

Segmentação de Imagens de Fundo de Olho e Identificação de Padrão ISNT para Auxílio ao Diagnóstico de Glaucoma

Maísa Fernandes Gomes

Programa de Pós-Graduação em Computação Aplicada
Universidade do Estado de Santa Catarina (UDESC)
Joinville, Brasil
gomesfmaisa@gmail.com

Rafael Stubs Parpinelli

Programa de Pós-Graduação em Computação Aplicada
Universidade do Estado de Santa Catarina (UDESC)
Joinville, Brasil
rafael.parpinelli@udesc.br

Abstract—O glaucoma é uma doença silenciosa e progressiva que pode levar os pacientes a total perda de visão, por essa razão o diagnóstico precoce é importante. Métodos de visão computacional podem servir como um auxílio aos especialistas na detecção precoce. Esse trabalho teve como objetivo utilizar um modelo de rede neural artificial para detectar a área de interesse para detecção do glaucoma, e também fazer anotações da área de escavação e disco óptico. O modelo de detecção da área de interesse alcançou um mAP de 99,2%, e o modelo de detecção das áreas de escavação e disco óptico alcançou um mAP de 92%.

Index Terms—Glaucoma. Redes neurais artificiais. Segmentação. Imagem de fundo de olho.

I. INTRODUÇÃO

Imagens de fundo de olho são imagens médicas que possibilitam a detecção de doenças oculares. Uma dessas doenças é o glaucoma que pode ser detectado através da análise médica na cabeça do nervo óptico, verificando o tamanho da escavação e do disco óptico [1]. O glaucoma é um conjunto de doenças que afetam a área do nervo óptico causando o aumento da escavação, que causa a perda de visão no paciente. A doença é progressiva e o diagnóstico precoce é importante para o tratamento, prevenindo que a doença chegue no seu estágio mais avançado [2].

A parte mais importante para o diagnóstico do glaucoma é a área do disco óptico. Nesta região são feitas as análises que identificam a área da escavação e a área do disco óptico comparando o tamanho do raio dessas duas regiões, ou analisando a distância do disco óptico em relação às regiões: Inferior, Superior, Nasal e Temporal (ISNT) [1].

O diagnóstico do glaucoma é uma tarefa desafiadora e custosa, uma vez que depende de muitas observações feitas por oftalmologistas especialistas na área. Com o avanço do uso de aprendizado profundo na área de visão computacional, alguns trabalhos tiveram como objetivo utilizar esses modelos para auxiliar os oftalmologistas na detecção do glaucoma em imagens de fundo de olho [3] [4].

As pesquisas feitas podem ser divididas em dois tipos: As que focam no treinamento dos modelos para classificação binária das imagens entre a classe glaucoma e a classe normal,

e as pesquisas focadas em utilizar métodos para segmentar a área do disco óptico e escavação [3] [4].

Tendo em vista a importância da segmentação da área do nervo óptico, o presente trabalho tem como objetivo utilizar um modelo de Rede Neural Convolucional para a detecção e segmentação das regiões do nervo óptico em imagem de fundo de olho. Foram treinados dois modelos, sendo o primeiro treinado para detectar a região do disco óptico e fazer o recorte desta área; O segundo modelo sendo treinado para a detecção da área do nervo óptico e escavação, para posteriormente fazer as anotações na imagem nestas áreas e calcular a relação ISNT.

As próximas sessões estão divididas da seguinte forma. Na Seção 2 está a parte de fundamentação teórica e os trabalhos relacionados, após a Seção 3 explica a metodologia usada para o desenvolvimento deste trabalho, os resultados obtidos estão apresentados na Seção 4 e a última seção apresenta a conclusão e os trabalhos futuros.

II. FUNDAMENTAÇÃO TEÓRICA

A. Imagens de Fundo de Olho e Glaucoma

Imagens de fundo de olho são imagens médicas obtidas através do exame de fundo de olho. Esse exame possibilita que o oftalmologista faça análises na saúde ocular do paciente. A Figura 1 apresenta uma imagem de fundo de olho. O disco óptico está localizado na região circundada pelo círculo branco, nesta região estão presentes diversos axônios interligados formando um conjunto fibroso responsáveis por transmitir a visão a área cerebral, esse conjunto é chamado de nervo óptico.

O disco óptico não é completamente preenchido pelo nervo óptico formando uma área chamada de escavação, na Figura 1 a região mais amarelada dentro do disco óptico representa a escavação. A escavação é presente em todas as pessoas, porém o aumento de seu tamanho pode ser um indício de doença, como o glaucoma.

O glaucoma é uma doença caracterizada por causar danos na cabeça do nervo óptico podendo levar o paciente à cegueira. A doença do glaucoma faz com que a escavação aumente devido à morte progressiva dos axônios presentes no disco óptico

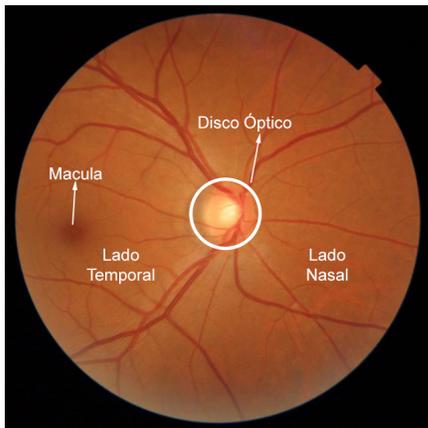


Fig. 1. Imagem de Fundo de Olho.

fazendo com que o paciente perca a visão gradativamente, isso acontece por diversos fatores como o aumento da pressão intraocular [2].

A doença é progressiva e silenciosa, por essa razão o diagnóstico precoce é importante para garantir a qualidade de vida do paciente, tratando a doença sem que ela esteja no estágio mais grave. Para o diagnóstico de glaucoma a parte mais significativa é a área do disco óptico, onde são feitas as análises [2] [4].

A região do disco é dividida em 4 setores, inferior, superior, nasal e temporal, conforme apresentado da figura 2. Em um disco óptico saudável as bordas tem a espessura seguindo a ordem decrescente: inferior, superior, bordas nasais e temporais, formando a relação ISNT onde $I \geq S \geq N \geq T$. A partir desta relação é criada a regra ISNT que calcula a margem do disco óptico em relação a escavação verificando se a relação está correta. A doença do glaucoma costuma danificar primeiramente as fibras do nervo óptico superior e inferior antes das temporal e nasal, isso faz com que o relacionamento ISNT se altere. Com isso, a regra verifica a distância das bordas nas quatro direções e ajuda no diagnóstico do glaucoma em seu estágio inicial [1].

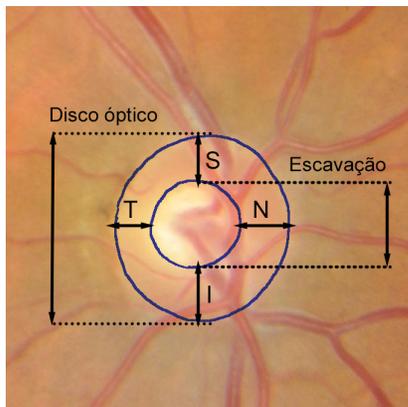


Fig. 2. Imagem da regra ISNT.

Outro indício de que o paciente possa estar com glaucoma

é a análise do crescimento da área de escavação, isso porque a doença aumenta o tamanho da escavação. Para essa análise é utilizada a métrica que calcula relação do escavação/disco óptico que pode indicar uma assimetria sendo um indício de glaucoma [1].

A análise feita nas imagens de fundo de olho para detectar o glaucoma é algo muito custoso, uma vez que os médicos têm que fazer diversas análises entre a escavação e o disco óptico. Nos últimos anos as técnicas de visões computacionais tem apresentado significativos avanços principalmente com o uso de técnicas de aprendizado profundo, as tarefas de visões computacionais estão relacionadas a classificação de imagens, detecção de objetos, segmentação [4].

Devido ao fato de muitas doenças serem detectadas a partir de imagens médicas a aplicação de técnicas de aprendizado profundo podem ser muito benéficas, como por exemplo a automatização da análise da escavação e disco óptico pode auxiliar o oftalmologista no diagnóstico do glaucoma. Existem dois tipos de pesquisa que focam nesse objetivo: as que focam em classificação binária, e as que focam em classificações a partir de indicadores médicos como a relação disco/escavação e a regra ISNT [3].

As que focam em classificação binária tem como objetivo utilizar modelos para classificar uma imagem médica, como imagens de fundo de olho, com glaucoma ou não. Dentre os modelos desenvolvidos, muitos trabalhos apresentaram resultados promissores utilizando as redes neurais convolucionais. O outro foco é as pesquisas que utilizam métodos para alcançar a identificação e a segmentação da área do nervo óptico e a área da escavação para então fazer os cálculos e conseguir identificar a presença do glaucoma [3].

B. Redes Neurais Convolucionais

Um dos modelos responsáveis pelo grande avanço do uso de aprendizado profundo em tarefas de visão computacional, são as Redes Neurais Convolucionais (CNN). São modelos de redes neurais artificiais que possuem camadas de convolução em sua arquitetura antes das camadas de classificação. Essas camadas de convolução são responsáveis por extrair características específicas da imagens, visando melhor o desempenho da classificação [5].

Em métodos supervisionados para a tarefa de classificação de imagens esses modelos aprendem a classificar uma entrada retornando a classe em que essa entrada pertence e a porcentagem de confiança do modelo naquela classificação. Outra tarefa é utilizar esse modelo para a detecção de objetos imagens, dessa forma o modelo recebe uma entrada e localiza o objeto na imagem e retorna a classificação, localização deste objeto, e a confiança da predição feita [6].

A You Only Look Once (YOLO), é um modelo desenvolvido para tarefas de visão computacional como a classificação de imagens e detecção de objetos, o nome do modelo está relacionado a proposta da arquitetura analisar somente uma vez a imagem e prever quais os objetos estão presentes na imagem, propondo assim um modelo preciso,

menos custoso em seu treinamento e mais rápido em suas predições [7].

A arquitetura YOLO possui 8 versões, sendo a última lançada no ano de 2023. Sendo essa versão, segundo os desenvolvedores, a versão mais rápida e precisa comparada aos outros modelos da YOLO disponíveis. Para esse teste as versões foram testadas com o COCO dataset, um grande conjunto de dados para a avaliação de modelos de detecção e segmentação de objetos.

Uma técnica utilizada em modelos de aprendizado de máquina é o *transfer learning*. Nesta técnica o modelo utilizado não começa com os pesos ajustados de forma aleatória e sim com os pesos adquiridos durante o treinamento com um grande conjunto de imagens. Com isso, os pesos são transferidos para o modelo, que é treinado para a resolução do problema desejado. O modelo YOLO é pré-treinado com o banco de dados COCO para detecção e segmentação, e com o conjunto ImageNet, que possui milhares de imagens para as tarefas de classificação, também foi utilizado para o pré-treinamento da classificação. [8].

C. Trabalhos Correlatos

Alguns trabalhos focaram em fazer a detecção e segmentação da área da escavação, na revisão feita pelo [4] em trabalhos que usam aprendizado de máquina e aprendizado profundo em segmentação do disco óptico. O trabalho aponta três tipos de modelos de aprendizado profundo mais utilizados para a segmentação, sendo as redes neurais generativas, U-net, e as que usam redes neurais convolucionais.

Das pesquisas que utilizam redes neurais convolucionais para segmentação do disco óptico, o trabalho [9] propôs o desenvolvimento de uma rede neural para detecção do disco óptico, com base no banco de dados ORIGA, o modelo alcançou um desempenho de 93.1% para a segmentação do disco óptico usando Faster R-CNN.

No trabalho apresentado em [10] foi desenvolvido um modelo para a segmentação do disco óptico e classificação das imagens, esse trabalho usou uma junção de banco de dados totalizando 2787 imagens. O trabalho foi dividido em 2 partes onde a primeira utilizou uma DeepLabv3+ combinada com uma MobileNet alcançando uma acurácia de 99.7%, para a detecção e extração do disco óptico e a segunda parte focou no modelo para classificação binária.

Um dos pontos apontados pelos pesquisadores é o fato de os modelos apresentados na pesquisa geralmente serem testados com o próprio conjunto de imagens usadas no treino. Isso dificulta a análise do desempenho do modelo em diferentes conjuntos de dados, essa análise é importante uma vez que exames de fundo de olho são executados de maneiras diferentes dependendo do local em que foram realizados.

A revisão feita por [3] apresenta trabalhos que utilizaram técnicas para a análise e segmentação do disco óptico e escavação. Nenhum dos trabalhos apresentados utiliza técnicas de redes neurais convolucionais para a segmentação. Entretanto, alguns trabalhos focam em trazer o cálculo do ISNT e

o cálculo da escavação utilizando métodos de processamento de imagens.

III. METODOLOGIA

Este trabalho traz a aplicação de uma CNN para a segmentação de imagens de fundo de olho, detecção do disco óptico e detecção da escavação. Para tal, foram utilizados dois modelos. Um modelo com a tarefa de detectar a área do disco óptico e fazer o recorte desta região de interesse; e outro modelo treinado para a tarefa de detecção do disco óptico e área de escavação, para então fazer as anotações nas imagens com a relação ISNT.

O fluxograma do primeiro modelo está representado na Figura 3, onde a primeira etapa é o pré-processamento que consiste nas anotações das imagens na área de interesse, para depois o treinamento no modelo escolhido e então a avaliação dos resultados obtidos. A segunda parte do teste consiste em utilizar o modelo treinado para prever as imagens de fundo de olho, a partir das coordenadas das áreas detectadas.

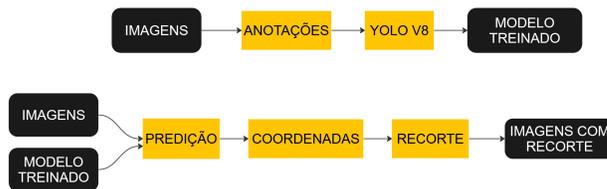


Fig. 3. Fluxograma do primeiro modelo.

O segundo modelo está representado na Figura 4, onde primeiramente se utiliza as imagens de fundo de olho recortadas e as anotações dos especialistas para fazer as anotações nas imagens e então treinar e avaliar o modelo. Com o modelo treinado é utilizado o conjunto de imagens de fundo de olho para fazer a predição da área e então a anotação das imagens.

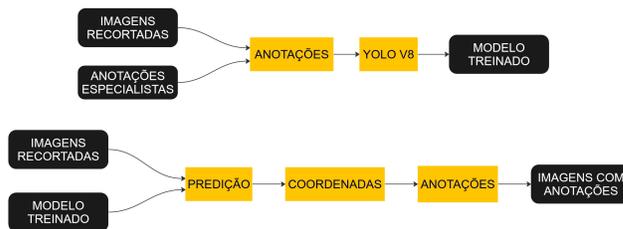


Fig. 4. Fluxograma do segundo modelo.

O modelo selecionado para o desenvolvimento da primeira parte do trabalho foi o YOLO v8. Foi utilizada a versão pré-treinada do modelo, que serviu como base para o treinamento da tarefa específica de identificação da área de interesse. Desta maneira, realizou-se transferência de aprendizado do modelo YOLO v8 para a aplicação em imagens de fundo de olho.

Para o treinamento deste modelo foi utilizado um conjunto de imagens com exames de fundo de olho de mais de 2071

pacientes. O conjunto utilizado possui imagens com grande variação de saturação e posições do disco óptico.

Desta base de dados foram selecionadas, de forma aleatória, 1033 imagens de exames de fundo de olho e então feita a anotação da área de interesse. O pré-processamento realizado nas imagens foi a conversão de formato DICOM para JPEG, e o redimensionamento das imagens para 640x640. Posteriormente as imagens foram divididas entre treino, validação e teste sendo 70% para treino, 20% para validação e 10% para teste. Dando o total de 718 imagens para treinamento, 193 para validação e 122 para teste.

A partir das configurações padrões do modelo algumas alterações de valores das configurações de treinamento foram feitas de forma empírica. Com isso foi atribuído ao modelo um treinamento por 200 épocas, com uma taxa de aprendizado inicial de 0.005, e *batch size* de tamanho 16. O otimizador escolhido foi o Adam.

Após feito o treinamento e validação, foi implementado o algoritmo para fazer o recorte da área de interesse a partir da predição. Apesar de o algoritmo YOLO v8 já possibilitar que o modelo faça o recorte a partir da região detectada, esse recorte pode ser retangular e o objetivo é que esse recorte seja no formato quadrado para facilitar a utilização desta imagem em outros modelos.

Após o treinamento, o modelo treinado foi utilizado para prever as imagens de teste, imagens que o modelo nunca teve contato durante o seu treinamento. Durante a predição foram salvas as coordenadas da caixa delimitadora detectada em cada entrada. Posteriormente, essas coordenadas foram utilizadas para fazer o recorte da área de interesse. Para garantir que o recorte da imagem fosse quadrado, foi utilizada a coordenada da lateral da caixa delimitadora que tivesse o maior comprimento para o recorte de todas as laterais. Para o recorte das imagens, foi utilizada a biblioteca Pillow.

Em seguida foi executado um segundo teste utilizando imagens providas de diferentes conjuntos de imagens. São eles:

- DR HAGIS [11]: Consiste em imagens de fundo de olho de pacientes que possuíam retinopatia diabética no Reino Unido, dentre esses pacientes alguns também possuíam glaucoma. Apenas as imagens de pacientes que possuíam glaucoma foram selecionadas.
- VEIRC [12] [13]: Consistem em imagens de fundo de olho coletadas em pacientes da Índia. A rotulação das imagens entre normal, glaucoma e suspeito foi feita por oftalmologistas com auxílio da relação escavação-disco (CDR).
- Data on OCT and Fundus Images [14] : O conjunto de dados contém imagens de fundo de olho, classificados em glaucoma e normal. Contendo imagens dos dois olhos de 26 pacientes, com anotação relação escavação-disco (CDR) feita por um oftalmologista.
- BEH [15]: Compreende imagens de pacientes de Bangladesh. O diagnóstico de glaucoma na cabeça do nervo óptico (ONH) foi realizado por dois profissionais de doenças oculares.

- Papila [16]: O conjunto de imagens de 244 pacientes, possui imagens dos dois olhos. Além disso possui anotação do disco óptico e da área da escavação.
- Riga [17] : Esse dataset é um junção de 3 bases, Magrabi, Mesidor e Bin Rushed Ophthalmic, contendo respectivamente 95,460 e 195 imagens. Todas as imagens possuem imagens com marcações médicas da área do disco óptico e da área da escavação.

Na Tabela I estão relacionados a quantidade de imagens presentes em cada base de dados.

TABELA I
QUANTIDADE DE IMAGENS EM CADA BASE DE DADOS

Nome	Quantidade de Imagens
[11]	10
[12] [13]	257
[14]	50
[15]	634
[16]	488
[17]	750

Para o segundo modelo proposto neste trabalho, que é a detecção do disco óptico e da escavação, foram utilizadas imagens das bases de dados Magrabi e Papila. Essas bases de dados possuem anotações médicas indicando a área do disco óptico e a área da escavação.

Esses conjuntos possuem imagens inéditas do fundo de olho, e para um melhor treinamento foi definido que as imagens passassem pelo primeiro modelo treinado para o recorte na área de interesse. Após as imagens terem sido cortadas, foi feita a parte de anotação das regiões do disco óptico e escavação. As anotações feitas pelos médicos serviram como base para as anotações feitas nas imagens para o treinamento deste segundo modelo.

A Figura 5 apresenta exemplos de anotações realizadas. Na linha A está demonstrada as anotações feitas para o treinamento do modelo 1 onde o interesse era a detecção da área do disco óptico. A linha B apresenta as marcações feitas nas imagens já recortada na região de interesse, as marcações da escavação e disco óptico foram realizadas com base nas anotações dos especialistas, essas imagens foram utilizadas para o treinamento do modelo 2.

Ao total foram empregadas 173 imagens de fundo de olho com a anotação do disco óptico e da escavação. Além do recorte e redimensionamento para 640x640, nenhum outro pré-processamento foi realizado. Nesse modelo, também foi mantida a divisão 70%,20%,10% para treino, validação e teste respectivamente. Com isso, foram 117 imagens para treinamento, 35 imagens da validação e 21 para teste.

Para essa parte do trabalho também foi escolhido o modelo YOLO na versão 8. Para este modelo também se empregou transferência de aprendizado e os seguintes parâmetros definidos empiricamente: 350 de épocas, a taxa de aprendizado inicial de 0.005 e o *batch size* de tamanho 16 e o otimizador escolhido foi o Adam.

Após o treinamento e validação, foi feita uma predição com as imagens de teste e as coordenadas da área de detecção foram

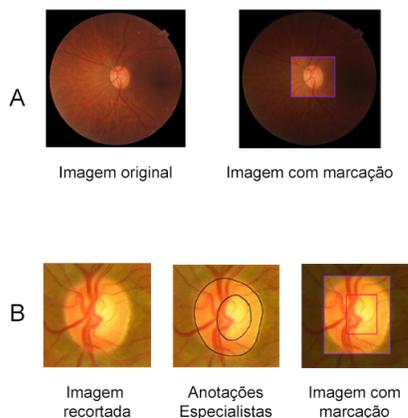


Fig. 5. Anotações feita para o treinamento do modelo. Linha A apresenta as anotações feitas nas imagens primeiro modelo. Linha B apresenta as anotações feitas para o treinamento do segundo modelo.

utilizadas para anotar um círculo na área do disco óptico e na área da escavação. Após isso foi anotado os pontos para o cálculo da relação ISNT.

IV. EXPERIMENTOS, RESULTADOS E ANÁLISES

A. Protocolo de Experimentos

Todo o projeto foi implementado no sistema operacional Ubuntu 20.04 com 16 GB RAM de memória, processador AMD Ryzen 7 2700X e com a placa de video NVIDIA GeForce RTX 2060 SUPER. A implementação foi realizada com a linguagem Python em sua versão 3.7, a biblioteca Ultralytics foi utilizada para o treinamento da rede YOLO v8. Para os recortes e anotações das imagens a biblioteca Pillow foi utilizada.

As métricas de avaliação empregadas foram *recall* e *precisão*. A métrica de precisão calcula o quão preciso o modelo está em suas classificações positivas, para isso é medido entre todas as classificações positivas quais estão certas. A métrica de *recall* mede dentre todas as predições verdadeiras esperadas, quantas estão corretas.

Modelos de detecção de objetos, como é o caso do modelo utilizado neste trabalho, utiliza uma métrica chamada mAP que é a média das precisões, para isso além de utilizar as métricas de *recall* e *precisão*, ele utiliza uma métrica chamada Intersecção sobre União (IoU). A métrica IoU, tem como base a caixa delimitadora, essa métrica tem como objetivo indicar o quanto a caixa delimitadora prevista se sobrepõem a caixa delimitadora exata.

A métrica mAP indica o quão preciso o modelo está para todas as classes. Na tarefa de detecção de objeto a precisão se utiliza da métrica IoU como a métricas mAP50, o números 50 refere ao grau de intersecção das caixas delimitadoras, ou seja, a mAP50 considera que se a caixa predita tem mais de 50% de intersecção com a caixa delimitadora real ela será considerada uma detecção correta.

Além da precisão, outra métrica usada para avaliar o modelo são os gráficos de perda que informam a discrepância entre

a caixa delimitadora predita e a caixa delimitadora real. São três tipos de perda, a cls, dlf, e box.

B. Resultados e Análises

1) *Modelo Para Detecção Da Área de Interesse:* O primeiro experimento se refere a detecção da área do disco óptico. Os resultados do treinamento e validação do primeiro experimento estão apresentados na Figura 6, onde é apresentado os gráficos de perda do modelo, dentre as perdas é observado que a perda da classificação é a que tem o menor valor, o que demonstra que o modelo executa bem a tarefa de classificar a área de interesse. A mAP50 do modelo ficou em 0.995, esse resultado demonstra que durante o treinamento o modelo conseguiu identificar a área de interesse nas imagens.

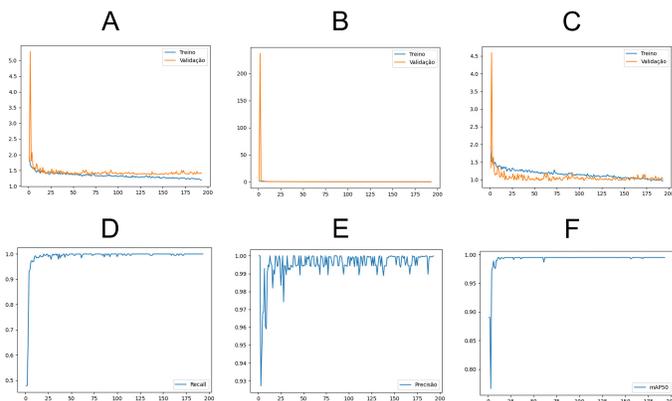


Fig. 6. Gráficos de desempenho do modelo 1, o eixo x representa a quantidade de épocas e o eixo y representa o valor alcançado em cada métrica. (A) Gráfico de perda dlf. (B) Gráfico de cls. (C) Gráfico de perda box. (D) Gráfico de Recall. (E) Gráfico de Precisão. (F) Gráfico de mAP50

Após foi feito o teste com as imagens da mesma base de dados, com esse conjunto de imagens o resultado foi de mAP50 0.992. Os resultados alcançados no teste e na validação estão apresentados na Tabela II e é possível visualizar que o modelo conseguiu um desempenho promissor tanto no treinamento quanto no teste. A Figura 7 apresenta o resultado do recorte utilizando as coordenadas das detecções feitas.

TABELA II
RESULTADOS OBTIDOS NA VALIDAÇÃO E NO TESTE DO MODELO 1

	Precision	Recall	mAP50	mAP50-95
Validação	0.99	0.992	0.995	0.7
Teste	0.99	0.992	0.992	0.698

O modelo treinado foi testado em imagens de diferentes conjuntos de dados. Neste teste foi analisado em quais imagens o modelo conseguiu detectar o disco óptico. Apesar dessas imagens não estarem no conjunto de dados do treinamento, o modelo conseguiu identificar a região do disco óptico na maioria das imagens. Foi analisado que em imagens que tinham a mácula mais aparente o modelo detectava a mácula como o disco óptico, por essa razão foi usado uma confiança de 0.5 na predição, ou seja apenas eram anotadas nas imagens

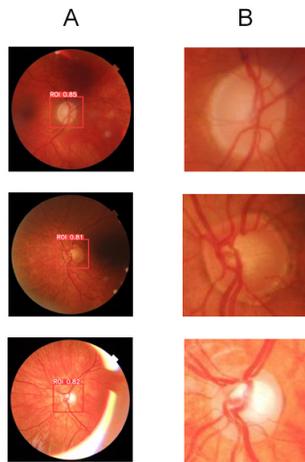


Fig. 7. Resultados do recorte das imagens usadas para o teste. Coluna A as imagem com a detecção feita pelo modelo. Coluna B o resultado do recorte feito.

às detecção que o modelo apresentava uma confiança acima de 0.50%. Os resultados alcançados neste teste estão apresentados na Tabela III.

TABELA III
RESULTADO DA DETECÇÃO EM DIFERENTES CONJUNTOS DE IMAGENS

Nome do Conjunto de Imagens Utilizado	Imagens	
	Detecção Correta	Detecção Incorreta
[11]	9	1
[12] [13]	254	3
[14]	49	1
[15]	630	4
[16]	483	5
[17]	747	3

São consideradas detecções incorretas quando um modelo não consegue identificar a área de interesse, quando o modelo identifica errado, ou quando a caixa delimitadora corta o disco óptico. Algumas classificações incorretas estão demonstradas na Figura 8. Nessas imagens é visualizado que a coloração escura impossibilita o modelo de fazer a detecção.

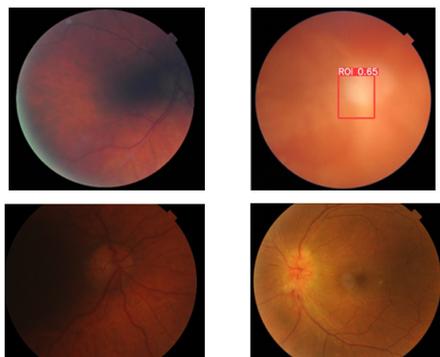


Fig. 8. Resultados de detecções incorretas.

Após esse teste foram realizados os recortes nas imagens utilizando as coordenadas da caixa delimitadora detectado pelos modelos, e os resultados obtidos estão demonstrados na Figura 9. Como mencionado acima, o modelo apresentou um desempenho promissor em recortar as imagens na região do disco óptico. Mesmo sendo imagens de distintos conjuntos de dados, o modelo detectou de forma correta 99% das imagens.

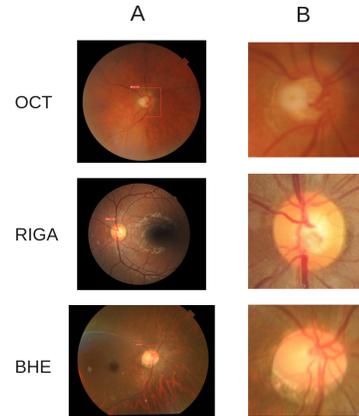


Fig. 9. Resultados dos recortes. Na primeira coluna estão os nomes do banco de dados utilizado. Coluna A as imagem com a detecção feita pelo modelo. Coluna B os resultado do recorte feito.

2) *Modelo Para Identificação do Disco Óptico e Escavação*: Para o treinamento do segundo modelo foram utilizadas imagens recortadas com as anotações da área do disco óptico e escavação. O resultado do treinamento e validação estão apresentados na Figura 10. Como esse modelo possui detecção para duas classes, temos a métrica para cada classe, para a detecção da escavação foi obtido uma mAP50 de 0.846 e para a detecção do disco óptico a mAP50 foi de 0.995, com esses resultados é possível constatar que modelo teve mais facilidade em identificar o disco óptico a escavação

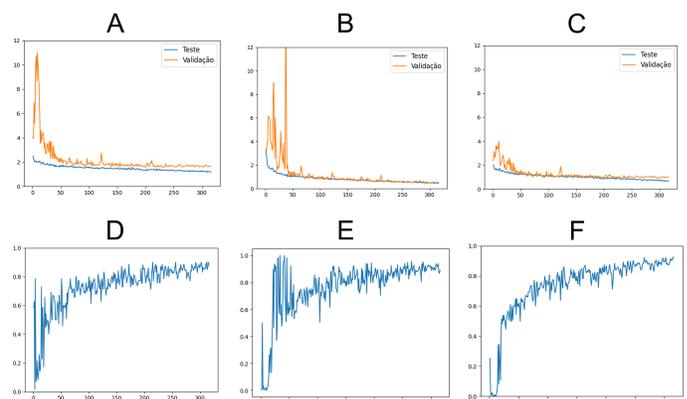


Fig. 10. Gráficos de desempenho do modelo 2, o eixo x representa a quantidade de épocas e o eixo y representa o valor alcançado em cada métrica. (A) Gráfico de perda dif. (B) Gráfico de cls. (C) Gráfico de perda box. (D) Gráfico de Recall. (E) Gráfico de Precisão. (F) Gráfico de mAP50

Após o treinamento, o modelo foi utilizado para analisar as imagens de teste. No teste o modelo conseguiu alcançar um

valor de precisão de 0.96. Todos os resultados obtidos estão apresentados na Tabela IV. No teste foi obtido um mAP50 de 0.961 para a detecção de escavação e mAP50 0.979 para a detecção do disco óptico.

TABELA IV
RESULTADOS OBTIDOS NA VALIDAÇÃO E NO TESTE

	Precision	Recall	mAP50	mAP50-95
validação - escavacao	0.895	0.714	0.846	0.465
validação - disco óptico	0.985	1	0.995	0.888
validação - geral	0.94	0.857	0.92	0.677
teste - escavacao	0.946	0.873	0.961	0.52
teste - disco óptico	0.974	0.955	0.979	0.865
teste - geral	0.96	0.914	0.97	0.693

As coordenadas obtidas foram utilizadas para fazer as anotações nas imagens e posteriormente as análises do ISNT. A Figura 11 apresenta uma comparação entre as anotações feitas pelo modelo e a anotação feita pelos especialistas. Pode-se perceber que as áreas detectadas com o modelo se assemelham com as anotações feitas pelos especialistas.

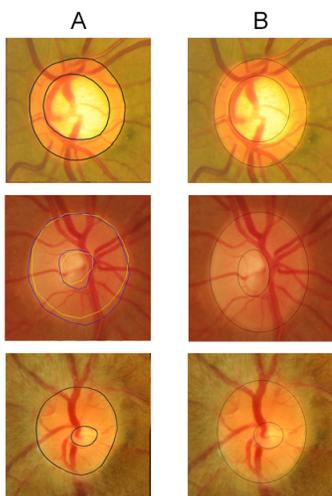


Fig. 11. Resultado das anotações do disco óptico e da escavação. Coluna A apresenta as anotações feitas pelos especialistas. Coluna B as anotações feitas pelo modelo.

A Figura 12 apresenta as imagens com a anotação do ISNT, resultado do modelo 2 com as imagens de teste. Nessas imagens foram utilizados a distância entre as duas elipses que circulavam o disco óptico e a escavação. Foi analisado que na maioria das imagens onde essa relação não foi alcançada a diferença entre a distância da região inferior e da região superior eram menores do que 10 unidades de diferença, o que indica que com mais ajustes do modelo em relação a segmentação do disco óptico e escavação existam mais exames que alcancem as regras.

Para o conjunto de dados Papila que possuía a separação das imagens do olho direito e do olho esquerdo, foi possível verificar a relação ISNT, comparando os valores calculados de forma automática. Além disso esse conjunto de dados

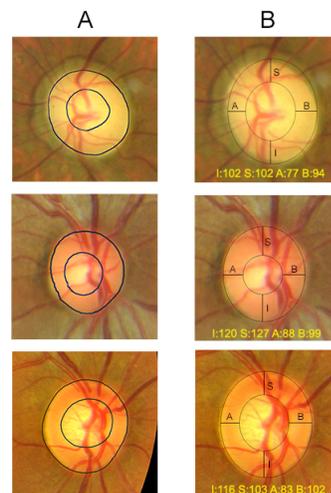


Fig. 12. Resultado da anotação do ISNT. Coluna A apresenta as anotações feitas pelos especialistas. Coluna B as anotações juntamente com a marcação e cálculo do ISNT feita pelo modelo.

continha o diagnóstico da imagem, a partir desta informação foi verificado a relação da regra ISNT com o diagnóstico do especialista, o modelo não encontrou nenhuma imagem que respeitasse a relação, entretanto existiam exames que não possuía glaucoma.

Devido ao fato do conjunto de dados Magrabia não informar o lado do olho, o modelo não pode fazer a análise da relação. Então conforme demonstrado na Figura 12 as laterais foram marcadas com a letra A e B informando os valores. Neste caso o modelo apenas consegue analisar a região superior e inferior ($I \geq S$), para a análise da relação entre os valores das laterais ($N \geq T$) seria necessário a análise humana.

Outro experimento foi realizado utilizando o conjunto de imagens Messidor, esse conjunto de imagens não estava contido nas imagens utilizadas para treinamento do modelo. Para a avaliação foi comparado a anotação feita pelos especialistas com a anotação feita pelo modelo treinado. A Figura 13 demonstra a comparação entre a detecção feita pelo modelo e a segmentação médica. Foi analisado que em alguns casos o modelo conseguiu identificar as áreas certas, entretanto em outros casos ele teve dificuldade em conseguir identificar a escavação, com isso é considerado que o modelo precisa ser ajustado para se adaptar melhor com a variabilidade de entradas.

V. CONCLUSÃO E TRABALHOS FUTUROS

O diagnóstico precoce do glaucoma é importante para prevenir que a doença cause a perda total da visão nos pacientes. Um dos exames utilizado para o diagnóstico da doença é o exame de fundo de olho que possibilita os especialistas analisarem e diagnosticarem a presença do glaucoma. O foco dessas análises é na área do disco óptico, nela são feitas observações entre a escavação e o disco óptico.

O objetivo deste trabalho foi automatizar o recorte da área de interesse utilizando uma CNN que teve como tarefa detectar

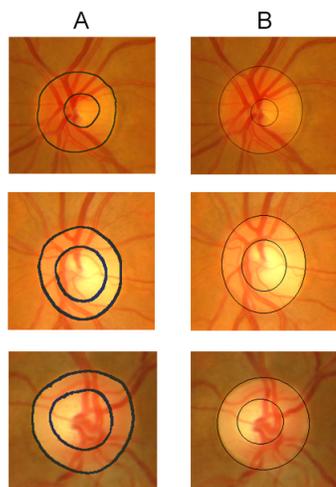


Fig. 13. Resultado das anotações em imagens da base diferente do treinamento e validação. Coluna A apresenta as anotações dos especialistas. Coluna B demonstra as anotações feitas pelo modelo.

a área do disco óptico. O modelo utilizado foi o YOLO v8, tendo sua camada densa treinada com imagens de fundo de olho anotadas com a região desejada. O modelo obteve uma mAP de 0.995 em seu treinamento, e 0.992 na sua validação com imagens do mesmo banco de dados que foi treinado.

Após o desenvolvimento desta etapa a segunda parte deste trabalho foi a implementação de um modelo para a detecção da escavação e disco óptico em imagens recortadas na área de interesse. Para isso foram usados para o treinamento imagens de banco de dados públicos que possuíam marcações de especialistas na área do disco óptico e escavação e essas anotações serviram para fazer a anotação nas imagens de treinamento. O modelo YOLO v8 teve sua camada densa treinada com essas imagens alcançando um desempenho de mAP para detecção escavação de 0.846 e mAP 0.995 para detecção do disco óptico no treinamento. Na validação os valores foram de mAP para escavação de 0.961 e mAP disco óptico de 0.979.

As coordenadas detectadas foram utilizadas para fazer a anotação da regra de ISNT nas imagens de fundo de olho. Os resultados obtidos se demonstraram promissores para o uso desta arquitetura para a anotação do disco óptico e da escavação. Entretanto, foi observado que o modelo apresentava mais facilidade em detectar o disco óptico. Foi executado um segundo teste utilizando o modelo para detectar a escavação e o disco óptico em imagens de conjunto de dados que o modelo não tinha sido treinado ou validado. Esse teste demonstrou que em algumas imagens o modelo consegue identificar a relação entre as duas áreas, entretanto em algumas imagens não conseguia detectar a escavação.

Como direcionamento para trabalhos futuros, pretende-se desenvolver um sistema completo agrupando o modelo de recorte e anotação desenvolvidos nesses trabalhos com um modelo de classificação de imagens com glaucoma, em conjunto com a análise do ISNT e relação da escavação do disco.

Com isso, o sistema terá etapas desde o recorte da área de interesse, a classificação e a análise da região óptica, podendo auxiliar oftalmologistas no diagnóstico do glaucoma.

AGRADECIMENTOS

Este trabalho recebeu apoio financeiro da CAPES - Brazil (PROAP/AUXPE - CONSOLIDA 3/4) e da FAPESC, Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina.

REFERENCES

- [1] R. N. Weinreb, T. Aung, and F. A. Medeiros, "The pathophysiology and treatment of glaucoma: a review," *Jama*, vol. 311, no. 18, pp. 1901–1911, 2014.
- [2] J. D. Stein, A. P. Khawaja, and J. S. Weizer, "Glaucoma in adults—screening, diagnosis, and management: a review," *Jama*, vol. 325, no. 2, pp. 164–174, 2021.
- [3] A. Almazroa, R. Burman, K. Raahemifar, and V. Lakshminarayanan, "Optic disc and optic cup segmentation methodologies for glaucoma image detection: a survey," *Journal of ophthalmology*, vol. 2015, 2015.
- [4] M. Alawad, A. Aljouie, S. Alamri, M. Alghamdi, B. Alabdulkader, N. Alkanhal, and A. Almazroa, "Machine learning and deep learning techniques for optic disc and cup segmentation—a review," *Clinical Ophthalmology*, pp. 747–764, 2022.
- [5] A. C. Thompson, A. A. Jammal, and F. A. Medeiros, "A review of deep learning for screening, diagnosis, and detection of glaucoma progression," *Translational vision science & technology*, vol. 9, no. 2, pp. 42–42, 2020.
- [6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.
- [8] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.
- [9] X. Sun, Y. Xu, W. Zhao, T. You, and J. Liu, "Optic disc segmentation from retinal fundus images via deep object detection networks," in *2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2018, pp. 5954–5957.
- [10] S. Sreng, N. Maneerat, K. Hamamoto, and K. Y. Win, "Deep learning for optic disc segmentation and glaucoma diagnosis on retinal images," *Applied Sciences*, vol. 10, no. 14, p. 4916, 2020.
- [11] S. Holm, G. Russell, V. Nourrit, and N. McLoughlin, "Dr hagsis—a fundus image database for the automatic extraction of retinal surface vessels from diabetic patients," *Journal of Medical Imaging*, vol. 4, no. 1, pp. 014 503–014 503, 2017.
- [12] A. Singh, M. K. Dutta, M. ParthaSarathi, V. Uher, and R. Burget, "Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image," *Computer methods and programs in biomedicine*, vol. 124, pp. 108–120, 2016.
- [13] A. Issac, M. P. Sarathi, and M. K. Dutta, "An adaptive threshold based image processing technique for improved glaucoma detection and classification," *Computer methods and programs in biomedicine*, vol. 122, no. 2, pp. 229–244, 2015.
- [14] H. Raja, M. U. Akram, S. G. Khawaja, M. Arslan, A. Ramzan, and N. Nazir, "Data on oct and fundus images for the detection of glaucoma," *Data in brief*, vol. 29, p. 105342, 2020.
- [15] M. T. Islam, S. T. Mashfu, A. Faisal, S. C. Siam, I. T. Naheen, and R. Khan, "Deep learning-based glaucoma detection with cropped optic cup and disc and blood vessel segmentation," *IEEE Access*, vol. 10, pp. 2828–2841, 2021.
- [16] O. Kovalyk, J. Morales-Sánchez, R. Verdú-Monedero, I. Sellés-Navarro, A. Palazón-Cabanes, and J.-L. Sancho-Gómez, "Papila: Dataset with fundus images and clinical data of both eyes of the same patient for glaucoma assessment," *Scientific Data*, vol. 9, no. 1, p. 291, 2022.
- [17] A. Almazroa, S. Alodhayb, E. Osman, E. Ramadan, M. Hummadi, M. Dlaim, M. Alkatee, K. Raahemifar, and V. Lakshminarayanan, "Retinal fundus images for glaucoma analysis: the riga dataset," in *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, vol. 10579. SPIE, 2018, pp. 55–62.