



Universidade Federal do Piauí
Centro de Ciências da Natureza
Programa de Pós-Graduação em Ciência da Computação

Metodologia para Identificação de Glaucoma em Imagens de Retina.

Maíla de Lima Claro

Número de Ordem PPGCC: M001

Teresina-PI, Março de 2018

Maíla de Lima Claro

Metodologia para Identificação de Glaucoma em Imagens de Retina.

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da UFPI (área de concentração: Sistemas de Computação), como parte dos requisitos necessários para a obtenção do Título de Mestre em Ciência da Computação.

Universidade Federal do Piauí – UFPI

Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação

Orientador: Rodrigo de Melo Souza Veras

Coorientador: André Macêdo Santana

Teresina-PI

Março de 2018

Maíla de Lima Claro

Metodologia para Identificação de Glaucoma em Imagens de Retina./ Maíla de
Lima Claro. – Teresina-PI, Março de 2018-

61 p. : il. (algumas color.) ; 30 cm.

Orientador: Rodrigo de Melo Souza Veras

Dissertação (Mestrado) – Universidade Federal do Piauí – UFPI

Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação, Março de 2018.

1. Classificação. 2. Descritores. 3. Glaucoma. I. Rodrigo Melo de Souza Veras.
II. Universidade Federal do Piauí. III. Metodologia para Identificação de Glaucoma
em Imagens de Retina.

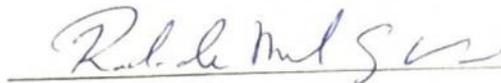
CDU 02:141:005.7

"Metodologia para Identificação de Glaucoma em Imagens de Retina"

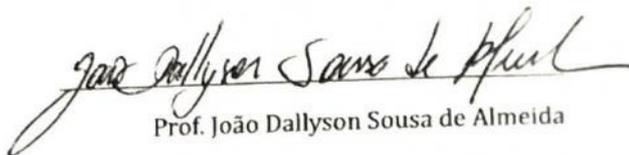
MAILA DE LIMA CLARO

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Natureza da Universidade Federal do Piauí, como parte integrante dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Aprovada por:

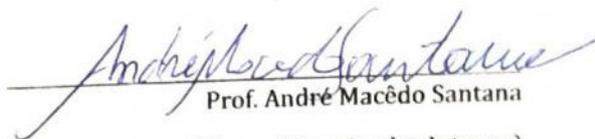


Prof. Rodrigo de Melo Souza Vêras
(Presidente da Banca Examinadora)



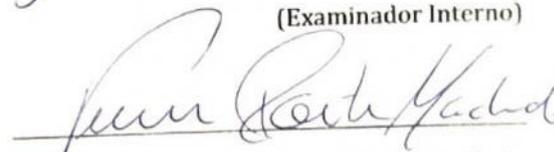
Prof. João Dallyson Sousa de Almeida

(Examinador Externo)



Prof. André Macêdo Santana

(Examinador Interno)



Prof. Vinícius Ponte Machado

(Examinador Interno)

Teresina, 22 de março de 2018

*Aos meus pais Maria Vilani de Lima Claro e José Claro Neto,
por sempre estarem comigo em todos os momentos.*

Agradecimentos

À Deus, meu Pai bondoso e misericordioso, meu Salvador e Senhor, por seu amor eterno e incondicional.

A toda minha família pelo amor, suporte e incentivo. Em especial aos meus pais, José Claro Neto e Maria Vilani de Lima Claro, pela dedicação e esforço em me proporcionar sempre o melhor e estarem ao meu lado em todos os momentos. Aos meus irmãos Maísa de Lima Claro e Maelson de Lima Claro pelos momentos de descontração.

Agradeço ao meu orientador, Rodrigo Melo de Souza Veras, pela paciência, compreensão, disponibilidade e grande contribuição para realização deste trabalho.

A todos os amigos. Em especial, Flávio Henrique Duarte de Araújo, Romuere Rodrigues Veloso e Silva, Alcilene Dalília de Sousa e Nayara Holanda de Moura, que colaboraram, direta ou indiretamente, para a realização deste trabalho, pelo carinho e aprendizado.

Aos professores do Departamento de Computação pelos ensinamentos e apoio.

À Universidade Federal do Piauí pelas oportunidades concedidas e apoio à qualificação profissional e desenvolvimento de pesquisa.

*“O Meu ou o seu caminho
não são muito diferentes
tem espinho, pedra, buraco
pra mode atrasar a gente.
Não desanime por nada
pois até uma topada
empurra você pra frente!”
(Bráulio Bessa)*

Resumo

O Glaucoma é uma doença ocular que danifica o nervo óptico causando a perda da visão, é considerada a segunda principal causa de cegueira no mundo, ficando atrás apenas da catarata. Com o objetivo de auxiliar especialistas no diagnóstico desta doença, sistemas de auxílio por computador são utilizados para reduzir a possibilidade da prescrição de tratamentos inadequados. Assim, vem sendo desenvolvidos métodos de diagnóstico automático de glaucoma. Contudo é possível realizar melhorias nestas técnicas, visto que, os métodos atuais demonstram não ser capazes de lidar com uma grande diversidade de imagens. Existem duas regiões das imagens da retina que são de suma importância para a detecção do glaucoma, a região do disco óptico e a região da escavação. Este trabalho apresenta uma nova metodologia de diagnóstico do glaucoma aplicada nas imagens da retina. É aplicado descritores de textura e Redes Neurais Convolucionais (CNNs) nas imagens da retina sem segmentação (imagem original) e em imagens segmentadas na região do disco óptico, com o intuito de avaliar o melhor cenário a ser aplicado os descritores. As características obtidas são selecionadas de acordo com a sua razão de ganho de informação. O método proposto tem como objetivo ser capaz de classificar corretamente imagens com diferentes características advindas de diferentes bases de imagens. Os resultados mostraram que a junção dos descritores GLCM e CNNs e a utilização do classificador *Random Forest* são promissores na detecção dessa patologia, obtendo uma acurácia de 93,35% em 873 imagens de quatro bases de dados públicas.

Palavras-chaves: Imagens Médicas. Diagnóstico de glaucoma. Características de textura. CNNs Pré-treinadas.

Abstract

Glaucoma is an eye disease that damages the optic nerve causing vision loss, it's considered the second leading cause of blindness in the world, falling behind only the cataract. In order to assist specialists in the diagnosis of this disease, computer aided systems are used to reduce the possibility of prescribing inappropriate treatments. Thus, automatic diagnostic systems for glaucoma have been developed. However, it is possible to make improvements in these techniques, since current systems do not deal with a great diversity of images. There are two regions of the retinal images that are of great importance for the detection of glaucoma, being the regions of the optic disc and the excavation. This work presents a new methodology for the creation of a diagnostic method for glaucoma applied to retinal images. It will be applied texture descriptors and Convolutional Neural Networks (CNNs) in the images of the retina without segmentation (original image) and in segmented images in the region of the optical disc. The characteristics obtained are selected according to their information gain ratio. The proposed system aims to be able to correctly classify images with different characteristics derived from different image bases. The results showed that the junction of the GLCM and CNNs descriptors and the use of the Random Forest classifier are promising in the detection of this pathology, obtaining an accuracy of 93.35% in 873 images from four public databases.

Keywords: Medical images. Glaucoma diagnosis. Texture features. Transfer learning.

Lista de ilustrações

| | |
|--|----|
| Figura 1 – Esquema de um corte longitudinal do olho humano. | 17 |
| Figura 2 – Principais estruturas de uma retina. | 18 |
| Figura 3 – Exemplos de imagens da base RIM-ONE-3 com DO (marcação branca) e <i>cup</i> (marcação verde) em evidência: (a) Imagem de uma retina saudável e (b) Imagem de um retina glaucomatosa. | 19 |
| Figura 4 – Exemplos de imagens da base DRISHTI (a) Imagem de uma retina saudável (b) Padrão Ouro da imagem de retina saudável (c) Imagem de um retina glaucomatosa e (d) Padrão ouro da imagem de retina glaucomatosa. | 26 |
| Figura 5 – Exemplos de imagens da base HRF (a) Imagem de uma retina saudável e (b) Imagem de um retina glaucomatosa. | 26 |
| Figura 6 – Exemplos de imagens da base RIM-ONE-1 (a) Imagem de uma retina saudável (b) Padrão Ouro da imagem de retina saudável (c) Imagem de um retina glaucomatosa e (d) Padrão ouro da imagem de retina glaucomatosa. | 27 |
| Figura 7 – Exemplos de imagens da base RIM-ONE-2 (a) Imagem de uma retina saudável (b) Padrão Ouro da imagem de retina saudável (c) Imagem de um retina glaucomatosa e (d) Padrão ouro da imagem de retina glaucomatosa. | 27 |
| Figura 8 – Exemplos de imagens da base RIM-ONE-3 (a) Imagem de uma retina saudável (b) Padrão Ouro da imagem de retina saudável (c) Imagem de um retina glaucomatosa e (d) Padrão ouro da imagem de retina glaucomatosa. | 28 |
| Figura 9 – Demonstração do funcionamento do descritor de textura LBP. | 28 |
| Figura 10 – Demonstração do cálculo entre a co-ocorrência dos níveis de cinza para geração da matriz de co-ocorrência. | 29 |
| Figura 11 – Representação do descritor HOG. | 31 |
| Figura 12 – Ilustração da arquitetura de uma CNN simplificada. | 35 |
| Figura 13 – Fluxograma do método proposto neste trabalho. | 43 |
| Figura 14 – Comparação dos resultados dos descritores GLCM, CNNs e GLCM+CNNs. | 51 |
| Figura 15 – Exemplos de sub e super-segmentação. | 53 |

Lista de tabelas

| | |
|--|----|
| Tabela 1 – Resumo dos Trabalhos Relacionados. | 24 |
| Tabela 2 – Matriz de confusão de duas classes (X,Y). | 41 |
| Tabela 3 – Nível de exatidão de uma classificação, conforme o valor do índice kappa. | 42 |
| Tabela 4 – Resultados da 1ª estratégia com o classificador SVM. | 46 |
| Tabela 5 – Resultados da 1ª estratégia com o classificador MLP. | 46 |
| Tabela 6 – Resultados da 1ª estratégia com o classificador KNN. | 47 |
| Tabela 7 – Resultados da 1ª estratégia com o classificador RT. | 47 |
| Tabela 8 – Resultados da 1ª estratégia com o classificador RF. | 47 |
| Tabela 9 – Resumo dos Melhores Resultados do Cenário 1. | 48 |
| Tabela 10 – Resultados da 2ª estratégia com o classificador SVM. | 48 |
| Tabela 11 – Resultados da 2ª estratégia com o classificador MLP. | 49 |
| Tabela 12 – Resultados da 2ª estratégia com o classificador KNN. | 49 |
| Tabela 13 – Resultados da 2ª estratégia com o classificador RT. | 50 |
| Tabela 14 – Resultados da 2ª estratégia com o classificador RF. | 50 |
| Tabela 15 – Resumo dos Melhores Resultados do Cenário 2. | 50 |
| Tabela 16 – Melhores resultados obtidos com a seleção de atributos. | 52 |
| Tabela 17 – Quantidade de atributos por descritor. | 53 |
| Tabela 18 – Resultados da classificação do método proposto em imagens com erros de segmentação. | 53 |
| Tabela 19 – Comparação do método proposto com os trabalhos relacionados. | 54 |
| Tabela 20 – Comparação do método proposto com o trabalho de Orlando et al. | 54 |

Lista de abreviaturas e siglas

| | |
|--------|--|
| CAD | <i>Computer Aided Diagnosis</i> |
| CDR | <i>Cup to Disc Ratio</i> |
| CNNs | <i>Convolutional Neural Networks</i> |
| DO | Disco Óptico |
| FP | Falso Positivo |
| FN | Falso Negativo |
| GLCM | <i>Gray Level Co-occurrence Matrix</i> |
| GLN | <i>Gray Level distribution</i> |
| GLRLM | <i>Gray Level Run Length Matrix</i> |
| HGLRE | <i>High Gray Level Run Emphasis</i> |
| HOG | <i>Histogram of Oriented Gradients</i> |
| HOS | <i>Higher Order Spectra</i> |
| HRF | <i>High-Resolution Fundus</i> |
| KNN | <i>K-Nearest Neighbors</i> |
| LBP | <i>Local Binary Pattern</i> |
| LGLRE | <i>Low Gray Level Run Emphasis</i> |
| LRE | <i>Long-Run Emphasis</i> |
| LRHGLE | <i>Long Run High Gray Level Emphasis</i> |
| LRLGLE | <i>Long Run Low Gray Level Emphasis</i> |
| MLP | <i>MultiLayer Perceptron</i> |
| NRR | <i>Neuro Retinal Rim</i> |
| OMS | Organização Mundial de Saúde |
| RF | <i>Random Forest</i> |

| | |
|--------|---|
| RP | <i>Run Percentage</i> |
| LRN | <i>Run Length distribution</i> |
| SRE | <i>Short-Run Emphasis</i> |
| SRHGLE | <i>Short Run High Gray Level Emphasis</i> |
| SRLGLE | <i>Short Run Low Gray Level Emphasis</i> |
| SVM | <i>Support Vector Machine</i> |
| VP | Verdadeiro Positivo |
| VN | Verdadeiro Negativo |

Sumário

| | | |
|----------|---|-----------|
| 1 | INTRODUÇÃO | 16 |
| 1.1 | Olho humano | 16 |
| 1.2 | Glaucoma | 17 |
| 1.3 | Objetivos | 19 |
| 1.3.1 | Objetivo Geral | 19 |
| 1.3.2 | Objetivo Específico | 19 |
| 1.4 | Organização da Dissertação | 20 |
| 2 | TRABALHOS RELACIONADOS | 21 |
| 2.1 | Cenário 1: Sem segmentação | 21 |
| 2.2 | Cenário 2: Com segmentação | 22 |
| 2.3 | Considerações Finais | 23 |
| 3 | MATERIAIS E MÉTODOS | 25 |
| 3.1 | Base de Imagens | 25 |
| 3.2 | Extração de características | 27 |
| 3.2.1 | LBP - <i>Local Binary Pattern</i> | 28 |
| 3.2.2 | GLCM - <i>Gray Level Co-occurrence Matrix</i> | 29 |
| 3.2.3 | HOG - <i>Histogram of Oriented Gradients</i> | 30 |
| 3.2.4 | <i>Tamura</i> | 31 |
| 3.2.5 | GLRLM - <i>Gray Level Run Length Matrix</i> | 32 |
| 3.2.6 | Morfologia | 34 |
| 3.2.7 | CNN - <i>Convolutional Neural Network</i> | 34 |
| 3.3 | Seleção de Atributos | 36 |
| 3.4 | Classificadores | 37 |
| 3.4.1 | SVM - <i>Support Vector Machine</i> | 37 |
| 3.4.2 | MLP - <i>MultiLayer Perceptron</i> | 38 |
| 3.4.3 | KNN - <i>K-Nearest Neighbors</i> | 38 |
| 3.4.4 | RT - <i>Random Tree</i> | 39 |
| 3.4.5 | RF - <i>Random Forest</i> | 39 |
| 3.5 | Métodos de Avaliação dos Resultados | 40 |
| 3.6 | Considerações Finais | 42 |
| 4 | MÉTODO PROPOSTO | 43 |
| 4.1 | Considerações Finais | 44 |

| | | |
|------------|---|-----------|
| 5 | RESULTADOS E DISCUSSÃO | 45 |
| 5.1 | Resultados sem segmentação | 45 |
| 5.2 | Resultados com a segmentação do disco óptico | 48 |
| 5.3 | Resultados com o descritor proposto | 50 |
| 5.4 | Considerações Finais | 54 |
| 6 | CONCLUSÕES E TRABALHOS FUTUROS | 55 |
| 6.1 | Conclusões | 55 |
| 6.2 | Trabalhos Futuros | 55 |
| | REFERÊNCIAS | 57 |

1 Introdução

O processo de globalização tem contribuído significativamente para avanços no campo da ciência, principalmente na área tecnológica. No contexto atual, vários são os benefícios que essa área tecnológica tem proporcionado ao ser humano, como meio de facilitar e auxiliar no desenvolvimento das tarefas diárias. Assim, pesquisas que utilizam abordagens computacionais tem se encaminhado para auxiliar no diagnóstico de doenças, e o processamento digital de imagens tem desempenhado papel relevante na detecção destas patologias (FENGSHOU, 2011). Com base nestas pesquisas, surgiu um considerável interesse no desenvolvimento de sistemas automáticos de diagnóstico médico. Esses sistemas são chamados de sistemas de Diagnóstico Auxiliado por Computador (CAD - *Computer Aided Diagnosis*).

Os sistemas CAD utilizados na análise de imagens médicas são definidos como sistemas computadorizados incorporados ao reconhecimento de padrões e habilidades na análise de dados, para que assim, possa alertar os especialistas para porções específicas de imagens médicas (LIN; HUANG; CHENG, 2016). Os estudos que abrangem esses sistemas comprovaram que eles podem auxiliar médicos em diversas tarefas como: medição de estruturas anatômicas, monitoração das mudanças pela comparação de imagens sequenciais e no diagnóstico e planejamento do tratamento. As aplicações de CAD mais estabelecidas nas áreas médicas envolvem o uso de sistemas automatizados em mamografia, tomografia computadorizada de tórax e radiografia (GONZALEZ; WOODS, 2011).

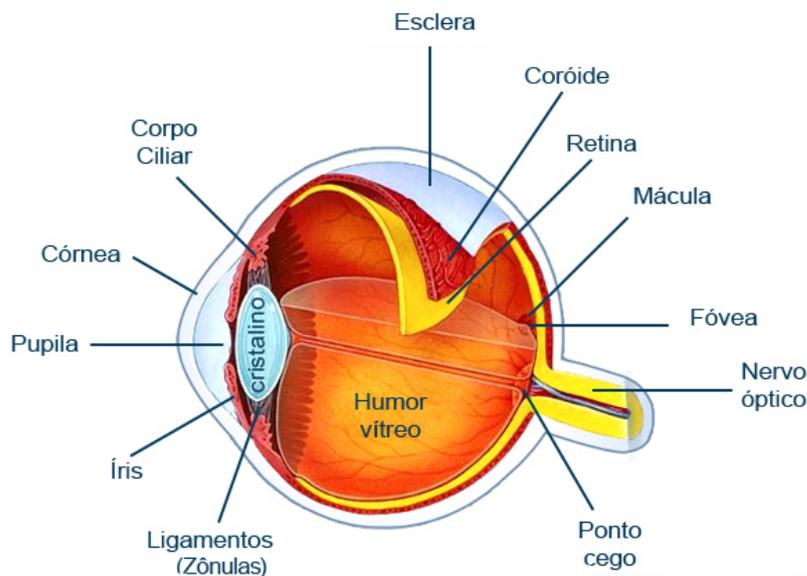
Foram desenvolvidos um grande número de sistemas CAD para o diagnóstico de vários tipos de doenças oculares, entre elas o glaucoma (USMAN; FRAZ; BARMAN, 2016). Estes sistemas têm potencial para fornecer uma solução alternativa aos programas de triagem em massa, que precisam examinar um grande número de imagens de fundo de olho de forma eficiente e robusta.

1.1 Olho humano

O olho humano é um sistema sensorial complexo que desempenha funções de formação e processamento de imagem em conexão com o cérebro. Ele capta a energia luminosa refletida pelos objetos e a transforma em impulsos nervosos transmitidos ao cérebro, que serão convertidos na imagem do objeto observado (PASSOS; ANDRADE-NETO; LEMAIRE, 2008). A Figura 1 demonstra o olho humano com as suas estruturas.

A retina é uma parte do olho humano responsável pela formação de imagens. Ela é como uma tela, onde as imagens se projetam, que arquiva as imagens e as traduz para o

Figura 1 – Esquema de um corte longitudinal do olho humano.



Fonte: www.optivista.com.br.

cérebro através de impulsos elétricos enviados pelo nervo óptico (RAMOS, 2006).

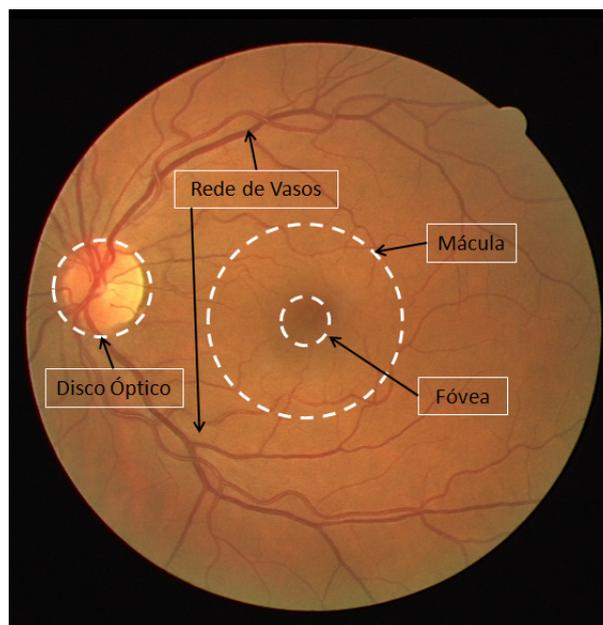
O Disco Óptico (DO) é a porção do nervo óptico vista no fundo dos olhos, formado pelo encontro de todos os axônios das células ganglionares da retina, assim que penetram no disco óptico. Essa região do DO pode ser enxergado pelo oftalmologista com o auxílio de um oftalmoscópio (aparelho oftalmológico para enxergar patologias nos olhos não vistas a olho nu) (PASSOS; ANDRADE-NETO; LEMAIRE, 2008), como pode ser observado na Figura 2. O indivíduo que possui glaucoma tem a pressão intra-ocular aumentada significativamente, o que pode causar danos irreversíveis ao disco óptico, acarretando na atrofia de modo progressivo do campo visual, podendo causar a cegueira. Lembrando que, se houver um tratamento de glaucoma adequado este risco é menor.

1.2 Glaucoma

O glaucoma é uma doença crônica, irreversível e neurodegenerativa (com perda progressiva), que pode ocasionar a perda da visão nos pacientes. O glaucoma, em sua fase inicial é assintomático, na maioria dos casos a doença progride lentamente sem que o paciente se dê conta de que está perdendo gradualmente a visão. As manifestações clínicas só aparecem em um estado mais avançado da doença (ZHANG et al., 2009). Portanto, a descoberta precoce desta patologia é imprescindível, pois existem tratamentos que evitam a sua progressão e, conseqüentemente, a perda da visão do paciente.

Os principais sinais de alerta desta doença é a dificuldade para focar objetos, visão embaçada, dificuldade para enxergar no escuro, diminuição da visão periférica, desconforto

Figura 2 – Principais estruturas de uma retina.



Fonte: <http://www.repositorio.ufc.br/handle/riufc/10851>.

ou dor nos olhos, entre outros. A perda da visão provocada pelo glaucoma é perfeitamente evitável, desde que o diagnóstico e o tratamento sejam feitos pelo oftalmologista no início do aparecimento da doença (LAMANI et al., 2014).

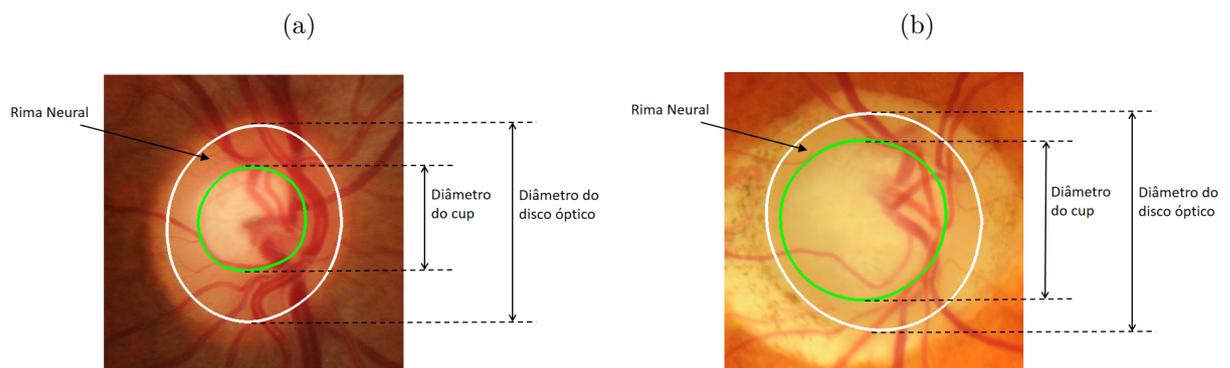
De acordo com a Organização Mundial de Saúde - OMS (*World Health Organization*), existem cerca de 60 milhões de glaucomatosos em todo o mundo, sendo que, a cada ano, surgem mais 2,4 milhões de casos. Uma estimativa sugere que, em 2020, cerca de 80 milhões de pessoas terão essa doença (QUIGLEY; BROMAN, 2006).

A incidência do glaucoma é estimada de 1% a 2% na população geral, aumentando após os 40 anos (2%), podendo chegar a 6% ou 7% após os 70 anos de idade. A OMS estima que 4,5 milhões de pessoas são cegas devido ao glaucoma. Dado o envelhecimento da população mundial, este número pode aumentar para quase 80 milhões até 2020. Em 2003, o Conselho Brasileiro de Oftalmologia estimava que no país houvesse 900 mil portadores dessa doença e que, provavelmente, 720 mil estavam assintomáticos, ainda necessitando de diagnóstico (ÁVILA; ALVES; NISHI, 2015). Um estudo populacional conduzido na região sul do Brasil selecionou aleatoriamente 1636 indivíduos acima de 40 anos, e destes, encontrou-se uma prevalência de Glaucoma em 3,4% dos casos. Ainda de acordo com essa pesquisa, 90% dos glaucomatosos desconheciam seu diagnóstico (SAKATA et al., 2007).

O diagnóstico de glaucoma pode ser realizado a partir da análise de imagens digitais de retina. O mesmo só é possível devido a quantidade de perda de fibras do nervo óptico ter um efeito direto na configuração da *Neuro Retinal Rim* (NRR) ou também denominado de rima neural. À medida que as fibras ópticas nervosas vão morrendo, a escavação (*cup*) torna-se mais larga em relação ao disco óptico (DO), o que acarreta em um valor da relação

Escavação/Disco (*Cup to Disc Ratio* - CDR) aumentado (MITTAPALLI; KANDE, 2016), como pode ser visto na Figura 3. A região do DO (representada por toda a área dentro da marcação branca), e a região da escavação (toda a área dentro da marcação verde) estão em evidência. A Figura 3a apresenta uma imagem de retina saudável, enquanto que a Figura 3b apresenta uma imagem com glaucoma.

Figura 3 – Exemplos de imagens da base RIM-ONE-3 com DO (marcação branca) e *cup* (marcação verde) em evidência: (a) Imagem de uma retina saudável e (b) Imagem de um retina glaucomatosa.



1.3 Objetivos

1.3.1 Objetivo Geral

O intuito deste trabalho consiste em dinamizar a rotina clínica nos exames de detecção do glaucoma através do uso de imagens digitais da retina. Assim, o objetivo principal desta dissertação é realizar um estudo, avaliação e desenvolvimento de um método de detecção automática do glaucoma em imagens digitais da retina. Para isso o método é aplicado em cinco bases de dados públicas uma estratégia de extração características com a junção dos descritores *Gray Level Co-occurrence Matrix* GLCM e redes neurais convolucionais, para que na sequência posso classificar a imagem como sendo glaucomatosa ou não.

1.3.2 Objetivo Específico

Durante a pesquisa foram estabelecidos os seguintes objetivos específicos:

- Construção de uma base de dados mais diversificada, a partir de base de imagens públicas do fundo de olho;
- Estudo e teste de algoritmos de extração de características em imagens (descritores de Redes Neurais Convolucionais-CNNs, textura e forma);

- Estudo e teste da combinação de descritores (descritores híbridos);
- Análise comparativa da combinação entre descritores e classificadores;
- Desenvolver um método para detecção automática do glaucoma a partir de imagens digitais da retina.

1.4 Organização da Dissertação

Este documento está estruturado em 6 capítulos. O Capítulo 1 apresentou uma introdução sobre o conteúdo que será abordado. No Capítulo 2 é realizada uma apresentação dos principais trabalhos encontrados na literatura sobre a detecção do glaucoma. O Capítulo 3 mostra as principais ferramentas e procedimentos computacionais utilizados, tais como: as bases de imagens, descritores, classificadores e as métricas de avaliação dos resultados. O Capítulo 4 mostra o detalhamento do método proposto e toda a abordagem utilizada para o desenvolvimento do mesmo. O Capítulo 5 mostra os resultados deste trabalho. Ao final, o Capítulo 6 apresenta as conclusões e trabalhos futuros.

2 Trabalhos Relacionados

Realizou-se uma revisão na literatura objetivando identificar quais técnicas estão sendo utilizadas no diagnóstico do glaucoma em imagens da retina, para viabilizar a percepção dos descritores e classificadores que estão sendo aplicados. Estes estudos serviram como base para a realização deste trabalho. Assim, este trabalho divide-se em dois cenários obtidos a partir da análise dos trabalhos, sendo eles: utilização da imagem original (Sem segmentação) para extração das características e a outra estratégia utiliza a região do disco óptico (DO) (Com segmentação).

2.1 Cenário 1: Sem segmentação

O trabalho de Acharya et al. (ACHARYA et al., 2011) apresentou a combinação de descritores de textura e de *Higher Order Spectra* (HOS) em 60 imagens digitais da retina, para extração de características. Essas imagens foram coletadas da *Kasturba Medical College* no Manipal na Índia onde 30 delas são normais e 30 possuem glaucoma. Foram utilizados alguns classificadores, sendo que o *Random Forest* (RF) (BREIMAN, 2001) foi o que apresentou melhor resultado, obtendo uma acurácia da ordem de 91%.

Simonthomas, Thulasi e Asharaf (2014), apresentaram um método de detecção do glaucoma utilizando o descritor de textura de Haralick (HARALICK; SHANMUGAM et al., 1973), sendo utilizado o classificador *K-Nearest Neighbors* (KNN). As imagens utilizadas neste trabalho foram as mesmas do trabalho de Acharya et al. (2011). As treze características de Haralick foram combinadas, para que tivessem um melhor desempenho, obtendo uma taxa de acurácia de 98%.

Balasubramanian et al. (2016) propuseram uma avaliação sobre a detecção do glaucoma utilizando o *Histogram of Oriented Gradients* (HOG) (DALAL; TRIGGS, 2005) para a extração de características. Antes da utilização deste descritor é realizada a remoção dos vasos sanguíneos utilizando o filtro de gabor, essa remoção visa melhorar a precisão. A imagem serve como entrada para o descritor HOG, onde logo após é utilizado o classificador *Support Vector Machine* (SVM) (CORTES; VAPNIK, 1995) para classificar a imagem como glaucomatosa ou não. Foi encontrada uma acurácia de 83,3% e uma sensibilidade de 75%.

2.2 Cenário 2: Com segmentação

Neste cenário são utilizadas para o processamento de imagens apenas as regiões de interesse, obtidas a partir da segmentação destas regiões. Nas imagens da retina existem duas principais regiões: disco óptico e escavação. A seguir é apresentado trabalhos encontrados na literatura com a utilização destas áreas.

Srinivasan, Dubey e Ganeshbabu (2016), aplicaram duas importantes técnicas de descritores de textura, a GLCM e *Local Binary Pattern* (LBP) (OJALA; PIETIKÄINEN, 1997) em 200 imagens da retina. A segmentação do DO é utilizada para a extração dessas características, onde logo após, o SVM é utilizado para a classificação da base de dados gerada por estas características. O sistema proposto neste trabalho apresentou resultados promissores com uma sensibilidade de 100%, especificidade de 99% e acurácia de 99%.

Nirmala e Venkateswaran (2016) apresentaram uma metodologia para o diagnóstico de glaucoma pela extração de características do descritor *Local Binary Pattern* (LBP) da imagem do fundo da retina. Primeiro, a imagem da retina é pré-processada utilizando o método de equalização de histograma e a região de interesse (disco óptico) é segmentada utilizando o padrão ouro. Em seguida, o LBP é utilizado na imagem melhorada. Por fim, essas características são classificadas com o classificador SVM. As bases de imagens utilizadas foram as bases de dados FAU e RIM-ONE. Este trabalho alcançou uma especificidade de 91,11%, sensibilidade de 90,50% e acurácia de 92%.

Claro et al. (2016) propuseram a detecção do glaucoma em imagens da retina. Os autores utilizaram informação de textura com o uso da GLCM juntamente com a entropia da imagem. Os modelos de cores utilizados foram o *Red, Green, Blue* - (RGB), *Hue, Saturation, Intensity* - (HSI) e *Luminosity*, cromaticidade U e cromaticidade V - (LUV). A base de dados utilizada foi a RIM-ONE-1 e obtiveram uma acurácia da ordem de 93,03% aplicando o classificador *MultiLayer Perceptron* (MLP) (AMARI, 1967).

Koty et al. (KOTYK et al., 2016), os autores propuseram um sistema semi-automatizado para detectar tanto o DO quanto a escavação. Essas duas regiões são utilizadas no cálculo do CDR. O canal vermelho do modelo de cor RGB é utilizado nas imagens da retina da base RIM-ONE-1. Em sequência, valores de limiar são utilizados para segmentação dessas regiões. As imagens são posteriormente processadas usando o descritor de textura de Haralick e os valores obtidos são treinados usando uma rede neural para determinar a exatidão do sistema. A acurácia do sistema proposto foi da ordem de 86,43%.

Outro estudo foi o de Salam et al. (SALAM et al., 2016), onde os autores combinaram os descritores estruturais (CDR) e descritores não estruturais (textura e intensidade) para uma melhor precisão do diagnóstico automatizado. Esse algoritmo utiliza 100 imagens adquiridas em uma base de dados local, essas imagens são pré-processadas e extraem

as regiões do DO e da escavação com a finalidade de aplicação no cálculo do CDR. Os descritores de textura utilizados foram: *Wavelets*, *Multiwavelets*, LBP, GLCM e as características de Haralick. Já para a descrição de intensidade foram utilizados o *Color Moments* e *Autocorrelogram*. Esses descritores foram aplicados nas imagens de treinamento e teste do classificador SVM, onde a sensibilidade, especificidade e acurácia média do sistema proposto são de 100%, 87% e 91%, respectivamente.

Sousa et al. (2017) propuseram um método que analisa a textura da região da imagem do disco óptico para diagnosticar o glaucoma. Esta análise é feita usando o Padrão Binário Local (LBP) para representar a região do disco óptico e funções geoestatísticas para descrever padrões de textura. É utilizada 455 imagens da retina e as características de textura obtidas são usadas para classificação baseada em SVM. O método proposto apresentou como melhores resultados uma sensibilidade de 95%, acurácia de 91% e especificidade de 88% no diagnóstico de glaucoma. O método demonstrou ser promissor na assistência ao diagnóstico de glaucoma.

Orlando et al. (2017) apresentaram resultados de uma análise da utilização das CNNs pré-treinadas, a partir de imagens da retina para detecção automatizada do glaucoma. Duas CNNs foram aplicadas nas imagens da base de dados DRISHTI para geração dos vetores de características, sendo elas a VGG-S e a Overfeat. São utilizadas técnicas de pré-processamento da imagem com o objetivo de melhorar a imagem para a etapa de extração de características, sendo estas técnicas a equalização de histograma, pintura dos vasos e um corte na imagem na região da cabeça do nervo óptico. A área sob a curva ROC (AUC) foi utilizada como métrica de avaliação tendo obtido resultado de 0,7626 no classificador *Logistic Regression* (LR).

2.3 Considerações Finais

Neste capítulo foram apresentados estudos encontrados na literatura que abordam a identificação do glaucoma em imagens da retina. Um resumo dos trabalhos apresentados está demonstrado na Tabela 1. Através desta pesquisa é possível verificar que existem trabalhos que utilizam tanto a imagem sem nenhum pré-processamento como também utilizam imagens com pré-processamento. Descritores de textura, como GLCM, LBP e HOG estão sendo utilizados na etapa de extração de características. Essas pesquisas que envolvem a aplicabilidade de texturas são motivadas, principalmente, pela alta eficiência do sistema visual humano em identificá-las e diferenciá-las.

Tabela 1 – Resumo dos Trabalhos Relacionados.

| Sem segmentação | | | | |
|---------------------------------------|-----------------------|---------------|---------------|----------------------|
| Trabalhos | Descritor | Classificador | Qt de Imagens | Métrica de Avaliação |
| Acharya et al. (2011) | HOS | RF | 60 | A = 91% |
| Simonthomas, Thulasi e Asharaf (2014) | GLCM e Haralick | KNN | 60 | A = 98% |
| Balasubramanian et al. (2016) | HOG | SVM | 30 | A = 83,3% |
| Com segmentação | | | | |
| Trabalhos | Descritor | Classificador | Qt de Imagens | Acurácia |
| Srinivasan, Dubey e Ganeshbabu (2016) | GLCM e LBP | SVM | 200 | A = 99% |
| Nirmala e Venkateswaran (2016) | LBP | SVM | - | A = 92% |
| Claro et al. (2016) | GLCM | MLP | 158 | A = 93,03% |
| Kotyky et al. (2016) | Haralick | MLP | 158 | A = 86,43% |
| Salam et al. (2016) | Intensidade e textura | SVM | 100 | A = 91% |
| Sousa et al. (2017) | LBP | SVM | 455 | A = 91% |
| Orlando et al. (2017) | CNNs | LR | 101 | AUC = 0,762% |

3 Materiais e Métodos

Nesta Seção são apresentados os materiais e métodos utilizados para a detecção automática do glaucoma, de modo a esclarecer os recursos utilizados para alcançar o resultado. Será apresentada e descrita as bases de imagens da retina adquiridas para este trabalho. Em sequência, serão expostos os descritores utilizados na etapa de extração de características e logo após os classificadores utilizados na classificação da imagem. Por fim, os métodos de avaliação dos resultados utilizados são apresentados.

3.1 Base de Imagens

Um dos desafios no processamento de imagens de glaucoma é a habilidade de diagnosticar a doença em bases de dados com características distintas. No entanto, vários autores testaram seus sistemas com apenas uma base de dados, ou então nas suas próprias bases. Nesse trabalho utilizamos cinco bases de imagens, a base DRISHTI (SIVASWAMY et al., 2014), a base *High-Resolution Fundus* (HRF) (FUNDUS, 2012) e a RIM-ONE (FUMERO et al., 2011) em suas 3 versões.

DRISHTI-GS - *Retinal image dataset for optic nerve head segmentation* é uma base de imagens que foi desenvolvida e distribuída com o intuito de fornecer meios para avaliação de sistemas CADs na detecção do glaucoma. As imagens foram previamente analisadas por quatro especialistas para a segmentação das regiões do DO e da escavação. A base é composta por 101 imagens de retina, todas foram tiradas com centro no DO, com um “campo de visão” de 30° e dimensões 2896×1944 pixels em formato PNG. Apenas 50 desse total possuem informações sobre seus padrões ouro, então normalmente apenas elas são utilizadas para testes. A Figura 4 apresenta exemplos de uma imagem saudável (a) e seu padrão ouro representado em (b) e uma imagem com glaucoma (c) e seu padrão ouro (d).

HRF - *High-Resolution Fundus* é um banco de dados que foi estabelecido por um grupo de pesquisa colaborativo para apoiar estudos comparativos sobre algoritmos de segmentação automática em imagens de fundo de retina (FUNDUS, 2012). Este banco de dados pode ser usado livremente para fins de pesquisa, ele contém no momento 15 imagens de pacientes saudáveis, 15 imagens de pacientes com retinopatia diabética e 15 imagens de pacientes glaucomatosos. Imagens binárias de segmentação de padrão ouro dos vasos estão disponíveis para cada imagem, contudo, esta base não possui o padrão ouro das regiões do disco óptico e da escavação. A Figura 5 apresenta exemplos de uma imagem saudável (a) e uma com glaucoma (b).

Figura 4 – Exemplos de imagens da base DRISHTI (a) Imagem de uma retina saudável (b) Padrão Ouro da imagem de retina saudável (c) Imagem de um retina glaucomatosa e (d) Padrão ouro da imagem de retina glaucomatosa.

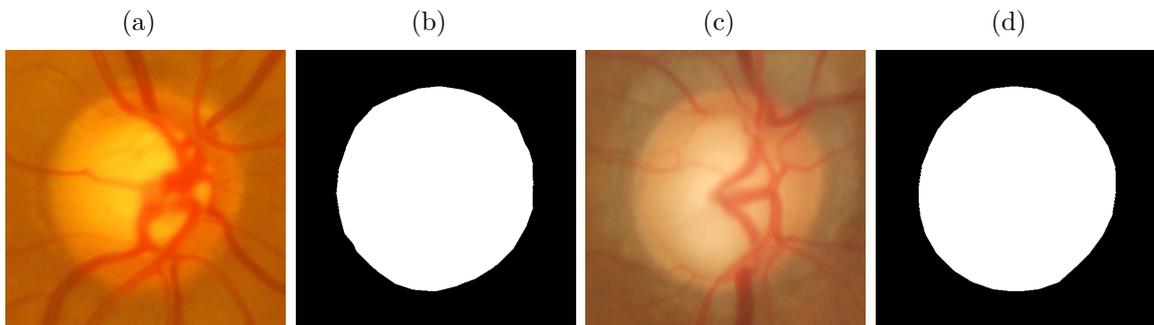
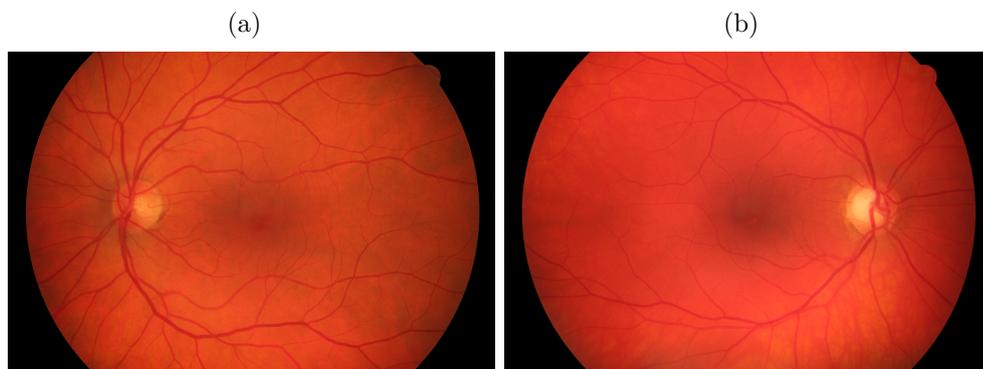


Figura 5 – Exemplos de imagens da base HRF (a) Imagem de uma retina saudável e (b) Imagem de um retina glaucomatosa.

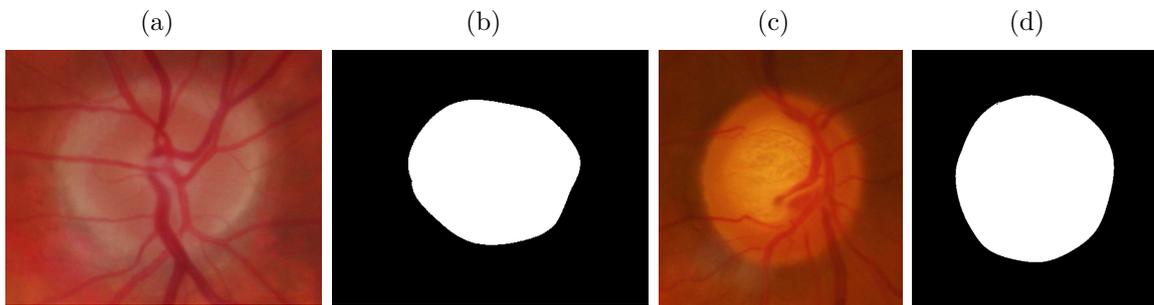


RIM-ONE - An Open Retinal Image Database for Optic Nerve Evaluation esta base de dados possui três versões sendo elas a RIM-ONE-1, RIM-ONE-2 e RIM-ONE-3, a seguir é descrito as informações destas bases.

A RIM-ONE-1 é uma base de imagens registrada com o intuito de auxiliar os trabalhos de pesquisa que são focados na análise de imagens de retina para localizar, avaliar e detectar o DO (FUMERO et al., 2011). A base é composta de 158 imagens de retina registradas em 3 diferentes regiões da Espanha, com tamanhos que variam de 342×316 à 869×723 pixels. As imagens foram capturadas por uma câmera Nidek AFC-210 com um corpo de uma Canon EOS 5D Mark II de 21,1 megapixels. O conjunto está dividido em 118 saudáveis, 12 com glaucoma inicial, 14 com glaucoma moderado e 14 com glaucoma profundo. A base foi criada com a colaboração de 4 oftalmologistas e 1 optometrista. Os padrões ouro de cada imagem estão disponíveis. A Figura 6 apresenta exemplos de uma imagem saudável (a) e seu padrão ouro representado em (b) e uma imagem com glaucoma (c) e seu padrão ouro (d).

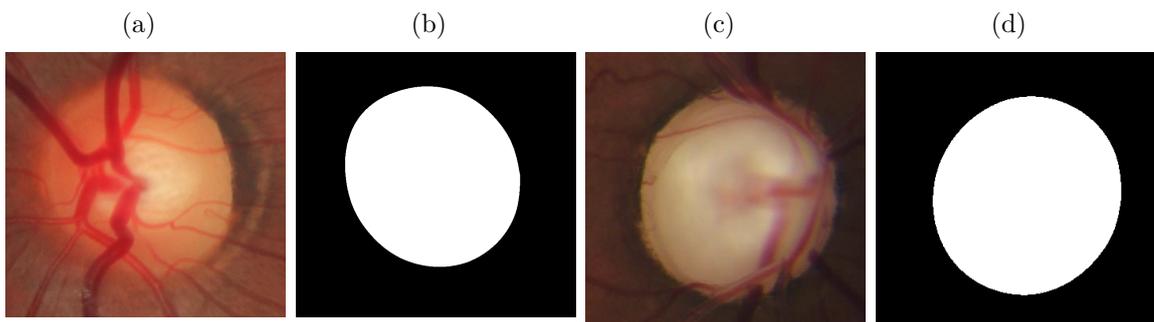
A RIM-ONE-2 possui um conjunto de 455 imagens da retina, sendo 200 imagens glaucomatosas e 255 imagens que não possuem a doença. As imagens variam de 345×386

Figura 6 – Exemplos de imagens da base RIM-ONE-1 (a) Imagem de uma retina saudável (b) Padrão Ouro da imagem de retina saudável (c) Imagem de um retina glaucomatosa e (d) Padrão ouro da imagem de retina glaucomatosa.



à 828×818 pixels. Todas as imagens possuem o padrão ouro da segmentação da região do disco óptico. A Figura 7 apresenta exemplos de uma imagem saudável (a) e seu padrão ouro representado em (b) e uma imagem com glaucoma (c) e seu padrão ouro (d).

Figura 7 – Exemplos de imagens da base RIM-ONE-2 (a) Imagem de uma retina saudável (b) Padrão Ouro da imagem de retina saudável (c) Imagem de um retina glaucomatosa e (d) Padrão ouro da imagem de retina glaucomatosa.

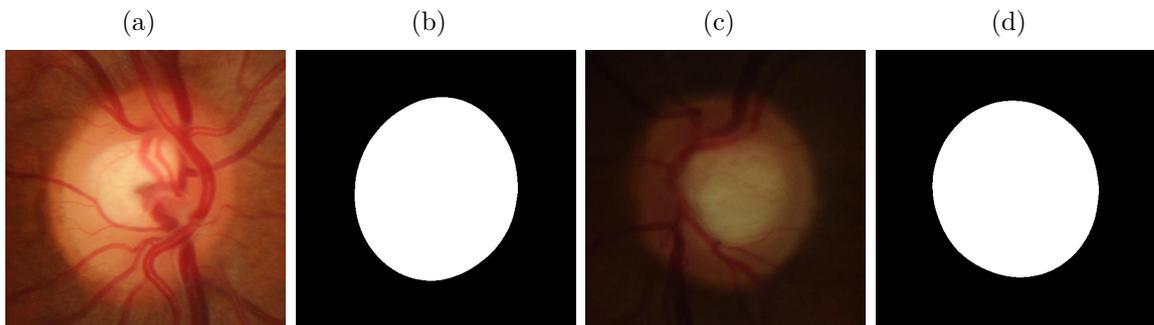


A RIM-ONE-3 é um conjunto de dados que consiste em 159 imagens de retina, sendo que, 85 dessas imagens não possui glaucoma e que 74 são glaucomatosas. Todas as imagens foram adquiridas por uma câmera de fundo não estéreo Kowa WX 3D. A imagens são centrados no DO com um “campo de visão” de 34° e as imagens estereoscópicas são capturadas no mesmo tiro da câmera, dando uma resolução de 2144×1424 pixels. O disco óptico e a escavação foram marcados por 2 especialistas, a fim de criar o padrão ouro. A segmentação média também está disponível para usar como referência. A Figura 8 apresenta exemplos de uma imagem saudável (a) e seu padrão ouro representado em (b) e uma imagem com glaucoma (c) e seu padrão ouro (d).

3.2 Extração de características

A extração de características tem por finalidade descrever imagens em função de atributos extraídos. Tais características são utilizadas para o reconhecimento de algum

Figura 8 – Exemplos de imagens da base RIM-ONE-3 (a) Imagem de uma retina saudável (b) Padrão Ouro da imagem de retina saudável (c) Imagem de um retina glaucomatosa e (d) Padrão ouro da imagem de retina glaucomatosa.



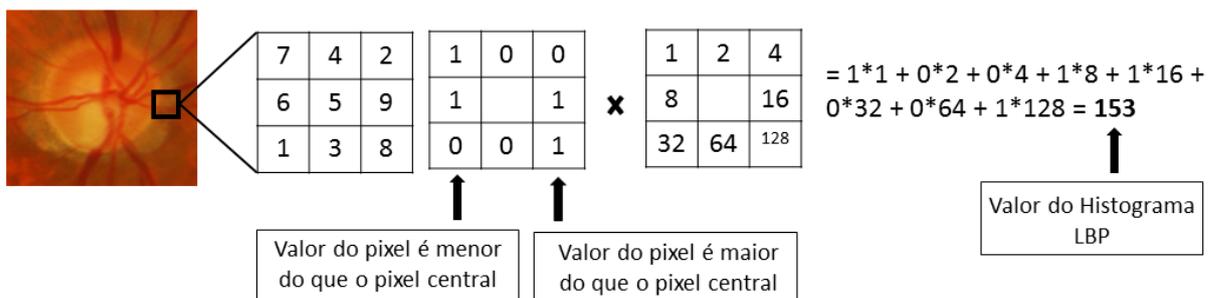
padrão, empregando certas propriedades para tornar o processo de decisão mais fácil durante a classificação. A identificação de imagens glaucomatosas pode ser eficientemente realizada usando os descritores de textura e forma de uma imagem (SALAM et al., 2016). As redes neurais convolucionais (DAHL; SAINATH; HINTON, 2013) também são utilizadas nesta etapa. Esta seção apresenta uma visão geral dos descritores utilizados neste trabalho.

3.2.1 LBP - Local Binary Pattern

O descritor LBP (OJALA; PIETIKÄINEN, 1997) rotula os pixels de uma imagem em níveis de cinza por limiarização. O valor do limiar é o pixel central da janela de vizinhança 3 x 3. Cada pixel é comparado com o central e o cálculo do resultado final é feito para todas as janelas escolhidas da imagem. Após todas as janelas serem computadas, o vetor de 256 características é montado com a junção dos valores obtidos. A Figura 9 mostra o funcionamento do descritor LBP.

Após a limiarização, um histograma da célula é calculado, ou seja, as quantidades de “0” e “1” são calculadas. As características mais relevantes do descritor LBP são a sua tolerância a mudanças de iluminação e sua simplicidade computacional.

Figura 9 – Demonstração do funcionamento do descritor de textura LBP.

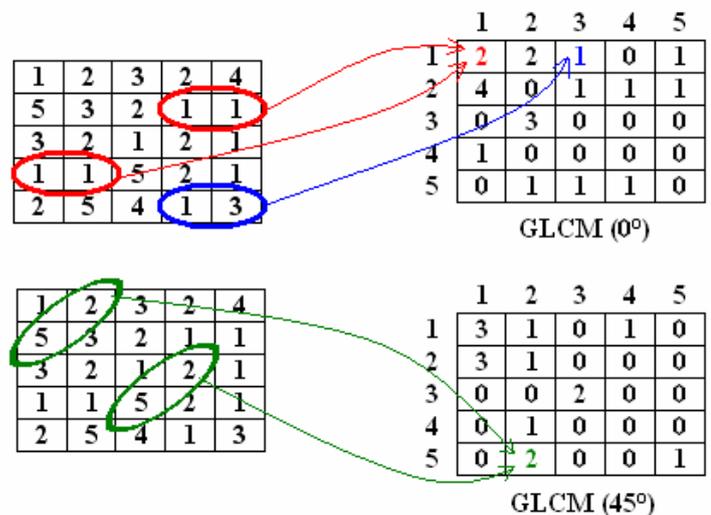


3.2.2 GLCM - Gray Level Co-occurrence Matrix

GLCM é uma técnica que tem como base a análise de textura em imagens, nela evidencia-se as co-ocorrências existentes entre pares de pixels através de algum padrão. Uma GLCM é sempre uma matriz quadrada, onde as informações de intensidades relativas dos pixels são armazenadas. Por este motivo, as imagens utilizadas são sempre em tons de cinza (HARALICK; SHANMUGAM et al., 1973).

As probabilidades de co-ocorrências são calculadas entre dois níveis de cinza i e j , utilizando uma orientação Θ (tetha) e uma distância conhecida como *pixel pair spacing*. Essa orientação pode assumir os valores de 0° , 45° , 90° e 135° . Para cada relacionamento especial possível (distância e orientação) existe uma matriz de co-ocorrência, a partir da qual calcula-se as características da imagem como: contraste, correlação, energia, homogeneidade nas quatro direções anteriormente citadas e a entropia da imagem. Ao todo foram extraídas 17 características. Uma representação de como é formada a matriz de co-ocorrência está representada na Figura 10.

Figura 10 – Demonstração do cálculo entre a co-ocorrência dos níveis de cinza para geração da matriz de co-ocorrência.



Nas equações abaixo, a quantidade de níveis de cinza na imagem é representada por G . A entrada da matriz GLCM é representada: linha i e coluna j por C_{ij} . A média e o desvio padrão são representados, respectivamente, por (μ_x, μ_y) e (σ_x, σ_y) para a linha i e coluna j da matriz.

Contraste - Corresponde à diferença entre os valores mais altos e mais baixos de um conjunto de pixels contíguos. Elevados valores de contraste correspondem a imagens compostas por valores digitais de grande amplitude, ou seja, com textura rugosa, mostrada

na Equação 3.1.

$$\text{Contraste} = \sum_{i,j=1}^G C_{ij}(i-j)^2 \quad (3.1)$$

Correlação - Retorna uma medida de quão correlacionado está um pixel com o seu vizinho. A comparação é feita em todos os pixels da imagem. A correlação é igual a 1 para uma imagem totalmente correlacionada e igual a 0 para uma imagem descorrelacionada, mostrada na Equação 3.2.

$$\text{Correlação} = \sum_{i,j=1}^G \frac{(i - \mu_x)(j - \mu_y)C_{ij}}{(\sigma_x)(\sigma_y)}. \quad (3.2)$$

Energia - Mede a uniformidade da textura, ou seja, a quantidade de repetição de pares de pixels. Valores elevados significam que a distribuição do valor digital é constante, isto é, ocorre grande repetição na variação dos valores digitais. Para uma matriz normalizada, os valores são positivos e menor ou igual a 1, mostrada na Equação 3.3.

$$\text{Energia} = \sum_{i,j=1}^G (C_{ij})^2. \quad (3.3)$$

Homogeneidade - Mede a homogeneidade da imagem. Valores elevados correspondem a pequenas diferenças de tonalidades entre os pixels vizinhos. Este parâmetro é inversamente proporcional ao contraste e à energia, e seu cálculo é demonstrada na Equação 3.4.

$$\text{Homogeneidade} = \sum_{i,j=1}^G \frac{C_{ij}}{1 + |i - j|}. \quad (3.4)$$

Entropia - É uma medida estatística que pode ser utilizada para caracterizar a textura da imagem de entrada, mostrada na Equação 3.5.

$$\text{Entropia} = - \sum_{i,j=1}^G C_{ij} \log C_{ij}. \quad (3.5)$$

3.2.3 HOG - *Histogram of Oriented Gradients*

O algoritmo HOG (DALAL; TRIGGS, 2005) é um descritor que calcula o histograma da orientação dos gradientes na imagem. O algoritmo se baseia na ideia de que a forma e a aparência de um objeto podem ser descritas pela intensidade dos gradientes ou da direção das bordas.

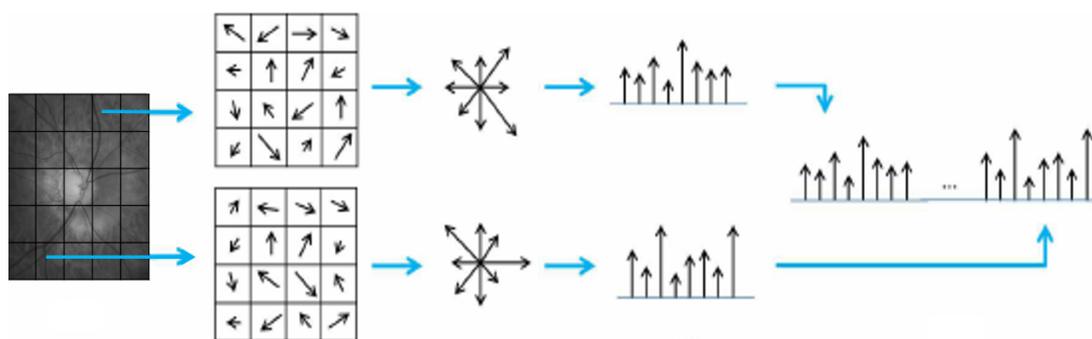
O primeiro passo do método consiste em converter a imagem para escala de cinzas. O próximo passo é a computação dos gradientes. Uma técnica que vem apresentando bons

resultados é a aplicação do operador de Sobel (SOBEL, 1970). Dada uma imagem I , os gradientes G_x e G_y são calculados, assim como a magnitude e a orientação de cada vetor.

Em seguida, é feita uma divisão da imagem em regiões espaciais chamadas blocos. Cada bloco é sub-dividido em regiões menores chamadas células. O tamanho das células e dos blocos são parâmetros essenciais do algoritmo. Para cada célula é calculado um histograma local 1-D das orientações sobre os pixels da célula. O descritor nada mais é do que uma lista dos histogramas calculados a partir de todas as células de todos os blocos. Após calculado, o descritor é normalizado.

Um exemplo de como é feita a extração de características é representada na Figura 11. Esta figura está dividida em células. Cada célula é formada por $N \times N$ pixels. É calculada a orientação de todos os pixels e construído o histograma de orientação da célula. Finalmente, todos os histogramas de orientação são concatenados para construir o vetor de características final.

Figura 11 – Representação do descritor HOG.



3.2.4 Tamura

Este descritor de textura imita a percepção visual humana, com escalares correspondentes ao tamanho das características das texturas (TAMURA; MORI; YAMAWAKI, 1978): *coarseness* (é o mais “grosso” elemento), contraste, orientação, semelhança com linhas, regularidade e rugosidade. Experimentos anteriores demonstraram que as características e os resultados apontam para os três primeiros recursos, como fortemente correlacionados com a percepção humana (ABBADENI; ZIOU; WANG, 2000).

Coarseness é a medida de granularidade da textura e é a característica mais fundamental. Os pesquisadores geralmente identificam a textura por *Coarseness*. Esta característica mede o tamanho das primitivas. Uma grande textura consiste em grandes primitivas e requer, portanto, um alto grau de uniformidade nos níveis de cinza. Uma textura fina é composta por pequenas primitivas e exige, portanto, um alto grau de variação

nos níveis de cinza. O seu cálculo é mostrado na Equação 3.6.

$$Coarseness = \frac{2^k}{n^2} \sum_i^n \sum_j^n p(i, j), \quad (3.6)$$

onde $n \times n$ denota o tamanho da imagem, a soma é realizada para cada pixel $p(i, j)$ e k é obtido como o valor que maximiza as diferenças das médias móveis $(1/2^{2k}) \sum \sum p(i, j)$, assumindo um valor de $2^k \times 2^k$ vizinhos, ao longo das direções horizontal e vertical.

Contraste mede os graus de clareza nos quais se pode distinguir entre as diferentes áreas de textura da imagem. Uma imagem é bem contrastada se as suas diferentes áreas são claramente visíveis. Alguns fatores que influenciam o contraste pode-se citar a relação entre o preto e branco na imagem e a frequência de mudanças de intensidade. O seu cálculo é mostrado na Equação 3.7.

$$Contraste = \frac{\sigma}{(\mu_4/\sigma^4)^{1/4}}, \quad (3.7)$$

onde σ é o desvio padrão da imagem e μ_4 é o quarto momento da imagem.

Orientação é uma propriedade global aplicada em uma determinada área da imagem. Essa característica mede a orientação dominante em uma área texturizada. Uma imagem pode ter uma, várias ou até nenhuma orientação dominante, neste último caso, ele é conhecido como isotrópico. A orientação implica a forma das primitivas, bem como suas regras de colocação. Duas primitivas que diferem apenas na orientação têm a mesma direção. O seu cálculo é mostrado na Equação 3.8.

$$Orientação = 1 - r \cdot n_p \cdot \sum_p^{n_p} \sum_{\phi \in w_p} (\phi - \phi_p)^2 \cdot H_D(\phi), \quad (3.8)$$

onde H_D é o histograma de direção local, n_p é o número de picos do histograma H_D , ϕ_p é a posição máxima do pico do histograma, w_p é a faixa entre os picos, r é o fator de normalização e ϕ é um código de direção quantizado.

3.2.5 GLRLM - Gray Level Run Length Matrix

Método de comprimento de primitiva a nível de cinza, é também uma técnica que tem por objetivo extrair as características de textura em uma determinada imagem. Para isso, ela faz uso de matrizes auxiliares e verifica as repetições dos pixels (primitivas). Segundo Galloway (GALLOWAY, 1975), a partir da matriz do GLRLM cinco medidas estatísticas podem ser extraídas. São elas: a ênfase em primitivas curtas (*Short-Run Emphasis* - SRE) (Equação 3.9), a ênfase em primitivas longas (*Long-Run Emphasis* - LRE) (Equação 3.10), a uniformidade do nível de cinza (*Gray Level distribution* - GLN) (Equação 3.11), a uniformidade do comprimento da primitiva (*Run Length distribution* -

RLN) (Equação 3.12) e a percentagem de primitivas (*Run Percentage* - RP) (Equação 3.13). Entretanto, mais duas medidas estatísticas foram determinadas por Chu et al. (CHU; SEHGAL; GREENLEAF, 1990), a ênfase em primitivas de baixo nível de cinza (*Low Gray Level Run Emphasis* - LGLRE) (Equação 3.14) e a ênfase em primitivas de alto nível de cinza (*High Gray Level Run Emphasis* - HGLRE) (Equação 3.15).

Outras quatro estatísticas podem ser extraídas da matriz GLRLM, sendo elas: a distribuição conjunta em primitivas curtas e baixos índices de nível de cinza (*Short Run Low Gray Level Emphasis* - SRLGLE) (Equação 3.16), a distribuição conjunta de primitivas curtas e altos valores de nível de cinza (*Short Run High Gray Level Emphasis* - SRHGLE) (Equação 3.17), a distribuição conjunta de primitivas longas e baixos valores de nível de cinza (*Long Run Low Gray Level Emphasis* - LRLGLE) (Equação 3.18) e a distribuição conjunta de primitivas longas e altos valores de nível de cinza (*Long Run High Gray Level Emphasis* - LRHGLE) (Equação 3.19).

As primitivas são um conjunto colinear maximal de pixels conectados, em que contêm o mesmo tom de cinza. O comprimento, inclinação e nível de cinza as caracterizam. Essa técnica funciona da seguinte forma: uma matriz GLRLM é criada para cada corte e possui o número de primitivas dos níveis de cinza na direção Θ e a uma distância d . Dessa forma, uma soma de todas as matrizes do GLRLM de cada corte é feita, e a partir do número de repetições encontrados, a matriz é dividida e normalizada.

$$SRE = \sum_{i=1}^G \sum_{j=1}^R \frac{p(i, j|\theta)}{j^2}. \quad (3.9)$$

$$LRE = \sum_{i=1}^G \sum_{j=1}^R j^2 p(i, j|\theta). \quad (3.10)$$

$$GLN = \sum_{i=1}^G \left(\sum_{j=1}^R p(i, j|\theta) \right)^2. \quad (3.11)$$

$$RLN = \sum_{j=1}^R \left(\sum_{i=1}^G p(i, j|\theta) \right)^2. \quad (3.12)$$

$$RP = \frac{1}{N} \sum_{i=1}^G \sum_{j=1}^R p(i, j|\theta). \quad (3.13)$$

$$LGLRE = \sum_{i=1}^G \left(\sum_{j=1}^R \frac{p(i, j|\theta)}{i^2} \right). \quad (3.14)$$

$$HGLRE = \sum_{i=1}^G \left(\sum_{j=1}^R i^2 p(i, j|\theta) \right). \quad (3.15)$$

$$SRLGLE = \sum_{i=1}^G \sum_{j=1}^R \frac{p(i, j|\theta)}{i^2 \cdot j^2}. \quad (3.16)$$

$$SRHGLE = \sum_{i=1}^G \sum_{j=1}^R \frac{i^2 \cdot p(i, j|\theta)}{j^2}. \quad (3.17)$$

$$LRLGLE = \sum_{i=1}^G \sum_{j=1}^R \frac{j^2 \cdot p(i, j|\theta)}{i^2}. \quad (3.18)$$

$$LRHGLE = \sum_{i=1}^G \sum_{j=1}^R i^2 \cdot j^2 \cdot p(i, j|\theta). \quad (3.19)$$

3.2.6 Morfologia

As medidas morfológicas são comumente aplicadas para descrição das imagens da retina. Essas características tentam quantificar atributos de forma e contorno. As mais importantes são: área, perímetro, área convexa, diâmetro equivalente e extensão. Outras características morfológicas foram testadas e analisadas, contudo, após a realização da etapa de seleção de características, utilizou-se o algoritmo de características utilizando o algoritmo de Razão de ganho (QUINLAN, 1986), observando que estas foram as que melhor se aplicaram nas imagens de retina.

3.2.7 CNN - *Convolutional Neural Network*

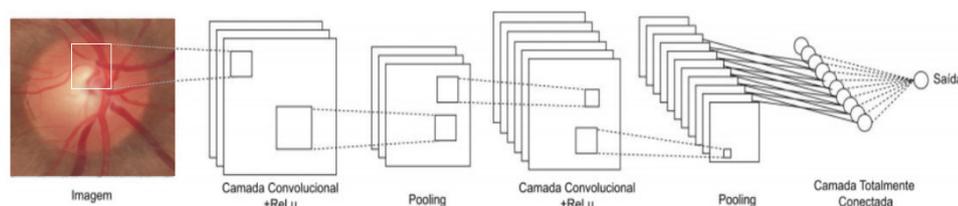
O processo de extração de características consiste na representação de um conjunto redundante de dados, por meio de características únicas que o façam se diferenciar de outros conjuntos. Existem inúmeras características que podem ser extraídas, dentre elas, as principais em um sistema de auxílio médico são a de cor, forma e textura. Dentre os descritores vistos no estado da arte, a utilização do descritor CNN's neste processo de extração de características vem ganhando destaque.

Uma CNN é uma rede formada por várias camadas que podem ser usadas no reconhecimento de objetos e na segmentação e classificação de imagens. Entre essas camadas, temos as camadas convolucionais que podem alterar a representação dos dados através de filtros. Normalmente, após uma camada convolucional, é utilizada uma função de ativação, essas funções realizam transformações não-lineares nos dados, de modo a gerar saídas linearmente separáveis. Uma das funções mais comuns utilizadas com este propósito é a *Rectified Linear units* (ReLU) (DAHL; SAINATH; HINTON, 2013), apresentadas na Equação 3.20, onde x é a entrada para um neurônio.

$$f(x) = \max(0, x) \quad (3.20)$$

A camada de *pooling* realiza uma redução no dados ao longo das dimensões espaciais. Ela sucede as camadas de convolução e permite um afinilamento na rede. As camadas totalmente conectadas são responsáveis por reunir todos as características dos descritores para que possam ser classificados pela camada final. Figura 12 apresenta uma CNN de forma simplificada.

Figura 12 – Ilustração da arquitetura de uma CNN simplificada.



As CNNs que utilizamos são pré-treinadas na base natural de imagens *ImageNet* e estão disponíveis no *website MatConv-Net*¹. Três arquiteturas propostas recentemente serão utilizadas na construção da abordagem, são elas:

AlexNet: É uma arquitetura desenvolvida recentemente e foi utilizada na competição LSVRC-2010 para classificar a base de imagens *ImageNet* (KRIZHEVSKY; SUTSKEVER; HINTON, 2012). Compreende oito camadas que precisam ser treinadas, cinco camadas convolucionais seguidas de três camadas totalmente conectadas, bem como camadas de agrupamento máximo.

CaffeNet: A arquitetura *CaffeNet* foi desenvolvida no Centro de Visão e Aprendizado de Bekeley, sendo considerada uma das mais populares no aprendizado profundo (JIA et al., 2014). Compreende 5 camadas convolucionais, cada uma seguida por uma camada de *pooling*, e 3 camadas inteiramente conectadas. Também é a mais fácil de manipular e alterar as propriedades dentre as analisadas neste trabalho, e possui um fácil acesso, disponibilizando APIs em C++, Matlab e Python.

VGGNet: A arquitetura *VGGNet*, proposta em 2014 por (CHATFIELD et al., 2014), foi projetada para aumentar a profundidade das redes. Sendo a considerada a mais difícil de se treinar, a rede requer muito mais memória e tempo computacional do que a *AlexNet*. A rede contém 8 camadas de aprendizado, são 5 convolucionais e as 3 últimas totalmente conectadas.

A *VGGMNet* e *VGGSNet* foram também desenvolvidas por (CHATFIELD et al., 2014). Estas duas redes tem a mesma profundidade, contudo a *VGGMNet* é mais larga em relação a *VGGSNet* e esta utiliza 5 camadas convolucionais e menos filtros em uma das camadas de convolução.

Simonyan e Zisserman (2014) desenvolveram duas novas redes convolucionais, com o objetivo de utilizar redes de maior profundidade usando uma arquitetura com filtros de

¹ <http://www.vlfeat.org/matconvnet/>

convolução muito pequenos. Assim a profundidade das camadas será de 16 e 19, sendo essas redes denominadas de *VGG16Net* e a *VGG19Net*, respectivamente.

3.3 Seleção de Atributos

O processo de seleção tem como objetivo eliminar atributos desnecessários e, conseqüentemente, simplificar o modelo de predição, melhorar a performance dos algoritmos, reduzir o custo computacional destes modelos, além de fornecer um melhor entendimento sobre os resultados encontrados. Complementarmente pode-se afirmar que a seleção de atributos desempenha um papel importante na preparação de dados para a mineração, pois quanto mais se remove dados redundantes e irrelevantes, mais se reduz a dimensionalidade do espaço de atributos (BOSCARIOLI, 2005).

Embora também possam ser utilizados todos os atributos de uma base de dados, a maioria dos autores consultados defende a seleção de atributos relevantes e a eliminação dos que são totalmente desnecessários, na tentativa de diminuir a complexidade do problema e alcançar um desempenho ótimo de aprendizagem (KIM; STREET; MENCZER, 2003).

Existem várias razões que justificam a redução do número de atributos, segundo (RAMOS et al., 1997) alguns atributos podem ser inúteis ou podem refletir informações que já estão embutidas em outros atributos. A redução do número de atributos reflete na criação de estruturas menores que permitem um melhor entendimento do domínio. Outra vantagem é que o custo tempo-espaço dos algoritmos de indução está diretamente relacionado ao número de atributos considerado, ou seja, a redução do número de atributos ocasiona na redução do custo de tempo-espaço do algoritmo.

O algoritmo de Razão de Ganho (KAREGOWDA; MANJUNATH; JAYARAM, 2010) foi escolhido para realizar essa seleção. Este algoritmo é baseado no Ganho de Informação, que utiliza a entropia como medida de impureza e árvores de decisão para representar os atributos do vetor. A sua performance é considerada melhor que o Ganho de Informação, ou seja, proporciona uma melhor acurácia em termos de complexidade das árvores geradas. Considerando um conjunto de dados S com m classes, a Equação 3.21 apresenta o cálculo da informação utilizada para classificar uma determinada amostra.

$$Info(S) = - \sum_{i=1}^m p_i \log_2(p_i), \quad (3.21)$$

onde p_i é a probabilidade da amostra pertencer a uma classe C_i . O cálculo da entropia de um determinado atributo A que possui v valores é demonstrado na Equação 3.22,

$$Entropia(A) = - \sum_{i=1}^m Info(S) \frac{s_1^i + s_2^i + \dots + s_m^i}{s}, \quad (3.22)$$

e a variável $s_{i,j}$ representa o número de amostras pertencentes a classe C_i do subconjunto S_i . O ganho do atributo A é representado pela Equação 3.23.

$$Gain(A) = Info(S) - Entropia(A). \quad (3.23)$$

A Equação 3.24 apresenta o valor de informação gerado pela divisão do conjunto de dados de S em v partições

$$SInfo(S) = - \sum_{i=1}^v \left(\frac{|S_i|}{|S|} \right) \log_2 \left(\frac{|S_i|}{|S|} \right). \quad (3.24)$$

Por fim, a Razão de Ganho é definida como a divisão dos resultados obtidos nas equações 3.23 e 3.24. Os atributos são ranqueados de acordo com o valor da sua razão de ganho, sendo selecionados os que n atributos com os maiores valores.

3.4 Classificadores

O objetivo de um classificador é dividir o espaço de atributos em regiões de decisão. Dessa forma, os vetores de atributos que estiverem contidos na mesma região de decisão compartilham a mesma classe. A entrada de um classificador é o vetor de atributos e a saída será a classe que a imagem pertence. Foram testados os classificadores apresentados a seguir.

3.4.1 SVM - *Support Vector Machine*

Os estudos sobre o classificador SVM foram introduzidos por Cortes e Vapnik (CORTES; VAPNIK, 1995). De um modo geral, a SVM faz um mapeamento do espaço de entrada para um espaço de dimensionalidade maior. Em seguida, é calculado um hiperplano de separação ótimo. Este hiperplano é escolhido de modo a maximizar a distância de separação entre as classes (HAYKIN, 2001). Diz-se que duas classes são linearmente separáveis se existe um hiperplano que separe as mesmas.

A SVM é uma técnica de aprendizagem supervisionada, ou seja, possui as fases de treinamento e teste. Na fase de treinamento, atributos e saídas desejadas são dadas a fim de projetar os vetores de suporte. Estes vetores de suporte são usados para obter um classificador. Na fase de teste, o classificador é utilizado para encontrar uma saída a partir de um vetor de entrada.

Ao utilizar SVM para reconhecimento de padrões se faz necessário tornar uma função não-linearmente separável em uma função linearmente separável. Para isso é necessário que seja aumentada a dimensionalidade do problema (COVER, 1965). As

funções que aumentam a dimensionalidade do espaço de entrada são chamadas de função de Kernel.

As funções de Kernel são aplicadas nos vetores de entrada. Ao aplicar uma função de Kernel em um vetor de entrada de dimensão N , é obtido um novo vetor de dimensão X , onde X é maior que N . Após esse cálculo são obtidos os vetores de suporte. Com os vetores calculados é possível definir o hiperplano de separação ótimo. O hiperplano estará a uma igual distância (d_0) dos vetores de suporte de cada classe. Um novo objeto é classificado utilizando o hiperplano ótimo, cada lado do hiperplano representa uma classe diferente.

3.4.2 MLP - *MultiLayer Perceptron*

Uma rede *MultiLayer Perceptron* (MLP) (AMARI, 1967) é um sistema de neurônios ligados por conexões sinápticas e dividido em neurônios de entrada, que recebem estímulos do meio externo, neurônios internos (ocultos) e neurônios de saída, que se comunicam com o exterior.

O MLP foi concebido para resolver problemas mais complexos, os quais não poderiam ser resolvidos pelo modelo de neurônio básico. Para isto são necessárias mais conexões, os quais só existem em uma rede de Perceptrons dispostos em camadas. Na camada oculta (ou camada intermediária), os neurônios são denominados escondidos porque eles não tem acesso direto à saída da rede MLP, onde os erros de aproximação são calculados.

O mais popular método para aprendizado em MLP é chamado de retro-propagação (*backpropagation*). A aprendizagem por retro-propagação do erro consiste em dois passos através das diferentes camadas da rede (HAYKIN, 2001). O primeiro passo é chamado de propagação. Nesse passo um padrão de atividades é aplicado aos nós sensoriais e é produzido a resposta real da rede. O segundo passo é a retro-propagação. Nesse passo os pesos sinápticos são ajustados de acordo com um sinal de erro. Este sinal é então propagado na direção contrária das conexões sinápticas. O objetivo é minimizar o erro médio global, ou seja, fazer com que a resposta real da rede se mova para mais perto da resposta desejada em um sentido estatístico.

3.4.3 KNN - *K-Nearest Neighbors*

K-Vizinhos Mais Próximos (KNN) é um classificador clássico baseado em memória, que frequentemente é utilizado em aplicações do mundo real devido a sua simplicidade. Apesar de simples, obteve considerável exatidão para classificação em diversas aplicações e é comumente utilizado na comparação com outros classificadores (AHA; KIBLER; ALBERT, 1991).

A fase de treinamento do KNN consiste simplesmente em armazenar todos os padrões de treinamento. O KNN tem um parâmetro k que é o número de vizinhos a ser considerado para a classificação. Para $K = 1$, KNN é também chamado de método do vizinho mais próximo (*Nearest Neighbor* - NN). O NN classifica um dado padrão como pertencente a mesma classe do padrão mais próximo a ele no conjunto de treinamento. Há um número de medidas de distâncias usadas neste processo, entretanto a distância euclidiana é a mais utilizada (AHA; KIBLER; ALBERT, 1991)

Quando $K > 1$, primeiramente são calculadas as distâncias do novo padrão a ser classificado com todos os padrões do conjunto de treinamento. Subsequentemente, o algoritmo considera os K padrões do conjunto de treinamento com as menores distâncias ao padrão a ser classificado. Finalmente, o novo padrão é classificado como pertencente a classe da maioria dos K padrões mais próximos do conjunto de treinamento.

3.4.4 RT - *Random Tree*

A *Random Tree* é uma árvore induzida aleatoriamente a partir de um conjunto de árvores possíveis, usando m atributos aleatórios em cada nodo. O termo “aleatoriamente” significa que cada árvore tem uma chance igual de ser amostrada. Uma das formas mais utilizadas de se construir uma *Random Tree*, é através do processo de ramificação Galton-Watson (RODRIGUEZ, 2012).

De modo geral, podemos pensar um processo de ramificação Galton-Watson da seguinte maneira. Suponha que temos árvores que dão nascimento a novas árvores segundo a variável aleatória X com valores no conjunto $0, 1, 2, \dots$ e função de distribuição de probabilidade expressa por $P(X = K) = pk$, para $K = 0, 1, 2, \dots, n$. Inicialmente suponha que, temos uma única árvore na iteração $n = 0$. Essa árvore dá nascimento, na iteração $n = 1$ a novas árvores segundo a variável aleatória X . Dizemos que essas novas árvores são descendentes diretas da anterior. Isto é, ela dá nascimento a k descendentes diretos com probabilidade pk . Em geral, se temos um certo número de árvores na iteração n , cada uma delas dá nascimento a novas árvores na iteração $n + 1$. Em cada caso, isto é feito de acordo uma variável aleatória independente e identicamente distribuída a variável aleatória X . Dizemos que a árvore da iteração $n = 0$ constitui a geração 0 e as árvores que nascem na iteração n constituem a n -ésima geração do processo, $n > 1$. Denotamos por Z_n a variável aleatória que conta o número de árvores da n -ésima geração (RODRIGUEZ, 2012).

3.4.5 RF - *Random Forest*

O algoritmo *Random Forest* (BREIMAN, 2001) é uma combinação de predições de diversas árvores em que cada árvore depende dos valores de um vetor independente, amostrados aleatoriamente e com a mesma distribuição para todas as árvores da floresta.

Aqui, floresta é o que se denomina para uma série de árvores de decisão. Após a geração de um grande número de árvores, as classes com maior número de votos são eleitas.

Em um *Random Forest*, cada nó é dividido usando o melhor dentre um subconjunto de indicadores escolhidos aleatoriamente naquele nó. Esta estratégia um tanto contraditória acaba por funcionar muito bem em comparação com muitos outros classificadores, além de ser robusto a superajuste nos parâmetros (BREIMAN, 2001). Além disso, é de fácil utilização, no sentido que ele tem apenas dois parâmetros (o número de variáveis no subconjunto aleatório em cada nó e o número de árvores da floresta) e, normalmente, não é muito sensível aos seus valores.

A partir de um vetor de atributos, são gerados outros vetores de atributos, que são embaralhados em relação ao vetor original. É gerado um vetor para cada árvore do *Random Forest*. Em seguida, os vetores de atributos são passados como parâmetro para as árvores de decisão. Cada árvore irá gerar um resultado para a classificação e, os resultados são combinados obtendo uma saída unificada.

3.5 Métodos de Avaliação dos Resultados

Para avaliar o desempenho dos algoritmos de classificação foi utilizada a matriz de confusão, onde a partir dela serão calculadas as métricas de avaliação dos resultados conhecidas na literatura.

A matriz de confusão é uma tabela que mostra o resultado da classificação, comparando-o com o resultado real. A matriz, nesse trabalho, é composta por 4 valores: Verdadeiro Positivo (VP), Falso Positivo (FP), Falso Negativo (FN) e Verdadeiro Negativo (VN), onde esses valores podem ser observados na Tabela 2. Tais valores são obtidos a partir do resultado da classificação. Supondo a existência de duas classes hipotéticas: classe X e classe Y, são computados os seguintes valores:

- VP: número de objetos da classe X classificados como da classe X;
- FP: número de objetos da classe Y classificados como X;
- FN: número de objetos da classe X classificados como Y;
- VN: número de objetos da classe Y classificados como sendo da classe Y.

A partir dessas quantidades os valores de Precisão (P), *Recall* (R), Acurácia (A), *F-Measure* (FM) e Índice Kappa (K) são calculados (POWERS, 2007).

Precisão (P) reflete a proporção de verdadeiros positivos em relação a todas as predições positivas. A precisão de uma classificação mostra a quantidade de objetos da

Tabela 2 – Matriz de confusão de duas classes (X,Y).

| Valor Previsto \ Valor verdadeiro | Classe X | Classe Y |
|-----------------------------------|----------|----------|
| | Classe X | VP |
| Classe Y | FN | VN |

classe X classificados corretamente em relação a todos os objetos classificados como sendo da classe X. Ela pode ser calculada a partir da Equação 3.25:

$$P = \frac{VP}{VP + FP}. \quad (3.25)$$

Recall (R) reflete a proporção de verdadeiros positivos em relação a suas previsões positivas e as suas incorretas previsões negativas. Essa medida mostra o comportamento dos objetos classificados como X, ou seja, de todos os objetos da classe X, quantos foram classificados como X. Ela pode ser calculada a partir da Equação 3.26:

$$R = \frac{VP}{VP + FN}. \quad (3.26)$$

Acurácia (A) de um classificador é a porcentagem de casos corretamente classificados em um conjunto de teste. Essa medida mede o quão bem um classificador reconhece instâncias de diversas classes. Ela pode ser calculada a partir da Equação 3.27:

$$A = \frac{VP + VN}{VP + FP + VN + FN}. \quad (3.27)$$

F-measure (FM) só produz bons resultados quando as taxas de precisão e *recall* são equilibradas. Assim, essa é uma medida bastante significativa (CHIMIESKI; FAGUNDES, 2013). O resultado da FM varia entre 0 e 1. Logo, quanto mais próximo de 1 melhor é o resultado da classificação. Ela pode ser calculada, a partir da Equação 3.28:

$$FM = \frac{2 * VP}{2 * VP + FN + FP}. \quad (3.28)$$

Área Sob a Curva ROC (AUC) é a representação da sensibilidade de um classificador dada pela taxa de VP em função da taxa de FP. Quanto maior a AUC mais preciso é o classificador.

O índice Kappa (K) é utilizado como uma medida apropriada da exatidão por representar inteiramente a matriz de confusão. Este índice toma todos os elementos da matriz em consideração, ao invés de apenas aqueles que retratam a quantidade de classificações verdadeiras, o que ocorre quando se calcula a exatidão global da classificação (ROSENFELD; FITZPATRICK-LINS, 1986).

O índice Kappa é um coeficiente de concordância para escalas nominais que mede o relacionamento entre a concordância, além da casualidade, e a discordância esperada (ROSENFELD; FITZPATRICK-LINS, 1986). O índice Kappa pode ser calculado com base na Equação 3.29:

$$K = \frac{(\text{observado} - \text{esperado})}{1 - \text{esperado}}. \quad (3.29)$$

Neste caso, “observado” é o valor global para a porcentagem correta, ou seja, o somatório da diagonal principal da matriz dividido pela quantidade de elementos e “esperado” são os valores calculados usando-se os totais de cada linha e cada coluna da matriz de confusão. A qualidade dos resultados podem ser avaliados pelo valor do índice Kappa observado na Tabela 3, conforme definido por Landis e Koch (LANDIS; KOCH, 1977).

Tabela 3 – Nível de exatidão de uma classificação, conforme o valor do índice kappa.

| Índice Kappa (K) | Qualidade |
|--------------------|-----------|
| $K \leq 0,2$ | Ruim |
| $0,2 < K \leq 0,4$ | Razoável |
| $0,4 < K \leq 0,6$ | Bom |
| $0,6 < K \leq 0,8$ | Muito Bom |
| $K > 0,8$ | Excelente |

3.6 Considerações Finais

Neste capítulo foram apresentados os principais materiais e métodos utilizados no desenvolvimento deste trabalho.

Descritores de imagens são algoritmos de visão computacional responsáveis por extrair atributos significativos das imagens. Algoritmos de classificação são utilizados em vastas áreas da computação, onde pode-se observar que na literatura, há uma grande variedade de algoritmos com este propósito. Neste trabalho aplicam-se os classificadores mais utilizados na literatura.

O capítulo seguinte descreve o método proposto para detecção automática do glaucoma em imagens da retina. É mostrado também como os algoritmos explanados nesse capítulo são utilizados para a construção do método.

4 Método Proposto

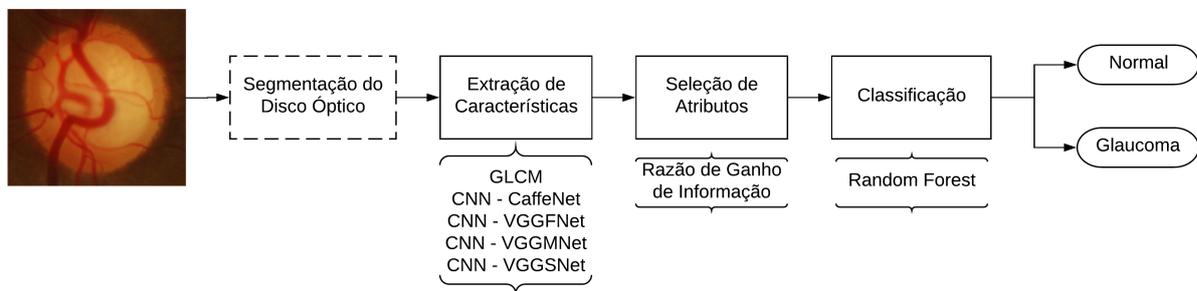
Neste capítulo é descrita a metodologia proposta neste estudo, que tem por finalidade detectar automaticamente o glaucoma a partir de imagens digitais da retina. Para a definição do método proposto foram realizados diversos testes com outros descritores e classificadores, bem como a utilização da imagem sem segmentação. Os resultados destes testes foram divididos em dois cenários, a saber:

- Resultados sem segmentação;
- Resultados com a utilização apenas da região do disco óptico;

Todos os algoritmos foram implementados utilizando a ferramenta computacional *MATLAB*[®] (THOMPSON; SHURE, 1995), a biblioteca *OpenCV*[®] (BRADSKI, 2000), além da ferramenta *WEKA*[®] (HALL et al., 2009).

Foram seguidas quatro fases, ilustradas no fluxograma da Figura 13. Inicialmente é utilizada a segmentação da região do DO nas imagens. O passo seguinte foi a extração das características da imagem, onde foram utilizados descritores de textura, forma e redes neurais convolucionais, e que a partir de uma visão mais ampla selecionou os cinco descritores que obtiveram um melhor desempenho, sendo eles o descritor de textura GLCM e quatro redes neurais convolucionais. Em seguida foram selecionadas as principais características que representavam as imagens, aplicando o algoritmo razão de ganho de informação. Por fim, as características selecionadas foram utilizadas para classificar as imagens em glaucomatosas ou saudáveis pelo classificador *Random Forest*.

Figura 13 – Fluxograma do método proposto neste trabalho.



Neste trabalho foram obtidas cinco bases de imagens públicas da retina. A segmentação da região do DO, que será aplicada neste estudo, vem disponível em quatro das cinco bases de dados adquiridas. A única que não possui essa segmentação é a base de dados HRF (FUNDUS, 2012), por isso ela não é utilizada no cenário 2, que corresponde ao cenário com segmentação.

A etapa de extração de características é muito significativa neste trabalho, visto que as características extraídas são utilizadas para o reconhecimento de algum padrão e torna o processo de decisão mais fácil durante a classificação. Nesta etapa foram utilizados os descritores a matriz de co-ocorrência (GLCM) e as redes neurais convolucionais (CNN-*CaffeNet*, CNN-*VGGFNet*, CNN-*VGGMNet* e CNN-*VGGSNet*). No descritor GLCM as suas características foram extraídas em cinco canais de cores diferentes. Para efetivação da utilização destes canais de cores foi necessário um pré-processamento, onde as imagens foram convertidas para o modelo de cor *Hue, Saturation, Value* (HSV) e o padrão de cor *Red, Green, Blue* (RGB). Nas CNNs foi utilizada apenas a imagens no modelo de cor RGB. A banda de cor *Value* do modelo de cor HSV não será utilizada, visto que, em imagens da retina ela é semelhante a banda de cor *Red* (CHENG et al., 2013).

Na seleção de atributos realizou-se testes, onde priorizou a aplicação do algoritmo Razão de Ganho de Informação (KAREGOWDA; MANJUNATH; JAYARAM, 2010). Estes testes foram aplicados em tamanho de vetores de características diferentes, variando o seu valor em 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% e 90%, para que sejam excluídos atributos menos significativos. Ou seja, se em um teste é utilizado apenas 80% dos atributos, significa que os 20% dos atributos menos significativos foram retirados do vetor, ficando apenas os 80% dos atributos mais significativos. Estes testes têm como intuito obter uma melhoria nos resultados encontrados.

Por fim, o classificador *Random Forest* é utilizado para a classificação das imagens em duas classes: saudável ou glaucoma. Este classificador foi utilizado com uma profundidade máxima das árvores ilimitada e o número de árvores foi 100. Outros valores foram testados, porém o melhor resultado foi com 100 árvores.

4.1 Considerações Finais

O método proposto é composto por um conjunto de ferramentas e procedimentos computacionais. Para detecção do glaucoma foram utilizadas cinco bases de dados públicas. Descritores de textura e redes neurais convolucionais foram utilizadas nas imagens segmentadas com o objetivo de extrair atributos das imagens. O classificador *Random Forest* é utilizado na classificação das mesmas.

5 Resultados e Discussão

Neste capítulo são apresentados e analisados os resultados obtidos com o método proposto. Ao final do capítulo é realizada uma análise comparativa com outros métodos encontrados na literatura de descrição de imagens e classificação realizadas nas imagens da retina para detecção do glaucoma.

5.1 Resultados sem segmentação

Para avaliação dos resultados obtidos foram utilizadas as métricas mostradas na Seção 3.5. Os parâmetros utilizados para a classificação foram os padrões de cada classificador no WEKA (HALL et al., 2009) e, o método de validação utilizado foi o *K-fold cross-validation* (com $k = 10$). No *10-fold cross-validation*, o conjunto de dados originais é aleatoriamente particionado em 10 subconjuntos de dados com o mesmo tamanho. A partir dos 10 subconjuntos, um único subconjunto é escolhido como conjunto de validação para testar o modelo, e os 9 subconjuntos restantes são utilizados como dados de treinamento. O processo de *cross-validation* é, então, repetido 10 vezes. Cada um dos 10 subconjuntos é utilizado somente uma vez para validação dos dados. É obtida a média dos 10 resultados gerados para produzir uma estimativa única. A vantagem desse método é que todos os subconjuntos são utilizados tanto para teste como para treino, e cada subconjunto é utilizado para teste somente uma vez. A seguir são mostrados os resultados obtidos (KOHAVI et al., 1995).

As Tabelas 4, 5, 6, 7 e 8 demonstram os resultados obtidos com a aplicação do primeiro cenário apresentado neste estudo, que neste caso é a utilização da imagem sem segmentação. Essas tabelas representam, respectivamente, a aplicação dos classificadores SVM, MLP, KNN, RT e RF nos descritores utilizados, onde em cada descritor foi aplicado 903 imagens do fundo de olho, disponíveis nas cinco bases de dados utilizadas, sendo 504 imagens saudáveis e 399 com glaucoma.

O SVM foi executado utilizando uma função de *kernel* linear. O resultado aplicando esse *kernel* foi melhor quando comparado com os *kernels* polinomial, de base radial e tangente sigmóide. A partir da análise da Tabela 4 podemos perceber que o maior resultado encontrado foi com a utilização das CNN-*VGGNet*, obtendo uma acurácia de 84,82% e kappa de 0,69, o que indica um resultado “Muito Bom”.

O classificador MLP, foi utilizado com uma camada oculta com 2 neurônios. Outros valores também foram testados, mas este foi o que retornou o melhor resultado. A Tabela 5 demonstra os resultados obtidos no primeiro cenário para esse classificador. O método

Tabela 4 – Resultados da 1ª estratégia com o classificador SVM.

| Descritor | SVM | | | | | | | | | |
|----------------------|------------|-----------|-----------|------------|--------------|--------------|--------------|--------------|-------------|-------------|
| | VP | FP | FN | VN | P(%) | R(%) | A(%) | FM(%) | AUC | K |
| LBP | 407 | 97 | 200 | 199 | 67,10 | 67,10 | 67,10 | 66,20 | 0,65 | 0,31 |
| GLCM | 422 | 82 | 241 | 158 | 64,60 | 64,20 | 64,23 | 62,20 | 0,61 | 0,24 |
| HOG | 401 | 103 | 85 | 314 | 79,30 | 79,20 | 79,18 | 79,20 | 0,79 | 0,58 |
| Tamura | 396 | 108 | 181 | 218 | 67,90 | 68,00 | 67,99 | 67,50 | 0,66 | 0,33 |
| GLRLM | 342 | 162 | 209 | 190 | 58,50 | 58,90 | 58,91 | 58,90 | 0,57 | 0,15 |
| CNN- <i>AlexNet</i> | 428 | 76 | 87 | 312 | 81,90 | 81,90 | 81,94 | 81,90 | 0,81 | 0,63 |
| CNN- <i>CaffeNet</i> | 427 | 77 | 82 | 317 | 82,40 | 82,40 | 82,39 | 82,40 | 0,82 | 0,64 |
| CNN- <i>VGGFNet</i> | 435 | 69 | 68 | 331 | 84,80 | 84,80 | 84,82 | 84,80 | 0,84 | 0,69 |
| CNN- <i>VGGMNet</i> | 421 | 83 | 87 | 312 | 81,20 | 81,20 | 81,17 | 81,20 | 0,80 | 0,61 |
| CNN- <i>VGGSNet</i> | 423 | 81 | 85 | 314 | 81,60 | 81,60 | 81,61 | 81,60 | 0,81 | 0,62 |
| CNN- <i>VGG16Net</i> | 422 | 82 | 90 | 309 | 80,90 | 81,00 | 80,95 | 80,90 | 0,80 | 0,61 |
| CNN- <i>VGG19Net</i> | 411 | 93 | 109 | 290 | 77,60 | 77,60 | 77,63 | 77,60 | 0,77 | 0,54 |

híbrido obteve o melhor resultado com 80,95% de acurácia.

Tabela 5 – Resultados da 1ª estratégia com o classificador MLP.

| Descritor | MLP | | | | | | | | | |
|----------------------|------------|-----------|-----------|------------|--------------|--------------|--------------|--------------|-------------|-------------|
| | VP | FP | FN | VN | P(%) | R(%) | A(%) | FM(%) | AUC | K |
| LBP | 365 | 139 | 134 | 265 | 69,80 | 69,80 | 69,76 | 69,80 | 0,74 | 0,38 |
| GLCM | 367 | 137 | 119 | 280 | 71,80 | 71,70 | 71,65 | 71,70 | 0,77 | 0,42 |
| HOG | 405 | 99 | 111 | 288 | 76,70 | 76,70 | 76,74 | 76,70 | 0,84 | 0,52 |
| Tamura | 368 | 136 | 122 | 277 | 71,60 | 71,40 | 71,42 | 71,50 | 0,78 | 0,42 |
| GLRLM | 381 | 123 | 117 | 282 | 73,50 | 73,40 | 73,42 | 73,40 | 0,77 | 0,46 |
| CNN- <i>AlexNet</i> | 413 | 91 | 84 | 315 | 80,70 | 80,60 | 80,62 | 80,60 | 0,87 | 0,60 |
| CNN- <i>CaffeNet</i> | 408 | 96 | 101 | 298 | 78,20 | 78,20 | 78,18 | 78,20 | 0,85 | 0,55 |
| CNN- <i>VGGFNet</i> | 393 | 111 | 93 | 306 | 77,60 | 77,40 | 77,40 | 77,50 | 0,85 | 0,54 |
| CNN- <i>VGGMNet</i> | 423 | 81 | 84 | 315 | 81,70 | 81,70 | 81,72 | 81,70 | 0,88 | 0,63 |
| CNN- <i>VGGSNet</i> | 424 | 80 | 74 | 325 | 83,00 | 82,90 | 82,94 | 83,00 | 0,88 | 0,65 |
| CNN- <i>VGG16Net</i> | 416 | 88 | 84 | 315 | 81,00 | 81,00 | 80,95 | 81,00 | 0,88 | 0,61 |
| CNN- <i>VGG19Net</i> | 417 | 87 | 91 | 308 | 80,30 | 80,30 | 80,28 | 80,30 | 0,86 | 0,59 |

O classificador *k-Nearest Neighbor* foi utilizado com o valor $k = 3$, ou seja, foram considerados 3 vizinhos para a classificação. A Tabela 6 retrata os resultados atingidos no primeiro cenário para esse classificador. Podemos visualizar que o maior resultado encontrado foi com o descritor CNN-*VGGSNet*, onde atingiu uma acurácia de 80,39%.

O classificador *Random Forest* foi utilizado com uma profundidade máxima das árvores ilimitada e o número de árvores foi 100. Outros valores foram testados, porém o melhor resultado foi com 100 árvores. A Tabela 8 apresenta os resultados obtidos no primeiro cenário para esse classificador. O melhor resultado encontrado, assim como os classificadores MLP e KNN, foi com a CNN-*VGGSNet*, onde encontrou uma acurácia de 80,39%.

Tabela 6 – Resultados da 1ª estratégia com o classificador KNN.

| Descritor | KNN | | | | | | | | | |
|----------------------|------------|-----------|-----------|------------|--------------|--------------|--------------|--------------|-------------|-------------|
| | VP | FP | FN | VN | P(%) | R(%) | A(%) | FM(%) | AUC | K |
| LBP | 378 | 126 | 118 | 281 | 73,00 | 73,00 | 72,98 | 73,00 | 0,78 | 0,45 |
| GLCM | 416 | 88 | 216 | 183 | 66,60 | 66,30 | 66,33 | 65,00 | 0,70 | 0,29 |
| HOG | 389 | 115 | 155 | 244 | 69,90 | 70,10 | 70,09 | 69,90 | 0,75 | 0,38 |
| Tamura | 397 | 107 | 146 | 253 | 71,90 | 72,00 | 71,98 | 71,80 | 0,78 | 0,42 |
| GLRLM | 372 | 132 | 188 | 211 | 64,30 | 64,60 | 64,56 | 64,20 | 0,69 | 0,27 |
| CNN- <i>AlexNet</i> | 430 | 74 | 112 | 287 | 79,40 | 79,40 | 79,40 | 79,30 | 0,84 | 0,57 |
| CNN- <i>CaffeNet</i> | 426 | 78 | 108 | 291 | 79,40 | 79,40 | 79,40 | 79,30 | 0,83 | 0,57 |
| CNN- <i>VGGFNet</i> | 434 | 70 | 115 | 284 | 79,60 | 79,50 | 79,51 | 79,30 | 0,85 | 0,58 |
| CNN- <i>VGGMNet</i> | 430 | 74 | 126 | 273 | 77,90 | 77,90 | 77,85 | 77,60 | 0,83 | 0,54 |
| CNN- <i>VGGSNet</i> | 422 | 82 | 95 | 304 | 80,40 | 80,40 | 80,39 | 80,40 | 0,86 | 0,60 |
| CNN- <i>VGG16Net</i> | 424 | 80 | 127 | 272 | 77,10 | 77,10 | 77,07 | 76,90 | 0,82 | 0,53 |
| CNN- <i>VGG19Net</i> | 415 | 89 | 139 | 260 | 74,70 | 74,80 | 74,75 | 74,50 | 0,80 | 0,48 |

Tabela 7 – Resultados da 1ª estratégia com o classificador RT.

| Descritor | RT | | | | | | | | | |
|----------------------|------------|------------|------------|------------|--------------|--------------|--------------|--------------|-------------|-------------|
| | VP | FP | FN | VN | P(%) | R(%) | A(%) | FM(%) | AUC | K |
| LBP | 342 | 162 | 154 | 245 | 65,10 | 65,00 | 65,00 | 65,00 | 0,64 | 0,29 |
| GLCM | 350 | 154 | 153 | 246 | 66,00 | 66,00 | 66,00 | 66,00 | 0,66 | 0,31 |
| HOG | 367 | 137 | 156 | 243 | 67,40 | 67,60 | 67,55 | 67,50 | 0,66 | 0,34 |
| Tamura | 356 | 148 | 141 | 258 | 68,10 | 68,00 | 67,99 | 68,00 | 0,67 | 0,35 |
| GLRLM | 342 | 162 | 157 | 242 | 64,70 | 64,70 | 64,67 | 64,70 | 0,64 | 0,28 |
| CNN- <i>AlexNet</i> | 361 | 143 | 163 | 236 | 66,00 | 66,10 | 66,11 | 66,00 | 0,65 | 0,31 |
| CNN- <i>CaffeNet</i> | 334 | 170 | 165 | 234 | 63,00 | 62,90 | 62,90 | 62,90 | 0,62 | 0,25 |
| CNN- <i>VGGFNet</i> | 364 | 140 | 138 | 261 | 69,20 | 69,20 | 69,21 | 69,20 | 0,68 | 0,37 |
| CNN- <i>VGGMNet</i> | 367 | 137 | 142 | 257 | 69,10 | 69,10 | 69,10 | 69,10 | 0,68 | 0,37 |
| CNN- <i>VGGSNet</i> | 374 | 130 | 167 | 232 | 66,90 | 67,10 | 67,11 | 66,90 | 0,66 | 0,32 |
| CNN- <i>VGG16Net</i> | 357 | 147 | 160 | 239 | 65,90 | 66,00 | 66,00 | 65,90 | 0,65 | 0,31 |
| CNN- <i>VGG19Net</i> | 361 | 143 | 153 | 246 | 67,10 | 67,20 | 67,22 | 67,10 | 0,66 | 0,33 |

Tabela 8 – Resultados da 1ª estratégia com o classificador RF.

| Descritor | RF | | | | | | | | | |
|----------------------|------------|-----------|------------|------------|--------------|--------------|--------------|--------------|-------------|-------------|
| | VP | FP | FN | VN | P(%) | R(%) | A(%) | FM(%) | AUC | K |
| LBP | 395 | 109 | 112 | 287 | 75,50 | 75,50 | 75,52 | 75,50 | 0,82 | 0,50 |
| GLCM | 388 | 116 | 114 | 285 | 74,50 | 74,50 | 74,52 | 74,50 | 0,81 | 0,48 |
| HOG | 402 | 102 | 97 | 302 | 78,00 | 78,00 | 77,96 | 78,00 | 0,86 | 0,55 |
| Tamura | 392 | 112 | 120 | 279 | 74,30 | 74,30 | 74,30 | 74,30 | 0,83 | 0,47 |
| GLRLM | 389 | 115 | 136 | 263 | 72,10 | 72,20 | 72,20 | 72,10 | 0,80 | 0,43 |
| CNN- <i>AlexNet</i> | 444 | 60 | 137 | 262 | 78,60 | 78,20 | 78,18 | 78,80 | 0,87 | 0,54 |
| CNN- <i>CaffeNet</i> | 445 | 59 | 114 | 285 | 81,00 | 80,80 | 80,84 | 80,60 | 0,83 | 0,60 |
| CNN- <i>VGGFNet</i> | 435 | 69 | 117 | 282 | 79,50 | 79,40 | 79,40 | 79,20 | 0,88 | 0,57 |
| CNN- <i>VGGMNet</i> | 438 | 66 | 110 | 289 | 80,60 | 80,50 | 80,51 | 80,30 | 0,88 | 0,60 |
| CNN- <i>VGGSNet</i> | 445 | 59 | 109 | 290 | 81,50 | 81,40 | 81,39 | 81,20 | 0,89 | 0,61 |
| CNN- <i>VGG16Net</i> | 430 | 74 | 117 | 282 | 78,90 | 78,80 | 78,84 | 78,70 | 0,87 | 0,56 |
| CNN- <i>VGG19Net</i> | 439 | 65 | 119 | 280 | 79,80 | 79,60 | 79,62 | 79,40 | 0,86 | 0,58 |

Assim, podemos observar que, o descritor CNN-*VGGSNet* obteve melhor resultado em três dos cinco classificadores testados. Contudo, o melhor resultado encontrado neste cenário foi com a rede neural convolucional CNN-*VGGFNet*, que obteve acurácia de 84,82%

e índice kappa de 0,69, o que é considerado um resultado “Muito Bom” de acordo com a Tabela 3. A Tabela 9 apresenta um resumo dos melhores resultados encontrados no cenário 1, o qual corresponde ao cenário sem segmentação.

Tabela 9 – Resumo dos Melhores Resultados do Cenário 1.

| Classificador | Descritor | P(%) | R(%) | A(%) | FM(%) | AUC | K |
|---------------|----------------|--------------|--------------|--------------|--------------|-------------|-------------|
| SVM | <i>VGGFNet</i> | 84,80 | 84,80 | 84,82 | 84,80 | 0,84 | 0,69 |
| MLP | <i>VGGSNet</i> | 83,00 | 82,90 | 82,94 | 83,00 | 0,88 | 0,65 |
| KNN | <i>VGGSnet</i> | 80,40 | 80,40 | 80,39 | 80,40 | 0,86 | 0,60 |
| RT | <i>VGGFnet</i> | 69,20 | 69,20 | 69,21 | 69,20 | 0,68 | 0,37 |
| RF | <i>VGGSNet</i> | 81,50 | 81,40 | 81,39 | 81,20 | 0,89 | 0,61 |

5.2 Resultados com a segmentação do disco óptico

Neste cenário, a segmentação da região do DO ou padrão-ouro é utilizada nos mesmos classificadores testados no cenário 1. Foram utilizados os mesmos parâmetros do cenário anterior para os classificadores. O número de imagens testadas foi de 873 imagens, sendo 489 saudáveis e 384 com glaucoma. Devido ao fato da base de dados HRF não possuir o padrão ouro do DO ela não é aplicada neste cenário.

A partir da análise das Tabelas 10, 11, 12, 13 e 14 podemos perceber que, em todos os cinco classificadores testados, os descritores GLCM e as redes neurais convolucionais obtiveram uma taxa de acerto maior em relação aos outros descritores. A GLCM obteve, nos classificadores MLP, KNN, RT e RF, o melhor resultado individual entre os descritores, com 84,99%, 82,24%, 81,40% e 83,16% de acurácia, respectivamente. No classificador SVM o melhor resultado encontrado foi no descritor CNN-*VGG19Net* com acurácia de 82,70%.

Tabela 10 – Resultados da 2ª estratégia com o classificador SVM.

| Descritor | SVM | | | | | | | | | |
|--------------|------------|-----------|-----------|------------|--------------|--------------|--------------|--------------|-------------|-------------|
| | VP | FP | FN | VN | P(%) | R(%) | A(%) | FM(%) | AUC | K |
| LBP | 489 | 0 | 0 | 384 | 31,40 | 56,00 | 56,01 | 40,20 | 0,50 | 0 |
| GLCM | 367 | 122 | 51 | 333 | 81,40 | 80,20 | 80,18 | 80,20 | 0,80 | 0,60 |
| HOG | 391 | 98 | 94 | 290 | 78,00 | 78,00 | 78,00 | 78,00 | 0,77 | 0,55 |
| Tamura | 381 | 108 | 157 | 227 | 69,50 | 69,60 | 69,64 | 69,30 | 0,68 | 0,37 |
| GLRLM | 346 | 143 | 174 | 210 | 63,40 | 63,70 | 63,68 | 63,50 | 0,62 | 0,25 |
| Morfologia | 224 | 265 | 153 | 231 | 53,80 | 52,10 | 52,11 | 52,10 | 0,53 | 0,05 |
| CNN-AlexNet | 419 | 70 | 83 | 301 | 82,50 | 82,50 | 82,47 | 82,50 | 0,82 | 0,64 |
| CNN-CaffeNet | 418 | 71 | 96 | 288 | 80,80 | 80,90 | 80,87 | 80,80 | 0,80 | 0,60 |
| CNN-VGGFNet | 414 | 75 | 93 | 291 | 80,70 | 80,80 | 80,50 | 80,70 | 0,80 | 0,60 |
| CNN-VGGMNet | 414 | 75 | 81 | 303 | 82,10 | 82,10 | 82,13 | 82,10 | 0,81 | 0,63 |
| CNN-VGGSNet | 411 | 78 | 89 | 295 | 80,80 | 80,90 | 80,87 | 80,80 | 0,80 | 0,61 |
| CNN-VGG16Net | 396 | 93 | 97 | 287 | 78,20 | 78,20 | 78,23 | 78,20 | 0,77 | 0,55 |
| CNN-VGG19Net | 414 | 75 | 76 | 308 | 82,70 | 82,70 | 82,70 | 82,70 | 0,82 | 0,65 |

Tabela 11 – Resultados da 2ª estratégia com o classificador MLP.

| Descritor | MLP | | | | | | | | | |
|--------------|------------|-----------|-----------|------------|--------------|--------------|--------------|--------------|-------------|-------------|
| | VP | FP | FN | VN | P(%) | R(%) | A(%) | FM(%) | AUC | K |
| LBP | 379 | 110 | 113 | 271 | 67,00 | 66,90 | 66,89 | 66,90 | 0,80 | 0,32 |
| GLCM | 415 | 74 | 57 | 327 | 85,10 | 85,00 | 84,99 | 85,00 | 0,92 | 0,69 |
| HOG | 376 | 113 | 125 | 259 | 72,70 | 72,73 | 72,70 | 72,70 | 0,81 | 0,44 |
| Tamura | 354 | 135 | 124 | 260 | 70,40 | 70,30 | 70,33 | 70,40 | 0,78 | 0,40 |
| GLRLM | 385 | 104 | 120 | 264 | 74,30 | 74,30 | 74,34 | 74,30 | 0,80 | 0,47 |
| Morfologia | 381 | 108 | 215 | 169 | 62,60 | 63,00 | 63,00 | 61,80 | 0,69 | 0,2 |
| CNN-AlexNet | 401 | 88 | 82 | 302 | 80,60 | 80,50 | 80,52 | 80,50 | 0,87 | 0,60 |
| CNN-CaffeNet | 413 | 76 | 92 | 292 | 80,70 | 80,80 | 80,75 | 80,70 | 0,87 | 0,60 |
| CNN-VGGFNet | 403 | 86 | 89 | 295 | 79,90 | 80,00 | 79,95 | 79,90 | 0,87 | 0,59 |
| CNN-VGGMNet | 417 | 72 | 74 | 310 | 83,30 | 83,30 | 83,27 | 83,30 | 0,89 | 0,66 |
| CNN-VGGSNet | 413 | 76 | 84 | 300 | 81,60 | 81,70 | 81,67 | 81,70 | 0,88 | 0,62 |
| CNN-VGG16Net | 404 | 85 | 89 | 295 | 80,00 | 80,10 | 80,06 | 80,10 | 0,87 | 0,59 |
| CNN-VGG19Net | 410 | 79 | 83 | 301 | 81,40 | 81,40 | 81,44 | 81,40 | 0,87 | 0,62 |

Tabela 12 – Resultados da 2ª estratégia com o classificador KNN.

| Descritor | KNN | | | | | | | | | |
|--------------|------------|-----------|-----------|------------|--------------|--------------|--------------|--------------|-------------|-------------|
| | VP | FP | FN | VN | P(%) | R(%) | A(%) | FM(%) | AUC | K |
| LBP | 386 | 103 | 130 | 254 | 73,20 | 73,30 | 73,31 | 73,20 | 0,79 | 0,45 |
| GLCM | 414 | 75 | 80 | 304 | 82,20 | 82,20 | 82,24 | 82,20 | 0,89 | 0,64 |
| HOG | 378 | 111 | 140 | 244 | 71,10 | 71,20 | 71,25 | 71,10 | 0,77 | 0,41 |
| Tamura | 366 | 123 | 151 | 233 | 68,40 | 68,60 | 68,61 | 68,50 | 0,73 | 0,35 |
| GLRLM | 393 | 96 | 142 | 242 | 72,60 | 72,70 | 72,73 | 72,50 | 0,77 | 0,44 |
| Morfologia | 334 | 155 | 170 | 214 | 62,60 | 62,80 | 62,77 | 62,70 | 0,68 | 0,24 |
| CNN-AlexNet | 420 | 69 | 100 | 284 | 80,60 | 80,60 | 80,64 | 80,50 | 0,86 | 0,60 |
| CNN-CaffeNet | 423 | 66 | 116 | 268 | 79,30 | 79,20 | 79,15 | 78,90 | 0,84 | 0,57 |
| CNN-VGGFNet | 433 | 56 | 108 | 276 | 81,40 | 81,20 | 81,21 | 81,00 | 0,86 | 0,61 |
| CNN-VGGMNet | 426 | 63 | 109 | 275 | 80,40 | 80,30 | 80,29 | 80,10 | 0,85 | 0,59 |
| CNN-VGGSNet | 418 | 71 | 96 | 288 | 80,80 | 80,90 | 80,87 | 80,80 | 0,86 | 0,61 |
| CNN-VGG16Net | 413 | 76 | 138 | 246 | 75,60 | 75,50 | 75,49 | 75,10 | 0,81 | 0,49 |
| CNN-VGG19Net | 428 | 61 | 139 | 245 | 77,50 | 77,10 | 77,09 | 76,60 | 0,81 | 0,52 |

Apesar do trabalho de [Srinivasan, Dubey e Ganeshbabu \(2016\)](#), ter obtido um alto desempenho utilizando o classificador SVM, neste estudo ele não obteve um desempenho considerável, onde apenas as redes neurais convolucionais e a GLCM obtiveram resultados considerados “Muito Bom”.

O maior resultado encontrado neste cenário foi com o descritor GLCM e o classificador MLP onde obteve uma acurácia de 84,99% e um índice kappa de 0,69. Neste estudo o classificador *Random tree* proporcionou resultados abaixo dos outros classificadores aplicados, tendo apenas a GLCM com o resultado considerado “Muito Bom” pois o índice kappa foi de 0,62.

Tabela 13 – Resultados da 2ª estratégia com o classificador RT.

| Descritor | RT | | | | | | | | | |
|--------------|------------|-----------|-----------|------------|--------------|--------------|--------------|--------------|-------------|-------------|
| | VP | FP | FN | VN | P(%) | R(%) | A(%) | FM(%) | AUC | K |
| LBP | 324 | 165 | 138 | 246 | 65,60 | 65,30 | 65,29 | 65,40 | 0,65 | 0,30 |
| GLCM | 412 | 77 | 85 | 299 | 81,40 | 81,40 | 81,44 | 81,40 | 0,81 | 0,62 |
| HOG | 327 | 162 | 153 | 231 | 64,00 | 63,90 | 63,91 | 64,00 | 0,63 | 0,27 |
| Tamura | 334 | 155 | 153 | 231 | 64,70 | 64,70 | 64,71 | 64,70 | 0,64 | 0,28 |
| GLRLM | 351 | 138 | 141 | 243 | 68,00 | 68,00 | 68,04 | 68,00 | 0,67 | 0,35 |
| Morfologia | 316 | 173 | 177 | 207 | 59,90 | 59,90 | 59,91 | 59,90 | 0,59 | 0,18 |
| CNN-AlexNet | 361 | 128 | 125 | 259 | 71,00 | 71,00 | 71,02 | 71,00 | 0,70 | 0,41 |
| CNN-CaffeNet | 373 | 116 | 122 | 262 | 72,70 | 72,70 | 72,74 | 72,70 | 0,72 | 0,44 |
| CNN-VGGFNet | 368 | 121 | 124 | 260 | 71,90 | 71,90 | 71,93 | 71,90 | 0,71 | 0,43 |
| CNN-VGGMNet | 349 | 140 | 117 | 267 | 70,80 | 70,60 | 70,56 | 70,60 | 0,70 | 0,40 |
| CNN-VGGSNet | 360 | 129 | 132 | 252 | 70,10 | 70,10 | 70,10 | 70,10 | 0,69 | 0,39 |
| CNN-VGG16Net | 347 | 142 | 133 | 251 | 68,60 | 68,50 | 68,50 | 68,50 | 0,68 | 0,36 |
| CNN-VGG19Net | 352 | 137 | 140 | 244 | 68,20 | 68,30 | 68,27 | 68,30 | 0,67 | 0,35 |

Tabela 14 – Resultados da 2ª estratégia com o classificador RF.

| Descritor | RF | | | | | | | | | |
|--------------|------------|-----------|-----------|------------|--------------|--------------|--------------|--------------|-------------|-------------|
| | VP | FP | FN | VN | P(%) | R(%) | A(%) | FM(%) | AUC | K |
| LBP | 386 | 103 | 120 | 264 | 74,40 | 74,50 | 74,45 | 74,40 | 0,82 | 0,48 |
| GLCM | 422 | 67 | 80 | 304 | 83,10 | 83,20 | 83,16 | 83,10 | 0,93 | 0,65 |
| HOG | 383 | 106 | 112 | 272 | 68,50 | 68,70 | 68,72 | 68,60 | 0,84 | 0,35 |
| Tamura | 385 | 104 | 150 | 234 | 70,80 | 70,90 | 70,90 | 70,60 | 0,79 | 0,40 |
| GLRLM | 393 | 96 | 116 | 268 | 75,60 | 75,70 | 75,71 | 75,60 | 0,83 | 0,50 |
| Morfologia | 352 | 137 | 157 | 227 | 66,20 | 66,30 | 66,32 | 66,20 | 0,72 | 0,31 |
| CNN-AlexNet | 414 | 75 | 97 | 287 | 80,30 | 80,30 | 80,29 | 80,20 | 0,89 | 0,59 |
| CNN-CaffeNet | 427 | 62 | 99 | 285 | 81,60 | 81,60 | 81,55 | 81,40 | 0,90 | 0,62 |
| CNN-VGGFNet | 431 | 58 | 98 | 286 | 82,20 | 82,10 | 82,12 | 82,00 | 0,89 | 0,63 |
| CNN-VGGMNet | 413 | 76 | 83 | 301 | 81,80 | 81,80 | 81,78 | 81,80 | 0,89 | 0,63 |
| CNN-VGGSNet | 427 | 62 | 89 | 295 | 82,70 | 82,70 | 82,70 | 82,60 | 0,90 | 0,64 |
| CNN-VGG16Net | 424 | 65 | 102 | 282 | 80,90 | 80,90 | 80,87 | 80,70 | 0,88 | 0,60 |
| CNN-VGG19Net | 420 | 69 | 97 | 287 | 81,00 | 81,00 | 80,98 | 80,90 | 0,87 | 0,61 |

Um resumo dos melhores resultados alcançados no cenário 2 para os classificadores aplicados está demonstrado na Tabela 15.

Tabela 15 – Resumo dos Melhores Resultados do Cenário 2.

| Classificador | Descritor | P(%) | R(%) | A(%) | FM(%) | AUC | K |
|---------------|-----------------|--------------|--------------|--------------|--------------|-------------|-------------|
| SVM | <i>VGG19Net</i> | 82,70 | 82,70 | 82,70 | 82,70 | 0,82 | 0,65 |
| MLP | <i>GLCM</i> | 85,10 | 85,00 | 84,99 | 85,00 | 0,92 | 0,69 |
| KNN | <i>GLCM</i> | 82,20 | 82,20 | 82,24 | 82,20 | 0,89 | 0,64 |
| RT | <i>GLCM</i> | 81,40 | 81,40 | 81,44 | 81,40 | 0,81 | 0,62 |
| RF | <i>GLCM</i> | 83,10 | 83,20 | 83,16 | 83,10 | 0,93 | 0,65 |

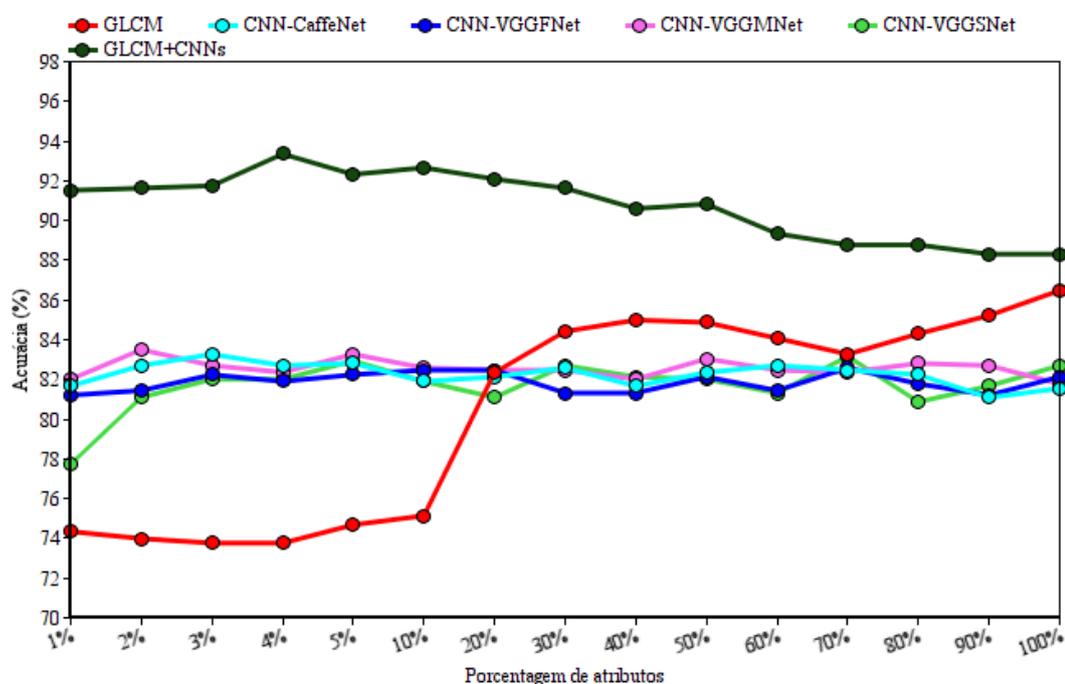
5.3 Resultados com o descritor proposto

Observando as Tabelas 10, 11 e 14 podemos perceber que os descritores GLCM, CNN-AlexNet, CNN-CaffeNet, CNN-VGGFNet, CNN-VGGMNet, CNN-VGGSNet, CNN-

VGG16Net e *CNN-VGG19Net* obtiveram os melhores resultados dentre os descritores utilizados. Podemos visualizar também que os classificadores SVM e *Random Forest* obtiveram um maior número de descritores com resultados considerados “Muito Bons”. Comparando os resultados dos descritores que alcançaram kappa acima de 0,60 nesses dois classificadores, podemos inferir que o *Random Forest* teve um melhor desempenho, assim pode-se concatenar os descritores com o objetivo de melhorar os resultados. Visto que cada CNN possui 4096 atributos, resolvemos utilizar apenas os cinco melhores descritores com o classificador *Random Forest* e assim concatená-los.

Com base nesta análise realizou-se 15 testes, onde a seleção de atributos é utilizada com o tamanho do vetor de características variando em 1%, 2%, 3%, 4%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% e 90%, para que possamos excluir os atributos menos significativos e tentar melhorar os resultados. Ao todo foram feitos 90 testes (15 testes de porcentagens de seleção de atributos \times 6 descritores) utilizando o classificador *Random Forest*, e seus resultados podem ser visualizados na Figura 14. Nesta figura, podemos analisar que, individualmente, o descritor GLCM teve resultados inferiores com a seleção de atributos, tendo obtido um resultado mais baixo com 3% e 4% dos seus atributos mais significativos.

Figura 14 – Comparação dos resultados dos descritores GLCM, CNNs e GLCM+CNNs.



Em todas as CNNs houveram resultados melhores com a seleção de atributos, onde na *CNN-CaffeNet*, *CNN-VGGFNet*, *CNN-VGGMNet* e *CNN-VGGSNet* os melhores resultados encontrados foram com 3%, 70%, 2% e 70% dos atributos mais significativos, respectivamente. Podemos perceber também que com a junção destes descritores, os resultados sempre estiveram acima dos seus valores obtidos individuais.

Assim, a Tabela 16 demonstra a porcentagem que apresentou o melhor resultado nos descritores GLCM, nas CNNs (CNN-*CaffeNet*, CNN-*VGGFNet*, CNN-*VGGMNet* e CNN-*VGGSNet*) e com a junção da GLCM com as CNNs.

Tabela 16 – Melhores resultados obtidos com a seleção de atributos.

| Descritor | Dados (%) | Qt de Atributos | P(%) | AUC | A(%) | K |
|----------------------|-----------|-----------------|--------------|-------------|--------------|-------------|
| GLCM | 100 | 85 | 86,50 | 0,95 | 86,48 | 0,72 |
| CNN- <i>CaffeNet</i> | 3 | 123 | 83,30 | 0,90 | 83,27 | 0,65 |
| CNN- <i>VGGFNet</i> | 70 | 2867 | 82,60 | 0,89 | 82,59 | 0,64 |
| CNN- <i>VGGMNet</i> | 2 | 82 | 83,50 | 0,90 | 83,50 | 0,66 |
| CNN- <i>VGGSNet</i> | 70 | 2867 | 82,60 | 0,89 | 82,59 | 0,64 |
| GLCM+CNNs | 4 | 659 | 93,40 | 0,98 | 93,35 | 0,86 |

Podemos visualizar que no descritor GLCM, a melhor quantidade de atributos foi de 100%, onde se obteve uma acurácia de 86,48%. A CNN-*CaffeNet* obteve um melhor resultado com 3%, ou seja, 123 atributos de 4096 são mais significativos. A CNN-*VGGFNet* assim como a CNN-*VGGSNet* obtiveram melhor resultado com 70% dos seus atributos, o que corresponde a 2867 atributos de 4096. A CNN-*VGGMNet* atingiu seu maior resultado com apenas 2% dos atributos mais relevantes, o que corresponde a 82 atributos de 4096. Contudo, com a junção destes, o resultado melhorou significativamente, alcançando uma acurácia da ordem de 93,35% com apenas 4% dos melhores atributos, o que equivale a 659 atributos de 16.469. Como o índice Kappa (K) foi de 0,86, este resultado é considerado “Excelente” e demonstra a eficiência do método proposto neste estudo.

A Tabela 17 demonstra a quantidade de atributos pertencentes a cada descritor durante o processo de seleção de atributos. Podemos inferir, a partir desta tabela, que com 4% dos atributos mais significativos o descritor GLCM, que possui 85 atributos extraídos, obteve 62 atributos, o que corresponde a uma taxa significativa nos atributos mais importantes em comparação com as CNNs.

Uma segmentação ótima ainda é um desafio no desenvolvimento de sistemas CAD. Assim, para analisar a robustez do método proposto, mesmo quando há erros de segmentação, foram realizados testes em dois outros cenários: sub-segmentação e sobre-segmentação. O baixo contraste entre o DO e o fundo da imagem pode ocorrer um nível de cor e variações de intensidade semelhantes o que acarreta esses tipos de erros de segmentação.

Figura 15 exemplifica como esses erros de segmentação foram simulados. Na figura 15a, o padrão ouro é ilustrado pelo contorno branco. Para simular o erro de sub-segmentação, diminuimos a área do padrão ouro em 10%, e o resultado pode ser visto na Figura 15b pelo contorno vermelho. Na Figura 15c apresenta o erro de super-segmentação onde o padrão ouro foi aumentado em 10% e está representado pelo contorno verde.

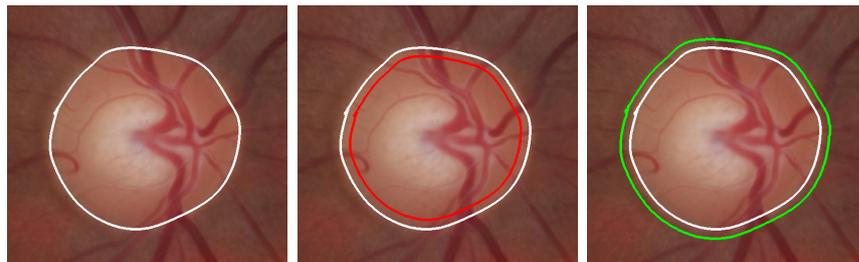
Tabela 18 apresenta os resultados dos testes obtidos com os erros de segmentação

Tabela 17 – Quantidade de atributos por descritor.

| (%) dos Dados | GLCM | CNN- <i>CaffeNet</i> | CNN- <i>VGGFNet</i> | CNN- <i>VGGMNet</i> | CNN- <i>VGGSNet</i> | Total |
|---------------|------|----------------------|---------------------|---------------------|---------------------|-------|
| 1 | 53 | 13 | 19 | 25 | 55 | 165 |
| 2 | 58 | 59 | 58 | 45 | 109 | 329 |
| 3 | 62 | 99 | 99 | 80 | 154 | 494 |
| 4 | 62 | 151 | 147 | 101 | 198 | 659 |
| 5 | 62 | 205 | 195 | 123 | 238 | 823 |
| 10 | 65 | 445 | 411 | 271 | 455 | 1647 |
| 20 | 66 | 873 | 870 | 584 | 901 | 3294 |
| 30 | 68 | 1287 | 1330 | 973 | 1283 | 4941 |
| 40 | 69 | 1682 | 1749 | 1382 | 1706 | 6588 |
| 50 | 71 | 2076 | 2184 | 1829 | 2075 | 8235 |
| 60 | 81 | 2477 | 2577 | 2298 | 2448 | 9881 |
| 70 | 84 | 2878 | 2965 | 2791 | 2810 | 11528 |
| 80 | 84 | 3236 | 3397 | 3491 | 2967 | 13175 |
| 90 | 84 | 3649 | 3764 | 3692 | 3633 | 14822 |
| 100 | 85 | 4096 | 4096 | 4096 | 4096 | 16469 |

Figura 15 – Exemplos de sub e super-segmentação.

(a) Padrão Ouro. (b) Sub-segmentação. (c) Super-segmentação.



(Sub e Super) usando a abordagem proposta. É possível observar que houve uma diminuição nas taxas relativas aos resultados anteriores. No entanto, o método obteve um índice Kappa “Muito Bom”. Esse resultado mostra que mesmo quando há erros de segmentação, o método proposto ainda obtém resultados promissores.

Tabela 18 – Resultados da classificação do método proposto em imagens com erros de segmentação.

| Erros de Segmentação | P(%) | AUC | A(%) | K |
|----------------------|-------|------|-------|------|
| Sub | 82,60 | 0,90 | 82,58 | 0,64 |
| Super | 87,60 | 0,94 | 87,62 | 0,74 |

Com a finalidade de validar o método proposto, foram realizadas comparações com métodos do estado da arte. A Tabela 19 apresenta uma comparação entre o método proposto e as técnicas de [Kotyk et al. \(2016\)](#) e [Claro et al. \(2016\)](#). Nessa comparação os algoritmos originais foram executados nas 873 imagens disponíveis. Nossa proposta obteve um melhor desempenho quando comparado aos demais. A acurácia obtida foi de 93,35% e

Kappa de 0,86 o que é um resultado considerado “Excelente”, já os outros dois trabalhos alcançaram índice Kappa considerado “Muito Bom”.

Tabela 19 – Comparação do método proposto com os trabalhos relacionados.

| Trabalhos | P(%) | AUC | A(%) | K |
|-------------------------------------|--------------|-------------|--------------|-------------|
| Kotyk et al. (2016) | 86,50 | 0,952 | 86,48 | 0,72 |
| Claro et al. (2016) | 87,00 | 0,955 | 87,05 | 0,73 |
| Método Proposto | 93,40 | 0,98 | 93,35 | 0,86 |

A Tabela 20 apresenta a comparação do método desenvolvido neste estudo com o do trabalho de [Orlando et al. \(2017\)](#). Este trabalho também utiliza redes neurais convolucionais como descritor. Contudo, essas redes foram diferentes das aplicadas nesse estudo. Como não tivemos acesso ao código original dos autores optamos por executar nosso método na base DRISHTI ([SIVASWAMY et al., 2014](#)) (a mesma utilizada em [Orlando et al. \(2017\)](#)). A única métrica de avaliação apresentada pelos autores é a AUC com valor 0,762. Nossa proposta obteve melhor desempenho na base DRISHTI com uma AUC de 0,814.

Tabela 20 – Comparação do método proposto com o trabalho de Orlando et al.

| Trabalhos | P(%) | AUC | A(%) | K |
|---------------------------------------|--------------|-------------|--------------|-------------|
| Orlando et al. (2017) | - | 0,762 | - | - |
| Método Proposto | 86,20 | 0,88 | 86,13 | 0,64 |

Os resultados apresentados nas Tabelas 19 e 20 evidenciam a validade da nossa proposta. Dado que nosso método se baseia na união de descritores de textura e transferência de aprendizado em CNNs conseguimos demonstrar que ele é mais eficiente que métodos baseados apenas em textura [Kotyk et al. \(2016\)](#) e [Claro et al. \(2016\)](#) e apenas em transferência de aprendizado [Orlando et al. \(2017\)](#).

5.4 Considerações Finais

Constatou-se que diversos métodos da literatura apresentaram resultados promissores, entretanto necessitam de maior robustez e de uma maior quantidade de imagens para uma melhor performance e confiabilidade nos seus resultados. O método apresentado neste trabalho descreve um novo método para diagnóstico do glaucoma em imagens da retina com a utilização da concatenação da GLCM com as redes neurais convolucionais. Com base na comparação dos resultados obtidos pela abordagem, podemos perceber que a proporção de quantidade de imagens utilizadas pelos resultados obtidos é possível validar a sua robustez perante outros trabalhos, atingindo taxa “Excelente” .

6 Conclusões e Trabalhos Futuros

6.1 Conclusões

O glaucoma causa danos no nervo óptico e, caso não haja um controle, a doença pode levar à cegueira. Assim, a descoberta antecipada desta doença é de fundamental importância pois previne a perda total da visão. Assim, a área de Processamento de Imagens Digitais visa desenvolver ferramentas que possibilitem às máquinas não apenas automatizar procedimentos realizados pelas funções visuais dos seres humanos, mas também superar seu desempenho, pois são capazes de processar grande quantidade de informação com muito mais rapidez. Como a extração de características e a classificação são considerados os passos mais importantes no diagnóstico automático de diferentes sistemas computacionais.

Dentro deste contexto, esta dissertação teve como objetivo o estudo, avaliação e desenvolvimento de um método computacional capaz de identificar automaticamente o glaucoma.

Um estudo foi realizado a partir dos achados da literatura, onde podemos perceber que os autores costumam utilizar suas próprias bases de dados ou apenas uma base de dados. Ao se utilizar cinco bases distintas nota-se que cada uma delas apresentam características específicas, o que torna o processo de descrição e classificação mais complexo.

Outro estudo realizado foi com os descritores de textura, forma e redes neurais convolucionais, onde avaliamos estes descritores em dois cenários (com ou sem segmentação) e com cinco classificadores diferentes. Assim, constatamos que o cenário com segmentação obteve melhores resultados, e que o classificador *Random Forest* teve um maior número de descritores com melhores resultados. Ao final concatenamos os cinco melhores descritores e realizamos uma seleção de atributos para melhorar o desempenho do método desenvolvido.

Assim, pode-se observar que o método proposto obteve taxas de acerto satisfatória, alcançando um desempenho considerado “Excelente” pelo índice kappa, tendo obtido um valor de 0,86 e uma acurácia de 93,35%. Outras métricas de avaliação também foram utilizadas. Essas métricas demonstram-se pertinentes para a avaliação dos resultados do método proposto.

6.2 Trabalhos Futuros

Os resultados logrados nesse estudo são prósperos, entretanto podem ser melhorados. Um prosseguimento importante para que os resultados possam melhorar consiste na implementação de um novo descritor que utilize além das características de textura e

forma a característica de cor, além da investigação de outros descritores que estão sendo utilizados.

A implementação de um novo método para a segmentação da região do DO e da escavação seria de fundamental importância, pois além de se extrair características destas regiões, estas duas regiões pode-se realizar cálculos que auxiliam no diagnóstico do glaucoma, como o cálculo da razão entre o raio da escavação e o raio do disco óptico *Cup to Disc Ratio* (CDR). Outro importante passo seria a averiguação e aplicação de novas imagens, assim o método poderá se tornar mais robusto.

Ainda como trabalhos futuros, propõe-se a implementação dos trabalhos do estado da arte, para fins comparativos com o método aqui proposto.

Referências

- ABBADENI, N.; ZIOU, D.; WANG, S. Perceptual textural features corresponding to human visual perception. In: *Proc. of the Thirteenth Vision Interface Conference, Montreal, Quebec (Canada)*. [S.l.: s.n.], 2000. p. 365–372. Citado na página 31.
- ACHARYA, U. R. et al. Automated diagnosis of glaucoma using texture and higher order spectra features. *IEEE transactions on Information Technology in Biomedicine*, IEEE, v. 15, n. 3, p. 449–455, 2011. Citado 2 vezes nas páginas 21 e 24.
- AHA, D. W.; KIBLER, D.; ALBERT, M. K. Instance-based learning algorithms. *Machine learning*, Springer, v. 6, n. 1, p. 37–66, 1991. Citado 2 vezes nas páginas 38 e 39.
- AMARI, S. A theory of adaptive pattern classifiers. *IEEE Transactions on Electronic Computers*, IEEE, n. 3, p. 299–307, 1967. Citado 2 vezes nas páginas 22 e 38.
- ÁVILA, M.; ALVES, M.; NISHI, M. *As condições de Saúde Ocular no Brasil-2015. São Paulo (SP): Conselho Brasileiro de Oftalmologia; 145p.* 2015. Online; Acessado: 01 de junho de 2017. Citado na página 18.
- BALASUBRAMANIAN, T. et al. Hog feature based svm classification of glaucomatous fundus image with extraction of blood vessels. In: IEEE. *IEEE Annual India Conference (INDICON)*. [S.l.], 2016. p. 1–4. Citado 2 vezes nas páginas 21 e 24.
- BOSCARIOLI, C. Pré-processamento de dados para descoberta de conhecimento em banco de dados: uma visão geral. In: *Anais do III CONGED-Congresso de Tecnologias para Gestão de Dados e Metadados do Cone Sul. Guarapuava: Unicentro Editora.* [S.l.: s.n.], 2005. v. 1, p. 101–120. Citado na página 36.
- BRADSKI, G. Dr dobb’s journal of software tools. *The OpenCV Library*, 2000. Citado na página 43.
- BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, n. 1, p. 5–32, 2001. Citado 3 vezes nas páginas 21, 39 e 40.
- CHATFIELD, K. et al. Return of the devil in the details: Delving deep into convolutional nets. *arXiv preprint arXiv:1405.3531*, 2014. Citado na página 35.
- CHENG, J. et al. Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. *IEEE Transactions on Medical Imaging*, IEEE, v. 32, n. 6, p. 1019–1032, 2013. Citado na página 44.
- CHIMIESKI, B. F.; FAGUNDES, R. D. R. Association and classification data mining algorithms comparison over medical datasets. *Journal of health informatics*, v. 5, n. 2, 2013. Citado na página 41.
- CHU, A.; SEHGAL, C. M.; GREENLEAF, J. F. Use of gray value distribution of run lengths for texture analysis. *Pattern Recognition Letters*, Elsevier, v. 11, n. 6, p. 415–419, 1990. Citado na página 33.

- CLARO, M. et al. Automatic glaucoma detection based on optic disc segmentation and texture feature extraction. *CLEI Electronic Journal*, Centro Latinoamericano de Estudios en Informática, v. 19, n. 2, p. 5–5, 2016. Citado 4 vezes nas páginas 22, 24, 53 e 54.
- CORTES, C.; VAPNIK, V. Support-vector networks. *Machine learning*, Springer, v. 20, n. 3, p. 273–297, 1995. Citado 2 vezes nas páginas 21 e 37.
- COVER, T. M. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE transactions on electronic computers*, IEEE, n. 3, p. 326–334, 1965. Citado na página 37.
- DAHL, G. E.; SAINATH, T. N.; HINTON, G. E. Improving deep neural networks for lvsr using rectified linear units and dropout. In: IEEE. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.], 2013. p. 8609–8613. Citado 2 vezes nas páginas 28 e 34.
- DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: IEEE. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.], 2005. v. 1, p. 886–893. Citado 2 vezes nas páginas 21 e 30.
- FENGSHOU, Y. *Extraction of features from fundus images for glaucoma assessment*. Tese (Doutorado) — National University of Singapore, 2011. Citado na página 16.
- FUMERO, F. et al. Rim-one: An open retinal image database for optic nerve evaluation. In: IEEE. *24th International Symposium on Computer-Based Medical Systems*. [S.l.], 2011. p. 1–6. Citado 2 vezes nas páginas 25 e 26.
- FUNDUS, H.-R. Image database [online].[cit. 2012-12-11]. *Dostupné z: <https://www5.cs.fau.de/research/data/fundus-images/>*, 2012. Citado 2 vezes nas páginas 25 e 43.
- GALLOWAY, M. M. Texture analysis using gray level run lengths. *Computer graphics and image processing*, Elsevier, v. 4, n. 2, p. 172–179, 1975. Citado na página 32.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento de imagens digitais*. 3. ed. [S.l.]: Pearson Education, 2011. Citado na página 16.
- HALL, M. et al. The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, ACM, v. 11, n. 1, p. 10–18, 2009. Citado 2 vezes nas páginas 43 e 45.
- HARALICK, R. M.; SHANMUGAM, K. et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, IEEE, v. 3, n. 6, p. 610–621, 1973. Citado 2 vezes nas páginas 21 e 29.
- HAYKIN, S. S. *Neural networks: a comprehensive foundation*. [S.l.]: Tsinghua University Press, 2001. Citado 2 vezes nas páginas 37 e 38.
- JIA, Y. et al. Caffe: Convolutional architecture for fast feature embedding. In: ACM. *Proceedings of the 22nd ACM international conference on Multimedia*. [S.l.], 2014. p. 675–678. Citado na página 35.
- KAREGOWDA, A. G.; MANJUNATH, A.; JAYARAM, M. Comparative study of attribute selection using gain ratio and correlation based feature selection. *International Journal of Information Technology and Knowledge Management*, v. 2, n. 2, p. 271–277, 2010. Citado na página 36.

- KIM, Y.; STREET, W. N.; MENCZER, F. Feature selection in data mining. *Data mining: opportunities and challenges*, IGI Global, v. 3, n. 9, p. 80–105, 2003. Citado na página 36.
- KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: MONTREAL, CANADA. *Ijcai*. [S.l.], 1995. v. 14, n. 2, p. 1137–1145. Citado na página 45.
- KOTYK, T. et al. Semi-automated system for cup to disc measurement for diagnosing glaucoma using classification paradigm. In: SPRINGER. *Proceedings of the Second International Afro-European Conference for Industrial Advancement AECIA 2015*. [S.l.], 2016. p. 653–663. Citado 4 vezes nas páginas 22, 24, 53 e 54.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2012. p. 1097–1105. Citado na página 35.
- LAMANI, D. et al. Early detection of glaucoma through retinal nerve fiber layer analysis using fractal dimension and texture feature. *International Journal of Research in Engineering and Technology*, 2014. Citado na página 18.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *biometrics*, JSTOR, p. 159–174, 1977. Citado na página 42.
- LIN, T.-W.; HUANG, P.-Y.; CHENG, C. W.-C. Computer-aided diagnosis in medical imaging: Review of legal barriers to entry for the commercial systems. In: IEEE. *IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*. [S.l.], 2016. p. 1–5. Citado na página 16.
- 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*, Elsevier, v. 24, p. 34–46, 2016. Citado na página 19.
- NIRMALA, K.; VENKATESWARAN, N. Local binary pattern features based detection of glaucoma using support vector machine classifier. *Journal of Medical Imaging and Health Informatics*, American Scientific Publishers, v. 6, n. 6, p. 1370–1378, 2016. Citado 2 vezes nas páginas 22 e 24.
- OJALA, T.; PIETIKÄINEN, M. Unsupervised texture segmentation using feature distributions. In: SPRINGER. *International Conference on Image Analysis and Processing*. [S.l.], 1997. p. 311–318. Citado 2 vezes nas páginas 22 e 28.
- ORLANDO, J. I. et al. Convolutional neural network transfer for automated glaucoma identification. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *12th International Symposium on Medical Information Processing and Analysis*. [S.l.], 2017. p. 101600U–101600U. Citado 3 vezes nas páginas 23, 24 e 54.
- PASSOS, E. C.; ANDRADE-NETO, A.; LEMAIRE, T. Comportamento ótico do olho humano e suas ametropias. *Caderno de Física da UEFS*, v. 6, n. 1-2, p. 7–18, 2008. Citado 2 vezes nas páginas 16 e 17.
- POWERS, D. M. *Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation*, School of Informatics and Engineering, Flinders University, Adelaide, Australia. [S.l.], 2007. Citado na página 40.

- 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. Citado na página 18.
- QUINLAN, J. R. Induction of decision trees. *Machine learning*, Springer, v. 1, n. 1, p. 81–106, 1986. Citado na página 34.
- RAMOS, A. Fisiologia da visão um estudo sobre o “ver” e o “enxergar”. *Rio de Janeiro*, 2006. Citado na página 17.
- RAMOS, D. R. et al. Automatic construction of descriptive rules. Citeseer, 1997. Citado na página 36.
- RODRIGUEZ, P. M. Processos de ramificação: teoria e aplicações. ICMC/USP, São Carlos, 2012. Citado na página 39.
- ROSENFELD, G. H.; FITZPATRICK-LINS, K. A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric engineering and remote sensing*, v. 52, n. 2, p. 223–227, 1986. Citado na página 41.
- SAKATA, K. et al. Prevalence of glaucoma in a south brazilian population: Projeto glaucoma. *Investigative ophthalmology & visual science*, The Association for Research in Vision and Ophthalmology, v. 48, n. 11, p. 4974–4979, 2007. Citado na página 18.
- SALAM, A. A. et al. Automated detection of glaucoma using structural and non structural features. *SpringerPlus*, Springer, v. 5, n. 1, p. 1519, 2016. Citado 3 vezes nas páginas 22, 24 e 28.
- SIMONTHOMAS, S.; THULASI, N.; ASHARAF, P. Automated diagnosis of glaucoma using haralick texture features. In: IEEE. *International Conference on Information Communication and Embedded Systems (ICICES)*. [S.l.], 2014. p. 1–6. Citado 2 vezes nas páginas 21 e 24.
- SIMONYAN, K.; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*, 2014. Citado na página 35.
- SIVASWAMY, J. et al. Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation. In: IEEE. *IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. [S.l.], 2014. p. 53–56. Citado 2 vezes nas páginas 25 e 54.
- SOBEL, I. *Camera models and machine perception*. [S.l.], 1970. Citado na página 31.
- SOUSA, J. A. de et al. Texture based on geostatistic for glaucoma diagnosis from fundus eye image. *Multimedia Tools and Applications*, Springer, p. 1–18, 2017. Citado 2 vezes nas páginas 23 e 24.
- SRINIVASAN, C.; DUBEY, S.; GANESHBABU, T. Complex texture features for glaucomatous image classification system using fundus images. v. 2, n. 13, 2016. Citado 3 vezes nas páginas 22, 24 e 49.
- TAMURA, H.; MORI, S.; YAMAWAKI, T. Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, IEEE, v. 8, n. 6, p. 460–473, 1978. Citado na página 31.

THOMPSON, C. M.; SHURE, L. Image processing toolbox [for use with matlab]. The Math Works, 1995. Citado na página 43.

USMAN, M.; FRAZ, M. M.; BARMAN, S. A. Computer vision techniques applied for diagnostic analysis of retinal oct images: A review. *Archives of Computational Methods in Engineering*, Springer, p. 1–17, 2016. Citado na página 16.

ZHANG, Z. et al. Neuro-retinal optic cup detection in glaucoma diagnosis. In: IEEE. *BMEI'09. 2nd International Conference on Biomedical Engineering and Informatics*. [S.l.], 2009. p. 1–4. Citado na página 17.