

Glaucoma Diagnosis Using Texture Attributes and Pre-trained CNN's

Diagnóstico de Glaucoma Utilizando Atributos de Textura e CNN's Pré-treinadas

Maíla de Lima Claro^{1*}, Rodrigo de Melo Souza Veras¹, André Macedo Santana¹, Luis Henrique Silva Vogado¹ and Leonardo Pereira de Sousa²

Resumo: Glaucoma is a disease that damages the optic nerve. It is considered the second leading cause of blindness in the world. Several automatic diagnostic systems have been proposed. However, such systems have not been shown to be able to handle a great diversity of images. Thus, this work proposes a method of detecting glaucoma, through the use of texture descriptors and Convolutional Neural Networks (CNNs). Tests were conducted in four public databases, making a total of 873 images. The results showed that the junction of GLCM and pre-trained CNN descriptors and the use of the Random Forest classifier are promising in the detection of this pathology, obtaining an accuracy of 91.06%.

Keywords: medical images — glaucoma diagnosis — texture features — transfer learning

Resumo: Glaucoma é uma doença que danifica o nervo óptico. Ela é considerada a segunda principal causa de cegueira no mundo. Vários sistemas de diagnóstico automático têm sido propostos, contudo, tais sistemas não demonstraram ser capazes de lidar com uma grande diversidade de imagens. Dessa forma, este trabalho propõe um método de detecção do glaucoma, através do uso de descritores de textura e Redes Neurais Convolucionais (CNNs). Testes foram conduzidos em quatro bases públicas, o que constitui um total de 873 imagens. Os resultados mostraram que a junção dos descritores GLCM e CNNs pré-treinadas e a utilização do classificador Random Forest são promissoras na detecção dessa patologia, obtendo uma acurácia de 91,06%..

Palavras-Chave: imagens médicas — diagnóstico de glaucoma — atributos de textura — transferência de aprendizagem

¹ Computer Department, Federal University of Piauí, Brazil

² Information Systems Course, Federal University of Piauí, Brazil

*Corresponding author: claromaila@gmail.com

DOI: <http://dx.doi.org/10.22456/2175-2745.76387> • Received: 19/07/2017 • Accepted: 12/02/2018

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introdução

Diversas pesquisas com utilização de abordagens computacionais têm se encaminhado para auxiliar o diagnóstico de doenças. Esses estudos objetivam apoiar o médico especialista na identificação de patologias, podendo prevenir sua evolução, desde que diagnosticada precocemente.

No campo oftalmológico, foram desenvolvidos um grande número de sistemas de Diagnóstico Auxiliado por Computador (CAD - *Computer Aided Diagnosis*) para o auxílio na detecção de vários tipos de doenças oculares, entre elas o glaucoma [1]. 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.

Nos estágios iniciais o glaucoma tende a ser assintomático. A descoberta precoce desta patologia é imprescindível, pois

existem tratamentos que evitam a sua progressão e, consequentemente, a perda da visão do paciente. De acordo com a Organização Mundial de Saúde, 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 [2].

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 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 [3], como pode ser visto na Figura 1. A região do DO (representada por toda a área dentro da marcação branca), e a região da escavação (toda a área

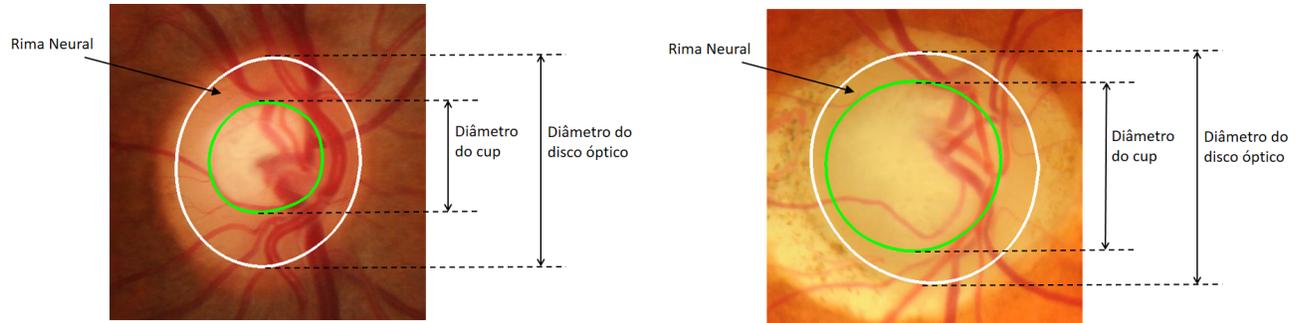


Figura 1. Exemplos de imagens da base RIM-ONE-3 com DO (marcação branca) e *cup* (marcação verde) em evidência: (a) Imagem de uma retina saudável e (b) Imagem de uma retina glaucomatosa.

dentro da marcação verde) estão em evidência. A Figura 1(a) apresenta uma imagem de retina saudável, enquanto que a Figura 1(b) apresenta uma imagem com glaucoma.

Existem vários algoritmos na literatura para o processamento e análise de imagens da retina visando o diagnóstico automático do glaucoma. Contudo, diferentes grupos de pesquisadores tendem a utilizar abordagens diferenciadas para que estes obtenham um melhor desempenho. O objetivo deste artigo é realizar a detecção automática do glaucoma em imagens da retina. Com o objetivo de desenvolver o nosso método foi realizado diversos experimentos que utilizam descritores de textura, forma e redes neurais convolucionais (*Convolutional Neural Networks* - CNNs) como atributos de entrada para os classificadores *Random Forest* (RF), *Support Vector Machine* (SVM) e *MultiLayer Perceptron* (MLP).

O presente trabalho tem a seguinte organização: Seção 2 apresenta os trabalhos relacionados referentes à identificação do glaucoma; Seção 3 descreve as bases de imagens utilizadas, os descritores utilizados neste trabalho e a apresentação da seleção de atributos e dos classificadores utilizados na detecção do glaucoma e as métricas de avaliação; Seção 4 demonstra os resultados obtidos e a Seção 5 discute estes resultados e por fim, a Seção 6 expõe a conclusão deste trabalho.

2. Trabalhos Relacionados

Realizou-se uma revisão na literatura objetivando identificar quais as principais técnicas estão sendo utilizadas no diagnóstico do glaucoma em imagens. Os trabalhos de Archarya et al. [4], Simonthomas et al. [5], Salam et al. [6] e Srinivasan et al. [7] apresentaram soluções para a mesma problemática. A Acurácia obtida nos quatro trabalhos foi superior a 91% com a utilização de descritores de textura. Sendo a *Gray Level Co-occurrence Matrix* (GLCM), um ponto em comum em todos os trabalhos. No trabalho de Salam et al. [6], os autores incluíram os descritores *Color Moments* e *Autocorrelogram* na etapa de extração de características, o que proporcionou 100% de sensibilidade do método proposto. Entretanto, a utilização de bases de dados privadas dificulta a recriação de experimentos e conseqüentemente possíveis comparações.

Entretanto os trabalhos de Kotyk et al. [8], Claro et al. [9] e Orlando et al. [10], os autores utilizaram bases públicas de fácil acesso.

Em Kotyk et al. [8], 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. Na 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 utilizados no classificador *MultiLayer Perceptron* (MLP) para determinar a exatidão do sistema. A acurácia do sistema proposto foi da ordem de 86,43%.

Claro et al. [9] 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 MLP.

Orlando et al. [10] apresentaram resultados de uma análise da utilização das CNNs pré-treinadas, a partir de imagens da retina para detecção automática 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).

3. Materiais e Métodos

O objetivo deste trabalho é realizar o diagnóstico do glaucoma em imagens da retina. Para isso aplicou-se na região do disco óptico experimentos com os descritores de textura e

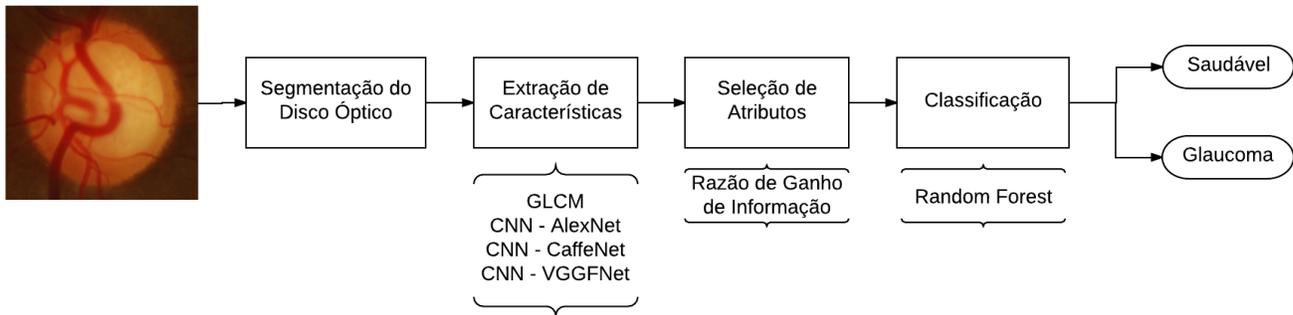


Figura 2. Fluxograma da metodologia proposta para diagnóstico automático do glaucoma.

forma, bem como *Convolutional Neural Networks* (CNNs). Nós realizamos uma avaliação do potencial de descrição, em separado, de cada um destes descritores e concluímos que a união de diferentes características leva a melhores taxas de classificação.

3.1 Método Proposto

O método proposto segue o fluxograma da Figura 2. Esta metodologia parte de uma imagem segmentada do DO. Os descritores utilizados na etapa de extração de características foram a concatenação da GLCM com a penúltima camada totalmente conectada das CNN-AlexNet [11], CNN-CaffeNet [12] e CNN-VGGFNet [13]. Na etapa de seleção de atributos é utilizado o algoritmo de razão de ganho de informação, com o intuito de ranquear as características mais significativas para as menos significativas. Por fim é aplicado o classificador *Random Forest* para informar se a imagem é saudável ou glaucomatosa.

3.2 Extração de Características

A extração de características tem por finalidade descrever imagens em função de atributos extraídos. Segundo Salam et al. [6] a identificação de imagens glaucomatosas pode ser eficientemente realizada utilizando os descritores de textura e forma. Neste estudo foram avaliados os descritores de textura *Local Binary Pattern* (LBP) [14]; *Gray Level Co-occurrence Matrix* (GLCM) [15]; *Histogram of Oriented Gradients* (HOG) [16]; Tamura [17]; *Gray Level Run Length Matrix* (GLRLM) [18], [19], [20] e descritores de forma. Adicionamos em nossa análise a descrição com CNNs através de três arquiteturas propostas recentemente: *AlexNet* [11], *CaffeNet* [12] e *VGGFNet* [13]. Estas arquiteturas são utilizadas em diversas aplicações, tais como a segmentação, descrição e classificação de imagens.

Todos os descritores de textura foram extraídos em cinco canais de cores diferentes, *Red*, *Green* e *Blue* do modelo RGB e os canais *Hue* e *Saturation* do modelo HSV. A banda *V* do modelo HSV não é utilizada, visto que, nas imagens retinianas esta banda de cor é quase a mesma que o canal de cor vermelho. Já para as CNNs, utilizamos como entrada a imagem original no sistema de cor RGB sem redimensionamento. A Tabela 1 apresenta um resumo dos descritores.

As CNNs utilizadas neste trabalho são pré-treinadas na base natural de imagens ImageNet e estão disponíveis no *website* MatConvNet¹. Devido a pequena quantidade de imagens obtidas neste trabalho, nós utilizamos a técnica de transferência de aprendizado, que consiste na extração da penúltima camada totalmente conectada da rede como vetor de características. A utilização de CNNs pré-treinadas na ImageNet proporciona a extração de características gerais da imagem, sendo elas: borda, cor, textura e contorno. Com o grande nível de abstração dos dados nas últimas camadas da CNN, não é possível identificar cada uma das características e sua origem a partir da imagem. Portanto, utilizamos a seleção dos atributos que possuem maior ganho de informação para assim treinar e validar o classificador.

3.3 Seleção de Atributos e Classificação

Após a criação do vetor de características, obtidos na etapa anterior, realizou-se uma seleção de atributos. Esta 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.

O algoritmo utilizado para seleção de atributos foi a razão de ganho, que é uma métrica de ganho de informação. Este algoritmo é um tipo de filtro, que tende a superestimar a qualidade dos atributos com muitos valores [21]. Neste contexto, realizou-se 10 testes, variando o tamanho do vetor de características em 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% e 90% do total de atributos, para que possamos excluir os atributos menos significativos.

O software utilizado para classificação foi o *Waikato Environment for Knowledge Analysis* (WEKA). Esta ferramenta oferece um conjunto de algoritmos de aprendizado de máquina para mineração de dados, nele são disponibilizados algoritmos de classificação, regressão, agrupamento, seleção de características e regras de associação [22]. Os algoritmos utilizados nos experimentos desse trabalho foram o *Support Vector Machine* (SVM) [23], *MultiLayer Perceptron* (MLP) [24] e *Random Forest* (RF) [25].

¹<http://www.vlfeat.org/matconvnet/>

Tabela 1. Informação dos descritores utilizado neste trabalho.

Descritores	Características	Qt de Atributos
LBP [14]	Histograma de textura	1280
GLCM [15]	Contraste, correlação, energia, homogeneidade (0°, 45°, 90° e 135°) e entropia	85
HOG [16]	Histograma de Gradientes Orientados	405
Tamura [17]	Contraste, Direcionalidade e <i>Coarseness</i>	15
GLRLM [18], [19] e [20]	<i>Short-Run Emphasis, Long-Run Emphasis, Gray Level distribution, Run Length distribution, Run Percentage, Low Gray Level Run Emphasis</i> textitHigh Gray Level Run Emphasis, <i>Short Run Low Gray Level Emphasis, Short Run High Gray Level Emphasis, Long Run Low Gray Level Emphasis e Long Run High Gray Level Emphasis</i> (0°, 45°, 90° e 135°)	220
Morfologia	Area, Diâmetro, Perímetro, Circularidade e Extensão	5
CNNs [11], [12] e [13]	Redes Neurais Convolucionais	4096

Foram realizados vários testes para identificar a melhor configuração de cada um dos classificadores avaliados. O SVM foi executado utilizando uma função de *kernel* linear. O classificador MLP, foi utilizado com uma camada oculta com 2 neurônios. O classificador *Random Forest* foi utilizado com uma profundidade máxima das árvores ilimitada e o número de árvores foi 100.

3.4 Métricas de Avaliação

Para a análise dos resultados de uma classificação, foram calculados os valores da matriz de confusão, sendo: Verdadeiro Positivo (VP), número de imagens corretamente classificadas como glaucomatosas; Falso Negativo (FN) é o número de imagens glaucomatosas classificadas incorretamente como saudáveis; Falso Positivo (FP) é o número de imagens saudáveis classificadas incorretamente como glaucomatosas; e Verdadeiro Negativo (VN), número de imagens corretamente classificadas como saudáveis [26]. A partir desses valores, são calculadas as métricas de desempenho da metodologia apresentada neste estudo, sendo elas as taxas de Precisão (P), Área Sob a Curva ROC (AUC) e Acurácia (A) [27].

Outra medida utilizada foi o índice Kappa (K), que vem sendo recomendado como uma medida apropriada da exatidão por representar inteiramente a matriz de confusão. Ele toma todos os elementos da matriz em consideração, ao invés de apenas aqueles que se situam na diagonal principal, o que ocorre quando se calcula a exatidão global da classificação [28]. De acordo com Landis e Koch [29] o valor de K assume valores entre 0 e 1. O resultado é qualificado de acordo com o valor de K da seguinte forma: $K \leq 0,2$: 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.

3.5 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. Assim, avaliamos a robustez da metodologia proposta em quatro bases de dados diferentes: DRISHTI [30] e RIM-ONE [31] em suas 3 versões. Ao todo, nossa proposta foi avaliada em 873 imagens, sendo 489 saudáveis e 384 glaucomatosas.

A base DRISHTI é composta por 101 imagens de retina coletadas no *Aravind Eye Hospital*, Madurai, na Índia. Todas as imagens foram tiradas com centro no DO, com um “campo de visão” de 30° e dimensões 2896×1944 pixels em formato PNG.

As imagens das bases RIM-ONE 1, 2 e 3 foram registradas em três diferentes regiões da Espanha. A RIM-ONE-1 é composta de 158 imagens de retina. O conjunto está dividido em 118 saudáveis, 12 com glaucoma inicial, 14 com glaucoma moderado e 14 com glaucoma profundo. As imagens foram capturadas por uma câmera Nidek AFC-210 com um corpo de uma Canon EOS 5D Mark II de 21,1 megapixels. A base foi criada com a colaboração de quatro oftalmologistas e um optometrista.

A RIM-ONE-2 possui um conjunto de 455 imagens de retina, sendo 255 imagens saudáveis e 200 glaucomatosas. Todas as imagens possuem o padrão ouro da segmentação da região do disco óptico. As imagens variam de 345×386 à 828×818 pixels.

A RIM-ONE-3 é um conjunto de dados que consiste em 159 imagens de retina, sendo que, 85 imagens saudáveis e 74 glaucomatosas. Todas as imagens foram tiradas 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 diagnóstico foi realizado por dois especialistas.

A Figura 3 apresenta exemplos de imagens disponíveis nas bases de dados utilizadas neste trabalho. A Figura 3(a) mostra um exemplo de imagem da base DRISHTI, com a sua respectiva marcação do DO demonstrada na Figura 3(b), ou comumente chamada de padrão ouro. A Figura 3(c) apresenta um exemplo de imagem da base RIM-ONE-1, e seu padrão ouro é demonstrado na 3(d).

4. Resultados

O Software utilizado para classificação foi o WEKA [22] e, o método de validação utilizado foi o *k-fold cross-validation* (com $k = 10$). A Tabela 2 demonstra os resultados obtidos, onde em cada descritor foi aplicado nas 873 imagens do

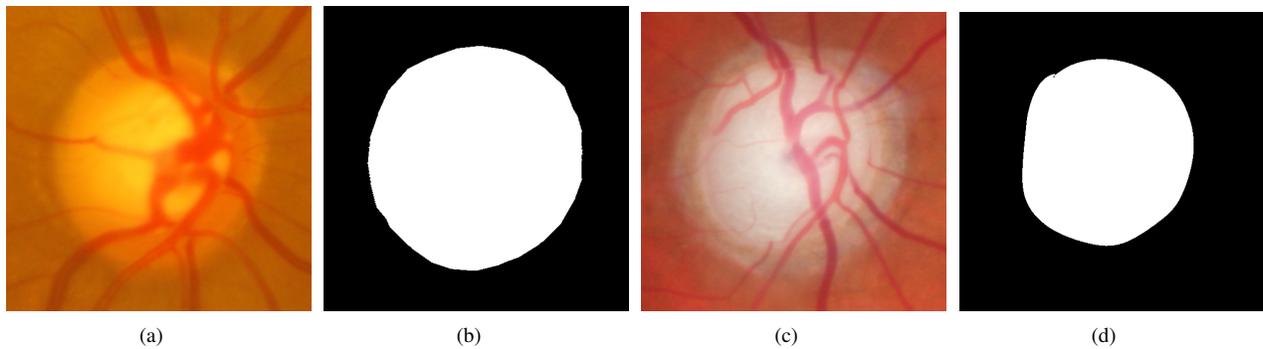


Figura 3. Exemplos de imagens das bases de dados: (a) DRISHTI, (b) Padrão Ouro DRISHTI, (c) RIM-ONE-1 e (d) Padrão ouro RIM-ONE-1.

fundo de olho disponíveis nas quatro bases de dados utilizadas. Na última linha desta tabela encontra-se o resultado da concatenação de todos os descritores testados (Híbrido).

A partir da análise da Tabela 2, podemos perceber que, em todos os três classificadores testados, GLCM e as CNNs obtiveram uma taxa de acerto maior em relação aos outros descritores. A GLCM obteve, no classificador MLP e *Random Forest*, o melhor resultado individual entres os descritores, com 82,93% e 83,27% de acurácia, respectivamente. No classificador SVM o melhor resultado encontrado foi com a CNN-AlexNet que obteve 82,47% de acurácia.

Pode-se observar também que com a concatenação de todos os descritores formou-se um descritor híbrido, onde foi possível obter uma melhoria em todas as métricas de avaliação nos classificadores MLP e *Random Forest*, com acurácia de 83,16% e 85,56%, respectivamente. O Kappa obtido demonstra que os resultados são considerados “Muito Bons” e que de acordo com o AUC alcançado, seu resultado é considerado satisfatório, pois possui um valor acima de 0,7.

Assim como no trabalho de Acharya et al., 2011 [4], este estudo teve o *Random Forest* e o MLP como os classificadores que apresentaram os melhores resultados na maioria dos descritores. Assim, com base nestes experimentos, podemos concluir que uma análise individual de todos estes descritores, a GLCM e as arquiteturas de CNNs obtiveram os melhores resultados. Com isso, investigamos o uso de um novo descritor composto por a concatenação estes quatro descritores: GLCM, CNN-AlexNet, CNN-CaffeNet e CNN-VGGNet.

Com base nesta análise realizou-se 10 testes, onde a seleção de atributos é utilizada com o tamanho do vetor de características variando em 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 total foram feitos 30 testes (10 testes de porcentagens de seleção de atributos \times 3 descritores) utilizando o classificador *Random Forest*, e seus resultados podem ser visualizados na Figura 4(a). Nesta figura também é demonstrado a quantidade de atributos pertencentes a cada descritor composto no descritor GLCM+CNNs (Figura 4(b)), e podemos inferir que o descritor GLCM, que possui 85 atributos extraídos, obteve uma taxa

significativa nos atributos mais importantes.

A Tabela 3 demonstra a porcentagem que apresentou o melhor resultado nos descritores GLCM, nas CNNs (CNN-AlexNet, CNN-CaffeNet e CNN-VGGNet) e com a junção da GLCM com as CNNs. Podemos visualizar que no descritor GLCM, a melhor quantidade de atributos foi de 90%, o que corresponde a 77 atributos, onde se obteve uma acurácia de 84,19%. As taxas de precisão individualmente descritas dos descritores da GLCM e CNNs variaram de 81,40% a 84,20%. Esses descritores conseguiram resultados considerados “muito bons” de acordo com a Seção 3.4, que demonstra o nível de exatidão da classificação a partir do índice Kappa. No entanto, ao concatená-los o resultado melhorou significativamente, atingindo uma precisão de 91,06% com apenas 5% dos atributos mais significativos, o que corresponde a 619 atributos de 12.373. Contudo, com a concatenação destes, o resultado melhorou significativamente, alcançando uma acurácia na ordem de 91,06% com apenas 5% dos melhores atributos, o que equivale a 619 atributos de 12.373. Como o índice Kappa (K) foi de 0,81, este resultado é considerado “Excelente” e demonstra a eficiência da união de características de textura com CNNs pré-treinadas na classificação de imagens glaucomatosas.

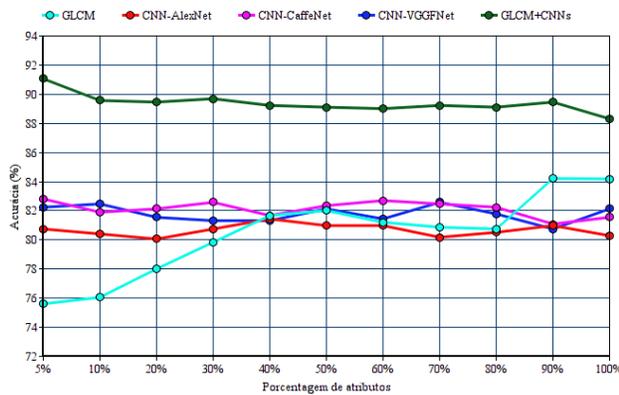
5. Discussão

Com a finalidade de validar o método proposto, foram realizadas comparações com métodos do estado da arte. A Tabela 4 apresenta uma comparação entre o método proposto e as técnicas de Kotyk et al. [8] e Claro et al. [9]. 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 91,06% e Kappa de 0,81 o que é um resultado considerado “Excelente”, já os outros dois trabalhos alcançaram índice Kappa considerado “Muito Bom”.

A Tabela 5 apresenta a comparação do método desenvolvido neste estudo com o do trabalho de Orlando et al. [10]. 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

Tabela 2. Resultados obtidos com todos os atributos.

Descritor	SVM				MLP				Random Forest			
	P(%)	AUC	A(%)	K	P(%)	AUC	A(%)	K	P(%)	AUC	A(%)	K
LBP	34,10	0,5	58,41	0	69,00	0,8087	68,61	0,36	71,50	0,8255	71,70	0,41
GLCM	81,20	0,8088	80,52	0,58	82,90	0,929	82,93	0,64	83,30	0,9353	83,27	0,65
HOG	67,00	0,7774	67,46	0,30	71,00	0,8107	71,02	0,40	69,40	0,8464	69,53	0,34
Tamura	63,17	0,6851	64,14	0,21	71,40	0,7887	71,59	0,41	69,40	0,7907	69,76	0,36
GLRLM	64,80	0,8034	65,06	0,27	65,30	0,6272	65,75	0,26	74,70	0,8392	74,80	0,47
Morfologia	53,80	0,5298	52,11	0,05	62,60	0,6974	63,00	0,22	66,20	0,7219	66,32	0,31
CNN-AlexNet	82,50	0,8204	82,47	0,64	80,60	0,8798	80,52	0,60	80,30	0,8934	80,29	0,59
CNN-CaffeNet	80,80	0,8024	80,87	0,60	80,70	0,8868	80,75	0,60	81,60	0,9016	81,55	0,62
CNN-VGGFNet	80,70	0,8022	80,50	0,60	79,90	0,881	79,95	0,59	82,20	0,8943	82,12	0,63
Híbrido	67,70	0,7923	66,20	0,33	83,10	0,9334	83,16	0,65	85,60	0,9382	85,56	0,70



(a)

(%) dos dados	GLCM	CNN-AlexNet	CNN-CaffeNet	CNN-VGGNet	Total
5	63	171	195	190	619
10	66	361	391	419	1237
20	69	738	811	857	2475
30	82	1083	1243	1304	3712
40	82	2025	1243	1599	4949
50	82	2808	1365	1931	6186
60	85	3666	1366	2307	7424
70	85	4096	1856	2624	8661
80	85	4096	2342	3375	9898
90	85	4096	3192	3762	11135
100	85	4096	4096	4096	12373

(b)

Figura 4. Comparação dos resultados dos descritores GLCM, CNNs e GLCM+CNNs com a seleção de porcentagem dos atributos: (a) Representação dos resultados obtidos e (b) Composição do descritor GLCM+CNNs

Tabela 3. Melhores resultados obtidos com a seleção de atributos.

Descritor	(%) dos Dados	Qt de Atributos	P(%)	AUC	A(%)	K
GLCM	90	77	84,20	0,9492	84,19	0,67
CNN-AlexNet	40	1638	81,40	0,89	81,44	0,62
CNN-CaffeNet	5	205	82,80	0,90	82,81	0,65
CNN-VGGFNet	70	2867	82,60	0,89	82,59	0,64
GLCM+CNNs	5	619	91,20	0,9633	91,06	0,81

Tabela 4. Comparação do método proposto com os trabalhos relacionados.

Trabalhos	P(%)	AUC	A(%)	K
Kotyk et al. [8]	86,50	0,952	86,48	0,72
Claro et al. [9]	87,00	0,955	87,05	0,73
Método Proposto	91,20	0,963	91,06	0,81

base DRISHTI [30] (a mesma utilizada em [10]). 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 5. Comparação do método proposto com os trabalhos relacionados.

Trabalhos	P(%)	AUC	A(%)	K
Orlando et al. [10]	-	0,762	-	-
Método Proposto	80,60	0,814	80,19	0,46

Os resultados apresentados nas Tabelas 4 e 5 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 [8] e [9] e apenas em transferência de aprendizado [10].

6. Conclusão

Muitos trabalhos identificados na literatura, demonstram que as pesquisas têm se direcionado para a área de processamento digital de imagens, na qual o diagnóstico automático do glaucoma é uma destas subáreas. Neste estudo, foi exposto uma estratégia para o diagnóstico automático do glaucoma, que utiliza descritores de textura, forma e redes neurais convolucionais nas imagens digitais de fundo de olho, para que estas sejam classificadas como glaucomatosas ou não.

De acordo com os achados na literatura, percebeu-se que os autores costumam utilizar suas próprias bases de dados ou apenas uma base de dados. Ao se utilizar 4 bases distintas nota-se que cada uma delas apresentam características específicas. A extração de características e a classificação são considerados as etapas mais importantes no diagnóstico automático de diferentes sistemas computacionais. Em razão disto foram utilizados os descritores de textura, forma e redes neurais convolucionais, além de um algoritmo de seleção de atributos para auxílio no desempenho do método. Com a concatenação das CNNs com a GLCM obteve-se um desempenho considerado “Excelente” pelo índice kappa, tendo uma acurácia de 91,06%.

Os resultados apresentados são promissores. Entretanto, podem ser melhorados. Como trabalhos futuros, propõe-se a extração de características na região da escavação (*cup*), além da região do disco óptico, pois, com 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).

Agradecimentos

Os autores agradecem a Fundação de Apoio à Pesquisa do Estado do Piauí (FAPEPI) por patrocinar nossa pesquisa.

Contribuição dos Autores

A autora Maíla Claro é a principal responsável pela execução do trabalho e escrita do artigo. Os professores Rodrigo Veras e André Santana são, respectivamente, orientador e co-orientador da autora principal. Eles orientaram ela e os demais alunos na execução dos testes, escrita e contribuíram com sua

experiência na versão final do artigo. Os autores Luis Vogado e Leonardo Sousa auxiliaram a autora principal na execução de testes, escrita inicial do artigo e revisão da versão final.

Referências

- [1] 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*, p. 1–17, 2016.
- [2] QUIGLEY, H. A.; BROMAN, A. T. The number of people with glaucoma worldwide in 2010 and 2020. *British Journal of Ophthalmology*, v. 90, n. 3, p. 262–267, 2006.
- [3] MITTAPALLI, P. S.; KANDE, G. B. Segmentation of optic disk and optic cup from digital fundus images for the assessment of glaucoma. *Biomedical Signal Processing and Control*, v. 24, p. 34–46, 2016.
- [4] ACHARYA, U. R. et al. Automated diagnosis of glaucoma using texture and higher order spectra features. *IEEE transactions on Information Technology in Biomedicine*, v. 15, n. 3, p. 449–455, 2011.
- [5] SIMONTHOMAS, S.; THULASI, N.; ASHARAF, P. Automated diagnosis of glaucoma using haralick texture features. In: *IEEE International Conference on Information Communication and Embedded Systems*. [S.l.: s.n.], 2014. p. 1–6.
- [6] SALAM, A. A. et al. Automated detection of glaucoma using structural and non structural features. *SpringerPlus*, v. 5, n. 1, p. 1519–1540, 2016.
- [7] SRINIVASAN, C.; DUBEY, S.; GANESHBABU, T. Complex texture features for glaucomatous image classification system using fundus images. v. 2, n. 13, 2016.
- [8] KOTYK, T. et al. Semi-automated system for cup to disc measurement for diagnosing glaucoma using classification paradigm. In: *Proceedings of the Second International Afro-European Conference for Industrial Advancement*. [S.l.: s.n.], 2016. p. 653–663.
- [9] CLARO, M. et al. Automatic glaucoma detection based on optic disc segmentation and texture feature extraction. *CLEI Electronic Journal*, v. 19, n. 2, p. 5–5, 2016.
- [10] ORLANDO, J. I. et al. Convolutional neural network transfer for automated glaucoma identification. In: *12th International Symposium on Medical Information Processing and Analysis*. [S.l.: s.n.], 2017. p. 101600U–101600U.
- [11] KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural

- networks. In: PEREIRA, F. et al. (Ed.). *Advances in Neural Information Processing Systems 25*. [S.l.: s.n.], 2012. p. 1097–1105.
- [12] JIA, Y. et al. Caffe: Convolutional architecture for fast feature embedding. In: *Proceedings of the 22Nd ACM International Conference on Multimedia*. New York, NY, USA: [s.n.], 2014. p. 675–678.
- [13] CHATFIELD, K. et al. Return of the devil in the details: Delving deep into convolutional nets. In: *British Machine Vision Conference*. [S.l.: s.n.], 2014.
- [14] OJALA, T.; PIETIKÄINEN, M. Unsupervised texture segmentation using feature distributions. In: *International Conference on Image Analysis and Processing*. [S.l.: s.n.], 1997. p. 311–318.
- [15] HARALICK, R. M.; SHANMUGAM, K. et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, v. 3, n. 6, p. 610–621, 1973.
- [16] DALAL NAVNEET E TRIGGS, B. Histograms of oriented gradients for human detection. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2005. v. 1, p. 886–893.
- [17] TAMURA, H.; MORI, S.; YAMAWAKI, T. Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 8, n. 6, p. 460–473, 1978.
- [18] GALLOWAY, M. M. Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, v. 4.
- [19] CHU, A.; SEHGAL, C. M.; GREENLEAF, J. F. Use of gray value distribution of run lengths for texture analysis. *Pattern Recognition Letters*, v. 11, n. 6, p. 415–419, 1990.
- [20] DASARATHY BELUR V E HOLDER, E. B. Image characterizations based on joint gray level—run length distributions. *Pattern Recognition Letters*, v. 12, n. 8, p. 497–502, 1991.
- [21] QUINLAN, J. R. Induction of decision trees. *Machine learning*, v. 1, n. 1, p. 81–106, 1986.
- [22] HALL, M. et al. The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, v. 11, n. 1, p. 10–18, 2009.
- [23] CORTES, C.; VAPNIK, V. Support-vector networks. *Machine learning*, v. 20, n. 3, p. 273–297, 1995.
- [24] RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. *Learning internal representations by error propagation*. [S.l.], 1985.
- [25] BREIMAN, L. Random forests. *Machine learning*, v. 45, n. 1, p. 5–32, 2001.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *Biometrics*, p. 159–174, 1977.
- [30] SIVASWAMY, J. et al. Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation. In: *IEEE 11th International Symposium on Biomedical Imaging*. [S.l.: s.n.], 2014. p. 53–56.
- [31] 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.: s.n.], 2011. p. 1–6.