *Applied Soft Computing*, 109:107512.
- Silva, I. M. (2016). Diagnóstico do glaucoma congênito—revisão sistemática. *Trabalho de Conclusão de Curso (Graduação) - Medicina (Faculdade de Medicina)*.
- Sivakumar, R. and Penkova, A. (2025). Enhancing glaucoma detection through multi-modal integration of retinal images and clinical biomarkers. *Engineering Applications of Artificial Intelligence*, 143:110010.
- Sułot, D., Alonso-Caneiro, D., Ksieniewicz, P., Krzyzanowska-Berkowska, P., and Iskander, D. R. (2021). Glaucoma classification based on scanning laser ophthalmoscopic images using a deep learning ensemble method. *Plos one*, 16(6):e0252339.
- Tamim, N., Elshrkawey, M., and Nassar, H. (2021). Accurate diagnosis of diabetic retinopathy and glaucoma using retinal fundus images based on hybrid features and genetic algorithm. *Applied Sciences*, 11(13):6178.
- Tan, O., Greenfield, D. S., Francis, B. A., Varma, R., Schuman, J. S., Huang, D., and Choi, D. (2024). A hybrid deep learning classification of perimetric glaucoma using peripapillary nerve fiber layer reflectance and other oct parameters from three anatomy regions. *ArXiv*.
- Wang, S., Kim, B., Kang, J., and Eom, D.-S. (2024). Precision diagnosis of glaucoma with vllm ensemble deep learning. *Applied Sciences*, 14(11):4588.
- Yunitasari, D. A., Sigit, R., and Harsono, T. (2021). Glaucoma detection based on cup-to-disc ratio in retinal fundus image using support vector machine. In *2021 International Electronics Symposium (IES)*, pages 368–373.