

Identificação de lesões periapicais em exames de radiografia utilizando clusterização baseada na φ -divergência e algoritmos de classificação

Joniel B. Barreto, Danilo A. Oliveira, Rui F. Vigelis, Iális C. de Paula Júnior

Resumo—O presente trabalho desenvolve uma metodologia para identificar lesões em imagens de radiografias periapicais utilizando os classificadores KNN e MLP em conjunto da clusterização baseada na φ -divergência como método de segmentação. Também é comparado duas técnicas de balanceamentos de dados, *Tomek Link* e o *Random Under Sampling*. Com base nos resultados obtidos observa-se que embora ambos classificadores tenham sido eficazes para identificação das lesões, o KNN se mostrou mais eficiente do que a MLP, com acurácias de 90,12% e 75,99%, respectivamente, para a técnica *Tomek Link* que se mostrou mais aplicável que o *Random Under Sampling*.

Palavras-Chave—Visão computacional, Segmentação, Lesões periapicais.

Abstract—The present work develops a methodology to identify lesions in images of periapical radiographs using the KNN and MLP classifiers together in the clustering based on φ -divergence as a segmentation method. Two data balancing techniques are also compared, *Tomek Link* and *Random Under Sampling*. Based on the results obtained, it is observed that although both classifiers have been effective in identifying lesions, KNN has been shown to be more efficient than MLP, with accuracy of 90.12% and 75.99%, respectively, for the technique *Tomek Link* which proved to be more applicable than *Random Under Sampling*.

Keywords—Computer Vision, Segmentation, Periapical lesions.

I. INTRODUÇÃO

Lesões periapicais são reações inflamatórias que ocorrem na maioria dos casos como resultado de infecção pulpar ou necrose causada por trauma ou cárie dentária [1]. Determinar a localização e o tamanho de lesões pode ser uma tarefa desafiadora. Algumas lesões são detectadas apenas quando realizadas avaliações radiográficas de rotina por serem assintomáticas.

A radiografia periapical é um exame que serve para visualizar radiograficamente a anatomia de um ou mais dentes e estruturas vizinhas. Esses exames possuem um papel importante na área da Odontologia, principalmente na obtenção de diagnósticos de patologias orais em pacientes, como as lesões periapicais. A utilização de imagens radiográficas odontológicas em trabalhos clínicos é bastante rotineiro, contudo, devido ao fator humano, podem ocorrer erros na interpretação do diagnóstico que resultariam em um tratamento incorreto [2].

Joniel B. Barreto, Danilo A. Oliveira, Rui F. Vigelis e Iális C. de Paula Júnior. Programa de Pós-Graduação em Engenharia Elétrica e da Computação - PPGEEC, Universidade Federal do Ceará, Campus Sobral, e-mails: joniell.bb@gmail.com, danilo.alves@alu.ufc.br, rfvigelis@gmail.com e ialis@sobral.ufc.br. Este trabalho foi financiado pela FUNCAP.

Um caminho que ajuda a melhorar os resultados do diagnóstico e pode ser usado para fornecer um bom diagnóstico precoce é o processamento de imagens [3]. Em vista disso, estudos voltados envolvendo Visão Computacional (VC), com técnicas de segmentação, e Inteligência Artificial estão sendo desenvolvidos para análise de imagens odontológicas visando otimizar o processo de diagnósticos [4].

Em VC aplicada a Odontologia, a segmentação é responsável por identificar regiões de interesse para o problema, como selecionar patologias, separar o cortical mandibular, entre outros. Isso torna esse processo uma etapa fundamental, visto que possibilita elevar a qualidade dos diagnósticos e consequentemente melhorar e otimizar o tratamento dos pacientes. Diante disso, este trabalho objetiva apresentar uma metodologia para identificar lesões periapicais em imagens de radiografias odontológicas para auxílio de diagnósticos. Este artigo propõe utilizar a divergência generalizada, ou entropia relativa generalizada, como medida de dissimilaridade no algoritmo de clusterização proposto por Banerjee *et al.* [5] para segmentar as imagens e utilizar um classificador para diferenciar as lesões de outras regiões.

Estudos voltados para auxílio de diagnósticos em radiografias dentárias já são apresentados em várias revisões, principalmente visando auxiliar o profissional de saúde no diagnóstico clínico. Hasan *et al.* [6] segmenta a área referente à mandíbula em imagens panorâmicas digitais utilizando o Fluxo do Vetor Gradiente (GFV) juntamente da técnica de clusterização *K-means*. Já Lin *et al.* [7] apresentou um método de segmentação de radiografias periapicais aplicando o limiar de Otsu e análise de componentes conectados para a identificação dos dentes.

Algoritmos de clusterização são comumente utilizados em sistemas envolvendo segmentação e análise de imagens gerais. Em radiografia dentária existem diversos na literatura, tais como Alsmadi [8] que desenvolveu um método novo, totalmente automático para segmentação de lesões mandibulares em imagens panorâmicas de raios-X usando uma abordagem híbrida de *Fuzzy C-Means* e *Neutrosophic* ou Son e Tuan [9] que combinam o algoritmo *Fuzzy C-Means* com limiar de Otsu e clusterização *fuzzy* semi-supervisionado como método de segmentação.

A organização do restante do artigo é apresentado como segue. Na Seção II é abordado equacionamentos importantes para o desenvolvimento da medida de dissimilaridade proposta e sobre a implementação da técnica de clusterização *hard* de Bregman aplicada para a segmentação. Em seguida, a Seção III aborda sobre os técnicas e processamentos utilizados para o

desenvolvimento da detecção das lesões nas imagens periapicais. Os experimentos e resultados obtidos são apresentados na Seção IV. Por fim, as considerações finais e propostas futuras são relatadas na Seção V.

II. FUNDAMENTAÇÃO TEÓRICA

Nessa seção é abordado definições e propriedades da entropia na teoria da informação e informações importantes usadas para determinar a entropia relativa generalizada. Também é discutido sobre o algoritmo de clusterização de Bregman em aplicações como esse trabalho proposto.

A. Entropia Relativa Generalizada

Dada uma distribuição de probabilidade $p = p_i$, em que $\Delta_n = \{(p_1, \dots, p_n)\} : \sum_{i=1}^n p_i = 1, p_i \geq 0 \forall i$, Shannon[10] definiu a expressão

$$H(p) = - \sum_i p_i \ln p_i \quad (1)$$

como uma forma de computar a medida de incerteza de uma variável aleatória. Para mesma distribuição de probabilidade em Δ_n a entropia de Tsallis[11] expressa

$$H_q(p) = \frac{1}{q-1} \left(1 - \sum_{i=1}^n p_i^q \right), \text{ para } q \in [0, \infty). \quad (2)$$

Essa entropia é considerada uma generalização da entropia de Shannon, uma vez que no limite de $q \rightarrow 1$ a entropia de Tsallis se resume a entropia de Shannon.

Na teoria da informação a entropia pode ser compreendida como uma medida de incerteza. Já a entropia relativa é interpretada como uma medida da distância estatística entre duas distribuições de probabilidade [12]. Kullback e Leibler[13] apresentaram a entropia relativa de Shannon, ou divergência de Kullback-Leibler (KL), que para as distribuições de probabilidade $p = (p_i)$ e $q = (q_i)$ em Δ_i é dada por

$$D(p||q) = \sum_{i=1}^n p_i \ln \left(\frac{p_i}{q_i} \right). \quad (3)$$

A entropia relativa de Tsallis[14] é uma generalização da entropia relativa de Shannon que é definida ao substituir o logaritmo natural pelo q -logaritmo na divergência KL. É expressa por

$$D_q(p||q) = - \sum_{i=1}^n p_i \ln_q \left(\frac{q_i}{p_i} \right). \quad (4)$$

Considerando que as entropias relativas de Shannon e Tsallis podem ser reescritas, respectivamente, como

$$D(p||q) = \sum_{i=1}^n \frac{\ln(p_i) - \ln(q_i)}{\ln'(p_i)} \quad (5)$$

e

$$D_q(p||q) = \sum_{i=1}^n \frac{\ln_q(p_i) - \ln_q(q_i)}{\ln'_q(p_i)}, \quad (6)$$

Vigelis e Cavalcante [15], [16], [17] usaram a similaridade entre essas duas expressões e apresentaram a entropia relativa generalizada substituindo as funções q -logaritmo $\ln_q(\cdot)$ e

logaritmo natural $\ln(\cdot)$ pelo inverso da exponencial deformada φ . A divergência generalizada é dada por

$$D_\varphi(p||q) = \sum_{i=1}^n \frac{\varphi^{-1}(p_i) - \varphi^{-1}(q_i)}{(\varphi^{-1})'(p_i)}, \quad (7)$$

para $\varphi : \mathbb{R} \rightarrow [0, \infty)$ e tendo $D_\varphi(p||q) = 0$ se $p = q$.

B. Clusterização Hard de Bregman

Algoritmos de clusterização de Bregman se baseiam em uma grande classe de funções de distorção, conhecidas como divergências de Bregman, como critério de dissimilaridade para o agrupamento de dados. A distância Euclidiana quadrática, distância de Mahalanobis, distância Itakura-Saito e a entropia relativa são exemplos de funções de distorção bastante utilizadas em problemas de clusterização.

Os métodos de clusterização podem ser divididos em duas principais categorias: *hard* e *soft*. Os algoritmos de clusterização *hard* particionam os dados de forma que cada ponto pertença a apenas um cluster, já na clusterização *soft* cada ponto possui uma certa probabilidade de pertencer a cada um dos clusters. Para segmentação de imagens o método de clusterização *hard* é mais aplicável.

Dada uma imagem $\mathcal{X} = \{x_1, \dots, x_n\} \subset \mathbb{R}^d$, em que cada x detona um pixel da imagem, a segmentação por clusterização *hard* de Bregman (Algoritmo 1) é feita particionando o conjunto \mathcal{X} em k agrupamentos disjuntos $\mathcal{C} = \{C_1, \dots, C_k\}$ de forma que:

$$C_1 \cup C_2 \cup \dots \cup C_k = \mathcal{X}; \quad (8)$$

$$C_i \neq \emptyset, \forall i, 1 \leq i \leq k; \quad (9)$$

$$C_i \cap C_j = \emptyset, \forall i \neq j, 1 \leq i \leq k, 1 \leq j \leq k. \quad (10)$$

Para isso é necessário primeiramente definir o número de clusters k . Em seguida é inicializado um conjunto de centroides $\mathcal{M} = \{\mu_1, \dots, \mu_k\} \subset \mathbb{R}^d$ com valores aleatórios, em que μ_j representa o cluster C_j . Posteriormente cada $x_i \in \mathcal{X}$ é atribuído ao cluster C_j que atenda $\min_{\mu \in \mathcal{M}} d_\phi(x_i, \mu)$. Por fim, são definidos novos representantes para o conjunto \mathcal{M} , em que cada μ_j é dado pela média dos dados em C_j .

As etapas de atribuição dos dados e a re-estimação dos centroides são repetidas até que alcance algum critério de parada, que pode ser um número pré-definido de iterações ou uma função objetivo. Na clusterização *hard* de Bregman o algoritmo é convergido quando não é mais possível reduzir a perda total de informação de Bregman, que é expressa por

$$L_\phi(\mathcal{M}) = I_\phi(\mathcal{X}) - I_\phi(\mathcal{M}) = E_\pi[I_\phi(X_j)], \quad (11)$$

sendo $L_\phi(\mathcal{M})$ e $I_\phi(\mathcal{M})$ a informação de Bregman intra cluster e inter cluster, respectivamente. $I_\phi(\mathcal{X})$ é a informação total de Bregman dada por

$$I_\phi(\mathcal{X}) = E_\nu[d_\phi(\mathcal{X}, \mu)] = \sum_{i=1}^n \nu_i d_\phi(x_i, \mu). \quad (12)$$

Algorithm 1: Clusterização *Hard* de Bregman

Input: Imagem $\mathcal{X} = \{x_i\}_{i=1}^n$; medida de probabilidade de ν sobre \mathcal{X} ; divergência de Bregman d_ϕ e o número de clusters k

Output: \mathcal{M} ; agrupamento *hard* $\mathcal{C} = \{C_j\}_{j=1}^k$ de \mathcal{X} ; minimizador local de $L_\phi(\mathcal{M})$

Método:

Inicializa $\{\mu_j\}_{j=1}^k$ com valores aleatórios

repetir

{Etapa de atribuição}

para $i = 1$ até n **fazer**

$C_j \leftarrow C_j \cup \{x_i\}$, onde $j = \operatorname{argmin} d_\phi(x_i, \mu_j)$

fim do para

{Etapa de re-estimação}

para $j = 1$ até k **fazer**

$\pi_j \leftarrow \sum_{x_i \in C_j} \nu_i$

$\mu_j \leftarrow \frac{1}{\pi_j} \sum_{x_i \in C_j} \nu_i$

fim do para

até convergir

$\mathcal{M} \leftarrow \{\mu_j\}_{j=1}^k$

III. METODOLOGIA

Para o desenvolvimento do método de segmentação proposto são necessárias aplicações de diversas técnicas. Estas estão divididas em quatro etapas principais: aquisição e corte das imagens, pré-processamento, segmentação, extração e classificação das características. A Figura 1 ilustra essas etapas.

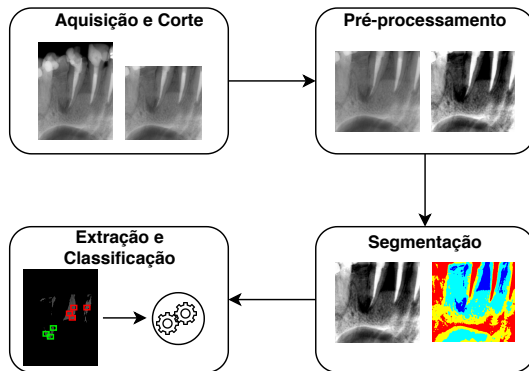


Fig. 1. Metodologia utilizada.

A. Base de imagens, técnica de corte e pré-processamento

O banco de imagens conta com 28 imagens de radiografias periapicais de pacientes diagnosticados com lesão periapical com resolução 476×620 pixels. Nessas imagens é possível observar a anatomia dentária (coroa e raiz), as estruturas ao redor dos dentes da região de incidência dos raios-x e do tecido ósseo adjacente. As imagens foram obtidas a partir da parceria realizada com o curso de Odontologia da UFC - Campus Sobral. Os *Ground truths* são marcados por especialistas cirurgiões-dentistas e professores do Curso de Odontologia da UFC - Campus Sobral.

Antes de realizar a segmentação dos cistos é realizado um procedimento de corte automático nas imagens, buscando remover regiões do interior da boca não relevantes para o diagnóstico.

Essa técnica consiste inicialmente na aplicação do filtro de média para suavização entre os *pixels*. Após isso, é aplicada a limiarização de Otsu, buscando separar a região externa das estruturas anatômicas. O resultado da limiarização é utilizado para rotacionar a imagem de entrada de forma que a área externa (região preta) fique na parte superior da imagem. Com a imagem resultante desse processo é selecionada a linha com maior valor e que possua uma porcentagem de *pixels* pretos maior que 5%. A Figura 2 apresenta um exemplo desse procedimento.

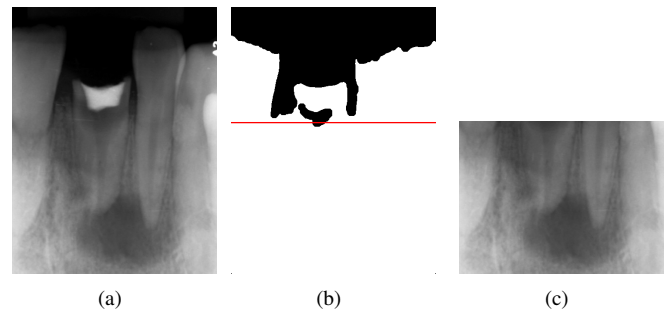


Fig. 2. (a) Imagem original. (b) Imagem binarizada e linha de corte. (c) Região cortada.

A Figura 2(a) apresenta um exame original, a Figura 2(b) apresenta o resultado da limiarização com a linha de corte destacada em vermelho e a Figura 2(c) apresenta a região da imagem cortada contendo a lesão. A partir da imagem obtida com o corte é aplicado o filtro gaussiano e uma operação de contraste morfológico, através de operações aritméticas da transformada *Top-hat* e *Bottom-hat*.

B. Segmentação e classificação

A segmentação dos cistos é feita por meio do algoritmo de clusterização *hard* de Bregman (Algoritmo 1) utilizando a divergência generalizada, Equação 7, como medida de dissimilaridade. Para implementar a clusterização baseada na φ -divergência primeiramente é necessário especificar a exponencial deformada φ . Assim como no trabalho desenvolvido por Barreto e Vigelis[18] é usado a função κ -exponencial[19], onde para cada $-1 < \kappa < 1$, é definida por

$$\exp_\kappa(x) = \exp\left(\int_0^x \frac{dt}{\sqrt{1 + \kappa^2 t^2}}\right), \quad (13)$$

em que para $\kappa = 0$, a função é reduzida à exponencial padrão. Assim sendo, para $\kappa \neq 0$, a exponencial pode ser reescrita da seguinte forma

$$\exp_\kappa(x) = (\kappa x + \sqrt{1 + \kappa^2 x^2})^{\frac{1}{\kappa}}, \quad (14)$$

cujas funções inversa e derivada da função inversa é dada, respectivamente, por

$$\ln_\kappa(y) = \frac{y^\kappa - y^{-\kappa}}{2\kappa}, \quad (15)$$

$$\ln'_\kappa(y) = \frac{y^\kappa + y^{-\kappa}}{2} \frac{1}{y}. \quad (16)$$

A escolha do valor de κ é realizada empiricamente, assim como o número de clusters k , para este trabalho foi definido $\kappa = 0,4$ e $k = 4$ para todas as imagens.

As imagens radiográficas são utilizadas em escala de cinza, assim sendo, cada ponto x_i do conjunto de entrada do algoritmo de clusterização é dado pelo valor de intensidade do brilho em escala de cinza de cada pixel. Assim sendo, visto que são usadas funções logarítmicas cujo domínio é $x > 0$, em seguida é feito um tratamento dos valores dos componentes do conjunto de entrada. Para cada $x_i = 0$ é somado um valor $\alpha \in \mathbb{R}$, sendo α um valor bem próximo de 0.

Realizada a clusterização, é definida a região em que consta a lesão dentre os quatro clusters segmentados, sendo ela a com maior valor de centroide. Após a segmentação, a região selecionada pelo algoritmo incluem com a lesão algumas áreas saudáveis. Acontecimento previsto pelas semelhanças entre lesões e outras estruturas anatômicas. A partir das regiões selecionadas pela segmentação são utilizados algoritmos de aprendizagem de máquina para separar os pixels referentes a região de interesse das demais.

O descritor utilizado é a entropia de Shannon. A extração dessa característica é feita de uma janela 17x17 que percorre no resultado obtido com a segmentação. Esse processo gera para cada imagem um banco de atributos. Esses bancos contam com amostras definidas por um vetor de característica de 12 atributos do descritor de Shannon para cada pixel e duas classes a serem separadas, indicando lesão e não-lesão. O número de amostras de cada classe varia para cada banco de dado, uma vez que são radiografias diferentes e logo possuem regiões com proporções diferentes.

Os algoritmos de classificação utilizados são, o *k-nearest neighbors* (KNN), com distância euclidiana e $K=1$, e o *multi-layer perceptron* (MLP), com duas camadas ocultas, uma com 100 neurônios e outra com 75. Como a presença de amostras de não-lesão é muito mais significativa que a de lesão, então também são empregadas técnicas de balanceamento de dados. Após a classificação, as imagens passam por uma etapa de pós-processamento, em que são aplicados o filtro da média, com janela 3x3, e a operação morfológica de fechamento, com elemento estruturante circular e dimensão 3x3.

As técnicas de balanceamento investigadas junto com a classificação são *Tomek Link* e o *Random Under Sampling*. O primeiro consiste em selecionar duplas de mostras com classes opostas em uma determinada vizinhança e remover os elementos de maioria nessas duplas. O segundo consiste em remover amostras aleatoriamente da classe majoritária para equilibrar as quantidades. Os resultados de classificação dos dois algoritmos com as duas técnicas de balanceamento são apresentados a seguir.

IV. RESULTADOS E DISCUSSÕES

Realizada a segmentação é possível observar que algumas regiões saudáveis também são segmentadas com as lesionadas. Em radiografias que a região referente à lesão não é adjacente à áreas saudáveis segmentadas como lesionadas, técnicas de pós-processamentos seriam o suficientes para diferenciá-las. No

entanto, em radiografias como a Figura 3 em que mesmo com a segmentação as áreas saudáveis e lesionadas são adjacentes torna necessário o uso dos classificadores para identificar as lesões. Também é possível observar no *Ground truth*, Figura 3(b), da radiografia a necessidade do uso das técnicas de balanceamentos, uma vez que a diferença das regiões é bastante significativa.

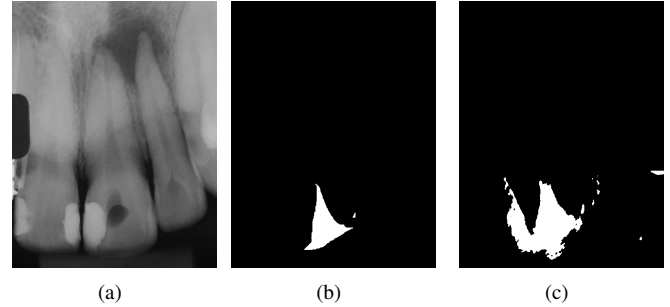


Fig. 3. (a) Imagem original. (b) *Ground truth*. (c) Imagem segmentada.

A classificação das regiões em lesão e não-lesão é feita utilizando o vetor de característica obtido a partir do descritor de entropia de Shannon como entrada dos classificadores. Para ambos os classificadores aplicados, o treinamento é realizado utilizando o método de validação *holdout*, em que é dividido aleatoriamente em 70% dos bancos de dados para treino e 30% para teste. Visando utilizar diferentes combinações para treino e teste, este processo é executado por 20 rodadas.

A partir dessas execuções são calculadas as métricas de Acurácia, Precisão, Sensibilidade e F1 - *score*. Em que, a acurácia representa o total de acertos obtidos pelo modelo, a precisão é a razão entre os verdadeiros positivos previstos e o total de observações positivas previstas, a sensibilidade é a razão entre verdadeiros positivos previstos e todas as observações positivas na classe real, já o F1 - *score* é uma métrica que combina as medidas de precisão e sensibilidade. Os resultados para as duas técnicas de balanceamentos obtidos pelos classificadores MLP e KNN são apresentados nas Tabelas I e II, respectivamente, sendo estes a média dos resultados das 20 rodadas.

TABELA I

MÉTRICAS OBTIDAS DO CLASSIFICADOR MLP

	Random Under Sampling	Tomek Link
Acurácia (%)	69,82	75,99
Precisão (%)	68,78	71,51
Sensibilidade (%)	71,25	65,84
F1-score (%)	67,96	67,19

TABELA II

MÉTRICAS OBTIDAS DO CLASSIFICADOR KNN

	Random Under Sampling	Tomek Link
Acurácia (%)	82,94	90,12
Precisão (%)	79,97	88,13
Sensibilidade (%)	83,82	88,05
F1-score (%)	81,10	88,09

Analisando a Tabela 1, observa-se que a técnica de Tomek Link possui acurácia de 75,99%, superior a da técnica de

Random Under Sampling com 69,82%. Em relação ao F1 - *score*, as duas técnicas apresentaram valores semelhantes, com a maior média de 67,96% pertencendo ao *Random Under Sampling*. Através desses dados é possível constatar também que a técnica de balanceamento que possibilita maior sensibilidade é a *Random Under Sampling*, com média de 71,25%, enquanto a Tomek Link possibilita maior precisão, com média de 71,51%.

Em relação a Tabela 2, observa-se que as duas técnicas de balanceamento obtiveram boas taxas de classificação, com exceção da precisão do *Random Under Sampling*, todas as demais taxas obtiveram médias superiores a 80%. Dentre elas, a técnica de Tomek Link possui valores superiores, com acurácia de 90,12%, precisão de 88,13%, sensibilidade de 88,05% e F1 - *score* de 88,09%. Com esses dados nota-se que a técnica de Tomek Link influencia positiva e significativamente na classificação das amostras.

Comparando os dois classificadores, é possível observar que o KNN possui maiores médias nas duas métricas avaliadas, representando melhor capacidade de decisão quanto as amostras do descritor de Shannon. Comparando as técnicas de balanceamento, a abordagem de seleção para excluir as amostras do Tomek Link resultou em melhores métricas na classificação.

A combinação do KNN com o balanceamento do Tomek Link providenciaram os melhores resultados de classificação da pesquisa. O resultado de classificação dessa melhor combinação é escolhido como o resultado final da metodologia e é apresentado na Figura 4.

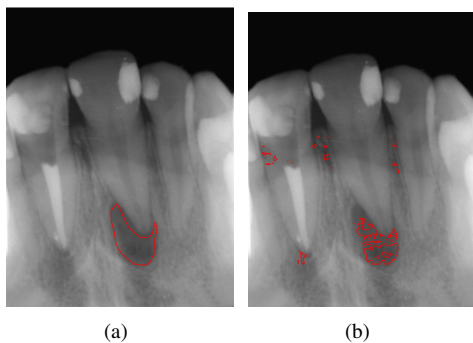


Fig. 4. (a) Imagem com *Ground Truth*. (b) Imagem com a classificação final.

A Figura 4(a) apresenta um exame do banco de imagens com a região do *Ground Truth* destacada em vermelho. A Figura 4(b) apresenta o mesmo exame com o destaque para o resultado final da metodologia. Observa-se que a metodologia proposta consegue indicar com boa incidência a região de lesão, no entanto também é possível observar a presença de falsos positivos em outras regiões do exame.

V. CONCLUSÕES

A principal contribuição desse trabalho é o desenvolvimento de uma metodologia para identificação de lesões periapicais em imagens de radiografia, buscando auxiliar profissionais de odontologia no diagnóstico clínico. Os resultados demonstraram que a melhor combinação de classificador com balanceamento

no banco é o KNN com o Tomek Link. Essas duas técnicas obtiveram acurácia de 90,12% e F1 - *score* de 88,09%. Embora haja possibilidades para melhoria, esse resultado cumpre o objetivo proposto.

Para trabalhos futuros, pretende-se expandir a base de imagens para melhor generalização da metodologia, bem como investigar outros descritores de representação das regiões. Além disso, também pretende-se realizar experimentos com algoritmos de *Deep Learning* e incorporar a metodologia.

REFERÊNCIAS

- [1] Z. B. Arslan, H. Demir, D. B. Yıldız e F. Yaşar, *Diagnostic accuracy of panoramic radiography and ultrasonography in detecting periapical lesions using periapical radiography as a gold standard*, *Dentomaxillofacial Radiology*, 20190290. v.49, pp. 1-7, 2020. doi: 10.1259/dmfr.20190290.
- [2] R. G. Birdal, E. Gumus1, A. Sertbas1 e I. S. Birdal, *Automated lesion detection in panoramic dental radiographs*, *Oral Radiol*, v. 32, pp. 111-118, 2016. doi: 10.1007/s11282-015-0222-8.
- [3] R. K. Devi et al., *Fuzzy based Regional Thresholding for Cyst Segmentation in Dental Radiographs*, 4th International Conference on Intelligent Computing and Control Systems (ICICCS 2020), pp. 544-549, doi: 10.1109/ICICCS48265.2020.9121104.
- [4] T. Iliou, C. N. Anagnostopoulos e G. Anastasopoulos, *Osteoporosis Detection Using Machine Learning Techniques and Feature Selection*, World Scientific Publishing Company. *International Journal on Artificial Intelligence Tools*. Grécia, v. 23, n. 5, pp. 1-9, 2014.
- [5] A. Banerjee, S. Merugu, I. S. Dhillon e J. Ghosh, *Clustering with Bregman Divergences*, *Journal of Machine Learning Research* 6, pp. 1705-1749, 2005.
- [6] M. M. Hasan et al., *Automatic segmentation of jaw from panoramic dental x-ray images using GVF snakes*, In: 2016 World Automation Congress (WAC). Institute of Electrical and Electronics Engineers (IEEE), 2016. Disponível em: <http://dx.doi.org/10.1109/WAC.2016.7583022>.
- [7] P. Lin et al., *Teeth segmentation of dental periapical radiographs based on local singularity analysis*, *Computer Methods and Programs in Biomedicine*, v. 113, n. 2, pp. 433 – 445, 2014. ISSN 0169-2607.
- [8] M. K. Alsmadi, *A hybrid Fuzzy C-Means and Neutrosophic for jaw lesions segmentation*, *Ain Shams Engineering Journal*. v.9, n. 4, pp. 697-706, 2018. ISSN 2090-4479, doi: 10.1016/j.asej.2016.03.016.
- [9] L. H. Son e T. M. Tuan, *A cooperative semi-supervised fuzzy clustering framework for dental X-ray image segmentation*, *Expert Systems with Applications*. v. 15, pp. 380-393, 2016.
- [10] C. E. Shannon, *A Mathematical Theory of Communication*, *The Bell System Technical Journal*, v. 27, pp. 379-423, 623-656, 1948.
- [11] C. Tsallis, *Possible Generalization of Boltzmann-Gibbs Statistics*, *Journal of Statistical Physics*, v. 52, n. 1-2, pp. 479-487, 1988.
- [12] T. M. Cover e J. A. Thomas, *Elements of information theory*, 2 ed, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2006.
- [13] S. Kullback e R. A. Leibler, *On Information and Sufficiency*, *Ann. Math. Statist.*, v. 22, n. 1, pp. 79-86, 1951.
- [14] L. Borland, A. R. Plastino e C. Tsallis, *Information Gain within Nonextensive Thermostatistics*, *Journal of Mathematical Physics*, v. 39, n. 12, pp. 6490-6501, 1998.
- [15] R. F. Vigelis e C. C. Cavalcante, *On φ -Families of Probability Distributions*, *Journal of Theoretical Probability*, v. 26, n. 3, pp. 870-884, 2013.
- [16] R. F. Vigelis e C. C. Cavalcante, *Information Geometry: An Introduction to New Models for Signal Processing*, *Signals and Images - CRC Press*, pp. 455-491, 2015.
- [17] D. C. de Souza, C. C. Cavalcante e R. F. Vigelis, *Modified Maximum Likelihood Estimator*, 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), pp. 1-5, 2016.
- [18] J. B. Barreto e R. F. Vigelis, *Clusterização Baseada na φ -Divergência Aplicada à Segmentação de Imagens*, XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais - SBRT2019, p. 5, Petrópolis, 2019.
- [19] G. Pistone, *κ -exponential models from the geometrical viewpoint*, *The European Physical Journal B*, v.70, pp. 29-37, 2009.