

Alexandre César Pinto Pessoa

**Diagnóstico Automático de Glaucoma em
Retinografias baseado no CDR usando U-Net**

São Luís - MA

2018

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Núcleo Integrado de Bibliotecas/UFMA

Pessoa, Alexandre César Pinto.

Diagnóstico Automático de Glaucoma em Retinografias
baseado no CDR usando U-Net / Alexandre César Pinto
Pessoa. - 2018.

52 f.

Orientador(a): João Dallyson Sousa de Almeida.

Monografia (Graduação) - Curso de Ciência da
Computação, Universidade Federal do Maranhão, São Luís-MA,
2018.

1. CDR. 2. Diagnóstico. 3. Glaucoma. 4.
Retinografias. 5. U-Net. I. Almeida, João Dallyson Sousa
de. II. Título.

Alexandre César Pinto Pessoa

Diagnóstico Automático de Glaucoma em Retinografias baseado no CDR usando U-Net

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. João Dallyson Sousa de Almeida

São Luís - MA

2018

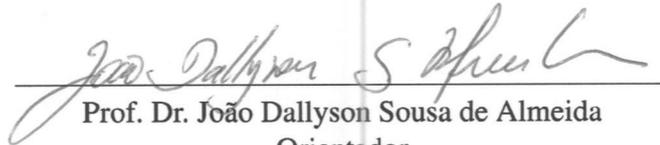
Alexandre César Pinto Pessoa

**Diagnóstico Automático de Glaucoma em Retinografias
baseado no CDR usando U-Net**

Monografia apresentada ao curso de Ciência da
Computação da Universidade Federal do Maranhão,
como parte dos requisitos necessários para obtenção
do grau de Bacharel em Ciência da Computação.

Data de Aprovação: 18/01/2018

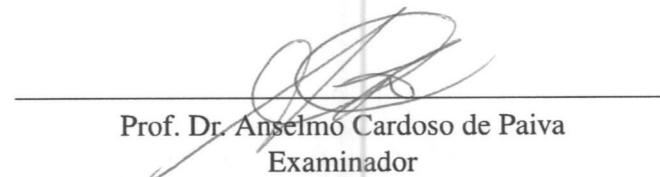
Banca Examinadora



Prof. Dr. João Dallyson Sousa de Almeida

Orientador

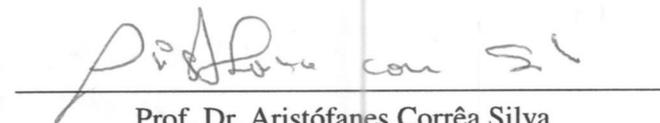
Universidade Federal do Maranhão



Prof. Dr. Anselmo Cardoso de Paiva

Examinador

Universidade Federal do Maranhão



Prof. Dr. Aristófanis Corrêa Silva

Examinador

Universidade Federal do Maranhão

São Luís - MA

2018

To the one who said I wouldn't make it.

Agradecimentos

Aos meus pais, por todo amor, carinho e apoio dado durante toda a minha vida, e por serem pessoas maravilhosas.

A minha namorada Polyana, por ser a melhor companheira que alguém poderia pedir, e por todos os momentos que compartilhamos nesses últimos 2 anos e 6 meses.

A minha avó, por ter sido sempre a pessoa mais amável e carinhosa possível comigo, inclusive em momentos onde eu não merecia.

Aos meus melhores amigos Thalles, Bárbara e Gabriel, por serem pessoas incríveis por mais tempo que eu consiga lembrar.

Ao meu irmão Maurício, por me fazer não desistir desse curso.

Ao professor Anselmo, por todos os ensinamentos e conselhos durante minha graduação, pela oportunidade de crescimento pessoal ao me convidar para o NCA, e por não ter desistido de mim ao ver meu histórico.

Ao professor João Dallyson, por todo conhecimento, conselhos, competência e paciência como professor e como meu orientador no desenvolvimento deste trabalho.

Aos professores Aristófanés, Geraldo e demais docentes do Departamento de Informática, pelos ensinamentos, experiências, críticas e conselhos fornecidos durante minha graduação.

Aos meus amigos Felipe, Mayara e Artur por toda ajuda fornecida durante o desenvolvimento deste trabalho, e por serem pessoas legais em geral.

Aos meus amigos Fernando, Gabriel, Lucas, Matheus, Rayanne, Rodrigo e Yasmin, pelas amizades fiéis que não cederam ao tempo.

A comunidade de Smash Bros. do Maranhão por ser a melhor comunidade de jogos de luta que existe.

A todas as pessoas que me ofereceram carona durante minha graduação, que foram pessoas demais para eu listar cada uma aqui, mas podem ter certeza que eu não me esqueci de vocês.

A qualquer pessoa que eu possa ter me esquecido porque essa página já está grande demais.

Você, por ter gastado um tempo do seu dia para ler isso.

*"Dear old friends, remember Navarro."
(Fallout: New Vegas)*

Resumo

O glaucoma é uma neuropatia óptica progressiva que causa danos irreversíveis ao nervo óptico, que levam a perda total da visão caso sua progressão não seja interrompida nos estágios iniciais. Um dos exames mais utilizados para o diagnóstico desta patologia é a Retinografia, que captura imagens das áreas do fundo do olho de um paciente, sendo utilizadas para efetuar o diagnóstico a partir de uma análise visual e interpretação. O principal objetivo deste trabalho é desenvolver um método para o diagnóstico automático de glaucoma em imagens de fundo de olho baseado no CDR, focando nas tarefas de segmentação das regiões do disco óptico e cavidade do disco óptico. Este método pré-processa as imagens de fundo de olho aplicando a técnica CLAHE para melhoramento de contraste, e então utilizando duas redes neurais convolucionais baseadas na arquitetura U-Net para segmentar as regiões do disco óptico e cavidade do disco óptico, respectivamente. O diagnóstico automático é feito com o cálculo do CDR baseado nas segmentações geradas, e a partir de um limiar definido, classificar as imagens como saudáveis ou glaucomatosas. Os testes foram efetuados utilizando a base pública DRISHTI, e a etapa de segmentação obteve um Dice score de 94.5% e 86.5% para o disco e a cavidade, respectivamente. A etapa do cálculo do CDR obteve um erro médio de 13.7%, e o diagnóstico alcançou 53.0% de Acurácia, 50.4% de Sensibilidade e 74.2% de Especificidade. Após a apresentação dos resultados, é feita uma comparação dos resultados obtidos com alguns trabalhos relacionados presentes na literatura.

Palavras-chaves: Glaucoma, Diagnóstico, Processamento de Imagens, Retinografias, CDR, U-Net.

Abstract

Glaucoma is a progressive optic neuropathy that causes irreversible damage to the optic nerve, which can cause total loss of vision if its progression isn't interrupted on its initial stages. One of the most used exams for the diagnosis of this pathology is the Retinography, which captures images from the patients eye fundus regions, which are utilized to make the diagnosis based on a visual analysis and interpretation. The main objective of this work is to develop a automatic glaucoma diagnosis method in eye fundus images based on the CDR, focusing on the tasks of segmenting the optic disc and optic cup regions. This method preprocesses the eye fundus images by applying the CLAHE technique for contrast enhancement and then, by utilizing two convolutional neural networks based on the U-Net architecture to segment the optic disc and optic cup regions, respectively. The automatic diagnosis is done by the computation of the CDR based on the generated segmentations, and from a defined threshold, classify the images as healthy or glaucomatous. The tests were performed using the DRISHTI public dataset, and the segmentation step obtained a Dice score of 94.5% and 86.5% for the segmentation of disc and cup, respectively. The CDR computation step reached a mean error of 13.7%, and the diagnosis reached a 53.0% of Accuracy, 50.4% of Sensitivity and 74.2% of Specificity. After the results are presented, it is made a comparison of the obtained results with some of the related works currently present in literature.

Keywords: Glaucoma, Diagnosis, Image Processing, Retinography, CDR, U-Net.

Lista de ilustrações

Figura 1 – Vista em corte do olho humano.	18
Figura 2 – Retinógrafo.	20
Figura 3 – Imagem de fundo de olho.	20
Figura 4 – Ilustrações das regras mais utilizadas para diagnóstico de glaucoma. . .	21
Figura 5 – Exemplo de olhos saudável e com glaucoma.	21
Figura 6 – Etapas de Aplicações de PDI.	22
Figura 7 – Espectro de Luz Visível.	23
Figura 8 – Modelo RGB.	24
Figura 9 – Exemplo de aplicação do CLAHE.	26
Figura 10 – Representação de um neurônio artificial.	27
Figura 11 – Funções de ativação.	27
Figura 12 – Rede MLP com duas camadas ocultas.	28
Figura 13 – Ilustração de uma CNN.	29
Figura 14 – Imagem da base DRISHTI e suas respectivas máscaras para DO e CDO. .	31
Figura 15 – Sequência de etapas do método proposto.	32
Figura 16 – Sequência de processos na etapa de pré-processamento.	33
Figura 17 – Arquitetura da rede utilizada para segmentação do DO.	34
Figura 18 – Arquitetura da rede utilizada para segmentação da CDO.	35
Figura 19 – Exemplo da segmentação do disco.	39
Figura 20 – Exemplo da segmentação do disco.	40
Figura 21 – Exemplo da segmentação da escavação.	41
Figura 22 – Exemplo da segmentação da escavação.	42
Figura 23 – Exemplo da segmentação da escavação.	42
Figura 24 – Curva de aprendizagem para a rede do DO.	43
Figura 25 – Curva de aprendizagem para a rede da CDO.	43

Lista de tabelas

Tabela 1 – Resultados para a segmentação do DO	39
Tabela 2 – Resultados para a segmentação da CDO utilizando a arquitetura proposta por Sevastopolsky (2017)	40
Tabela 3 – Resultados para a segmentação da CDO utilizando a arquitetura descrita na seção 4.3.2	41
Tabela 4 – Desempenho do cálculo do CDR e diagnóstico automático	44
Tabela 5 – Comparação dos resultados	45

Lista de abreviaturas e siglas

CDR	Cup-to-Disc Ratio
PIO	Pressão Intra-Ocular
DO	Disco Óptico
CDO	Cavidade do Disco Óptico
PDI	Processamento Digital de Imagens
RGB	Red, Green, Blue
CLAHE	Contrast Limited Adaptive Histogram Equalization
RNA	Rede Neural Artificial
MLP	Multilayer Perceptron
CNN	Convolutional Neural Networks
CSD	Coefficiente de Similaridade de Dice
IoU	Intersection over Union
ROI	Region of Interest

Sumário

1	INTRODUÇÃO	13
1.1	Justificativa	14
1.2	Objetivos	14
1.2.1	Objetivos específicos	14
1.3	Organização do trabalho	14
2	TRABALHOS RELACIONADOS	16
3	FUNDAMENTAÇÃO TEÓRICA	18
3.1	Glaucoma	18
3.2	Retinografia	19
3.3	Processamento Digital de Imagens	21
3.3.1	Modelos de representação de cor	23
3.3.2	Histograma de imagens	25
3.4	Redes Neurais Artificiais	26
3.4.1	Redes Neurais Convolucionais	28
4	MATERIAIS E MÉTODOS	31
4.1	Base de imagens de retinografia	31
4.2	Ferramentas	32
4.3	Método proposto	32
4.3.1	Pré-processamento	32
4.3.2	Segmentação utilizando U-Net	33
4.3.2.1	Segmentação do DO	33
4.3.2.2	Segmentação da CDO	35
4.3.3	Diagnóstico baseado no CDR	36
4.4	Métricas para avaliação dos resultados	36
5	RESULTADOS E DISCUSSÃO	38
5.1	Apresentação dos resultados	38
5.2	Comparação de resultados	45
6	CONCLUSÃO	46
	REFERÊNCIAS	47

1 Introdução

Glaucoma é uma neuropatia óptica progressiva que causa danos estruturais irreversíveis ao nervo óptico, causando redução do campo visual. Estima-se que o número de casos mundiais da doença seja de 64,3 milhões, e é previsto que esse número aumente para 76 milhões em 2020, e que fique próximo dos 112 milhões em 2040, devido ao envelhecimento da população mundial (THAM et al., 2014).

Como os danos causados por esta patologia são progressivos e irreversíveis, ela eventualmente leva a perda total da visão. Por conta disso, essa doença é uma das principais causas de cegueira, com aproximadamente 5,2 milhões de casos no mundo, com a previsão de aumentar para 11,2 em 2020 (QUIGLEY; BROMAN, 2006). Apesar de não existir cura para o glaucoma, sua progressão pode ser interrompida se for detectada e tratada nos seus estágios iniciais, tornando o diagnóstico precoce essencial para prevenir a perda total da visão.

Um dos exames médicos mais utilizados para o diagnóstico de glaucoma é a Retinografia (Fotografia de Fundo de Olho). Com a imagem adquirida a partir deste exame, é feito o diagnóstico a partir de uma análise e interpretação visual.

Uma dessas análises consiste na localização da região do disco óptico (DO), da cavidade ou escavação do disco óptico (CDO), que é uma região mais interna do disco, e na delimitação das bordas dessas duas regiões. Um dos indicadores da presença de glaucoma é o aumento da região da cavidade. Esse aumento pode ser percebido através do *Cup-to-Disc Ratio* (CDR), que é a razão entre os diâmetros verticais do disco e de sua escavação (ALMAZROA et al., 2015).

Devido a um grande número de detalhes nessas imagens, além da sua própria interpretação ser subjetiva, vários especialistas podem fornecer diferentes interpretações de uma mesma imagem analisada. Além disso, o processo de segmentação manual das regiões do disco e sua cavidade e o cálculo do CDR são tarefas subjetivas que requerem bastante tempo, levando aproximadamente 8 minutos para serem efetuadas por avaliadores experientes (LIM et al., 2015).

Diante destes fatores, soluções que consigam avaliar retinografias e detectar a presença de glaucoma nas mesmas de maneira automática são desejáveis. Neste contexto, este trabalho objetiva propor um novo método computacional para realização do diagnóstico de glaucoma em imagens de retinografia, focando nas tarefas de segmentação automática das regiões do disco óptico e da escavação do disco óptico, além de calcular o CDR para efetuar o diagnóstico.

1.1 Justificativa

O processo de avaliação de retinografias buscando detectar glaucoma através do cálculo do CDR é apenas feito por especialistas. Por conta disso, a segmentação do disco óptico e escavação do disco óptico para encontrar o CDR é uma atividade custosa, demorada e consideravelmente subjetiva, onde diferentes especialistas podem fornecer segmentações e diagnósticos distintos, dependendo de sua experiência (ZILLY; BUHMANN; MAHAPATRA, 2017).

Para acelerar o processo, reduzir os custos e diminuir a subjetividade, a utilização de métodos automáticos para efetuar estes procedimentos é desejável. A aplicação de abordagens automáticas podem reduzir a carga de trabalho dos oftalmologistas, além de também ser capaz de oferecer uma segunda opinião para diagnósticos efetuados, para reduzir a subjetividade dos exames.

Após validado, o método poderá ser incorporado em uma ferramenta de triagem de pacientes, com o objetivo de detectar o glaucoma precocemente.

1.2 Objetivos

O objetivo deste trabalho é propor um método automático para detectar glaucoma em imagens de retinografias baseando-se no CDR.

1.2.1 Objetivos específicos

Para alcançar o objetivo geral pretendido, foram desenvolvidos os seguintes objetivos específicos:

- Estudar e aplicar técnicas para detectar e segmentar o disco óptico
- Estudar e aplicar técnicas para segmentar a região da escavação do disco óptico, a partir da região do disco segmentado previamente
- Efetuar o cálculo do CDR de maneira automática para a detecção de glaucoma, a partir das regiões segmentadas do disco e cavidade
- Avaliar a efetividade da solução proposta em bases públicas de imagens de retina

1.3 Organização do trabalho

Este trabalho está organizado em mais 5 capítulos, sendo que o Capítulo 2 apresentará trabalhos na literatura que também tinham como objetivo a segmentação

das regiões do DO e da CDO, para detectar posteriormente patologias em imagens de retinografias.

Já o Capítulo 3 apresenta a fundamentação teórica utilizada como base para o desenvolvimento do trabalho proposto. Este capítulo é dividido em 4 seções, onde a primeira apresenta um breve resumo da anatomia do olho humano, seguido de informações sobre o glaucoma. A segunda seção desenvolve sobre retinografias, resumindo como é feito o exame e exemplificando como são as imagens geradas por ele, e como normalmente é feito o diagnóstico de glaucoma nessas imagens. Por sua vez, a terceira seção apresenta conceitos básicos de processamento digital de imagens, e também resume alguns métodos específicos dessa área utilizados durante o desenvolvimento deste trabalho. Por fim, a quarta seção aborda redes neurais artificiais, que inicialmente explica os conceitos gerais, e logo após discorre sobre redes neurais convolucionais.

O Capítulo 4 inicialmente apresenta os materiais utilizados durante o desenvolvimento deste trabalho, e logo após discorre sobre os passos do método proposto. Os resultados obtidos pela aplicação da abordagem proposta são apresentados e discutidos, bem como comparados com os resultados de alguns trabalhos relacionados no Capítulo 5. Por fim, o Capítulo 6 conclui este trabalho com as considerações finais.

2 Trabalhos relacionados

Neste capítulo, são apresentados trabalhos que utilizaram imagens de fundo de olho para segmentação do DO e da CDO, a fim de detectar patologias nas imagens.

Um desses trabalhos foi proposto por [Joshi, Sivaswamy e Krishnadas \(2011\)](#), onde desenvolveram um método que integra informações locais da imagem dos pontos de interesse em um espaço de características multidimensional para segmentar o disco óptico. Para a segmentação da CDO, foram detectados e selecionados pontos de dobras nos vasos sanguíneos presentes na imagem de fundo de olho, e esses pontos foram utilizados como entrada para o método de interpolação *spline* 2D para delimitar a região da cavidade. Este método foi testado em uma base fornecida por um hospital local, e obteve 97.0% de *Dice score* para a segmentação do DO, e 84.0% para a CDO.

Por sua vez, [Sedai et al. \(2016\)](#) inicia seu método por uma aproximação inicial do DO utilizando Transformada Circular de Hough. O disco aproximado é então utilizado para computar as formas iniciais do DO e da CDO, que são utilizados como entrada de um método de regressão de formas em cascata, para iterativamente aprender as formas do disco e escavação. O método também foi testado na base DRISHTI inteira, e alcançou um *Dice score* de 95.0% para o disco, e 85.0% para a cavidade.

[Mittapalli e Kande \(2016\)](#) propuseram um método que, primeiramente segmenta os vasos sanguíneos usando o filtro *Gaussian-Hermite*, e então utiliza um modelo baseado no *Local Binary Fitting* para a segmentação do DO, e uma técnica de agrupamento chamada *Spatially Weighted Fuzzy C-Means* para segmentar a CDO. Este método foi testado em uma base construída por 40 imagens fornecidas por um hospital local, e 19 imagens selecionadas aleatoriamente de outras bases públicas, e obteve um *Dice score* de 97.5% para a segmentação do disco e 89.0% para a escavação.

Já [Zilly, Buhmann e Mahapatra \(2017\)](#) propôs um método onde, primeiramente as imagens são convertidas para o modelo de representação de cor $L^*a^*b^*$, e então seus valores de intensidade de *pixel* são subtraídos pela intensidade média, e divididos pelo desvio padrão de intensidade, deixando os valores de *pixel* entre 0 e 1. Em seguida, é aplicada uma técnica de amostragem chamada *Entropy Sampling* para diminuir o custo computacional da etapa seguinte, onde é utilizada uma rede neural convolucional de apenas duas camadas completamente conectadas para aprender as características de cada imagem. Por fim, as características extraídas pela rede são utilizadas para treinar um classificador chamado *Softmax Logistic Regression Classifier*, que prediz se cada *pixel* de cada imagem pertence as regiões do DO ou CDO. Este método obteve um *Dice score* de 97.3% para o DO e 87.1% para a CDO para 50 imagens da base DRISHTI-GS ([SIVASWAMY et al.](#),

2014).

Outro método que utiliza redes neurais convolucionais para a segmentação do DO e CDO foi proposto por [Sevastopolsky \(2017\)](#). O autor utilizou *Contrast Limited Adaptive Histogram Equalization* para melhorar o contraste das imagens de entrada, e então treinou uma rede neural convolucional para segmentar a região do DO utilizando as próprias imagens de fundo de olho no espaço de cor RGB equalizadas, e uma outra rede para segmentar a escavação, sendo que essa foi treinada com as máscaras resultantes da primeira rede. A arquitetura dessa rede foi uma modificação da rede U-Net, e obteve um *Dice score* de 94.0% e 95.0% nas bases DRIONS-DB e RIM-ONE, respectivamente, e 85.0% para a segmentação da CDO na base DRISHTI.

Ao contrário dos trabalhos de [Joshi, Sivaswamy e Krishnadas \(2011\)](#), [Sedai et al. \(2016\)](#) e [Mittapalli e Kande \(2016\)](#), e da mesma maneira que [Zilly, Buhmann e Mahapatra \(2017\)](#) e [Sevastopolsky \(2017\)](#), o método descrito neste trabalho utiliza redes neurais convolucionais para segmentar as regiões do DO e da CDO. Diferentemente da abordagem proposta por [Zilly, Buhmann e Mahapatra \(2017\)](#), que utilizou uma rede neural convolucional para extrair características que foram utilizadas por um outro classificador para efetuar a segmentação, o método proposto utiliza redes que já efetuam a segmentação por conta própria, que foi baseada na arquitetura U-Net ([RONNEBERGER; FISCHER; BROX, 2015](#)). O método proposto por este trabalho se difere da abordagem de [Sevastopolsky \(2017\)](#), pois o autor apenas efetua a segmentação das regiões do DO e da CDO, enquanto que no método proposto, além de segmentar essas regiões, calcula-se o CDR baseando-se no resultado dessas segmentações, a fim de efetuar um diagnóstico de glaucoma. Além disso, a arquitetura utilizada para a rede dedicada a segmentação da CDO no método proposto utiliza a mesma quantidade de filtros de convolução utilizados na proposta inicial da U-Net, onde [Sevastopolsky \(2017\)](#) utilizou uma modificação da mesma, com menos filtros por camada.

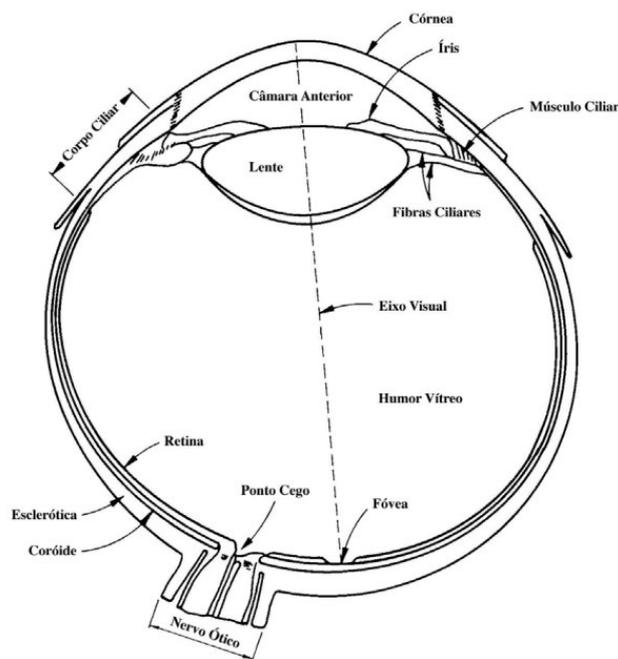
3 Fundamentação teórica

Neste capítulo são discutidos os principais fundamentos da oftalmologia e técnicas computacionais aplicados no trabalho.

3.1 Glaucoma

O olho humano possui um formato aproximadamente esférico, e um diâmetro de cerca de 20 mm. Ele é envolvido por 3 membranas, uma camada externa formada pela córnea e pela esclerótica, a coróide e a retina, que é a camada interna (FILHO; NETO, 1999). Sua estrutura interna é ilustrada na Figura 1.

Figura 1 – Vista em corte do olho humano.



Fonte – Filho e Neto (1999)

Glaucoma é um termo que engloba um amplo conjunto de doenças com inúmeras causas. Porém, elas contêm várias características em comum, como provocar o aumento da pressão intra-ocular (PIO), e da escavação e atrofia do nervo óptico e causar a perda do campo visual, comprometendo a visão periférica. Caso o glaucoma não seja tratado, os danos causados eventualmente levam a cegueira irreversível (LAURETTI; FILHO, 1997). Além da PIO elevada, existem outros fatores de risco para o glaucoma, como idade superior a 40 anos; aumento da região da escavação do disco óptico; etnia (maior incidência em

peessoas negras) e histórico familiar (SAUDE, 2014). Também existem correlações com outras condições como enxaquecas, diabetes, hipertensão, miopia e hipermetropia alta.

O glaucoma possui 5 classificações que ocorrem com mais frequência, o Glaucoma Primário de Ângulo Aberto (GPAA), o Glaucoma de Pressão Normal (GPN), o Glaucoma Primário de Ângulo Fechado (GPAF), o Glaucoma Congênito (GC) e o Glaucoma Secundário (GS). A forma mais comum de glaucoma é o GPAA, que tem como principais características o seu desenvolvimento lento e assintomático, que eventualmente ocasiona a perda progressiva da visão, e aumento da PIO (SAUDE, 2014). O GPN possui características similares a do GPAA, com exceção do aumento da PIO.

Já o GPAF é caracterizado por um aumento repentino da PIO, o que leva a dores intensas na região dos olhos, vermelhidão, vômitos e a uma perda total da visão em poucos dias. O GC é aquele onde a criança nasce com a patologia, devido a uma anomalia no desenvolvimento ocular, e possui como principais sintomas visão embaçada, sensibilidade a luz, lacrimejamento excessivo, globo ocular aumentado e córnea larga e opacificada. Por fim, o GS é aquele causado por outras doenças como cataratas avançadas, lesões oculares ou até mesmo a utilização de remédios para o tratamento de inflamações oculares ou outras doenças (SAUDE, 2014).

O diagnóstico desta patologia normalmente é feito baseado no histórico clínico do paciente, juntamente a PIO e a testes de perda do seu campo visual e uma avaliação manual do disco óptico (DO, também chamado de cabeça do nervo óptico). O DO é o local onde os axônios das células ganglionares deixam o olho, formando o nervo óptico, que é o local onde a informação fotovisual é transmitida ao cérebro (ZHANG et al., 2014).

3.2 Retinografia

A retinografia é um exame clínico que consiste na observação da retina e registro de uma imagem do fundo do olho. Este exame é aplicado para efetuar o diagnóstico ou acompanhamento de diversas doenças oculares, como retinopatia diabética, neurites, glaucoma, degeneração macular, etc. Para efetuar o exame, primeiramente o paciente tem sua pupila dilatada, e então seu olho é iluminado pela fonte de luz interna do retinógrafo (Figura 2), e então a retina é filmada, fotografada ou observada, dependendo do equipamento utilizado pelo médico (CARVALHO, 2006).

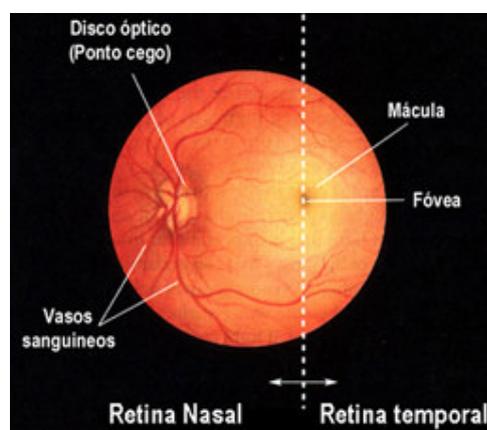
Figura 2 – Retinógrafo.



Fonte – Royo (2015)

As imagens capturadas pelo retinógrafo permitem visualizar estruturas como o disco óptico, a retina, a mácula e a fóvea, como ilustrado na Figura 3.

Figura 3 – Imagem de fundo de olho.



Fonte – UNESP (2008)

Na fotografia de fundo de olho (Retinografia), o DO pode ser dividido em duas zonas distintas: a escavação do disco óptico (também chamada de cavidade), que é uma zona central brilhante; e uma região periférica chamada anel neuroretinano. Um dos sintomas do glaucoma é a ampliação da região da escavação em relação ao DO, e por isso que ele é um dos indicadores mais importantes, e um dos vários parâmetros utilizados para detectar e diagnosticar o glaucoma (ZHANG et al., 2014).

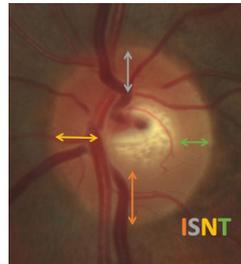
Figura 4 – Ilustrações das regras mais utilizadas para diagnóstico de glaucoma.

(a) Atrofia peripapilar.



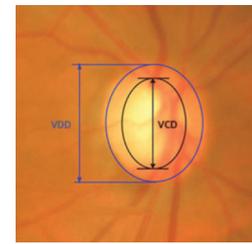
Fonte OPTICIAN (2016)

(b) Regra ISNT.



Fonte OPTICIAN (2016)

(c) Razão Escavação/Disco.

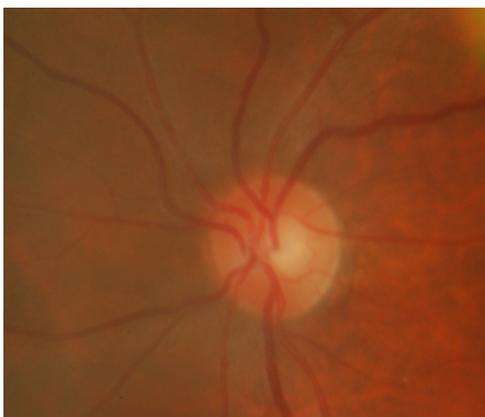


Fonte ScienceDaily (2013)

Alguns outros parâmetros para o diagnóstico de glaucoma são a atrofia peripapilar (PPA), ilustrada na Figura 4a, a regra ISNT, que pode ser visualizada na Figura 4b, e a razão do diâmetro vertical da escavação (VDD) com o diâmetro vertical do disco óptico (VDD), demonstrados na Figura 4c. Essa medida, chamada de CDR é a mais popular entre elas (ZHANG et al., 2014), além de ser considerada uma das características mais representativas para a detecção de glaucoma, e, segundo Spaeth (1981), olhos que apresentam esse índice com valor de 0.65 ou superior são normalmente considerados como glaucomatosos (SPAETH, 1981).

Figura 5 – Exemplo de olhos saudável e com glaucoma.

(a) Imagem de fundo de olho saudável



(b) Imagem de fundo olho glaucomatoso



Fonte – Sivaswamy et al. (2014)

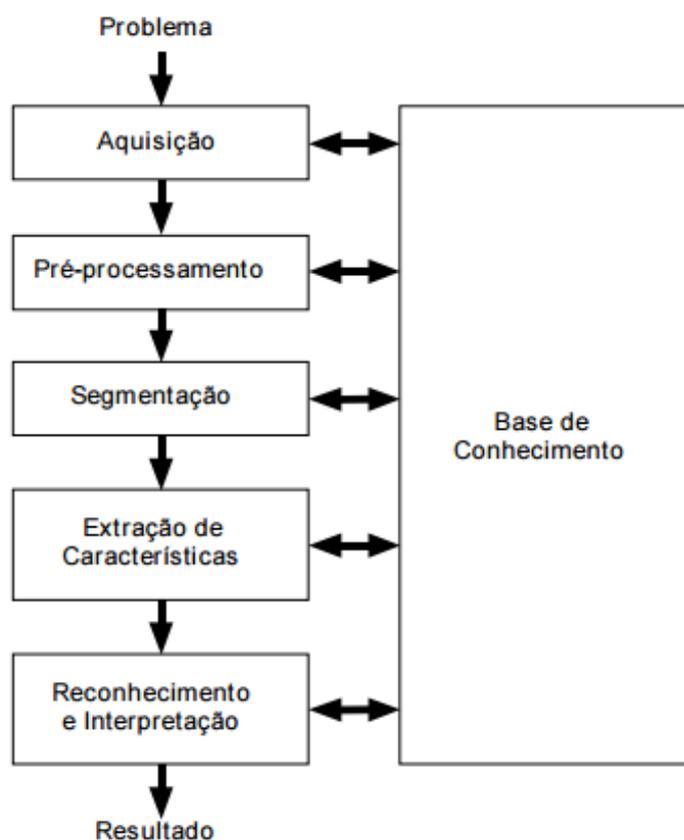
3.3 Processamento Digital de Imagens

Uma imagem qualquer pode ser compreendida como uma representação visual de um objeto (HOSHIRO; GOLDSCHMIDT, 2008). Ela também pode ser descrita matematicamente como uma função bidimensional $f(x, y)$, onde x e y são coordenadas em R^2 , e a amplitude de f em qualquer par (x, y) de coordenadas representa a intensidade

ou nível de cinza da imagem nesse ponto. Se a amplitude de f faz parte de um conjunto de valores finitos e discretos, a imagem é chamada de imagem digital (GONZALEZ; WOODS, 2010). As imagens digitais também são discretizadas espacialmente, onde elas são representadas por uma matriz cujas linhas e colunas identificam pontos da imagem, e seus valores correspondem ao valor de intensidade de *pixel* em cada ponto (FILHO; NETO, 1999).

A área de Processamento Digital de Imagens (PDI) abrange tanto análise de objetos ou características em uma imagem quanto as operações que podem ser realizadas na mesma. Boa parte das aplicações de PDI, como reconhecimento ou segmentação, segue as etapas expostas na Figura 6.

Figura 6 – Etapas de Aplicações de PDI.



Fonte – Filho e Neto (1999)

A primeira etapa trata de aquisição de imagens, onde as imagens que serão processadas pelo método proposto são adquiridas. Para aplicações que visam detectar e diagnosticar automaticamente glaucoma (ou outras patologias oculares) em imagens de fundo de olho, essas imagens normalmente são adquiridas previamente e armazenadas em uma base de dados.

A segunda etapa é o pré-processamento, onde são aplicadas técnicas de melhora-

to de imagens a fim de melhorar os resultados durante a aplicação do método proposto. Comumente, nesta etapa são utilizadas técnicas de remoção de ruído, realce de contraste ou até isolamento das regiões de interesse (GONZALEZ; WOODS, 2010).

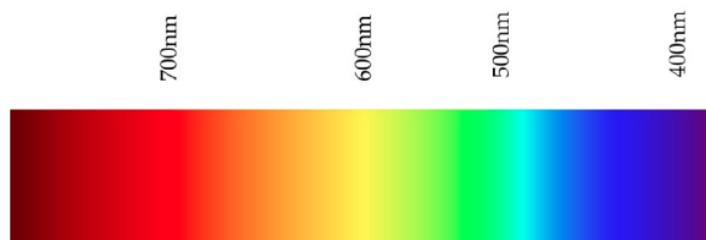
Em seguida, normalmente ocorre o processo de segmentação, onde apenas os objetos de interesse para a aplicação são mantidos na imagem, e ela pode ser feita de maneira automática, semi-automática ou manual. Para métodos que visam diagnosticar glaucoma, normalmente procura-se segmentar as regiões da imagem de fundo de olho que são mais afetadas pela patologia: o disco óptico e a escavação.

Na imagem pré-processada e segmentada, aplica-se a etapa de extração de características (ou *features*), onde bons descritores dos objetos presentes na imagem são detectados e extraídos, de tal maneira que os mesmos possam ser comparados com outros objetos, podendo discriminá-los em classes. Essas *features* normalmente são modelos matemáticos que descrevem algum comportamento da imagem, mas também podem ser cálculos diretos como o próprio CDR. Por fim, segue a etapa de reconhecimento e interpretação, onde as características extraídas são interpretadas, e esse resultado é utilizado pelo resto da aplicação para a tomada de decisões.

3.3.1 Modelos de representação de cor

A cor é uma sensação criada em resposta a uma excitação do sistema visual humano causada pela luz. Mais especificamente, a cor é o resultado perceptível da luz em uma região visível do espectro eletromagnético, com comprimentos de onda na região de $400nm$ até $700nm$, incidentes na retina. A retina humana possui 3 tipos de células fotorreceptoras, chamadas de cones, que respondem a luz com curvas de respostas espectrais relativamente diferentes (PLATANIOTIS; VENETSANOPOULOS, 2013).

Figura 7 – Espectro de Luz Visível.



Fonte – Plataniotis e Venetsanopoulos (2013)

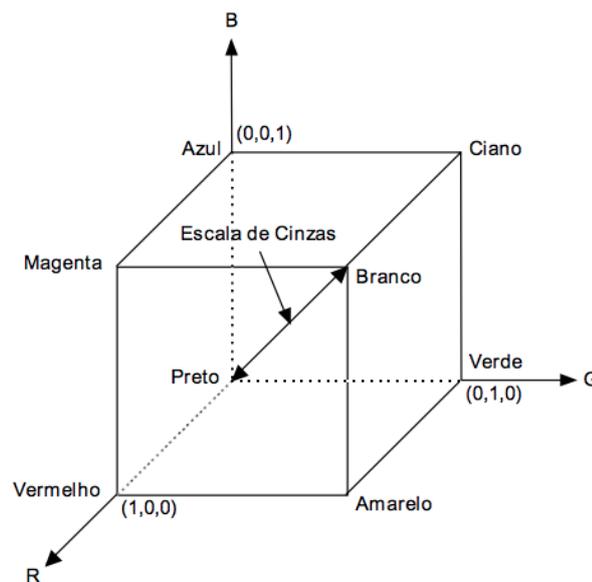
A utilização de cores em PDI possui duas motivações principais: o fato de que elas são poderosos descritores das propriedades de um objeto, facilitando sua identificação ou segmentação em aplicações de análise automática de imagens; e também auxilia em aplicações semi-automáticas, pois o olho humano é capaz de discernir milhares de cores,

enquanto o mesmo só consegue diferenciar algumas poucas dezenas de níveis de cinza (FILHO; NETO, 1999).

Um modelo de cor é, de maneira geral, uma representação tridimensional onde cada cor é representada por um ponto em um sistema de coordenadas 3D. Os modelos mais utilizados são o RGB (*Red, Green, Blue*), o CMY (*Cyan, Magenta, Yellow*), CMYK (Variação do CMY, onde K representa a cor preta) e o HSI (*Hue, Saturation, Intensity*) que também é chamado de HSV (onde o V significa *Value*) (FILHO; NETO, 1999).

O modelo RGB é baseado em um sistema de coordenadas cartesianas. Ele pode ser representado como um cubo, cujos 3 de seus vértices são as cores primárias (Vermelho, Verde e Azul), outros 3 sendo as cores secundárias (Magenta, Ciano e Amarelo), o vértice posicionado na origem correspondendo a cor preta, e o mais afastado da origem correspondendo ao branco, como ilustrado na Figura 8.

Figura 8 – Modelo RGB.



Fonte – Filho e Neto (1999)

Neste modelo, a escala de cinza se estende por uma linha que sai da origem do cubo (o ponto representando a cor preta) e segue em direção ao ponto mais distante da origem (a cor branca). Por conveniência, normalmente os valores máximos possíveis de R, G e B estão normalizados entre 0 e 1, mas em imagens digitais, eles geralmente são representados como valores discretos variando de 0 ao valor máximo que pode ser representado com a quantidade de *bits* utilizada para representar a cor de um *pixel* em uma imagem digital (JUNEJA; SANDHU; WALIA, 2009).

No processamento digital de imagens de fundo de olho, normalmente utiliza-se o canal R para aplicações que visam segmentar o disco óptico, pois essa região possui

um forte contraste com o resto da imagem no canal vermelho (SALEM; NANDI, 2007). Já para técnicas que buscam segmentar a região da escavação, comumente o canal G é escolhido, pois essa região tende a possuir intensidades maiores nesse canal (GOWRI; BABU; MAMANDUR, 2015).

3.3.2 Histograma de imagens

O histograma de uma imagem digital com valores de intensidade variando de 0 até $L - 1$ é uma função discreta $h(r_k) = n_k$, onde r_k é o k -ésimo valor de intensidade e n_k é a quantidade de *pixels* com aquela intensidade na imagem (GONZALEZ; WOODS, 2010). Histogramas são a base para várias técnicas de PDI de domínio espacial. A manipulação de histogramas pode ser usada para melhoramento de imagens, e os seus valores fornecem estatísticas da imagem, além de suas informações inerentes serem úteis em outras aplicações de PDI, como compressão ou segmentação.

Uma das manipulações possíveis de histograma é sua equalização, onde se procura redistribuir os valores de intensidade de *pixel* de uma imagem, de tal maneira que o seu histograma fique uniforme e o percentual de *pixels* para cada nível de intensidade fique praticamente o mesmo (FILHO; NETO, 1999). Para efetuar esta equalização, uma função auxiliar, chamada função de transformação é utilizada. A maneira mais comum de equalizar um histograma é utilizando a função de distribuição acumulada (CDF, do inglês *Cumulative Distribution Function*) da distribuição de probabilidades original, podendo ser expressa pela Equação 3.1.

$$S_k = T(r_k) = \sum_{j=0}^k \frac{n_j}{n} \quad (3.1)$$

Onde:

$$0 \leq T(r_k) \leq 1$$

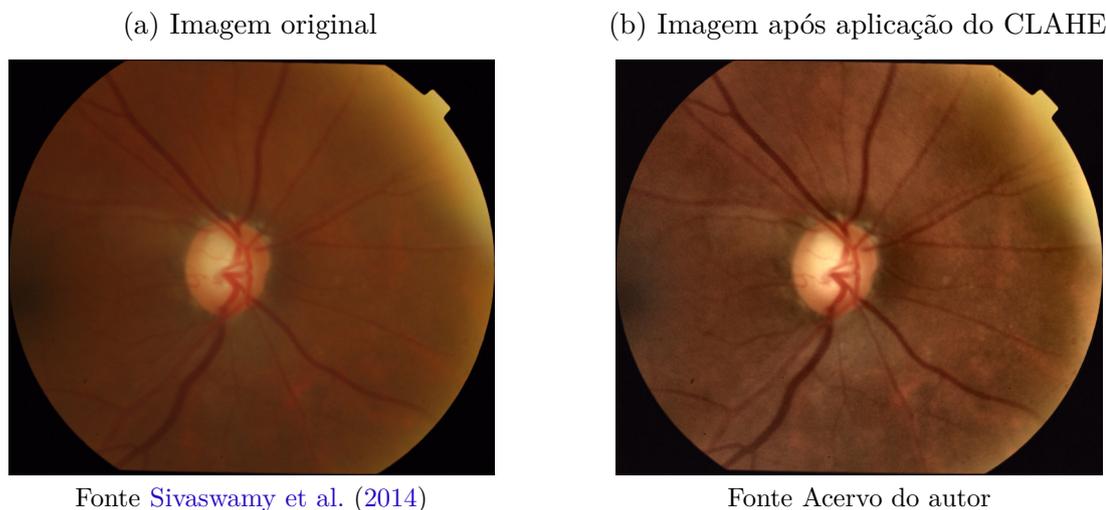
$$k = 0, 1, \dots, L - 1$$

A equalização de histograma pode ser usada para melhoramento de contraste de uma imagem (FILHO; NETO, 1999). Existe uma variação para equalização de histograma chamada de Equalização de Histograma Adaptativa (AHE, do inglês, *Adaptive Histogram Equalization*), que gera vários histogramas para diferentes regiões da imagem, e os utiliza para redistribuir os valores de intensidade da imagem. Por conta disso, ela é utilizada para melhoramento local de contraste (PIZER et al., 1987).

Porém, a utilização da AHE possui uma tendência de amplificar ruído em regiões relativamente homogêneas de uma imagem. Para evitar esta amplificação, foi desenvolvida uma variante do AHE, chamada de *Contrast Limited Adaptive Histogram Equalization* (CLAHE), proposta por Zuiderveld (1994), que limita o melhoramento de contraste da

AHE (PIZER et al., 1987). Essa limitação é efetuada pelo corte de uma região predefinida do histograma antes do cálculo de seu CDF para a sua normalização. O CLAHE mostrou-se útil para o melhoramento de imagens médicas (PIZER et al., 1987) e um exemplo de sua utilização em imagens de fundo de olho pode ser visto na Figura 9.

Figura 9 – Exemplo de aplicação do CLAHE.



3.4 Redes Neurais Artificiais

Uma rede neural pode ser definida como um modelo de raciocínio baseado no cérebro humano. O cérebro consiste em um conjunto de células nervosas, ou unidades básicas de processamento de informação, densamente interconectadas chamadas de neurônios. O cérebro humano possui cerca de 10 bilhões de neurônios e 60 trilhões de conexões entre eles, chamadas de sinapses (NEGNEVITSKY, 2005).

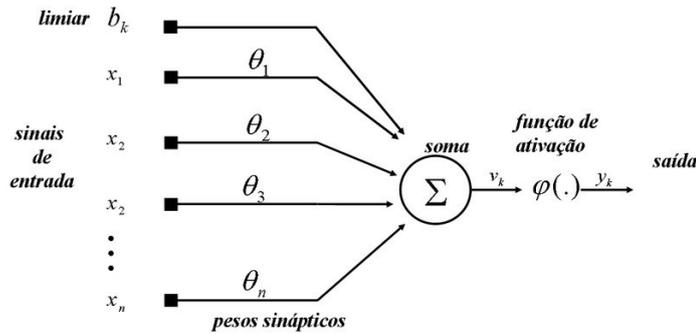
Uma rede neural artificial (RNA) consiste em várias unidades de processamento simples e fortemente conectadas chamadas de neurônios, análogos aos neurônios biológicos do cérebro humano. Os neurônios são ligados entre si por conexões ponderadas, e cada um recebe uma quantidade de sinais de entrada através de suas conexões, e produz um único sinal de saída, que pode ser propagado para vários outros neurônios (NEGNEVITSKY, 2005).

Cada conexão entre neurônios possui um peso numérico associado a ela, que são a maneira básica de representar memória de longo prazo em redes neurais artificiais. Esses pesos representam a importância da entrada de cada neurônio, e uma rede é capaz de aprender através de vários ajustes desses pesos (NEGNEVITSKY, 2005).

Um neurônio recebe vários sinais de entrada através de suas conexões, e propaga pelas suas ligações de saída um nível de ativação calculado através do uso de uma função de ativação aplicada em um somatório ponderado dos sinais que ele recebeu de entrada.

Os sinais de entrada podem ser dados puros, ou sinais de saída de outros neurônios, e o sinal de saída desse neurônio pode ser ou uma das entradas de outros neurônios ou a própria solução final de um dado problema (NEGNEVITSKY, 2005). Um exemplo de neurônio artificial é ilustrado na Figura 10.

Figura 10 – Representação de um neurônio artificial.



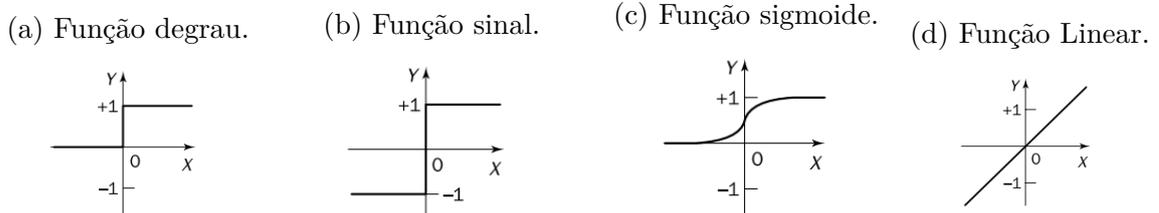
Fonte – Furtado, Velho e Macau (2011)

Existem várias possíveis funções de ativação, porém apenas algumas possuem aplicações práticas (NEGNEVITSKY, 2005). As 4 opções mais comuns são apresentadas nas Equações 3.2 e 3.3, e seus respectivos gráficos são ilustrados na Figura 11.

$$Y^{Degrau} = \begin{cases} 1 & \text{se } X \geq 0 \\ 0 & \text{se } X < 0 \end{cases} \quad Y^{Sinal} = \begin{cases} +1 & \text{se } X \geq 0 \\ -1 & \text{se } X < 0 \end{cases} \quad (3.2)$$

$$Y^{Sigmoide} = \frac{1}{1 + e^{-x}} \quad Y^{Linear} = X \quad (3.3)$$

Figura 11 – Funções de ativação.

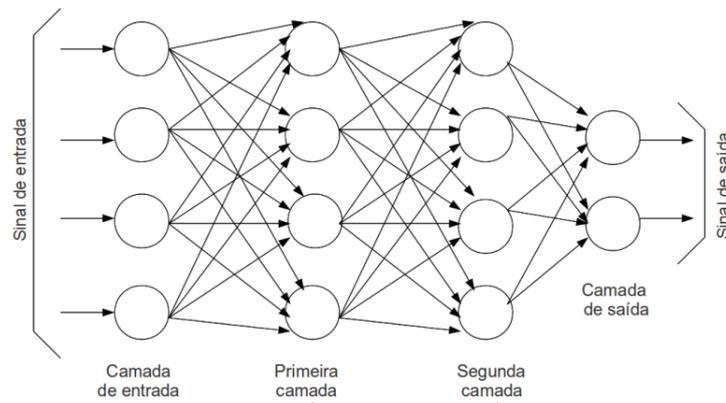


Fonte – Negnevitsky (2005)

A função sigmoide transforma qualquer valor de entrada em um valor no intervalo entre 0 e 1. Neurônios com essa função de ativação são utilizados em redes do tipo *back-propagation*. Uma rede neural artificial com um único neurônio toma decisões da mesma maneira, independente da função de ativação utilizada, de tal maneira que ela só é capaz de classificar apenas padrões linearmente separáveis (SHYNK, 1990).

Para tratar esse tipo de problema, utilizam-se redes neurais multicamadas, como as redes *Multilayer Perceptron* (MLP) treinadas com o algoritmo de *back-propagation*. A MLP é uma rede que possui uma camada de neurônios de entrada, uma ou mais camadas de neurônios intermediárias chamadas de camadas escondidas ou ocultas, e uma camada de neurônios de saída. Os sinais são sempre propagados de uma camada para a sua próxima.

Figura 12 – Rede MLP com duas camadas ocultas.



Fonte – Volpi (2015)

O método mais popular para tornar as redes MLP capazes de aprender é o *back-propagation*, proposto por (BRYSON, 1975). Esse algoritmo de aprendizagem possui duas etapas. Na primeira, um conjunto de padrões de treino é apresentado para a camada de entrada da rede, e os propaga de camada em camada até um padrão de saída ser gerado pela última camada. Se essa saída for diferente da esperada, um erro é calculado e então é propagado na direção contrária, até alcançar a camada de entrada. Os pesos da rede são reajustados a medida que esse erro é propagado. Assim como em qualquer outra RNA, uma que utiliza o algoritmo de *back-propagation* é determinada pela conectividade entre os neurônios (a arquitetura da rede), a função de ativação utilizada e o procedimento para o ajuste dos pesos (NEGNEVITSKY, 2005).

3.4.1 Redes Neurais Convolucionais

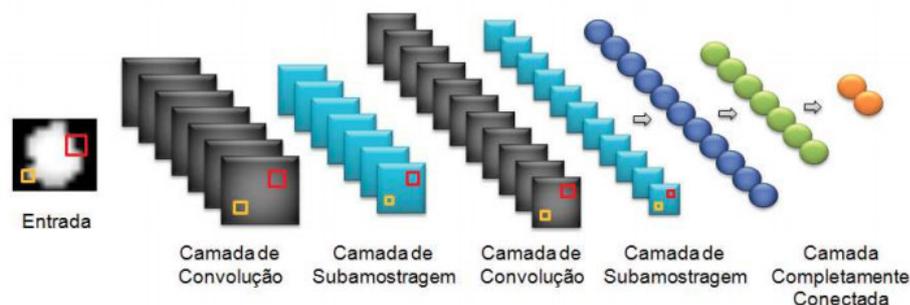
Os humanos e os animais tendem a interpretar e entender conceitos mais simples primeiro, para eventualmente aprenderem abstrações de mais alto nível, baseando-se nos conceitos previamente observados. Esse processo de aprendizagem que sugere uma divisão em camadas hierárquicas com diferentes responsabilidades no cérebro é chamada de aprendizagem profunda (FERNANDES, 2013).

Os métodos de aprendizagem profunda são técnicas de aprendizagem com múltiplos níveis de representação, compostos por modelos não-lineares simples que transformam a representação do nível da própria entrada para um nível de abstração mais alto (LECUN;

BENGIO; HINTON, 2015). Esses métodos são capazes de automaticamente descobrir características a partir de um conjunto de imagens sem a necessidade de um *pipeline* complexo de etapas de processamento de imagens e reconhecimento de padrões (HUA et al., 2015).

As redes neurais convolucionais (CNN, do inglês *Convolutional Neural Networks*) são modelos de aprendizagem profunda biologicamente inspirados que buscam reproduzir os mecanismos de percepção visual com o objetivo de aprender os padrões locais e espaciais de objetos de interesse (HUA et al., 2015). Uma CNN é tipicamente composta de várias camadas de convoluções e subamostragem (que são responsáveis pelo processo implícito de extração de características), seguidas por uma rede MLP completamente conectada (responsável pela classificação dos padrões de entrada).

Figura 13 – Ilustração de uma CNN.



Fonte – Silva et al. (2017)

As características extraídas em cada camada são combinadas através de seus campos receptivos locais, que conectam uma camada com sua anterior. Isso permite que características de mais alto nível sejam extraídas a partir das características combinadas da camada anterior, como bordas ou cantos (FERNANDES, 2013).

Uma camada de convolução é composta por filtros treináveis que são aplicados individualmente em toda a extensão da imagem de entrada, gerando vários mapas de características (LECUN; KAVUKCUOGLU; FARABET, 2010). Cada filtro será responsável pela detecção de uma característica que ocorre em qualquer parte da imagem após a rede ser treinada (SILVA et al., 2017). A Equação 3.4 mostra uma definição do cálculo do mapa de características, onde: k é o índice do neurônio; h_k é o mapa de características com saída k ; w_k é o filtro treinável k ; x é o mapa de características de entrada do neurônio; b_k é o valor de *bias* k ; e o operador $*$ representa o processo de convolução espacial bi-dimensional (LECUN; KAVUKCUOGLU; FARABET, 2010).

$$(h_k)_{ij} = (w_k * x)_{ij} + b_k \quad (3.4)$$

Uma convolução espacial é a aplicação de uma janela $N \times M$ (onde normalmente $N = M$) em uma imagem, de tal maneira que essa janela percorre a imagem inteira, sempre sobrepondo uma região da imagem. Essa aplicação gera uma nova imagem onde cada valor de *pixel* Z_{ij} é calculado pela Equação 3.5.

$$(Z)_{ij} = \sum_{k=1}^{N \cdot M} Z_k W_k \quad (3.5)$$

Onde W_k são os valores da janela de convolução, e Z_k são os valores de *pixel* da imagem original que estão sendo sobrepostos por W_k . A janela percorrerá a imagem original, desde o seu canto superior esquerdo até seu canto inferior direito (FILHO; NETO, 1999).

As camadas de subamostragem são responsáveis por reduzir a resolução espacial dos mapas de características e selecionar as características invariantes a translações e distorções (LECUN et al., 1998). Um dos tipos de subamostragem mais comuns é a ativação máxima (*Max Pooling*), onde apenas o *pixel* de maior intensidade do campo receptivo anterior é mantido.

Por fim, as características extraídas pelas camadas de convolução e subamostragem (os *pixels* de todos os mapas de características da camada anterior) são apresentadas a rede MLP completamente conectada, que é responsável por classificar os padrões de entrada (FERNANDES, 2013).

Foram utilizadas 2 CNNs neste trabalho, uma treinada para a segmentação da região do disco óptico, e outra para a segmentação da escavação do disco óptico. A arquitetura da rede utilizada para este propósito será apresentada no capítulo seguinte.

4 Materiais e Métodos

Neste capítulo, é apresentada a sequência de etapas adotada para efetuar o diagnóstico automático de glaucoma em retinografias. Primeiramente, é apresentada a base de imagens de retinografia utilizada neste trabalho. Em seguida, comenta-se sobre as ferramentas utilizadas para o desenvolvimento da metodologia proposta. Logo após, cada etapa que constitui o método proposto é detalhada. Por fim, são apresentadas as métricas utilizadas para a avaliação do método proposto para segmentação.

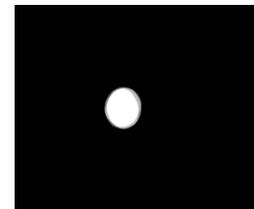
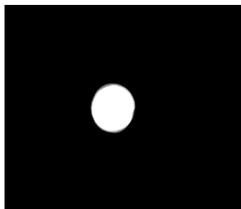
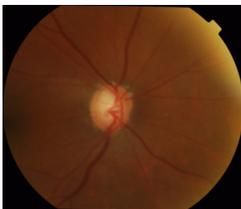
4.1 Base de imagens de retinografia

A base de imagens de retinografia utilizada neste trabalho foi a base pública DRISHTI-GS1, inicialmente apresentada por [Sivaswamy et al. \(2014\)](#). As imagens desta base foram adquiridas e marcadas no hospital Aravind, em Madurai, Índia. Nesta base, existe um número de exames aproximadamente igual para homens e mulheres, sendo que todos os pacientes eram indianos, com idades variando entre 40 a 80 anos.

O propósito principal da base DRISHTI é a validação de métodos para a segmentação do DO e da CDO. Esta base consiste em 101 imagens, adquiridas o campo visual de 30 graus, e resolução de 2896x1944 *pixels*, armazenadas no formato PNG sem compressão. Para cada imagem, a base também possui marcações de 4 especialistas, que possuíam experiência clínica de 3, 5, 9 e 20 anos, e essa marcação foi efetuada com uma ferramenta dedicada para marcações.

Figura 14 – Imagem da base DRISHTI e suas respectivas máscaras para DO e CDO.

(a) Imagem de fundo de olho. (b) Marcação do DO. (c) Marcação da CDO.



Fonte – [Sivaswamy et al. \(2014\)](#)

A Figura 14 ilustra um exemplo de uma imagem de fundo de olho da base DRISHTI, assim como as suas respectivas máscaras para o DO e a CDO. Essas máscaras foram construídas a partir da junção das marcações dos 4 especialistas, onde, quanto maior for a intensidade de um *pixel* nas máscaras, maior a quantidade de especialistas que marcou aquele determinado *pixel* como OD ou CDO. Além das máscaras e marcações, a base

também possui os valores de CDR calculados a partir das marcações dos 4 especialistas, assim como o diagnóstico de cada retinografia (normal ou glaucomatoso) baseado na opinião majoritária dos especialistas. Como existe um número par de especialistas, casos de empate foram resolvidos com o diagnóstico de um especialista adicional.

4.2 Ferramentas

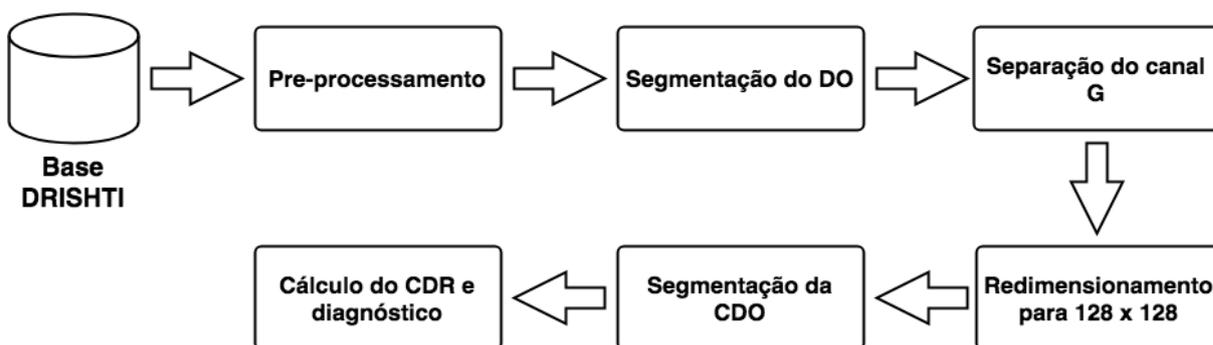
A linguagem de programação utilizada no desenvolvimento deste trabalho foi Python (versão 3.5.4). Foram utilizadas as bibliotecas *OpenCV* (BRADSKI; KAEHLER, 2008), versão 3.3.1, e *Keras* (CHOLLET et al., 2015), versão 2.1.1. Todas as operações de PDI na etapa de pré-processamento e do cálculo do CDR foram feitas utilizando a primeira biblioteca, e a implementação da CNN foi feita utilizando a segunda. O *Keras* foi configurado de tal maneira que a etapa de treinamento foi executada utilizando a placa gráfica.

O computador onde o método proposto foi implementado e testado possui a seguinte configuração: processador Intel Core i7-6700; 16GB de memória RAM; sistema operacional Windows 10 de 64 bits; e placa gráfica GeForce GTX 1060, com 6GB de memória dedicada.

4.3 Método proposto

Esta seção apresenta o método proposto para o diagnóstico de glaucoma a partir de imagens de fundo de olho. A Figura 15 exibe a sequência de processos da abordagem proposta, que serão explicados nas subseções seguintes.

Figura 15 – Sequência de etapas do método proposto.



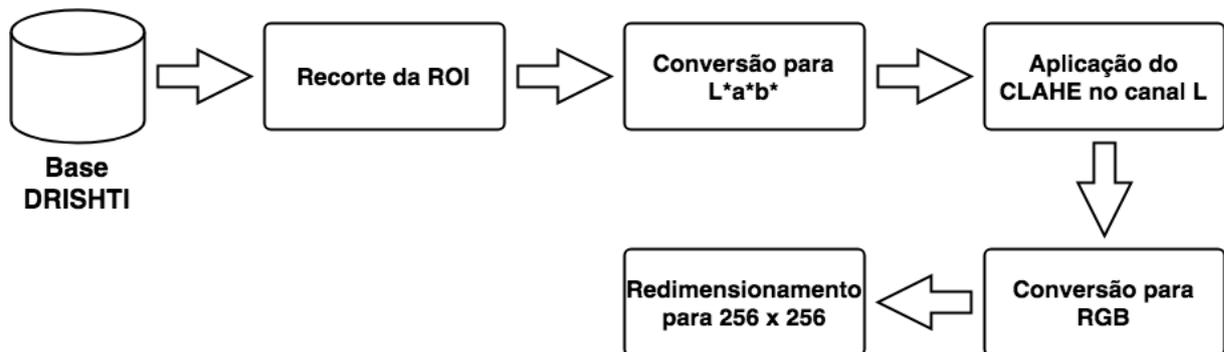
Fonte – Acervo do autor

4.3.1 Pré-processamento

Nesta etapa do método proposto, as imagens são tratadas de tal maneira que o desempenho das etapas seguintes seja melhorado. Os processos da etapa de pré-processamen-

to são apresentados na Figura 16.

Figura 16 – Sequência de processos na etapa de pré-processamento.



Fonte – Acervo do autor

O primeiro processo nesta etapa foi o recorte de uma região de interesse (ROI, do inglês *Region of Interest*). Essa extração foi feita através da remoção de aproximadamente $\frac{1}{5}$ das regiões laterais de cada imagem. Esse recorte foi feito dessa maneira a fim de reduzir o custo computacional para as etapas seguintes. Além disso, em todas as imagens da base utilizada neste trabalho, o DO sempre estava localizado na região extraída.

O processo seguinte é a conversão da imagem para o modelo de cor $L^*a^*b^*$ (HUNTER, 1958), e então é aplicado o CLAHE no canal L da imagem convertida, para melhorar o contraste no canal de luminância da imagem (ZHOU et al., 2017). Após a equalização do histograma, a imagem é convertida novamente para RGB, e é redimensionada para a resolução de 256×256 *pixels*, a fim de diminuir o custo computacional do processamento pela CNN.

4.3.2 Segmentação utilizando U-Net

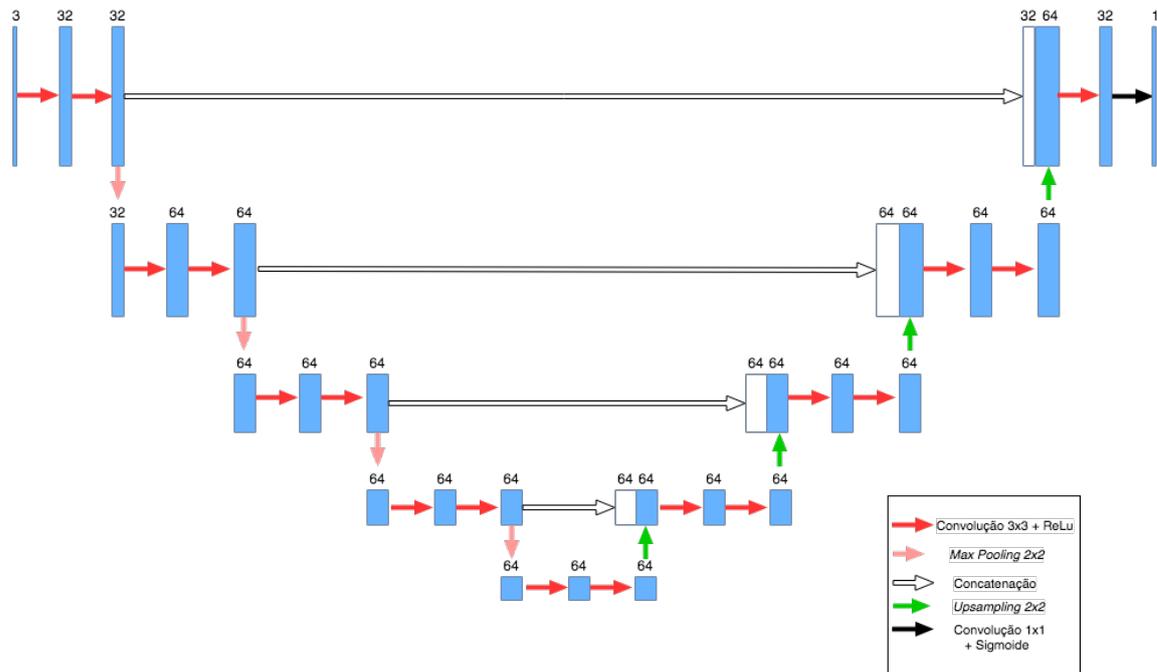
4.3.2.1 Segmentação do DO

A próxima etapa da metodologia é a segmentação do DO, utilizando uma CNN que foi treinada para este fim. A arquitetura da CNN utilizada foi a U-Net que foi proposta inicialmente por Ronneberger, Fischer e Brox (2015). Ela é uma rede completamente convolucional utilizada para a segmentação de regiões específicas em imagens, ou seja, a classificação é feita para cada *pixel*. Essa rede é capaz de produzir segmentações precisas com poucas imagens de treinamento.

A principal característica dessa CNN é, além dela possuir as usuais camadas redutoras (onde a resolução das imagens de entrada é reduzida), ela também possui camadas expansivas, onde os operadores de subamostragem são substituídos por camadas que aumentam a resolução das imagens de entrada. Esse tipo de camada é chamada de *upsampling*, que pode ser traduzido como "superamostragem". Além disso, as características

de alto nível extraídas pelas camadas redutoras são combinadas com as camadas de *upsampling* para que as informações de localidade não sejam perdidas. A U-Net também não possui uma camada completamente conectada, e utiliza apenas a parte válida de cada convolução, ou seja, o mapa de segmentação contém apenas quais *pixels* foram segmentados, e não quais informações estavam presentes neles.

Figura 17 – Arquitetura da rede utilizada para segmentação do DO.



Fonte – Acervo do autor

A Figura 17 ilustra a arquitetura da CNN utilizada para a segmentação do DO. Nessa Figura, cada retângulo representa uma camada, onde a seta ligada a cada um representa o tipo da camada, e o número acima representa o número de filtros de cada camada. A rede recebe os 3 canais de uma imagem de fundo de olho preprocessada pela etapa anterior como entrada, e depois são aplicadas duas camadas de convolução sucessivas, com *kernel* 3x3. Cada camada de convolução é seguida por uma *Rectified Linear Unit* (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), ou ReLu, que é descrita pela Equação 4.1. As camadas redutoras são seguidas por uma operação de *Max Pooling* com tamanho de *kernel* 2x2, que reduz a resolução dos mapas de características recebidos pela metade.

$$g(x) = \max(x, 0) \quad (4.1)$$

Já as camadas expansivas são seguidas de operações de *upsampling*, efetuadas através de de-convoluções de *kernel* 2x2 (DUMOULIN; VISIN, 2016). Bordas com valores iguais a 0 foram adicionadas aos mapas de característica de saída dessa operação, para que a resolução deles fosse exatamente o dobro do tamanho da entrada. Cada camada

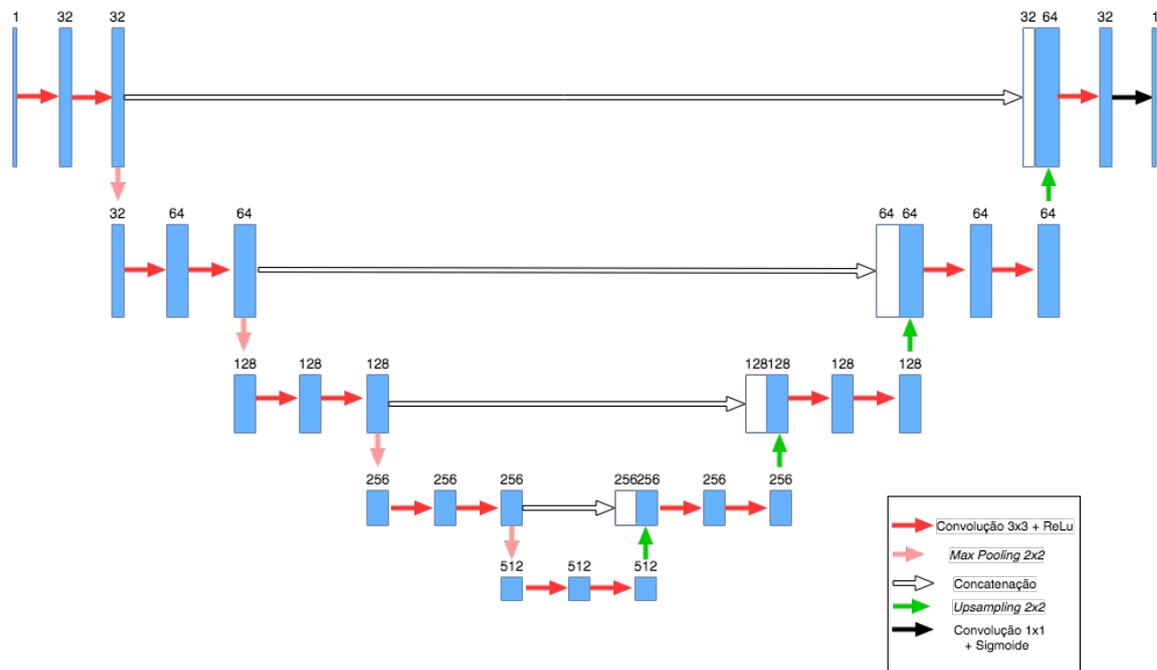
que é seguida de uma operação de *upsampling* também é concatenada com a saída da sua camada redutora correspondente, que é a camada diretamente a esquerda, como pode ser visualizado na Figura 17. Por fim, uma última camada de convolução com *kernel* 1x1 é usada para mapear cada um dos mapas de características em cada *pixel* da imagem de saída, com a função de ativação sigmoide, o que resulta em uma máscara binária para segmentação.

A rede U-Net original dobra o número de mapas de características a cada operação de *Max Pooling*, mas na arquitetura utilizada para a segmentação do disco, isso ocorre apenas uma vez. Experimentos mostram que a redução de filtros não afetam a qualidade da segmentação, e reduzem o custo computacional do treinamento da rede (SEVASTOPOLSKY, 2017).

4.3.2.2 Segmentação da CDO

Com a máscara binária gerada, o próximo processo envolve utilizar essa máscara para segmentar o DO na imagem, e utilizar o canal G do resultado como entrada para a etapa seguinte. Para a segmentação da CDO, a imagem de entrada é redimensionada para 128x128, e depois ela é enviada a uma outra CNN que foi treinada especificamente para a segmentação da CDO.

Figura 18 – Arquitetura da rede utilizada para segmentação da CDO.



Fonte – Acervo do autor

A arquitetura dessa rede é apresentada na Figura 18, e como pode-se perceber, as únicas diferenças dela para a rede utilizada na etapa anterior são a quantidade de

mapas de características. Como é utilizado apenas o canal G, a camada de entrada possui apenas um mapa de características, e como a segmentação da CDO é uma tarefa mais complexa, optou-se por utilizar a arquitetura original da U-Net, onde a cada operação de *Max Pooling*, o número de filtros aplicados em cada camada dobra; e a cada operação de *Upsampling*, a quantidade de mapas de característica é reduzida pela metade.

Com ambos disco e escavação segmentados, suas máscaras binárias são utilizadas na etapa final do método proposto.

4.3.3 Diagnóstico baseado no CDR

A etapa final do método proposto é o cálculo do CDR a partir das máscaras binárias do DO e da CDO geradas pela etapa descrita na subseção anterior. Sendo o CDR a razão dos diâmetros verticais das regiões da escavação pelo disco, para calcular o CDR, é necessário inicialmente computar esses diâmetros.

Os diâmetros foram calculados a partir da diferença do índice da última linha da imagem que possui um valor de intensidade não nulo pelo índice da primeira linha com valor de intensidade maior que 0. Com esses dois valores encontrados, a razão do diâmetro vertical da CDO pelo diâmetro do DO é calculado, e o valor resultante será o CDR.

A partir desse valor, é dado um diagnóstico. Baseado nas estatísticas levantadas por [Tatham et al. \(2013\)](#), os seguintes limiares foram determinados: se um exame possuir um CDR menor que 0.46, ele é considerado saudável, caso contrário, ele é considerado glaucomatoso.

4.4 Métricas para avaliação dos resultados

Todas as métricas utilizadas para a avaliação da etapa de segmentação deste trabalho são calculadas em termos de Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN). Em problemas de segmentação, um VP é um *pixel* que faz parte da região de interesse, que normalmente é uma marcação manual feita por um especialista, e que foi corretamente segmentado pela técnica proposta. Já os VN são aqueles *pixels* que não foram segmentados pela técnica avaliada, e de fato não fazem parte da região de interesse. Analogamente, FP são todos os *pixels* que foram segmentados pelo método proposto que não pertencem a região desejada, e FN são os *pixels* que pertencem a região de interesse procurada, mas que não foram segmentados pela técnica utilizada.

Uma das métricas utilizadas para validação do método proposto foi o Coeficiente de Similaridade de Dice (CSD) *Dice Score* ([DICE, 1945](#)). Esse coeficiente é comumente utilizado para avaliar técnicas de segmentação, onde a saída de um método de segmentação

automática é comparada com uma marcação manual feita por um especialista (ZIJDENBOS et al., 1994).

O *Dice score* é calculado pela Equação 4.2:

$$CSD = \frac{2VP}{(2VP + FP + FN)} \quad (4.2)$$

Ele também pode ser escrito como uma combinação da Sensibilidade e Precisão, outras métricas também utilizadas na validação de métodos de segmentação automática, como apresentado na Equação 4.3:

$$CSD = 2 \cdot \frac{\text{Sensibilidade} \cdot \text{Precisão}}{(\text{Sensibilidade} + \text{Precisão})} \quad (4.3)$$

Sendo Sensibilidade e Precisão calculados pela Equação 4.4:

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad \text{Precisão} = \frac{VP}{VP + FP} \quad (4.4)$$

Um valor de Sensibilidade alto indica que a técnica de segmentação consegue retornar boa parte da região de interesse, ou seja, ela retorna poucos FN. Já um índice de Precisão alto indica que o método de segmentação retornou uma quantidade consideravelmente maior de pixels que pertencem a região de interesse, quando comparada com o número de FP. Como uma combinação dessas duas métricas, é possível verificar o desempenho geral de um método de segmentação automática utilizando o *Dice score*.

A última métrica utilizada para a avaliação da etapa de segmentação do método proposto foi o Índice de Jaccard, proposto por Jaccard (1901). Esse índice, também conhecido como *Intersection over Union* (IoU), é calculado pela Equação 4.5. A diferença desse índice para o CSD é o fato de que o fator VP é multiplicado por 2 no CSD. A ausência desse produto no IoU faz com que erros sejam mais penalizados nesse índice que no CSD.

$$IoU = \frac{VP}{(VP + FP + FN)} \quad (4.5)$$

A etapa de diagnóstico automático de glaucoma foi avaliada utilizando as métricas de Sensibilidade, Especificidade e Acurácia. Sensibilidade é a razão do número de diagnósticos corretos para exames glaucomatosos pela quantidade de exames glaucomatosos presentes. Especificidade é a razão de diagnósticos corretos para exames saudáveis pela quantidade de exames saudáveis presentes. Acurácia é o número de diagnósticos corretos em geral sobre a quantidade de exames totais.

5 Resultados e discussão

Neste capítulo, são apresentados os resultados obtidos nos testes efetuados do método proposto, e em seguida os resultados são comparados com alguns trabalhos relacionados da literatura.

5.1 Apresentação dos resultados

A avaliação do método proposto para segmentação foi feita através das métricas de Sensibilidade, Precisão, CSD e o Índice de Jaccard. O método de separação da base de retinografias em treino e teste foi a Validação Cruzada *k-folds* (KOHAVI et al., 1995), onde a base é separada em k divisões. Uma dessas divisões é separada, enquanto o resto é utilizada para treinar a CNN, enquanto a divisão separada é utilizada para avaliar o desempenho da rede com base nas métricas citadas anteriormente. Depois dessa avaliação, uma nova divisão é escolhida para teste, e uma outra rede é treinada e avaliada, e este processo acaba quando todas as divisões forem utilizadas para teste.

Durante os testes, o método de otimização utilizado para o treinamento da rede foi o *Adam* (KINGMA; BA, 2014), com taxa de aprendizado 10^{-6} . A função de perda utilizada foi o valor do CSD multiplicado por -1. A CNN utilizada para segmentar o DO foi treinada utilizando as imagens de fundo de olho pré-processadas conforme explicado no capítulo anterior, enquanto a rede destinada a segmentar a CDO foi treinada com imagens de retinografias que foram previamente segmentadas para conter apenas a região do DO, utilizando a marcação do especialista disponível na base utilizada.

Para a Validação Cruzada *k-folds*, o valor de k escolhido foi 5, ou seja, 20 imagens foram separadas para teste enquanto o resto foi selecionado para treino. Das imagens separadas para treino, 60 foram utilizadas para o treinamento, enquanto 21 foram separadas para validação, sendo que essa separação foi feita de maneira aleatória. Cada rede foi treinada durante 150 épocas, isto é, todas as imagens de treino foram processadas pela rede 150 vezes.

Tabela 1 – Resultados para a segmentação do DO

<i>Fold</i>	Precisão	Sensibilidade	<i>Dice score</i>	Índice de Jaccard
<i>Fold 1</i>	97.1%	94.6%	95.8%	92.0%
<i>Fold 2</i>	93.0%	96.6%	94.7%	90.0%
<i>Fold 3</i>	93.7%	94.0%	94.0%	88.5%
<i>Fold 4</i>	95.8%	93.8%	94.7%	90.0%
<i>Fold 5</i>	95.8%	91.1%	93.4%	87.6%
Média	95.1%	94.0%	94.5%	89.6%
Desvio Padrão	1.4%	1.7%	0.8%	1.4%

A Tabela 1 apresenta os resultados obtidos na etapa de segmentação do DO. São apresentados os resultados individuais para cada *Fold* (divisão), bem como a média e o desvio padrão dos valores obtidos em cada métrica. Os valores dos índices de Precisão e Sensibilidade se mantiveram consistentemente altos durante todas as divisões, o que indica que uma grande parte das regiões segmentadas pelo método vão pertencer a região desejada, além de garantir que boa parte da região do DO vai ser segmentada pela rede.

Como a Precisão e Sensibilidade foram consideravelmente altos, os valores do CSD e IoU também foram bastante satisfatórios. Isso indica que o método foi bastante satisfatório para a segmentação do DO. Essa ideia é reforçada ainda mais pelo fato do desvio padrão ser consistentemente baixo para todas as métricas, o que mostra que provavelmente não ocorreu sobre-ajuste, isto é, a rede não se ajustou fortemente para as instâncias de treino.

Figura 19 – Exemplo da segmentação do disco.



Fonte – Acervo do autor

A Figura 19 apresenta um exemplo da aplicação do método de segmentação do DO. É perceptível que a máscara envolve a marcação do especialista praticamente por completo, e poucos *pixels* que não fazem parte da região do DO foram incluídos na máscara.

Figura 20 – Exemplo da segmentação do disco.



Fonte – Acervo do autor

A Figura 20 demonstra outro exemplo de aplicação da segmentação da região do disco. Nessa instância, é notável que a segmentação não foi adequada na região superior esquerda do DO. Acredita-se que isso ocorreu devido ao fato da borda nessa região do DO não ser bem definida. Porém, essa segmentação ainda é adequada ao método proposto pois o diâmetro vertical da máscara continua sendo praticamente o mesmo do DO.

Tabela 2 – Resultados para a segmentação da CDO utilizando a arquitetura proposta por Sevastopolsky (2017)

<i>Fold</i>	Precisão	Sensibilidade	<i>Dice score</i>	Índice de Jaccard
<i>Fold 1</i>	86.4%	88.5%	87.4%	77.7%
<i>Fold 2</i>	83.0%	94.4%	88.3%	79.0%
<i>Fold 3</i>	79.6%	92.0%	85.3%	74.4%
<i>Fold 4</i>	95.0%	72.7%	82.2%	70.0%
<i>Fold 5</i>	73.4%	90.0%	80.8%	68.0%
Média	83.4%	87.5%	84.8%	73.8%
Desvio Padrão	7.1%	7.6%	2.8%	4.3%

A Tabela 2 mostra os resultados obtidos para a etapa seguinte, a segmentação da CDO, utilizando a arquitetura proposta por Sevastopolsky (2017). Os valores de *Dice score* obtidos foram adequados, contudo, ao analisar a Precisão e Sensibilidade de cada *fold* individualmente, percebe-se uma grande variação entre eles, com os menores valores sendo 73.4% e 72.7% e os maiores 95.0% e 94.4% para Precisão e Sensibilidade, respectivamente. Isso tornou o desvio padrão de ambas medidas consideravelmente elevados, e também é notável que nas divisões que alcançaram valores altos em uma medida, foram encontrados valores baixos para outra.

Tabela 3 – Resultados para a segmentação da CDO utilizando a arquitetura descrita na seção 4.3.2

<i>Fold</i>	Precisão	Sensibilidade	<i>Dice score</i>	Índice de Jaccard
<i>Fold 1</i>	87.1%	88.9%	88.0%	78.6%
<i>Fold 2</i>	87.0%	84.8%	85.9%	75.3%
<i>Fold 3</i>	93.0%	75.9%	83.6%	71.8%
<i>Fold 4</i>	83.7%	76.0%	86.3%	76.0%
<i>Fold 5</i>	89.3%	87.9%	88.7%	79.6%
Média	88.0%	85.3%	86.5%	76.2%
Desvio Padrão	2.9%	2.9%	1.7%	2.7%

A Tabela 3 apresenta o desempenho obtido na etapa de segmentação da CDO utilizando a arquitetura para segmentação da CDO descrita na seção 4.3.2. O *Dice score* alcançado foi satisfatório, e relativamente consistente, com um desvio padrão de 1.7%. O desempenho desta arquitetura com a proposta por Sevastopolsky (2017) foi bem similar, onde também é perceptível que houve uma variação notável entre eles, com os menores valores sendo 83.7% e 75.9% e os maiores 93.0% e 88.9% para Precisão e Sensibilidade, respectivamente. Porém, ao comparar com os resultados apresentados na Tabela 2, percebe-se que o desvio padrão para essas medidas foi significativamente menor, o que mostra que a arquitetura utilizada pelo método proposto foi consideravelmente mais consistente.

Figura 21 – Exemplo da segmentação da escavação.



Fonte – Acervo do autor

A Figura 21 apresenta um exemplo da execução da segmentação da CDO. A máscara gerada envolveu uma parte considerável da região da escavação, porém houve uma quantidade notável de falsos positivos principalmente na região inferior, e na superior direita da máscara. Apesar disso, o diâmetro vertical da máscara gerada e da CDO ainda é bastante similar.

Figura 22 – Exemplo da segmentação da escavação.



Fonte – Acervo do autor

Já a Figura 22 apresenta um caso onde a segmentação foi errônea. Apesar da máscara gerada envolver completamente a região da CDO, ela também incluiu uma grande quantidade de *pixels* que não fazem parte da escavação. Além disso, a máscara possuiu um diâmetro vertical consideravelmente diferente da CDO, o que tornou essa segmentação completamente inadequada para o cálculo do CDR.

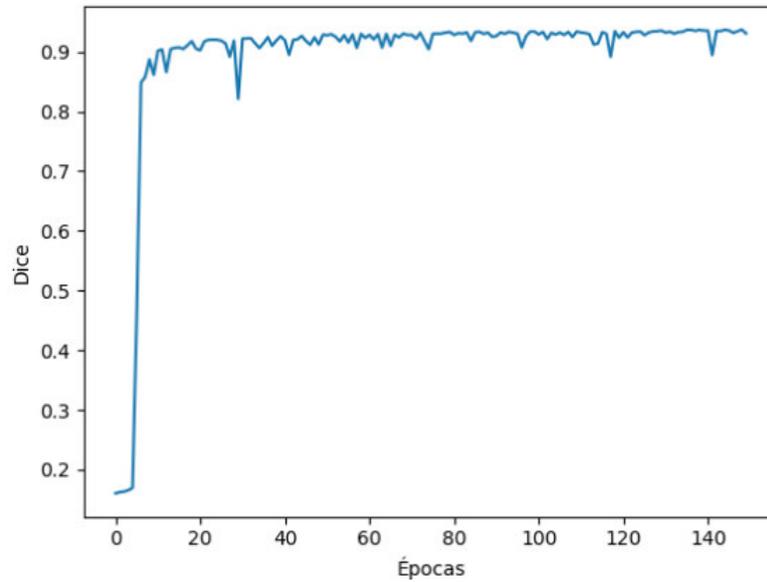
Figura 23 – Exemplo da segmentação da escavação.



Fonte – Acervo do autor

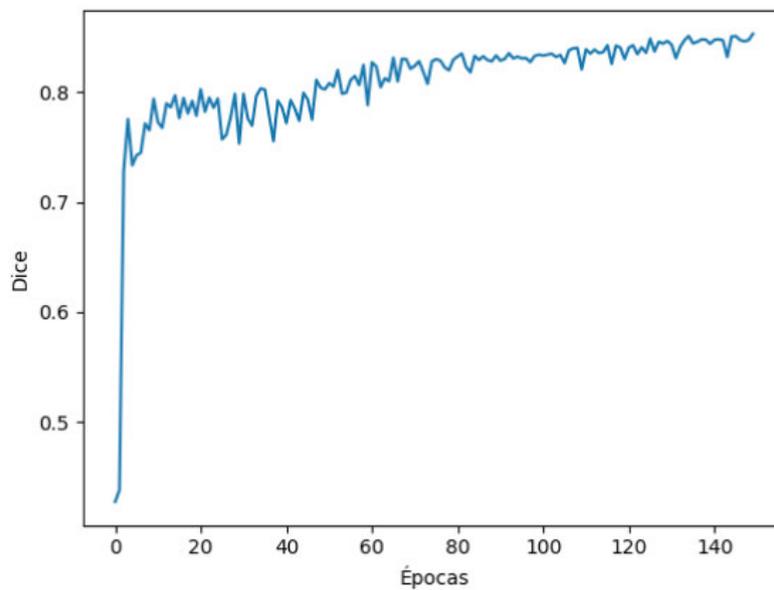
Um último exemplo de segmentação é apresentado na Figura 23. Boa parte dos *pixels* presentes na máscara são de fato da CDO. Entretanto, uma região significativa da CDO, da sua parte inferior direita, não está inclusa na máscara gerada. Supõe-se que isso ocorreu devido ao fato desta região estar oclusa pelos vasos sanguíneos. Apesar da ausência desta região, é notável que os diâmetros verticais da máscara e da CDO ainda são bem similares.

Figura 24 – Curva de aprendizagem para a rede do DO.



Fonte – Acervo do autor

Figura 25 – Curva de aprendizagem para a rede da CDO.



Fonte – Acervo do autor

As Figuras 24 e 25 ilustram as curvas de aprendizagem das redes responsáveis pela segmentação do DO e da CDO, respectivamente. Esses gráficos representam o desempenho da rede aplicada na divisão de validação, medido pelo *Dice score*, a cada época durante seu treinamento. Verifica-se que a primeira rede conseguiu alcançar um valor de *Dice score* relativamente alto em menos de 20 épocas, e seu desempenho ainda melhorava lentamente

durante o restante das épocas. A segunda também alcançou um *Dice score* relativamente alto em poucas épocas, contudo, o aprendizado dela foi significativamente mais instável.

Como não se sabe a orientação na qual a imagem de fundo de olho foi adquirida, não se pode determinar exatamente o que seria o diâmetro vertical das regiões do DO e CDO (MITTAPALLI; KANDE, 2016), então uma comparação direta com o CDR clínico disponível na base não seria adequada, pois não se sabe como que esse diâmetro foi calculado pelos especialistas. Para tentar contornar este problema, foram calculados os valores de CDR da maneira descrita na seção 4.3.3 nas máscaras binárias geradas pelos especialistas disponíveis na base utilizada, e esses valores foram comparados com as medidas encontradas pelo método proposto. Esses testes foram efetuados em 5 *fold*s de 20 imagens de teste diferentes.

Tabela 4 – Desempenho do cálculo do CDR e diagnóstico automático

<i>Fold</i>	Erro médio do CDR	Acurácia	Sensibilidade	Especificidade
<i>Fold 1</i>	0.10	55.0%	35.7%	100.0%
<i>Fold 2</i>	0.12	60.0%	50.0%	75.0%
<i>Fold 3</i>	0.09	50.0%	33.3%	100.0%
<i>Fold 4</i>	0.09	35.0%	33.3%	50.0%
<i>Fold 5</i>	0.23	65.0%	100.0%	46.1%
Média	0.13	53.0%	50.4%	74.2%
Desvio Padrão	0.05	10.2%	25.5%	23.2%

A Tabela 4 apresenta os resultados obtidos nesta etapa final. Considerando que o desvio médio do cálculo do CDR entre os especialistas da base DRISHTI foi de 0.12 (SIVASWAMY et al., 2014), o cálculo automático do CDR não se diferiu muito da medida calculada nas máscaras geradas pelos especialistas. É notável que o *fold* de teste onde ocorreu o maior erro médio foi o que possuía a maior concentração de de imagens saudáveis (13 das 20 imagens). Porém, o diagnóstico automático foi completamente inadequado, com uma Acurácia média alcançando apenas 53.0%, com um dos casos de teste alcançando apenas 35.0%. Isso pode ser atribuído aos casos onde a segmentação da escavação não foram satisfatórios, ou ao fato de que apenas o CDR não é o suficiente para diagnosticar glaucoma, ou que a maneira como o CDR foi calculado foi errônea, o que dificultou a realização de um diagnóstico correto a partir dele.

Vale ressaltar que existem várias instâncias presentes na base DRISTHI onde vários diagnósticos obtiveram valores de CDR não condizentes com as estatísticas levantadas por Tatham et al. (2013). Isto é, vários casos onde o CDR calculado foi superior a 0.6 foram diagnosticados como saudáveis, e vice-versa.

5.2 Comparação de resultados

Nesta seção, os resultados obtidos durante os testes do método proposto são comparados com alguns trabalhos relacionados. Na Tabela 5, são expostos os *Dice scores* obtidos pela segmentação do DO e da CDO, o erro médio obtido no cálculo do CDR, bem como as bases de dados utilizadas durante os testes de vários trabalhos.

Tabela 5 – Comparação dos resultados

Autor	<i>Dice score</i>		Erro médio do CDR	Base de dados utilizada
	DO	CDO		
(JOSHI; SIVASWAMY; KRISHNADAS, 2011)	97.0%	84.0%	0.12	Base montada pelo autor
(MITTAPALLI; KANDE, 2016)	97.5%	89.0%	0.13	Base montada pelo autor
(SEDAI et al., 2016)	95.0%	86.0%	0.09	DRISHTI
(ZILLY; BUHMANN; MAHAPATRA, 2017)	97.3%	87.1%	0.08	DRISHTI
(SEVASTOPOLSKY, 2017)	95.0%	85.0%	-	DRIONSDB, RIM-ONE, DRISHTI
Método Proposto	94.5%	85.6%	0.13	DRISHTI

Para a segmentação do DO, o trabalho que obteve o maior CSD foi o de [Mittapalli e Kande \(2016\)](#). Porém, este trabalho utilizou uma base própria e privada, então uma comparação direta não pode ser feita. O trabalho que utilizou exatamente a mesma base que o método proposto foi o de [Zilly, Buhmann e Mahapatra \(2017\)](#), que obteve um *Dice score* que é equivalente ao do método proposto.

Já para a segmentação da CDO, o método que obteve o melhor desempenho utilizando a base DRISHTI foi o de [Zilly, Buhmann e Mahapatra \(2017\)](#). Contudo, os piores casos encontrados nos testes desse trabalho são bem similares ao ilustrado na Figura 22, o que mostra que situações onde a CDO é consideravelmente pequena são casos desafiadores.

Analisando o desempenho dos trabalhos que segmentam o DO e a CDO em retinografias, é possível afirmar que o método proposto se equipara aos melhores trabalhos encontrados durante o desenvolvimento deste trabalho, principalmente a etapa da segmentação do DO. Entretanto, o erro médio na etapa do cálculo do CDR foi superior a todos os trabalhos relacionados analisados, o que mostra que ainda há espaço para melhorias nesta etapa do trabalho, bem como na segmentação da CDO.

6 Conclusão

Este trabalho apresentou uma proposta de método para a detecção automática de glaucoma em imagens de retinografias baseado no CDR. O método proposto utilizou técnicas de processamento de imagens envolvendo espaços de representação de cores e equalização de histograma para o melhoramento dessas imagens. O foco deste trabalho foi na utilização de redes neurais convolucionais para a segmentação das regiões do disco óptico e cavidade do disco óptico, cujas máscaras binárias geradas foram utilizadas para o cálculo do CDR.

A etapa de segmentação utilizou duas redes baseadas na arquitetura U-Net, que foram treinadas com o objetivo de segmentar o DO e a CDO, respectivamente. O desempenho da segmentação não conseguiu superar alguns dos trabalhos relacionados, porém pode-se afirmar que a performance foi similar e bastante satisfatória. Já a etapa do diagnóstico automático não obteve resultados adequados, um fato que pode ser atribuído tanto aos casos onde a segmentação da CDO ou o cálculo do CDR não foram adequados, quanto a noção de que o CDR é um fator inconsistente para se diagnosticar glaucoma (MITTAPALLI; KANDE, 2016).

Como trabalhos futuros, sugere-se explorar outras técnicas de pré-processamento, como abordagens que removam os vasos sanguíneos, e que melhor represente as estruturas da retina, utilizando outros modelos de representação de cor, como o HSV ou o $L^*a^*b^*$. Para a etapa de segmentação, recomenda-se a modificação da utilização de uma técnica baseada em algoritmos genéticos para a otimização dos parâmetros das redes utilizadas, como a quantidade de filtros utilizadas em cada camada, similar ao que foi feito por Silva et al. (2017). Também indica-se a utilização de mais bases de dados, além de aplicar técnicas que consigam artificialmente aumentar a quantidade de imagens disponíveis, como aplicar transformações de rotação nas imagens disponíveis.

Por fim, para a etapa de diagnóstico automático, aconselha-se a utilização de outras características além do CDR para efetuar o diagnóstico. Além disso, sugere-se a utilização de uma outra arquitetura de CNN especializada na classificação de imagens de fundo de olho, que receberia como entrada as regiões segmentadas pelo método proposto.

Referências

- ALMAZROA, A.; BURMAN, R.; RAAHEMIFAR, K.; LAKSHMINARAYANAN, V. Optic disc and optic cup segmentation methodologies for glaucoma image detection: a survey. *Journal of ophthalmology*, Hindawi Publishing Corporation, v. 2015, 2015. Citado na página 13.
- BRADSKI, G.; KAEHLER, A. *Learning OpenCV: Computer vision with the OpenCV library*. [S.l.]: "O'Reilly Media, Inc.", 2008. Citado na página 32.
- BRYSON, A. E. *Applied optimal control: optimization, estimation and control*. [S.l.]: CRC Press, 1975. Citado na página 28.
- CARVALHO, É. G. d. *Desenvolvimento de um sistema óptico para retinografia e angiografia digital*. Tese (Doutorado) — Universidade de São Paulo, 2006. Citado na página 19.
- CHOLLET, F. et al. *Keras*. [S.l.]: GitHub, 2015. <<https://github.com/keras-team/keras>>. Acessado em 24 de Novembro de 2017. Citado na página 32.
- DICE, L. R. Measures of the amount of ecologic association between species. *Ecology*, Wiley Online Library, v. 26, n. 3, p. 297–302, 1945. Citado na página 36.
- DUMOULIN, V.; VISIN, F. A guide to convolution arithmetic for deep learning. *arXiv preprint arXiv:1603.07285*, 2016. Citado na página 34.
- FERNANDES, B. J. T. *Redes neurais com extração implícita de características para reconhecimento de padrões visuais*. Tese (Doutorado) — Universidade Federal de Pernambuco, 2013. Citado 3 vezes nas páginas 28, 29 e 30.
- FILHO, O. M.; NETO, H. V. *Processamento digital de imagens*. [S.l.]: Brasport, 1999. Citado 5 vezes nas páginas 18, 22, 24, 25 e 30.
- FURTADO, H.; VELHO, H. de C.; MACAU, E. Assimilação de dados com redes neurais artificiais em equações diferenciais. *Anais do DINCON*, 2011. Citado na página 27.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento digital de imagens*. [S.l.]: Pearson Prentice Hall, 2010. Citado 3 vezes nas páginas 22, 23 e 25.
- GOWRI, K.; BABU, T. G.; MAMANDUR, M. Segmentation of optic disc and optic cup using spatially weighted fuzzy c mean clustering and superpixel algorithm. In: IJARTET. *International Conference on Recent Advances in Communication Systems and Technologies (ICRACST'15)*. [S.l.], 2015. v. 2. ISSN 2394-3785. Citado na página 25.
- HOSHIRO, S.; GOLDSCHMIDT, R. R. *Processamento de imagens*. Rio de Janeiro: Seção de Engenharia de Computação, p. 1, 2008. Citado na página 21.
- HUA, K.-L.; HSU, C.-H.; HIDAYATI, S. C.; CHENG, W.-H.; CHEN, Y.-J. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *OncoTargets and therapy*, Dove Press, v. 8, 2015. Citado na página 29.

- HUNTER, R. S. Photoelectric color difference meter*. *J. Opt. Soc. Am.*, OSA, v. 48, n. 12, p. 985–995, Dec 1958. Disponível em: <<http://www.osapublishing.org/abstract.cfm?URI=josa-48-12-985>>. Citado na página 33.
- JACCARD, P. Étude comparative de la distribution florale dans une portion des alpes et des jura. *Bull Soc Vaudoise Sci Nat*, v. 37, p. 547–579, 1901. Citado na página 37.
- JOSHI, G. D.; SIVASWAMY, J.; KRISHNADAS, S. R. Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment. *IEEE Transactions on Medical Imaging*, v. 30, n. 6, p. 1192–1205, June 2011. ISSN 0278-0062. Citado 3 vezes nas páginas 16, 17 e 45.
- JUNEJA, M.; SANDHU, P. S.; WALIA, E. Application of lsb based steganographic technique for 8-bit color images. *World Academy of Science, Engineering and Technology*, v. 50, p. 423–425, 2009. Citado na página 24.
- KINGMA, D.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. Citado na página 38.
- KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: STANFORD, CA. *Ijcai*. [S.l.], 1995. v. 14, n. 2, p. 1137–1145. Citado na página 38.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F.; BURGESS, C. J. C.; BOTTOU, L.; WEINBERGER, K. Q. (Ed.). *Advances in Neural Information Processing Systems 25*. Curran Associates, Inc., 2012. p. 1097–1105. Disponível em: <<http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>>. Citado na página 34.
- LAURETTI, C. R.; FILHO, A. L. Glaucomas. *Medicina (Ribeirao Preto. Online)*, v. 30, n. 1, p. 56–65, 1997. Citado na página 18.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, Nature Research, v. 521, n. 7553, p. 436–444, 2015. Citado na página 29.
- LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, IEEE, v. 86, n. 11, p. 2278–2324, 1998. Citado na página 30.
- LECUN, Y.; KAVUKCUOGLU, K.; FARABET, C. Convolutional networks and applications in vision. In: IEEE. *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*. [S.l.], 2010. p. 253–256. Citado na página 29.
- LIM, G.; CHENG, Y.; HSU, W.; LEE, M. L. Integrated optic disc and cup segmentation with deep learning. In: *2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI)*. [S.l.: s.n.], 2015. p. 162–169. ISSN 1082-3409. Citado na página 13.
- MITTAPALLI, P. S.; KANDE, G. B. Segmentation of optic disk and optic cup from digital fundus images for the assessment of glaucoma. *Biomedical Signal Processing and Control*, v. 24, n. Supplement C, p. 34 – 46, 2016. ISSN 1746-8094. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1746809415001512>>. Citado 5 vezes nas páginas 16, 17, 44, 45 e 46.

NEGNEVITSKY, M. *Artificial intelligence: a guide to intelligent systems*. [S.l.]: Pearson Education, 2005. Citado 3 vezes nas páginas 26, 27 e 28.

OPTICIAN. *C51251: Glaucoma management part 2 - Optic disc assessment in glaucoma*. Optician, 2016. Acessado em 19 de Dezembro de 2017. Disponível em: <<https://www.opticianonline.net/cet-archive/118>>. Citado na página 21.

PIZER, S. M.; AMBURN, E. P.; AUSTIN, J. D.; CROMARTIE, R.; GESELOWITZ, A.; GREER, T.; ROMENY, B. ter H.; ZIMMERMAN, J. B.; ZUIDERVELD, K. Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, v. 39, n. 3, p. 355 – 368, 1987. ISSN 0734-189X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0734189X8780186X>>. Citado 2 vezes nas páginas 25 e 26.

PLATANIOTIS, K. N.; VENETSANOPOULOS, A. N. *Color image processing and applications*. [S.l.]: Springer Science & Business Media, 2013. Citado na página 23.

QUIGLEY, H. A.; BROMAN, A. T. The number of people with glaucoma worldwide in 2010 and 2020. *British Journal of Ophthalmology*, BMJ Publishing Group Ltd, v. 90, n. 3, p. 262–267, 2006. ISSN 0007-1161. Disponível em: <<http://bj.o.bmj.com/content/90/3/262>>. Citado na página 13.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. [S.l.], 2015. p. 234–241. Citado 2 vezes nas páginas 17 e 33.

ROYO, D. S. B. *Centro de Oftalmología Bonafonte. Cirugía de la Retina y del Vítreo*, 2015. Acessado em 19 de Dezembro de 2017. Disponível em: <<http://centrodeoftalmologiabonafonte.blogspot.com.br/2015/09/cirurgia-de-la-retina-y-del-vitreo.html>>. Citado na página 20.

SALEM, N. M.; NANDI, A. K. Novel and adaptive contribution of the red channel in pre-processing of colour fundus images. *Journal of the Franklin Institute*, Elsevier, v. 344, n. 3, p. 243–256, 2007. Citado na página 25.

SAUDE, M. da. *Protocolo Clínico e Diretrizes Terapêuticas: Glaucoma*. Ministério da Saúde, 2014. Disponível em: <<http://www.saudedireta.com.br/docsupload/1340495619PCDT20Glaucoma.pdf>>. Citado na página 19.

SCIENCEDAILY. *Ophthalmology: Automated method to detect glaucoma in its early stages*. ScienceDaily, 2013. Acessado em 20 de Dezembro de 2017. Disponível em: <<https://www.sciencedaily.com/releases/2013/09/130911125309.htm>>. Citado na página 21.

SEDAI, S.; ROY, P. K.; MAHAPATRA, D.; GARNAVI, R. Segmentation of optic disc and optic cup in retinal fundus images using shape regression. In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.: s.n.], 2016. p. 3260–3264. ISSN 1557-170X. Citado 3 vezes nas páginas 16, 17 e 45.

- SEVASTOPOLSKY, A. Optic disc and cup segmentation methods for glaucoma detection with modification of u-net convolutional neural network. *CoRR*, abs/1704.00979, 2017. Disponível em: <<http://arxiv.org/abs/1704.00979>>. Citado 6 vezes nas páginas 10, 17, 35, 40, 41 e 45.
- SHYNK, J. J. Performance surfaces of a single-layer perceptron. *IEEE Transactions on Neural Networks*, IEEE, v. 1, n. 3, p. 268–274, 1990. Citado na página 27.
- SILVA, G. L. F. d. et al. *Diagnóstico de nódulos pulmonares em imagens de tomografia computadorizada usando redes neurais convolucionais evolutivas*. Dissertação (Mestrado) — Universidade Federal do Maranhão, 2017. Citado 2 vezes nas páginas 29 e 46.
- SIVASWAMY, J.; KRISHNADAS, S.; JOSHI, G. D.; JAIN, M.; TABISH, A. U. S. Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation. In: *IEEE Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on*. [S.l.], 2014. p. 53–56. Citado 5 vezes nas páginas 17, 21, 26, 31 e 44.
- SPAETH, G. Appearances of the optic disc in glaucoma: a pathogenetic classification. In: *Trans Acad Ophthalmol Symposium on Glaucoma*. CV Mosby, St. Louis. [S.l.: s.n.], 1981. Citado na página 21.
- TATHAM, A. J.; WEINREB, R. N.; ZANGWILL, L. M.; LIEBMANN, J. M.; GIRKIN, C. A.; MEDEIROS, F. A. The relationship between cup-to-disc ratio and estimated number of retinal ganglion cells and retinal ganglion cell number. *Investigative ophthalmology & visual science*, The Association for Research in Vision and Ophthalmology, v. 54, n. 5, p. 3205–3214, 2013. Citado 2 vezes nas páginas 36 e 44.
- THAM, Y.-C.; LI, X.; WONG, T. Y.; QUIGLEY, H. A.; AUNG, T.; CHENG, C.-Y. Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis. *Ophthalmology*, v. 121, n. 11, p. 2081 – 2090, 2014. ISSN 0161-6420. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0161642014004333>>. Citado na página 13.
- UNESP. UNESP, 2008. Acessado em 19 de Dezembro de 2017. Disponível em: <<http://www.ibb.unesp.br/Home/MuseuEscola/>>. Citado na página 20.
- VOLPI, A. *Ferramentas de Python para Aprendizado de Máquina*. Alexandre Volpi, 2015. Acessado em 20 de Dezembro de 2017. Disponível em: <<https://alexandrevolpi.wordpress.com/tag/machine-learning/>>. Citado na página 28.
- ZHANG, Z.; SRIVASTAVA, R.; LIU, H.; CHEN, X.; DUAN, L.; WONG, D. W. K.; KWOH, C. K.; WONG, T. Y.; LIU, J. A survey on computer aided diagnosis for ocular diseases. *BMC Medical Informatics and Decision Making*, v. 14, n. 1, p. 80, Aug 2014. ISSN 1472-6947. Disponível em: <<https://doi.org/10.1186/1472-6947-14-80>>. Citado 3 vezes nas páginas 19, 20 e 21.
- ZHOU, M.; JIN, K.; WANG, S.; YE, J.; QIAN, D. Color retinal image enhancement based on luminosity and contrast adjustment. *IEEE Transactions on Biomedical Engineering*, PP, n. 99, p. 1–1, 2017. ISSN 0018-9294. Citado na página 33.
- ZIJDENBOS, A. P.; DAWANT, B. M.; MARGOLIN, R. A.; PALMER, A. C. Morphometric analysis of white matter lesions in mr images: method and validation.

IEEE Transactions on Medical Imaging, v. 13, n. 4, p. 716–724, Dec 1994. ISSN 0278-0062. Citado na página 37.

ZILLY, J.; BUHMANN, J. M.; MAHAPATRA, D. Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. *Computerized Medical Imaging and Graphics*, v. 55, n. Supplement C, p. 28 – 41, 2017. ISSN 0895-6111. Special Issue on Ophthalmic Medical Image Analysis. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0895611116300775>>. Citado 4 vezes nas páginas 14, 16, 17 e 45.

ZUIDERVELD, K. Graphics gems iv. In: HECKBERT, P. S. (Ed.). San Diego, CA, USA: Academic Press Professional, Inc., 1994. cap. Contrast Limited Adaptive Histogram Equalization, p. 474–485. ISBN 0-12-336155-9. Disponível em: <<http://dl.acm.org/citation.cfm?id=180895.180940>>. Citado na página 25.