

Detecção e Segmentação Automática de Estruturas em Imagens de Exames Oftalmológicos

Filipe A. Sampaio¹, Rodrigo M. S. Veras¹

¹Universidade Federal do Piauí (UFPI) - Departamento de Computação
Teresina - PI - Brasil

felipealvessampaio@hotmail.com, rveras@ufpi.edu.br,

Abstract. *This paper proposes a semi-automatic method for segmenting lesions in corneal images to assist specialists in monitoring the evolution of the lesion. Such lesions affect the ocular structure in the form of ulceration, infection, erosion, or some other type of trauma. The proposed method uses expert-marked regions to train the Random Forest classifier and the semi-supervised Seed Fuzzy C-means clustering (SFC-means). In the feature extraction, we made groups of pixels with similar texture, called superpixel. We achieved a sensitivity of 99.48% and similarity Dice: of 80.03. Finally, we concluded that the results showed that the Random Forest classifier had a better performance than SFC-Means.*

Resumo. *Este artigo propõe um método semiautomático para segmentar lesões em imagens da córnea, visando auxiliar os especialistas no monitoramento da evolução da lesão. Tais lesões podem ser categorizadas como patologias que afetam a estrutura ocular na forma de ulceração, infecção, erosão ou algum outro tipo de trauma. Assim, o método proposto utiliza regiões marcadas pelo especialista para treinar o classificador Random Forest e o agrupador semi-supervisionado Seed Fuzzy C-means. Na extração de atributos, foi realizado agrupamentos de pixels com textura similar, chamado superpixel. Assim, foi obtido bons resultados com as métricas sensibilidade: 99,48% e similaridade Dice: 80,03%. Por fim, concluímos que os resultados mostraram que o classificador Random Forest obteve um melhor desempenho.*

1. Introdução

Recentes avanços na aquisição e processamento de imagens digitais tornaram possível o uso de imagens médicas de maneiras novas e revolucionárias [Santos et al. 2018]. Isso levou a um considerável interesse no desenvolvimento de sistemas automáticos de diagnóstico médico, denominados como Sistemas de Diagnóstico Auxiliado por Computador (CAD – *Computer Aided Diagnosis*).

No campo da oftalmologia clínica, imagens digitais de retina podem prover informações sobre mudanças patológicas causadas por doenças oculares locais e sinais recentes de doenças sistêmicas como a hipertensão, a arteriosclerose e o diabetes melitus (DM). O tratamento para essas doenças se torna mais eficaz quando descobertas em seu estado inicial, sendo que os exames oftalmológicos periódicos são os principais responsáveis pela identificação da doença nessa etapa [Hajer et al. 2006].

Neste trabalho propomos a segmentação semiautomática de imagens de lesões de córnea utilizando os algoritmos *Seeded Fuzzy C-means (SFC-means)* e *Random Forest (RF)*, com *superpixels*. Nosso principal objetivo era mensurar a eficiência da aplicabilidade do algoritmo *SFC-means* durante a etapa de classificação do modelo proposto, comparando-o com um algoritmo bastante utilizado para o mesmo fim e com resultados comprovados. Além disso, os resultados obtidos são comparados com método do estado da arte anteriores. Tal comparação é realizada com o intuito de avaliar nosso método.

2. Trabalhos Relacionados

No trabalho de Wang *et al.* [Wang et al. 2017] foi realizado um estudo comparativo de métodos de classificação para diagnóstico automático de imagens oftalmológicas, mais especificamente a catarata pediátrica. O trabalho reuniu diferentes imagens da região ocular, com a finalidade de avaliar a complexidade referente ao processo de análise dessas imagens. Os autores utilizaram como descritores o espaço de cores RGB (iniciais das cores vermelho, verde e azul em inglês), métodos de extração de texturas, transformação *Wavelet* [Nayak et al. 2016] e *Local Binary Pattern* [Tang et al. 2015]. Os classificadores e descritores foram agrupados em oito esquemas para gerar diferentes resultados. Estes foram satisfatórios para os esquemas relacionados com os classificadores *Support Vector Machine (SVM)* e *K-Nearest Neighbors (KNN)*, mas foram insatisfatórios quando relacionados ao classificador *Extreme Learning Machine (ELM)*.

No trabalho de Santos *et al.* [Santos et al. 2018] é proposto o algoritmo *Seeded Fuzzy C-means*, que usa poucas informações fornecidas por um médico para segmentar regiões demarcadas. Seu trabalho foi avaliado por uma base de imagens contendo 2200 exemplares de leucemia, câncer de pele, câncer cervical e imagens de glaucoma. Os resultados ilustram a capacidade do algoritmo de auxiliar eficientemente os médicos na detecção de regiões de interesse.

Já no trabalho de Lima *et al.* [Lima et al. 2018] desenvolveram um método semi-automático para segmentação de lesões em imagens da córnea para auxiliar o médico no acompanhamento da evolução da lesão. Para isso, é utilizado algumas regiões marcadas pelo especialista para treinar o classificador *Random Forest* e, posteriormente, classificar as demais regiões da imagem entre lesão e não lesão. Na etapa de extração de características, é avaliado o uso de dois modelos de descrição, onde o primeiro usa *pixels* individualmente e o segundo utiliza agrupamento de *pixels* similares, denominados *superpixels*. Com o modelo de *pixels* e uma base de dados com 19 imagens, obtiveram 97,57% de acurácia e 0.89 de índice de similaridade Dice.

Neste trabalho propomos uma abordagem semiautomática para segmentação de lesões na córnea utilizando aprendizado de máquina supervisionado a partir de dois algoritmos que realizam a classificação de regiões não marcadas pelo especialista. Além disso, testes com diversas quantidades de segmentos de *superpixels* para treino foram realizados, onde buscamos uma quantidade mínima para garantir uma boa performance do sistema. O trabalho contribui com uma abordagem utilizando poucos cálculos durante o processo de classificação, obtendo resultados tão bons quanto usando segmentações com *pixels* de trabalhos anteriores.

3. Base de Dados

A base de dados utilizada no trabalho contém 19 imagens da superfície da córnea e foi adquirida em procedimentos médicos, sob condições diversas de iluminação e equipamentos. A base possui diferentes tipos de lesões na córnea, o que beneficia o trabalho pelo aumento da variabilidade de dados. Essa variabilidade nos levou a definir análises específicas na textura, procurando processar dados relacionados aos níveis de cinza próximos, comparações de forma e irregularidades na imagem. É importante salientar que essa base foi obtida do trabalho de Lima et. al. [Lima et al. 2018]. Após a captura das imagens, um médico especialista, com auxílio de uma caneta para toque em telas digitais, demarca a área correspondente à lesão no segmento anterior do olho em vermelho e a área total que será analisada é demarcada em azul. Na Figura 1 é possível visualizar a configuração da base de dados.

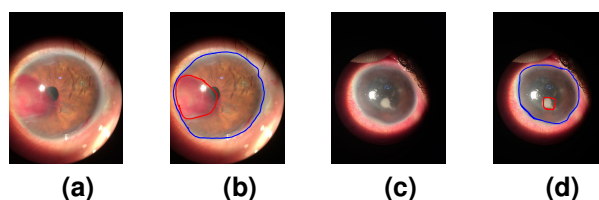


Figura 1. Exemplo das imagens presentes na base de dados. Em (a) e (c): imagens originais. (b) e (d): imagens marcadas pelo especialista para realização de treino dos classificadores.

4. Métricas de avaliação

Neste trabalho decidimos utilizar cálculos de uma matriz de confusão para avaliar o desempenho dos algoritmos de classificação utilizados. A matriz de confusão, por sua vez, possui a relação entre dados reais e aqueles preditos pelo classificador. O uso dessa relação é bastante usada para avaliações de modelos de classificação em aprendizado de máquina.

Na metodologia desse trabalho, o conceito de Positivo se aplica ao que se caracteriza como lesão, assim como o conceito de Negativo é caracterizado como não lesão. Portanto, de posse dos valores de VP (verdadeiro positivo para *pixels* preditos como positivo), VN (verdadeiro negativo para *pixels* preditos como negativo), FP (falso positivo para *pixels* originalmente negativos mas preditos como positivo) e FN (falso negativo para *pixels* originalmente positivos mas preditos como negativo) a matriz de confusão é gerada e, utilizando os valores de VP, VN, FP, FN podemos calcular 3 métricas: acurácia, sensibilidade e especificidade.

A acurácia mede o acerto geral do classificador, entre aquilo que foi corretamente predito como lesão e não lesão. Enquanto a sensibilidade mede a capacidade do classificador em prever que um pixel representa uma área lesionada, a especificidade mede se os valores preditos pelo classificador equivalem a *pixels* da área não lesionada.

Além dessas métricas de avaliação, foi utilizado também o índice de similaridade Dice [Dice 1945], que é um índice utilizado para comparar a igualdade entre duas amostras. Neste trabalho, foi considerado a comparação, em relação à área lesionada, entre a imagem marcada pelo médico e a imagem gerada pelo método proposto.

5. Método Proposto

Nesse trabalho são utilizadas cinco quantidades diferentes de segmentos de *superpixel*: 1500, 2000, 2500, 3000 e 4000. Esses segmentos são aglomerados de *pixels* que juntas formam o segmento, que por sua vez é utilizado para extração de características. Essas quantidades são utilizadas a fim de determinar qual seria a melhor configuração para obter o melhor resultado de classificação possível, tendo como foco dois quesitos de qualidade: velocidade de processamento e qualidade de aproximação de acerto das regiões lesionadas. Para cada algoritmo de classificação foram utilizadas todas as quantidades de segmentos, sendo executados isoladamente, comparando-os ao final do processo.

Para obter o modelo de *superpixels* foi utilizado o algoritmo *Simple Linear Iterative Clustering* (SLIC) [Achanta et al. 2012], um algoritmo de clusterização baseado no *K-means* que consiste em analisar as informações de cor dos *pixels*, dividindo-os em grupos. Com os *superpixels* selecionados, prosseguiu-se para o passo de extração de características de textura utilizando a *Gray Level Co-Occurrence Matrix* (GLCM), conhecida como matriz de co-ocorrência de níveis de cinza [Haralick et al. 1973]. Neste trabalho foram selecionados cinco características de textura: contraste, energia, homogeneidade, correlação e entropia. Assim, os *superpixels* restantes dentro da região ocular terão suas características extraídas e submetidas aos classificadores, no qual será informada se a segmentação feita possui ou não a lesão. Houve interesse em verificar o desempenho do *SFc-means*, em comparação com o *Random Forest* [Ho 1995], pois queríamos comprovar a eficiência desse agrupador presente no trabalho de Santos *et. al.* [Santos et al. 2018], aplicando-o em nossa base de dados.

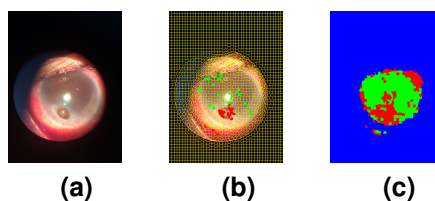


Figura 2. Etapas do modelo de descrição utilizando *superpixels*. (a) Captura, (b) *superpixels* em verde e vermelho representam a marcação do especialista e (c) classificação.

Após a classificação, o pós-processamento é aplicado na área correspondente a lesão, com o intuito de melhorar a imagem resultante do processo, aprimorando os resultados da matriz de confusão. Esta etapa consiste na utilização de morfologia matemática, seleção do maior componente e uma técnica de contorno para realizar a suavização das bordas da região selecionada. Finalizando, com a aplicação do operador morfológico de preenchimento de buracos (*Binary Fill Holes*), usado para fazer com que a lesão classificada se assemelhasse a área real da lesão. Na Figura 3 é possível visualizar a etapa de pós-processamento do método proposto aplicada a uma das imagens da base de dados.

6. Resultados e Discussão

No trabalho, os classificadores foram submetidos a testes usando as diferentes quantidades de segmentos de *superpixels*, junto a diferentes porcentagens de dados para treino. Cada uma das imagens do banco de dados passou pelos testes com esses diferentes valores. As porcentagens utilizadas de dados de treino foram: 1%, 5%, 10%, 15%, 20%, 25%,

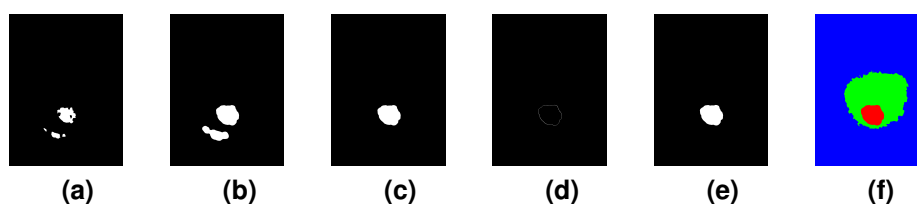


Figura 3. Etapas do pós-processamento. (a) resultado da classificação. (b) dilatação e erosão. (c) seleção do maior componente. (d) contorno. (e) preenchimento do buraco feito pelo contorno. (f) resultado do pós-processamento.

30%, 35%, 40% e 50%. Além disso, realizamos cinco execuções em cada classificação para obter uma média e um desvio padrão. Após análise, foram organizados na Tabela 1 os melhores resultados em ambos os algoritmos de classificação utilizados neste trabalho, obtidos com 2000 (1), 2500 (2) e 4000 (3) segmentos, além das seguintes porcentagens de treino: 10% (4), 25% (5) e 50% (6). Destacamos que na tabela abaixo também é realizado comparações com os melhores resultados obtidos usando o método de *pixels* proposto por Lima *et. al.* [Lima et al. 2020], com 50% de dados de treino.

Modelo	A (%)	S (%)	E (%)	D
RF	75.70 (3) e (6)	99.48 (2) e (6)	80.40 (3) e (6)	0.80 (1) e (5)
SFc-means	76.00 (2) e (6)	96.31 (3) e (4)	80.52 (2) e (6)	0.76 (3) e (4)
<i>Lima et. al.</i>	97.57	99.42	97.29	0.89

Tabela 1. Tabela comparativa dos valores obtidos pela classificação, utilizando as métricas A (acurácia), S (sensibilidade), E (especificidade) e D (índice de similaridade dice).

Observamos que ambos os algoritmos de classificação utilizados neste trabalho obtiveram resultados próximos. Os resultados obtidos no modelo de *pixels* de Lima *et. al.* [Lima et al. 2018] mostram resultados superiores, como observado na Tabela 1, devido ao fato da utilização de *superpixels* nesse trabalho facilitar a perda de características importantes para análise durante a classificação. No entanto, é importante salientar que a sensibilidade e Dice obtidos pelo RF proposto em nossa metodologia se mostrou promissora devido a proximidade dos resultados com o trabalho de Lima *et. al.* [Lima et al. 2018]. Essa proximidade de resultados deu-se ao melhoramento no pós-processamento utilizado neste trabalho. O método de *superpixels* com RF é capaz de diminuir drasticamente o custo computacional usado durante a classificação, onde no modelo de Lima *et. al.* [Lima et al. 2018] se mostra bastante custoso.

7. Conclusão e Trabalhos Futuros

Este trabalho propôs um método para segmentação de imagens de lesões da córnea por meio de um sistema semiautomático que utiliza técnicas de aprendizado de máquina semisupervisionado para classificar os *superpixels* entre lesão e não lesão. Comparando os melhores resultados obtidos para ambos os algoritmos de classificação utilizados, o RF utilizando 2500 segmentos e 50% de dados de treino apresentou melhor desempenho em comparação com o *SFc-means* com a mesma configuração.

Tinhamos como objetivo utilizar o *SFc-means* como algoritmo de classificação principal, devido aos ótimos resultados que o mesmo obteve no trabalho de Santos *et. al.*

[Santos et al. 2018], mas após testes neste trabalho, o mesmo não mostrou bons resultados possivelmente devido a utilização de *superpixels*, ou mesmo a incapacidade do algoritmo de realizar o agrupamento das diferentes texturas da base de imagens. O que não notou-se utilizando o RF, o qual obtém consideráveis resultados utilizando poucos dados (com 1500 segmentos de *superpixels* e 25% de dados de treino já era possível obter resultados melhores com o RF, comparando esses mesmos valores com o *SFc-means*).

Como proposta para trabalho futuro pretendemos implementar um método que consiga unir a utilização de *pixels* e *superpixels* para descrição de regiões, passando a analisar *pixels* isoladamente juntamente com sua região de vizinhança. Além disso, pretendemos utilizar somente o *SFc-means*, afim de realizar novos testes com esta nova proposta e ampliação da base de dados utilizada.

Referências

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11):2274–2282.
- Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302.
- Hajer, J., Kamel, H., and Nouredine, E. (2006). Blood vessels segmentation in retina image using mathematical morphology and the stft analysis. In *International Conference on Information & Communication Technologies*, pages 1130–1134. IEEE.
- Haralick, R. M., Shanmugam, K., et al. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621.
- Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, pages 278–282. IEEE.
- Lima, P., Veras, R., Leite, D., Tales, W., and Mendes, F. (2018). Avaliação de classificadores como método de segmentação de lesões na córnea. In *Anais da IV Escola Regional de Informática do Piauí*, pages 91–96. SBC.
- Lima, P. V., de MS Veras, R., Vogado, L. H., Portela, H. M., de Almeida, J. D., Aires, K. R., and Leite, D. (2020). A semiautomatic segmentation approach to corneal lesions. *Computers & Electrical Engineering*, 84:106625.
- Nayak, D. R., Dash, R., and Majhi, B. (2016). Brain mr image classification using two-dimensional discrete wavelet transform and adaboost with random forests. *Neurocomputing*, 177:188–197.
- Santos, L., Veras, R., Aires, K., et al. (2018). Medical image segmentation using seeded fuzzy c-means: A semi-supervised clustering algorithm. In *2018 International Joint Conference on Neural Networks*, pages 1–7. IEEE.
- Tang, Z., Su, Y., Er, M. J., Qi, F., Zhang, L., and Zhou, J. (2015). A local binary pattern based texture descriptors for classification of tea leaves. *Neurocomputing*, 168:1011–1023.
- Wang, L., Zhang, K., Liu, X., Long, E., Jiang, J., An, Y., Zhang, J., Liu, Z., Lin, Z., Li, X., et al. (2017). Comparative analysis of image classification methods for automatic diagnosis of ophthalmic images. *Scientific reports*, 7:41545.