

Diagnóstico de Glaucoma em Retinografias Usando Índices Taxonômicos e Aprendizado de Máquina

Layane M Azevedo
Núcleo de Computação
Aplicada
Universidade Federal do
Maranhão
Brasil
layane@nca.ufma.br

João D S Almeida
Núcleo de Computação
Aplicada
Universidade Federal do
Maranhão
Brasil
jdallyson@nca.ufma.br

Anselmo C Paiva
Núcleo de Computação
Aplicada
Universidade Federal do
Maranhão
Brasil
paiva@nca.ufma.br

Geraldo Braz Junior
Núcleo de Computação
Aplicada
Universidade Federal do
Maranhão
Brasil
geraldo@nca.ufma.br

Rodrigo M S Veras
Departamento de
Computação
Universidade Federal do Piauí
Brasil
rveras@ufpi.edu.br

ABSTRACT

Glaucoma is an irreversible disease that degrades the optic nerve until the complete loss of the visual field. The World Health Organization estimates that glaucoma accounts for approximately 5.2 million cases of blindness. The symptoms only manifest when the disease is at an advanced stage. In this context, this paper aims to present an automated computational method using texture descriptors to detect glaucoma in eye fundus images. To describe the texture, taxonomic indexes were used. To define the best image classification model, automated machine learning techniques were used along with the Bayesian optimization approach. The proposed method is organized into five steps: (1) Preprocessing, (2) Color Conversions, (3) Feature Extraction, (4) Classification, and (5) Validation. The experiments were performed on 922 public retinography images databases: RIM-ONE version 2, iChallenge-GON and Kaggle1000fundusimages. The method obtained 90.20% of sensitivity, 94.03% of specificity, f1-score of 0.82 and 92.97% of accuracy.

RESUMO

O glaucoma é uma doença óptica irreversível que degrada o nervo óptico até a completa perda do campo visual. A Organização Mundial de Saúde estima que o glaucoma seja responsável por aproximadamente 5,2 milhões de casos de cegueira que equivale a 15% do total de cegueira mundial, e até 2020 pode chegar a 11,2 milhões de casos. Os sintomas só

se manifestam quando a doença está em estágio avançado. Nesse contexto, este trabalho apresenta um método computacional utilizando descritores de textura para detectar o glaucoma automaticamente em imagens de fundo de olho. Para descrever a textura foram utilizados índices taxonômicos e na definição do melhor modelo de classificação das imagens utilizou-se técnicas de aprendizado de máquina automatizado com otimização Bayesiana. O método proposto está organizado em cinco etapas: (1) Pré-processamento, (2) Conversão em Espaços de Cores, (3) Extração de características, (4) Classificação e (5) Validação de resultados. Os experimentos foram realizados em 922 imagens de retinografia de bases públicas, sendo 455 da RIM-ONE versão 2, 400 da iChallenge-GON e 67 da base Kaggle1000fundusimages. O método apresentou resultados promissores obtendo 90,20% de sensibilidade, 94,03% de especificidade, f1 score de 0,82 e 92,97% de acurácia.

CCS Concepts

•Computing methodologies → Computer vision representations; •Applied computing → Health informatics; •General and reference → Experimentation;

Keywords

Glaucoma; Diagnosis; Retinography; Taxonomic Indexes; Machine Learning.

Palavras-chave

Glaucoma; Diagnóstico; Retinografia; Índices Taxonômicos; Aprendizado de Máquina.

1. INTRODUÇÃO

Segundo o Conselho Brasileiro de Oftalmologia, o glaucoma é a maior causa de cegueira irreversível no mundo e estima-se que o número de pessoas com glaucoma em todo o

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

mundo aumentará para 111,8 milhões até 2040 [19]. O glaucoma é uma doença ocular, inicialmente assintomática, que causa dano ao nervo óptico ocasionando, posteriormente, a perda irreversível do campo visual e possui como maior fator de risco o aumento da pressão intraocular (pressão dentro do olho) [26].

Ademais, o glaucoma tem a característica de ser hereditário, sendo extremamente relacionado a pressão intraocular. Normalmente pessoas desse grupo de risco fazem acompanhamento da pressão intraocular, mas o diagnóstico é feito por meio de alguns exames, sendo a retinografia um dos principais. A retinografia é uma forma muito comum de se diagnosticar algumas doenças oculares, onde faz-se a utilização da imagem do fundo de olho para analisar as estruturas da retina: disco, cavidade e nervo óptico [8].

O diagnóstico do glaucoma através de imagens requer análise e interpretação de um especialista da área. Por outro lado, ao longo dos anos pesquisas na área de diagnóstico assistido por computador (CAD) vem crescendo, onde é feita a análise de imagens computadorizadas com a finalidade de auxiliar a interpretação de exames com imagens médicas fornecendo uma automatização nas análises e uma segunda opinião para o especialista [13]. As principais formas encontradas na literatura para a detecção e diagnóstico do glaucoma são: por segmentação da região do disco óptico e da cavidade; e por extração de características com técnicas de classificação.

Na detecção por segmentação é calculado o CDR (*cup-to-disc ratio*), que é a razão do DO (Disco Óptico) com a CDO (escavação do disco óptico) e, segundo Spaeth (1981), se o valor gerado for maior ou igual a 0,65 normalmente indica que o paciente possui glaucoma. Um problema desse método é encontrar as regiões de interesse (ROI). Essa abordagem, por utilizar algoritmos de segmentação, pode apresentar problemas de localização para determinação de valores de limiar ou demarcação que levam a erros no diagnóstico [24]. Uma abordagem diferente desta é a extração de características que consideram a imagem como um todo descrevendo-as baseadas em cor, textura e/ou forma. Igualmente, alguns autores [7, 2, 3] obtiveram resultados relevantes com esta abordagem.

Este artigo tem como objetivo apresentar a avaliação da utilização dos índices taxonômicos, comumente utilizados na ecologia, como descritores de textura, com a finalidade de serem aplicados no diagnóstico do glaucoma em retinografias. Para tal, são utilizadas bases públicas de imagens de retinografia com casos de glaucoma, nas quais, são aplicadas técnicas de pré-processamento de imagens, seguida da extração dos índices taxonômicos adequados para a proposta e da classificação.

O restante deste artigo está organizado da forma descrita a seguir. Na Seção ??, são apresentados o objetivo geral e específicos deste trabalho. Na Seção 3, são descritos os materiais e métodos utilizados. Os resultados obtidos, as discussões sobre esses resultados e um comparativo com outros trabalhos equivalentes ao nosso são apresentados na Seção 4. Concluindo, na Seção 5 são apresentados as conclusões da pesquisa e os trabalhos futuros em andamento ou a realizar.

2. TRABALHOS RELACIONADOS

Esta seção apresenta trabalhos existentes na literatura sobre a utilização de descritores de textura em imagens médicas para diagnosticar o glaucoma em retinografias. Sousa et

al. (2017) propuseram uma metodologia para diagnosticar o glaucoma utilizando funções geoestatísticas como: semivariograma, semimadograma, covariograma e correlograma. As imagens passaram por um pré-processamento sendo transformadas para nível de cinza e em seguidas equalizadas. Para a representação da região do disco óptico foi aplicado o Padrão Binário Local (LBP) e funções geoestatísticas para descrever os padrões de textura.

Al-Bander et al. (2017) apresentam uma abordagem para o diagnóstico automático de glaucoma usando redes neurais convolucionais (CNN). A metodologia consiste em três etapas: pré-processamento com redimensionamento de imagem, recursos de extração usando CNN e classificação usando *Support Vector Machine*. Foi utilizado um modelo CNN pré-treinado (AlexNet) com 23 camadas. Este método alcançou uma acurácia de 88,22%, especificidade de 90,8% e sensibilidade de 85%.

Araújo et al. (2018) propuseram uma metodologia para o diagnóstico de glaucoma em imagens de retinografia, usando como descritor os índices de diversidade. O método é dividido em pré-processamento; extração de características usando os índices de diversidade de Shannon e McIntosh; na terceira etapa o algoritmo genético é usado para seleção de recursos e para a classificação foi utilizado o SVM (*Support Vector Machine*). Os resultados alcançaram uma acurácia de 93,4%, uma sensibilidade de 92,8% e especificidade de 93,6%.

Semelhante ao citado anteriormente, neste trabalho é proposta uma metodologia para a utilização de índices ecológicos como descritores de textura, os índices taxonômicos. Os dois trabalhos se diferem na etapa de pré-processamento, onde Araújo et al. (2018) propõem a utilização de filtros e equalização de histograma; e na etapa de extração de características, onde este utiliza os índices de Shannon, de McIntosh e de Diversidade Total. No método proposto, o pré-processamento é feito utilizando redução de ruído, normalização de cor, CLAHE e Especificação de Histograma, é realizada a transformação de espaços de cores e são utilizados oito descritores (índices taxonômicos), diferentes dos utilizados por Araújo et al., além de utilizar uma base de imagens maior.

3. MATERIAIS E MÉTODOS

Esta seção apresenta as tecnologias, as bases de imagens e o método proposto neste estudo. Para a implementação das técnicas utilizadas no pré-processamento e extração de características, foi utilizada a linguagem de programação *Python*. A manipulação das imagens foi realizada através da biblioteca *OpenCV* [16], de código livre. Os classificadores foram estimados através da biblioteca *Auto Sklearn*, também de código livre.

O método proposto está estruturado em cinco etapas: pré-processamento, transformação em cinco espaços de cores, extração de características, classificação e validação de resultados. A Figura 1 apresenta as etapas do método proposto.

Os primeiros testes foram realizados apenas na base RIM-ONE versão 2. Conforme foram sendo observados os resultados, percebeu-se a necessidade de realizar a especificação do histograma na etapa de pré-processamento, com o objetivo de tentar corrigir as diferenças de contraste e iluminação que elas possuíam entre si por serem adquiridas de forma diferente.

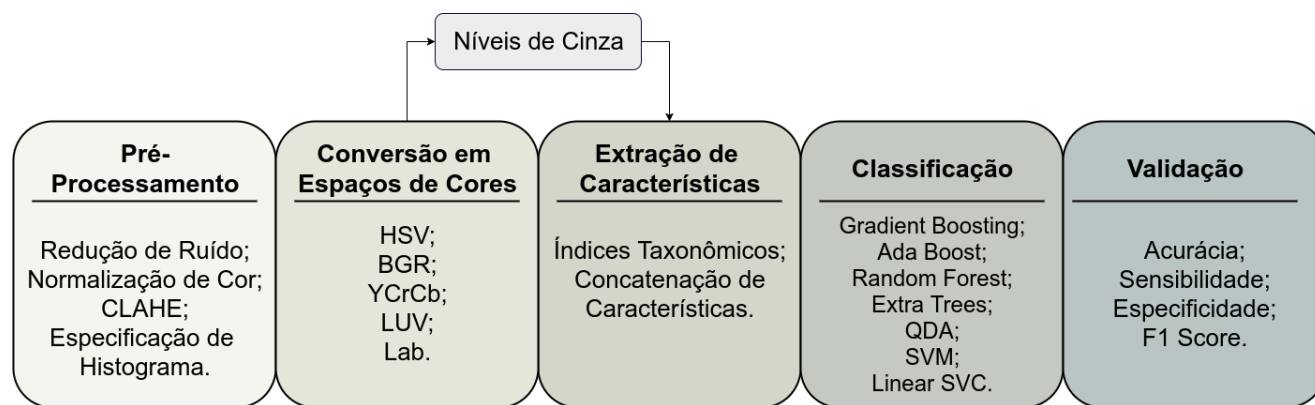


Figure 1: Etapas do Método Proposto.

3.1 Base de Imagens

Antes de realizar o estudo, são necessárias a definição e aquisição de bases de imagens de fundo de olho com pacientes com glaucoma e sem glaucoma. A primeira base utilizada neste trabalho foi a RIM-ONE [27]. A RIM-ONE é uma base pública e está disponível na internet. Foi utilizada a versão 2, que possui 455 imagens no formato RGB, e são separadas em duas classes, *Glaucoma and suspicious* (200 imagens) e *Normal* (255 imagens).

Além da base RIM-ONE, foram adquiridas mais duas bases: iChallenge-GON [31] e Kaggle1000fundusimages [17]. As duas bases são públicas e estão disponíveis. A base iChallenge-GON possui 400 imagens, sendo destas 360 sinalizadas como pacientes sem glaucoma e 40 glaucomatosos. A base Kaggle1000fundusimages é uma base que possui 1000 imagens de retinografia, estando 13 destas sinalizadas como pacientes com glaucoma e 54 sem glaucoma. A Figura 2 demonstra a diversidade de imagens encontradas nas bases adquiridas.

Portanto, para o desenvolvimento desta pesquisa, foram utilizadas um total de 922 imagens, sendo 669 delas de pacientes sem glaucoma e 253 glaucomatosos.

3.2 Pré-Processamento

A etapa de pré-processamento teve como objetivo melhorar a imagem e evidenciar características da região do disco óptico e da cavidade do disco óptico, para a fase de extração de características. Para tal foram utilizados alguns métodos: a Redução de Ruído por *Non-local Means*, Equalização de Histograma Adaptativa (CLAHE) e a Normalização de Cor.

O propósito principal da utilização do algoritmo *Non-Local Means* é de remover o ruído de uma imagem – que pode ter sido inserido durante a sua aquisição – a fim de recuperar a imagem original. Comumente, os algoritmos de remoção de ruído realizam esse procedimento de forma local, como sugerem o filtro Gaussiano e o filtro da Mediana. A principal diferença do algoritmo *Non-Local Means* em relação aos filtros locais ou aos filtros no domínio da frequência é o uso sistemático de todas as informações possíveis que a imagem pode oferecer [6].

Um método comum em processamento de imagens é a equalização do histograma por ser um método simples possuir baixo custo computacional. O CLAHE (*Contrast Limited Adaptive Histogram Equalization*) tem como objetivo principal melhorar a imagem de retinografia colorida usando

a manipulação de histograma para obter uma distribuição uniforme da intensidade [21]. Esse método de aprimoramento é muito utilizado na oftalmologia, como, por exemplo, em detecção de micro aneurismas. Esse método, por sua vez, funciona espalhando a distribuição de intensidade no histograma e ajustando a intensidade da imagem original.

A informação de cor na imagem também pode ser muito útil na etapa de classificação, justificando assim a utilização da técnica de normalização de cor [32]. Para ser usada como pré-processamento, no entanto, essa informação deve estar intrinsecamente relacionada às propriedades dos objetos captados, nesse caso, deve oferecer informações a respeito da saúde do paciente fotografado. Assim, a cor da imagem deve possuir correlação com a saúde do paciente apresentado nela.

Para as bases iChallenge-GON e Kaggle1000fundusimages, além do pré-processamento passado, antes da etapa de conversão em espaços de cores, foi aplicada também a especificação de histograma (*matching*), para que apresentassem características mais semelhantes às imagens da RIM ONE v2. Após a especificação de histograma, toda a metodologia é análoga à realizada apenas na etapa anterior.

3.3 Transformação de Espaços de Cores

Após o pré-processamento, as imagens são transformadas para outros espaços de cores, o HSV, LUV, Lab [14], YCrCb e RGB. Esses espaços de cores foram utilizados para avaliar se existem características relevantes para distinguir as imagens de pacientes glaucomatosos de pacientes saudáveis. O espaço de cor HSV, por exemplo, é um sistema de cores formado pela matiz, saturação e valor de intensidade (brilho) da imagem [29], enquanto o RGB é um sistema constituído das cores vermelho, azul e verde.

Primeiramente, foram realizados testes em cada espaço de cor separadamente. Observando-se os resultados, resolveu-se utilizar todos os espaços de cores para a geração de características. Dessa forma, para cada espaço de cor foram separados os canais gerando três imagens em níveis de cinza para cada espaço de cor. Assim, para cada espaço de cor, tem-se 3 níveis de cinza correspondentes.

3.4 Extração de Características

Após o pré-processamento, as imagens são submetidas a extração de características. Essas características são definidas como propriedades presentes em objetos registrados nas

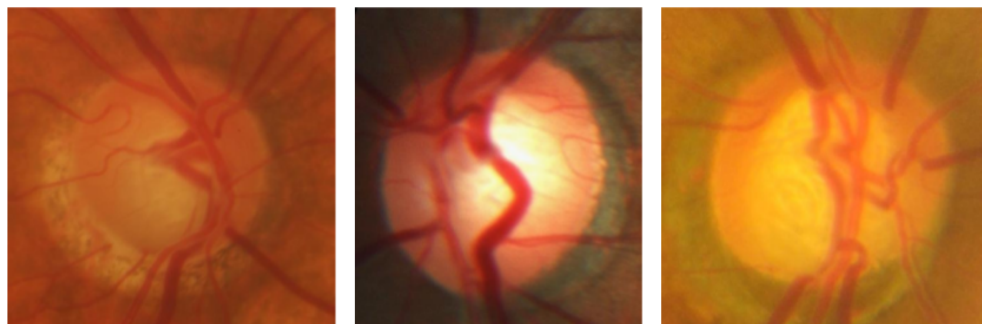


Figure 2: Exemplos de imagens das bases RIM-ONE v2 (esquerda), iChallenge-GON (meio) e Kaggle1000fundusimages (direita).

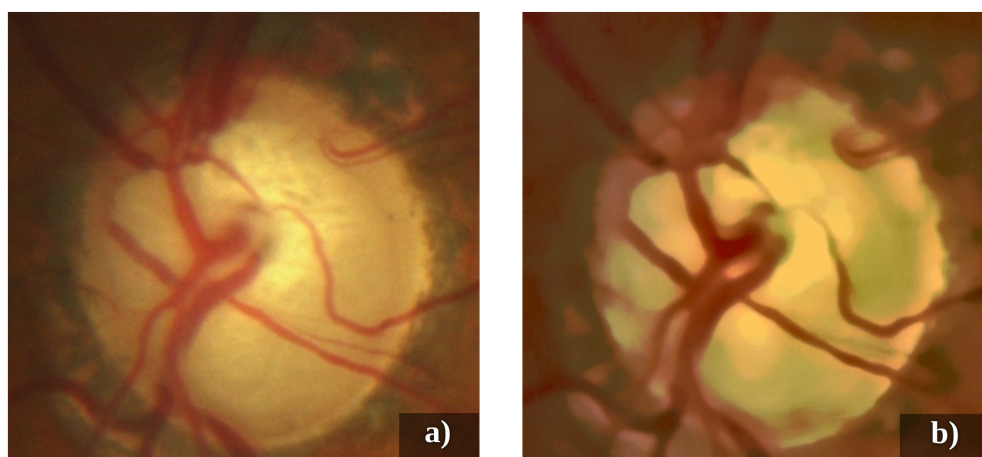


Figure 3: Exemplo de imagem da base iChallenge-GON antes e após a especificação de histograma.

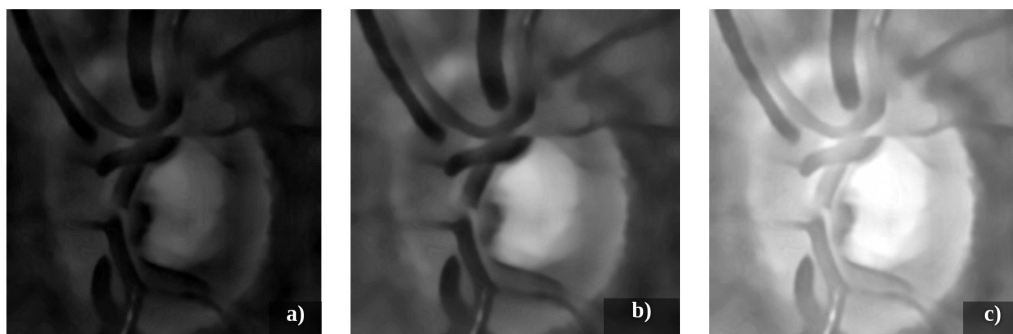


Figure 4: Exemplo de imagem da base RIM ONE v2 após a divisão em níveis de cinza para o espaço de cores BGR. (a) B - Azul; (b) G - Verde; (c) R - Vermelho

imagens, que podem ser usados como descritores da imagem [23]. Para este trabalho, o método de extração de características utiliza índices taxonômicos.

Os índices taxonômicos, ou índices de diversidade, são índices comumente utilizados na ecologia para descrever a variedade de espécies em uma determinada comunidade ou região. É possível medir a diversidade de duas formas: pela riqueza de espécies, representada pelo número de espécies encontradas em uma região, e pela abundância relativa, ou seja, o número de indivíduos de uma espécie específica em uma região.

Para que os índices de diversidade possam ser utilizados na análise de textura em imagens, é necessário fazer uma analogia entre o que eles representam na Ecologia e o que representam em Processamento de Imagens. Dessa forma, a comunidade é abstraída como uma imagem (ou parte dela), as espécies são os níveis de cinza encontrados na imagem e os indivíduos são os *pixels*. A Tabela 3.4 sumariza a correspondência entre os termos utilizados quando se trata de índices de diversidade taxonômica em relação ao termos em processamento de imagens.

A Figura 5 demonstra um exemplo contendo três espécies (níveis de cinza): preto (P), cinza (C) e branco (B). A quantidade de indivíduos (pixels) de P é 4, C é 2 e B é 3.



Figure 5: Exemplo de uma imagem contendo 3 espécies (níveis de cinza).

Uma representação genérica de uma árvore taxonômica e sua matriz de distâncias para a Figura 5 está representada a seguir.

A diversidade taxonômica é uma abordagem baseada no conjunto de distâncias entre pares de espécies a partir de árvores taxonômicas [20]. Para este artigo foi escolhida uma forma de representar a árvore taxonômica a partir de uma matriz. Assim, as distâncias entre uma espécie e outra é representada pela distância entre os pixels.

Oito índices foram utilizados neste trabalho, sendo estes:

- Índice de Diversidade Taxonômica:

Considera a abundância das espécies e a relação taxonômica entre elas [18].

$$\Delta = \frac{\sum \sum_{i < j} \omega_{ij} x_i x_j}{n(n-1)/2} \quad (1)$$

Onde x_i ($i, j = 1, \dots, n$) é a abundância da i -ésima espécie; n é o número total de espécies e ω é a distância da espécie i à espécie j .

- Índice de Distinção Taxonômica:

Representa a distância taxonômica média entre dois indivíduos [18].

$$\Delta^* = \frac{\sum \sum_{i < j} \omega_{ij} x_i x_j}{\sum \sum_{i < j} x_i x_j} \quad (2)$$

Onde x_i ($i, j = 1, \dots, s$) é a abundância da i -ésima espécie; ω é a distância da espécie i à espécie j e s é o número de espécies.

- Soma das Distâncias Filogenéticas:

Representa a soma das distâncias filogenéticas entre pares de espécies [28].

$$SPD = \left(\frac{s(s-1)}{2} \right) \left(\frac{\sum \sum_{m < n} d_{mn} a_m a_n}{\sum \sum_{m < n} a_m a_n} \right) \quad (3)$$

Onde d_{mn} ($m, n = 1, \dots, s$) é a distância da espécie m à espécie n ; a é a abundância da espécie e s é o número de espécies.

- Distância Média do Vizinho Mais Próximo:

Representa a distância média até o táxon mais próximo [28].

$$MNND = \sum_{m=1}^s \min(d_{mn}) a_m \quad (4)$$

Onde d_{mn} ($m, n = 1, \dots, s$) é a distância da espécie m à espécie n ; a é a abundância da espécie e s é o número de espécies.

- Entropia Quadrática Intensiva:

Representa o número de espécies e suas relações taxonômicas [1].

$$J = \frac{(\sum d_{ij})}{s^2} \quad (5)$$

Onde d_{ij} ($i, j = 1, \dots, s$) é a distância da espécie i à espécie j e s é o número de espécies.

- Entropia Quadrática Extensiva:

Representa a soma das distâncias entre as espécies [1].

$$F = \sum d_{ij} \quad (6)$$

Onde d_{ij} ($i, j = 1, \dots, s$) é a distância entre as espécies i e j ; e s é o número de espécies.

- Distinção Taxonômica Média:

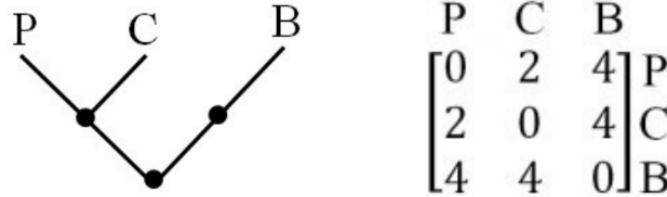
Representa a distância taxonômica média entre quaisquer duas espécies escolhidas aleatoriamente [1].

$$AvTD = \frac{\sum \sum_{i < j} d_{ij}}{s(s-1)/2} \quad (7)$$

Onde d_{ij} ($i, j = 1, \dots, s$) é a distância entre as espécies i e j ; e s é o número de espécies.

Table 1: Correspondência entre termos na biologia e os utilizados neste artigo. Adaptado de [22].

Biologia	Método Proposto
Comunidade	Imagem
Espécie	Níveis de Cinza
Riqueza de Espécie	Número de níveis de cinza em uma imagem (intensidades)
Indivíduos	Pixels (em 2D) ou Voxels (3D)
Abundância da Espécie	Número de níveis de cinza dada uma área

**Figure 6: Exemplo representando uma árvore taxonômica (à esquerda) e sua matriz de distâncias (à direita). Adaptado de [18].**

- Distinção Taxonômica Total: Representa a distinção média filogenética somada ao longo de todas as espécies [1].

$$TTD = \sum_i \frac{\sum_{i \neq j} d_{ij}}{s-1} \quad (8)$$

Onde d_{ij} ($i, j = 1, \dots, s$) é a distância entre as espécies i e j ; e s é o número de espécies.

Sendo os dois primeiros índices baseados na riqueza de espécie e os outros seis, na distância entre pares de espécies. É importante destacar que os índices se complementam, ou seja, um grupo pode ser capaz de complementar uma propriedade da imagem que outro talvez não possa.

Baseado na Figura 6, a aplicação dos índices é demonstrada abaixo:

$$\Delta = \frac{\omega_{pc}x_px_c + \omega_{pb}x_px_b + \omega_{bc}x_bx_c}{n(n-1)/2} = \frac{2.4.3 + 4.4.2 + 4.3.2}{3(3-1)/2} = 26,67$$

$$\Delta^* = \frac{\omega_{pc}x_px_c + \omega_{pb}x_px_b + \omega_{bc}x_bx_c}{x_px_c + x_px_b + x_bx_c} = \frac{2.4.3 + 4.4.2 + 4.3.2}{4.3 + 4.2 + 3.2} = 3,08$$

$$SPD = \left(\frac{s(s-1)}{2} \right) \left(\frac{d_{pc}x_px_c + d_{pb}a_pa_b + d_{bc}a_ba_c}{a_pa_c + a_pa_b + a_ba_c} \right) = \left(\frac{3(3-1)}{2} \right) \left(\frac{2.4.3 + 4.4.2 + 4.3.2}{4.3 + 4.2 + 3.2} \right) = 9,24$$

$$MNND = d_{pc}a_p + d_{cp}a_c + d_{pb}a_p + d_{bp}a_b + d_{cp}a_c + d_{bc}a_b + d_{cb}a_c = 2.4 + 2.2 + 4.3 + 4.4 + 4.3 + 4.2 = 60$$

$$J = \frac{(d_{pc} + d_{pb} + d_{bc})}{s^2} = \frac{(2 + 4 + 4)}{3^2} = 1,11$$

$$F = d_{pc} + d_{pb} + d_{bc} = 2 + 4 + 4 = 10$$

$$AvTD = \frac{d_{pc} + d_{pb} + d_{bc}}{s(s-1)/2} = \frac{2 + 4 + 4}{3(3-1)/2} = 3,33$$

$$TTD = (p + b + c) \frac{d_{pc} + d_{pb} + d_{bc}}{s-1} = (4 + 3 + 2) \frac{2 + 4 + 4}{3-1} = 45$$

Analogamente, todos os índices foram aplicados a cada imagem da RIM ONE v2 gerada no pré-processamento e a cada imagem da iChallenge-GON e Kaggle1000fundusimages após a especificação de histograma. Assim, para cada um dos 5 espaços de cores, tem-se 3 níveis de cinza, originando um total de 15 imagens resultantes da transformação do espaço de cor. Após essa etapa, para cada imagem gerada são aplicados todos os 8 índices, gerando um total de 120 características por imagem, que são passadas para a etapa de classificação.

3.5 Classificação

Esta etapa tem como objetivo realizar o reconhecimento das imagens classificando-as em saudáveis ou glaucomatosas. Após realização do pré-processamento e da extração de características, separou-se aleatoriamente 80% da base para treino e 20% para teste. Na classificação foi utilizada uma abordagem de aprendizado de máquina automatizada, baseada em *Auto Machine Learning* (Auto ML) [10] para obter o melhor classificador para as características dadas. A classificação foi realizada segundo rótulos atribuídas na base de imagens.

No *Auto Sklearn* [10], cada conjunto de treinamento foi treinado por cerca de 10 horas, o tempo usado tanto para

gerar um modelo (estimar classificadores) quanto para treiná-lo. Essa ferramenta oferece a funcionalidade de utilizar seleção e otimização de hiperparâmetros (110 ao todo) de 15 algoritmos de classificação com *Machine Learning* por meio de otimização Bayesiana, meta-aprendizado e construção de um ensemble[9].

É importante destacar que, para o desenvolvimento deste método, foram desabilitadas as configurações de pré-processamento de características da ferramenta, assim como foi configurado de forma que ela retorne apenas um classificador por conjunto de características.

Para a construção do método, foram utilizados sete classificadores, sendo estes: *Gradient Boosting* [11], *Ada Boost* [30], *Random Forest* [5], *Extra Trees* [12], *Quadratic Discriminant Analysis* (QDA) [4], *Support Vector Machine* (SVM) e *Linear Support Vector Classification* (Linear SVC) [15].

3.6 Validação dos Resultados

A validação dos resultados tem como objetivo medir o desempenho e promover a avaliação dos resultados obtidos. Neste trabalho foram utilizadas as estatísticas: Acurácia, Sensibilidade (SEN), Especificidade (ESP) e F1 Score.

Segundo [24] para a análise de amostra de determinada doença com caso positivo e negativo, os resultados dos testes de classificação dos casos analisados podem ser divididos em quatro grupos:

- VP : Verdadeiro Positivo - o teste é positivo e o paciente tem glaucoma.
- FP : Falso Positivo - o teste é positivo, mas o paciente não tem glaucoma.
- VN : Verdadeiro Negativo - o teste é negativo e o paciente não tem glaucoma.
- FN : Falso Negativo - o teste é negativo, mas o paciente tem glaucoma.

A Sensibilidade é o número de imagens corretamente classificadas como glaucomatosas.

$$SEN = \frac{VP}{VP + FN} \quad (9)$$

Especificidade é o número de imagens classificadas como saudáveis corretamente.

$$ESP = \frac{VN}{VN + FP} \quad (10)$$

Acurácia é a proporção total dos resultados corretos. Definida como a razão entre o número de casos na amostra em estudo que foram classificados corretamente e o número total de amostras.

$$ACU = \frac{VP + VN}{VP + FP + FN + VN} \quad (11)$$

Precisão é a proporção entre as imagens que foram classificadas como glaucomatosas corretamente e o total de imagens glaucomatosas.

$$PREC = \frac{VP}{VP + FP} \quad (12)$$

Recall é a proporção entre as imagens que foram classificadas como glaucomatosas corretamente e o total de imagens que foram classificadas como glaucomatosas.

$$REC = \frac{VP}{VP + FN} \quad (13)$$

F1 Score é a combinação entre a precisão e o recall de forma a trazer um número único que indique a qualidade geral do modelo gerado.

$$F1 = \frac{2 * PREC * REC}{PREC + REC} \quad (14)$$

4. RESULTADOS

Esta seção descreve os resultados obtidos nos experimentos realizados utilizando apenas a bases públicas de retinografia descritas nas Seção 3.1. Os resultados alcançados pelo método proposto foram avaliados pelas seguintes métricas: Acurácia (ACU), Sensibilidade (SEN), Especificidade (ESP) e F1-Score (F1). Para este artigo, escolheu-se a sensibilidade como o número de imagens corretamente classificadas como glaucomatosas e especificidade, como saudáveis.

Para validar o método proposto, inicialmente, foram utilizadas todas as 455 imagens da base RIM ONE v2. A Tabela 2 mostra os melhores resultados obtidos nos experimentos realizados com a base RIM ONE v2. Comparando com os resultados para cada espaço de cor, é possível notar que os índices taxonômicos apresentaram o melhor resultado quando foi feita a concatenação das características para todos os espaços de cores, alcançando acurácia de 98,90%, sensibilidade de 100% e especificidade de 98,25% com o classificador Extra Trees [12]. Considerando o F1-score, a sensibilidade e a especificidade, este também obteve melhores resultados que os demais experimentos.

A Tabela 3 mostra uma comparação dos resultados encontrados neste trabalho com os resultados dos trabalhos relacionados que também utilizam características de textura para diagnosticar o glaucoma usando imagens de retinografia. Todos os trabalhos utilizam a base RIM-ONE v2 com a mesma quantidade de imagens.

A partir da Tabela 3, observa-se que as métricas de acurácia igual a 98,9%, sensibilidade igual a 100% e a especificidade de 98,25% obtidas pelo método proposto foram superiores aos resultados obtidos pelos autores citados, indicando que a metodologia apresentada é promissora.

Após testes iniciais com as imagens da base RIM-ONE v2, foram realizados testes incluindo as imagens das bases iChallenge-GON e Kaggle1000fundusimages. A Tabela 4 apresenta os resultados para o método utilizando base gerada com 922 imagens.

Para a geração dos resultados na Tabela 4, foram extraídas características para todos os canais de cor e, ainda, a ferramenta *Auto Sklearn* foi utilizada mais de uma vez, justificando, assim, a quantidade de classificadores (*Quadratic Discriminant Analysis*, *Support Vector Machine* e *Linear Support Vector Classification*) encontrados para a mesma configuração da base. Observa-se, ainda, que a especificidade é sempre superior ao valor da sensibilidade, o que se deve ao fato de que as bases possuem mais imagens de pacientes saudáveis do que de glaucomatosos. O resultado geral de 92,97% de acerto é promissor e próximo ao resultado obtido nos trabalhos relacionados que utilizaram apenas a base RIM-ONE v2.

Na Tabela 4 também nota-se o impacto positivo da aplicação da especificação de histograma (*matching*). Comparando com o resultado obtido com o classificador Quadratic

Table 2: Resultado da classificação para a base RIM ONE v2.

CLASSIFICADOR	CANAL	ACU. (%)	SEN. (%)	ESP. (%)	F1
Gradient Boosting	HSV	81,32	85,00	78,43	0,80
Ada Boost	LUV	90,11	90,00	90,20	0,89
Gradient Boosting	Lab	83,52	77,50	88,24	0,81
Random Forest	YCrCb	86,81	90,00	84,31	0,86
Extra Trees	BGR	92,31	85,00	98,04	0,91
Extra Trees	Todos	98,90	100,00	98,25	0,99

Table 3: Comparação com os trabalhos relacionados.

Classificador	Técnicas	ACU. (%)	SEN. (%)	ESP. (%)
Sousa et al. [7]	Funções Geoestatísticas	91,20	95,00	88,20
Al-Bander et al. [2]	CNN	88,20	85,00	90,80
Araújo et al. [3]	Índices de diversidade	93,41	92,83	93,69
Método proposto	Índices Taxonômicos	98,90	100,00	98,25

Table 4: Resultado da classificação nas bases RIM-ONE v2, iChallenge-GON e Kaggle1000fundusimages.

CLASSIFICADOR	Matching	ACU. (%)	SEN. (%)	ESP. (%)	F1
Gradient Boosting	Não	90,81	76,47	96,26	0,82
Quadratic Discriminant Analysis	<i>Sim</i>	92,97	90,20	94,03	0,88
Support Vector Machine	Sim	85,94	72,55	91,04	0,74
Linear Support Vector Classification	Sim	91,35	78,43	96,27	0,83

Discriminant Analysis, nota-se, apesar de haver uma perda em especificidade (relativa à classe de pacientes saudáveis), o ganho em sensibilidade é maior, justificando o uso da especificação de histograma. Isso se dá porque, como dito anteriormente, as bases inicialmente eram muito distintas, logo, geravam características que não representavam a classe de pacientes glaucomatosos – classe com número inferior de amostras. Assim, para classificar esses pacientes, a escolha pode muitas vezes ter sido feita de forma aleatória. Isso pode ser observado no resultado gerado para o classificador *Gradient Boosting*. Nele, apesar de a especificidade ser alta, a sensibilidade é insuficiente para a proposta.

É possível melhorar este resultado fazendo um balanceamento das bases — fazendo com que a quantidade de imagens de pacientes glaucomatosos seja próxima a quantidade de imagens de pacientes saudáveis — e testando métodos de seleção de características.

É possível ver pela Figura 7 que as imagens, apesar de visualmente possuírem características de cor próximas, parecem dispor de texturas diferentes, tornando pertinente a sua classificação em classes distintas. O mesmo pode ser notado na Figura 8, onde nota-se que na imagem à esquerda, existem vasos mais espessos que na imagem da direita, estabelecendo assim características muito distintas para ambas as imagens.

5. CONCLUSÃO

A utilização de processamento de imagens está cada vez mais presente em várias áreas, como na área da saúde, oferecendo um meio de auxílio ao profissional da área. Este artigo apresentou uma abordagem capaz de fornecer um descritor que possa servir como meio para auxiliar os especialistas na elaboração do diagnóstico do glaucoma em imagens do fundo de olho.

Para o objetivo deste artigo, foram comparadas técnicas

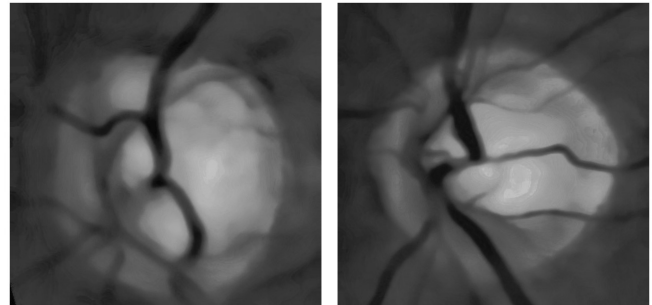


Figure 7: Retinografias classificadas como glaucomatosas. À esquerda, uma imagem corretamente classificada e à direita, incorretamente.

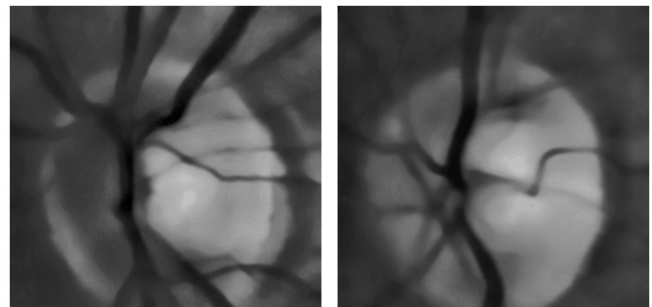


Figure 8: Retinografias classificadas como saudáveis. À esquerda, uma imagem corretamente classificada e à direita, incorretamente.

de extração de textura — a proposta e as já presentes na literatura —, visando o uso no diagnóstico de imagens glaucomatosas. Os resultados mostraram que os índices taxonômicos são descritores de textura promissores na análise de imagens glaucomatosas, alcançando uma taxa de sucesso de 92,97% nos experimentos realizados em uma base formada por 922 exames de retinografia.

Como trabalhos futuros, a fim de ampliar e melhorar os resultados, sugere-se:

1. Buscar e incorporar ao método outras bases de retinografia, como a Drishti-GS;
2. Usar outros tipos de pré-processamento para verificar se ainda existem formas de potencializar algumas características das imagens de forma a melhor diferir os pacientes saudáveis dos glaucomatosos;
3. Utilizar métodos de seleção de características, como o Select K Best, visando que algumas características que foram passadas pro classificador podem não ser tão relevantes para o modelo.

6. AGRADECIMENTOS

Esta pesquisa foi inteiramente financiada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), através do Programa Institucional de Bolsas de Iniciação Científica (PIBIC). Agradecemos também aos nossos colegas do Núcleo de Computação Aplicada (NCA) da Universidade Federal do Maranhão, por todo o suporte oferecido durante a realização desta pesquisa.

7. REFERENCES

- [1] *Métodos para sistemas CAD e CADx de nódulo pulmonar baseada em tomografia computadorizada usando análise de forma e textura*. PhD thesis, 2016. DEPARTAMENTO DE ENGENHARIA DA ELETRICIDADE/CCET.
- [2] B. Al-Bander, W. Al-Nuaimy, M. A. Al-Tae, and Y. Zheng. Automated glaucoma diagnosis using deep learning approach. In *Systems, Signals & Devices (SSD), 2017 14th International Multi-Conference on*, pages 207–210. IEEE, 2017.
- [3] J. D. L. Araújo, J. C. Souza, O. P. S. Neto, J. A. de Sousa, J. D. S. de Almeida, A. C. de Paiva, A. C. Silva, G. B. Junior, and M. Gattass. Glaucoma diagnosis in fundus eye images using diversity indexes. *Multimedia Tools and Applications*, pages 1–18, 2018.
- [4] S. Bose, A. Pal, R. SahaRay, and J. Nayak. Generalized quadratic discriminant analysis. *Pattern Recogn.*, 48(8):2676–2684, Aug. 2015.
- [5] L. Breiman. Random forests. *Mach. Learn.*, 45(1):5–32, Oct. 2001.
- [6] B. Coll and J.-M. Morel. A non-local algorithm for image denoising. volume 2, pages 60–65 vol. 2, 07 2005.
- [7] J. A. de Sousa, A. C. de Paiva, J. D. S. de Almeida, A. C. Silva, G. B. Junior, and M. Gattass. Texture based on geostatistic for glaucoma diagnosis from fundus eye image. *Multimedia Tools and Applications*, pages 1–18, 2017.
- [8] T. Dorion. *Manual de Exame Do Fundo de Olho*. Manolen, 2002.
- [9] M. Feurer, A. Klein, K. Eggenberger, J. Springenberg, M. Blum, and F. Hutter. Efficient and robust automated machine learning. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 2962–2970. Curran Associates, Inc., 2015.
- [10] M. Feurer, A. Klein, K. Eggenberger, J. T. Springenberg, M. Blum, and F. Hutter. Auto-sklearn: Efficient and robust automated machine learning. In *Automated Machine Learning*, pages 113–134. Springer, 2019.
- [11] J. Friedman. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29, 11 2000.
- [12] P. Geurts, D. Ernst, and L. Wehenkel. Extremely randomized trees, Apr 2006.
- [13] M. L. Giger and K. Suzuki. Computer-aided diagnosis. In *Biomedical information technology*, pages 359–XXII. Elsevier, 2008.
- [14] R. C. E. Gonzalez, S. L. Woods, R. E. R. E. R. C. Gonzalez, R. E. Woods, and S. L. Eddins. *Digital image processing using MATLAB*. Number 04; TA1637, G6. 2004.
- [15] M. A. Hearst. Support vector machines. *IEEE Intelligent Systems*, 13(4):18–28, July 1998.
- [16] INTEL. Opencv, open computer vision library, 2016. Disponível em: <<http://www.sourceforge.net/projects/opencvlibrary/>>.
- [17] Linchundan. 1000 fundus images with 39 categories, Jun 2019.
- [18] F. S. S. d. Oliveira. Classificação de tecidos da mama em massa e não-massa usando índice de diversidade taxonômico e máquina de vetores de suporte, 2013.
- [19] J. A. A. Ottaiano, M. P. de Ávila, C. C. Umbelino, and A. C. Taleb. As condições de saúde ocular no brasil. *Conselho Brasileiro de Oftalmologia*, 1, 2019.
- [20] C. Ricotta. A parametric diversity measure combining the relative abundances and taxonomic distinctiveness of species. *Diversity and Distributions*, 10:143 – 146, 02 2004.
- [21] A. Setiawan, T. Mengko, O. Santoso, and A. Sukmono. Color retinal image enhancement using clahe. pages 1–3, 06 2013.
- [22] G. Silva, A. Filho, A. Silva, A. Paiva, and M. Gattass. Taxonomic indexes for differentiating malignancy of lung nodules on ct images. *Research on Biomedical Engineering*, 32, 09 2016.
- [23] R. R. Silva, J. G. Lopes, and F. H. Araújo. Visão computacional em python utilizando as bi-bliotecas scikit-image e scikit-learn. pages 407–428, 2017.
- [24] J. A. d. Sousa. Diagnóstico de glaucoma em retinografias utilizando funções geoestatística, 2017. Dissertação (Mestrado) - Universidade Federal do Maranhão, Curso de Pós Graduação em Ciência da Computação, São Luís, MA.
- [25] G. Spaeth. Appearances of the optic disc in glaucoma: a pathogenetic classification. In *Trans Acad Ophthalmol Symposium on Glaucoma. CV Mosby, St. Louis*, 1981.
- [26] A. Taleb, M. A. R. d. Faria, M. Ávila, and P. Mello.

- As condições de saúde ocular no brasil–2012. *Conselho Brasileiro de Oftalmologia. São Paulo*, 2012.
- [27] E. Trucco, A. Ruggeri, T. Karnowski, L. Giancardo, E. Chaum, J. P. Hubschman, B. Al-Diri, C. Y. Cheung, D. Wong, M. Abramoff, et al. Validating retinal fundus image analysis algorithms: Issues and a proposal validating retinal fundus image analysis algorithms. *Investigative ophthalmology & visual science*, 54(5):3546–3559, 2013.
- [28] M. Vellend, W. Cornwell, K. Magnuson-Ford, and A. Mooers. Measuring phylogenetic biodiversity. *Magurran McGill*, pages 194–207, 01 2011.
- [29] J. L. Vilaça. Manual processamento de imagem., Jan 2014.
- [30] R. Wang. Adaboost for feature selection, classification and its relation with svm, a review. *Physics Procedia*, 25:800–807, 12 2012.
- [31] Y. Xu, J. I. Orlando, H. Bogunovic, H. Fu, and X. a. F. Zhang. Refuge: Retinal fundus glaucoma challenge., 2018.
- [32] A. Youssif, A. Ghalwash, and A. Ghoneim. A comparative evaluation of preprocessing methods for automatic detection of retinal anatomy. 03 2007.