Automatic Method for Glaucoma Classification Using Texture Analysis, Xgboost and Grid search

Nonato R. de S. Carvalho
Universidade Federal do Piauí
nrdesales@ufpi.edu.br

Antônio Oseas de Carvalho Filho
Universidade Federal do Piauí
antoniooseas@ufpi.edu.br

Francisco das Chagas dos A. C. Júnior
Universidade Federal do Piauí
franciscojn49@gmail.com

Thiago José Barbosa Lima
Universidade Federal do Piauí
thiagojb12@ufpi.edu.br

Maria da Conceição L. C. Rodrigues
Instituto Federal do Piauí
mariacarvalho@ifpi.edu.br

Pablo Vieira de Abreu
Universidade Federal do Piauí
pablolukan@hotmail.com

Rafael Luz Araújo
Universidade Federal do Piauí
rafaluzaraujo@ufpi.edu.br

Deusimar Damião de Sousa
Universidade Federal do Piauí
deusimar@ufpi.edu.br

ABSTRACT

Glaucoma is an irreversible pathology, generated by increased intraocular pressure. Early detection is critical and can prevent total vision loss. Clinical examinations are commonly used to detect the disease. Still, the time and cost of identification is quite high. This paper presents a computational methodology that aims to assist specialists in the discovery of glaucoma through Computer Vision techniques. The proposed methodology consists in the application of several texture descriptors combined with a parameter optimization done through Grid search with the XGBoost classifier. A result was obtained with accuracy of 82.37% and ROC of 82.08%.

Keywords

Glaucoma; Grid search; XGBoost

1. INTRODUCTION

Glaucoma can be defined as an optic neuropathy, which can compromise the entire optic nerve structure, cause damage to the nervous system and also increase intraocular pressure [16]. The World Health Organization (WHO) reports that glaucoma is the second leading cause of irreversible blindness in the world [22]. Studies show that by 2020 more than 79 million people could be carriers of glaucoma worldwide [2].

Glaucoma is an asymptomatic disease in its early stages, which leads many patients to be unaware of the disease until it reaches advanced stages, which is why an early diagnosis that enables timely treatment is important [7]. Statistically, the increase of intraocular pressure does not indicate the presence of glaucoma, but normal intraocular pressure does not exclude the presence of the disease. This implies that the diagnosis of glaucoma should be based primarily on the analysis of the optic disc through fundus images [17].

Diagnoses of diseases in the field of ophthalmology require analysis of large quantities of images. The image analysis process for the diagnosis of glaucoma is still done manually in a slow and subjective manner. In addition, results may vary among professionals with a very high workload [24].

In this context, this work presents an automatic method that applies the descriptors Gray-Level Co-Occurrence Matrix (GLCM) and Local Binary Pattern (LBP), combined with Grid search parameter estimator and Extreme Gradient Boosting (Xgboost) classifier, with to provide an improvement in the classification of healthy glaucomatous images.

2. RELATED WORKS

Known studies on the problem of classification of fundus images have become relevant over the years and have become popular in the scientific world. Table 1 shows a summary of the main information from related works for the classification of glaucomatous images.

3. METHODOLOGY

In this section we expose the methodology used to develop the proposed method. Figure 1 provides a summary of the steps taken: Image Acquisition, Feature Extraction, Classification, and Validation Metrics.

3.1 Image Acquisition

The base used for this work was the RIM-ONE R2 [23], composed of 455 retinal images. The previous diagnostic process was performed by specialists, who determined a set of 200 glaucoma images and 255 healthy images. In order to using the proposed method, the Optical Disc regions were segmented from masks made available by specialists from the base itself.

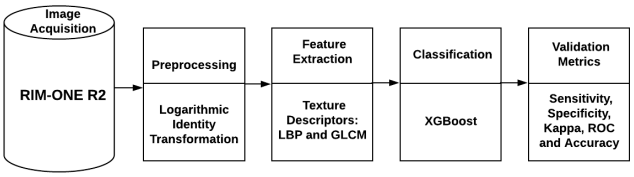

Figure 1: Methodology.

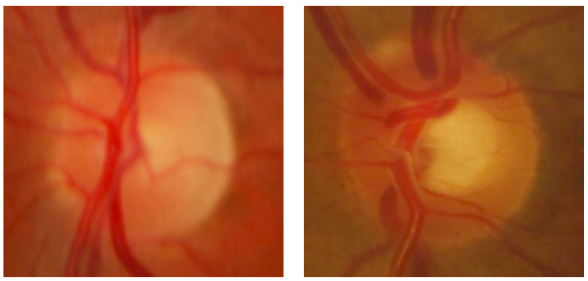

Figure 2: Examples of RIM-ONE R2 base images. (a) Normal, (b) Glaucoma.

3.2 Preprocessing

At this stage, logarithmic intensity transformation was used for image enhancement, increasing the contrast of intensity ranges or even binarizing the images. This transformation maps a narrow range of low input intensity values to a wider range of output levels, which can be expressed in Equation 1 [11] where, in the proposed method, $C = 80$

$$S = C * log(1 + r) \quad (1)$$

in which C is a constant and r is the pixel value.

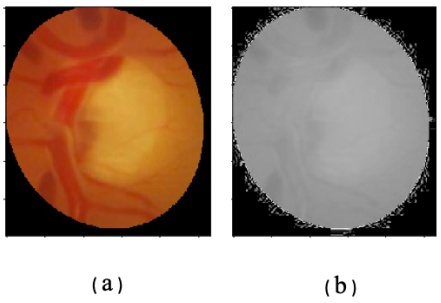

Figure 3: Examples of RIM-ONE R2 base images. (a) Without preprocessing, (b) Image (a) With preprocessing using logarithmic intensity transformation.

3.3 Feature Extraction

In the proposed method, we used descriptors that extract information related to: spatial distribution, luminosity variation, smoothness, roughness, regularity, surface structures and the relationships between nearby pixels [21].

GLCM is a texture descriptor that analyzes competitions between pixel pairs, storing their relative intensities in a square matrix. Co-occurrence probabilities are calculated between two gray levels i and j, using the angles (0^o, 45, 90^o e 135^o) and a distance named as pixel pair spacing [12].

LBP is used for feature extraction in the image recognition and classification process. In this method the intensity of each pixel of an image is replaced by a binary vector, determined by comparing each neighboring pixel with the central pixel, where the size of the neighborhood is 3x3, constituting the default limit value. The values obtained by neighbors are concatenated and the binary number generated is converted to the decimal base to replace the central value [20].

3.4 CLASSIFICATION

After the image characterization process, the extracted features were classified. This process is done with and without parameter optimization through Grid search and the Extreme Gradient Boosting classifier, using the cross validation method to give more consistency to the results.

Grid search.

It implements a fit and score method. It also implements "predict", "predict proba", "decision function", "transform" and "inverse transform" if they are implemented in the estimator used. The estimator parameters used to apply these methods are optimized by cross-validated grid search on a parameter grid [14].

XGBoost.

The XGBoost or Extreme Gradient Boosting algorithm is based on the gradient boosting machine (GBM) technique [10] and can be applied in the context of supervised learning. The method uses predictors (trees) that minimize the chosen loss function. The loss function consists of two factors: a calculated error rate on validation and a smoothing factor. XGBoost is scalable and its details are described in [4].

Cross validation.

It is a computationally intensive technique that uses all available examples of training and test samples. This type of underperformance estimation lacks computational efficiency because the training process is repetitive, but the ultimate goal is to decrease the estimate variance [25].

3.5 Validation Metrics

To evaluate the performance of the proposed method will be used the evaluation metrics: sensitivity, specificity, accuracy [3], Kappa [6] and Receiver Operating Charac teristic (ROC) [19].

For evaluation of the test four situations are possible:

1. True Positive (TP): An unhealthy image is correctly classified;
2. False Positive (FP): A healthy image is considered ill;
3. True Negative (TN): A healthy image is considered classified;
4. False Negative (FN): An unhealthy image is considered healthy.

Sensitivity (S) indicates the percentage of times the test

Table 1: RELATED WORKS.

Working	Used Techniques	Base	Sample	Accuracy
[15]	*Deep learning*	Various Chinese Clinics	48.116	92.9%
[5]	Use of texture descriptors And Convolutional Neural (CNN)	DRISHTI e RIM-ONE in its 3 versions	873	91,06%
[1]	Diversity Indexes	RIM-ONE version 2	455	93,41%
[9]	Image Decomposition and Bayesian Optimization	RIM-ONE version 2	455	91,24%

identify sick images, as defined by Equation (2).

$$S = \frac{TP}{TP + FN}. \quad (2)$$

Specificity (P) is the percentage of times that the test identify healthy images, being defined by Equation (3).

$$P = \frac{TN}{TN + TP}. \quad (3)$$

Accuracy (A) is the percentage of cases classified correctly, regardless of what is positive and what is negative, being defined by Equation (4).

$$A = \frac{TP + TN}{TP + FP + TN + FN}. \quad (4)$$

The Kappa index (K) can be classified according to Table 2, suggested by [13]. The closer the Kappa value is to 1, the greater the agreemen.

Table 2: Levels of classification accuracy according to the Kappa index.

Índice Kappa (k)	Quality
$K < 0.2$	Poor
$0.2 \leq K < 0.4$	Reasonable
$0.4 \leq K < 0.6$	Good
$0.6 \leq K < 0.8$	Very good
$K \geq 0.8$	Excellent

Source: [13]

The receiver operating curve, also called Receiver Operating Characteristic (ROC), is the area of TPR versus FPR varying the threshold. These metrics are summarized in equations 5 and 6 [18].

$$TPR = S = \frac{TP}{TP + TN} \quad (5)$$

$$FPR = (1 - S) = \frac{TP}{FP + TN} \quad (6)$$

,

in which, TPR represents true positive rate and FPR false positive rate.

The ROC chart provides an elegant way to present various confusion matrices produced at different boundaries. A ROC traces the relationship between the true positive rate and the false positive rate, as shown in Figure 4 [18].

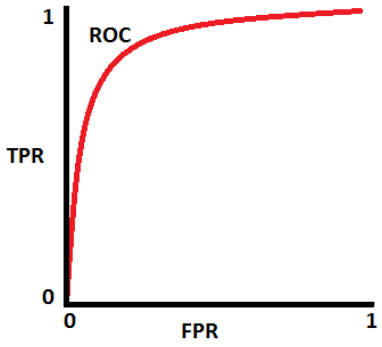

Figure 4: ROC chart illustration
Source: [18]

4. RESULTS

After the acquisition of the images, indicated in Section 3.1, and the preprocessing exposed in Section 3.2, tests were performed to evaluate the proposed method. The tests were performed on a set of 455 images, 200 glaucomatous and 255 healthy. Tables 3 and 4 present the results.

Based on the data presented in Table 3, using the GLCM descriptor the best result obtained was in the red channel with the parameters of: distance = 1 and angle = 90. Without applying the Grid search, it was obtained an accuracy of 75,77% and ROC of 75,73%. With the Grid search, it was obtained an accuracy of 76,87% and ROC of 76,68%.

Based on the data presented in Table 4, using the LBP descriptor the best result was the green channel with the parameters: radius = 5 and number of neighbors = 10. Without applying the Grid search, an accuracy of 81,71% and ROC of 81,43%. With the Grid search, an accuracy of 82,37% and ROC of 82,08%.

The results obtained with the proposed method indicate that Grid search makes it possible to achieve better results in the green and red channels.

Cross-validation brings greater assurance to these results. Each channel of the Optical Disc (OD) image is capable of producing a different representation of the anatomy of the fundus of the eye. The Optical Disc and its edges are most visible in the red channel. This property is used to check for changes caused by glaucoma that are close to the edges of the OD. The green channel has better contrast and, as in the blue channel, the excavation is more visible because its pixels have a higher intensity (brighter) than the pixels in the rest of the OD [8].

5. CONCLUSION

This paper presents a computational methodology that aims to assist specialists in the quantification of glaucoma through Computational Vision techniques. The proposed methodology uses LBP and GLCM to characterize the tex-

Table 3: Classification results applying the GLCM descriptor and Logarithmic Intensity Transformations.

	Without Grid Search					With Grid Search				
Channel	A (%)	S (%)	E (%)	ROC (%)	K (%)	A (%)	S (%)	E (%)	ROC (%)	K (%)
Green	70,27	67,53	73,00	70,27	40,28	71,80	68,18	74,60	71,39	42,74
	69,85	66,00	73,70	69,85	39,77	70,04	65,84	73,41	69,62	39,29
	69,37	66,15	72,58	69,37	38,61	68,94	65,44	71,48	68,46	36,69
Red	75,55	73,29	77,18	75,24	50,16	**76,87**	**75,67**	**77,69**	**76,68**	**52,70**
	75,77	**75,56**	**75,89**	**75,73**	**50,21**	73,78	71,20	75,66	73,34	46,56
	75,77	75,56	75,89	75,73	50,21	73,78	71,20	75,66	73,24	46,56
Blue	70,04	66,84	72,34	69,59	38,90	71,14	69,71	72,04	70,87	40,67
	68,28	65,21	70,37	67,79	35,10	70,48	67,18	72,90	70,04	39,86
	68,28	65,21	70,37	67,79	35,10	68,72	65,42	71,05	68,23	36,13

Table 4: Classification results applying the LBP descriptor and Logarithmic Intensity Transformations.

	Without Grid Search					With Grid Search				
Channel	A (%)	S (%)	E (%)	ROC (%)	K (%)	A (%)	S (%)	E (%)	ROC (%)	K (%)
Green	**81,71**	**77,61**	**85,24**	**81,43**	**63,09**	**82,37**	**78,74**	**85,42**	**82,08**	**64,37**
	80,61	77,33	83,26	80,30	60,72	82,37	79,60	84,58	82,09	64,25
	80,61	80,00	81,04	80,52	60,32	80,39	76,69	83,46	80,08	60,34
Red	76,43	72,54	79,60	76,07	52,26	75,55	72,68	77,69	75,18	50,20
	76,43	72,54	79,60	76,07	52,26	75,11	71,68	78,17	74,73	49,53
	75,77	71,49	79,35	75,42	51,00	74,88	70,33	78,77	74,55	49,28
Blue	77,75	72,68	82,35	77,51	55,23	77,97	72,81	82,70	77,75	55,70
	75,99	72,50	78,74	75,62	51,26	77,75	76,06	78,94	77,50	54,54
	75,55	70,37	80,25	75,31	50,80	77,75	75,78	79,16	77,47	54,59

ture and parameter optimization using the grid search in the Xgboost classifier. The tests were performed on the RIM-ONE R2 database and yielded promising results in which the grid search showed a considerable improvement in the performance of the glaucoma classification. For a future work, we intend to evaluate the impact that grid search has on the performance of the results of other classifiers and descriptors.

6. REFERENCES

[1] J. D. L. Araujo et al. Diagnóstico de glaucoma a partir de imagens de fundo de olho utilizando índices de diversidade. 2018.

[2] J. Ayub, J. Ahmad, J. Muhammad, L. Aziz, S. Ayub, U. Akram, and I. Basit. Glaucoma detection through optic disc and cup segmentation using k-mean clustering. In *2016 International Conference on Computing, Electronic and Electrical Engineering (ICE Cube)*, pages 143–147. IEEE, 2016.

[3] M. Bland. *An introduction to medical statistics*. Oxford University Press (UK), 2015.

[4] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794. ACM, 2016.

[5] M. L. Claro, R. d. M. S. Veras, A. M. Santana, L. H. S. Vogado, and L. P. Sousa. Diagnóstico de glaucoma utilizando atributos de textura e cnn's pré-treinadas. *Revista de Informática Teórica e Aplicada*, 25(1):82–89, 2018.

[6] J. Cohen. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, 70(4):213, 1968.

[7] A. S. V. de Carvalho Junior, E. D. Carvalho, A. O. de Carvalho Filho, A. D. de Sousa, A. C. Silva, and M. Gattass. Automatic methods for diagnosis of glaucoma using texture descriptors based on phylogenetic diversity. *Computers & Electrical Engineering*, 71:102–114, 2018.

[8] J. D. de Lima Araujo. Diagnóstico de glaucoma a partir de imagens de fundo de olho utilizando Índices de diversidade. Pós-graduação em ciência da computação, Universidade Federal do Maranhã, São Luís, 2018.

[9] A. G. S. Fernandes, C. M. da Silva Martins, A. C. de Moura Lima, G. B. Junior, J. D. S. de Almeida, and A. C. de Paiva. Meta aprendizagem de extração de caracterısticas aplicada ao diagnóstico de glaucoma. In *Anais do XIX Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 342–347. SBC, 2019.

[10] J. H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232, 2001.

[11] R. C. Gonzalez and R. C. Woods. *Processamento digital de imagens*. Pearson Educación, 2010.

[12] R. M. Haralick, K. Shanmugam, et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621, 1973.

[13] J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *biometrics*, pages 159–174, 1977.

[14] P. Lerman. Fitting segmented regression models by grid search. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 29(1):77–84, 1980.

[15] Z. Li, Y. He, S. Keel, W. Meng, R. T. Chang, and M. He. Efficacy of a deep learning system for detecting

glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology*, 125(8):1199–1206, 2018.

[16] R. Lim and I. Goldberg. Glaucoma in the twenty-first century. In *The Glaucoma Book*, pages 3–21. Springer, 2010.

[17] N. A. Loewen and A. P. Tanna. Glaucoma risk factors: Intraocular pressure. *The Glaucoma Book: A Practical, Evidence-Based Approach to Patient Care*, page 35, 2010.

[18] C. Maklin. Metrics for evaluating machine learning classification models. Disponível em: https://towardsdatascience.com/metrics-for-evaluating-machine-learning-classification-models-python-example-59b905e079a5, 2019. Acessado em: 19-12-2019.

[19] C. E. F. MATOS. Diagnosis of breast cancer in images mammography through local features and invariants, 2017.

[20] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.

[21] E. B. e. F. S. R. Oliveira, W. Classificacao de padroes em imagens usando descritores de textura. 2016.

[22] S. Resnikoff, D. Pascolini, D. Etya'Ale, I. Kocur, R. Pararajasegaram, G. P. Pokharel, and S. P. Mariotti. Global data on visual impairment in the year 2002. *Bulletin of the world health organization*, 82:844–851, 2004.

[23] M. I. A. G. RIMONE. Rim-one r2, 2015.

[24] E. B. Severo. Quantificação automática do glaucoma utilizando imagens de fundo do olho. 2014.

[25] M. Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(2):111–133, 1974.