

VIABILIZANDO O *BENCHMARK* DE TÉCNICAS DE DETECÇÃO AUTOMÁTICA DE BACIOS DA TUBERCULOSE EM IMAGENS BACIOSCÓPICAS DE CAMPO CLARO

M. G. F. Costa*, L. B. Fujimoto** e C. F. F. Costa Filho*

*Centro de Tecnologia Eletrônica e da Informação, CETELI –UFAM, Manaus-AM, Brasil

**Faculdade de Medicina –UFAM, Manaus-AM, Brasil

e-mail: mcosta@ufam.edu.br

Resumo: Este trabalho descreve o projeto, a construção e a disponibilização do primeiro banco de imagens de microscopia de campo claro de esfregaços baciloscópicos, o *TbImages*, viabilizando, dessa forma, o *benchmark* de algoritmos de detecção automática de bacilos em baciloscopia de campo claro. O *TbImages* é constituído de um conjunto de imagens baciloscópicas de lâminas cujas secreções são oriundas de 12 pacientes anônimos. Foram capturadas imagens de 10 campos distintos por paciente e de cada campo, 10 imagens em diferentes profundidades focais. Foi realizada uma caracterização das imagens, com respeito a uniformidade do fundo e a densidade de bacilos. Todas as imagens foram avaliadas por um especialista que identificou e quantificou os bacilos presentes. O banco está disponível para a comunidade acadêmica

Palavras-chave: Banco de imagens, tuberculose, bacilos, detecção automática.

Abstract: *This paper describes the design and construction of the first image database of sputum smear microscopy, the TBImages, enabling thus the benchmarking of algorithms for automatic detection of bacilli in bright field microscopy. The TBImages consists of a set of images of sputum smear slides from 12 patients. Images of 10 different fields per patient and 10 images at different focal depths per field were captured. A characterization of the images with respect to uniformity of the background, and density of bacilli was performed. All images were evaluated by an expert who identified and quantified the bacilli present. The database is available to the research community.*

Keywords: *image database, tuberculosis, automatic detection, bacillus*

Introdução

O desenvolvimento de ferramentas de auxílio a diagnóstico que realizam a detecção automática de achados diagnósticos em imagens médicas ou microscópicas tem sido objeto de pesquisa e desenvolvimento nas últimas décadas. É possível identificar, nas bases de dados literárias, inúmeros trabalhos científicos sobre detecção e reconhecimento de achados clínicos que propõem algoritmos novos ou

aperfeiçoamentos em algoritmos já existentes. Entretanto, parte do sucesso relatado decorre de trabalhos cujas técnicas propostas foram validadas com bancos de imagens proprietários e, muitas vezes, pouco representativos das condições reais de aplicação das técnicas. Quando cada autor faz uso de uma casuística proprietária, a comparação de desempenho entre as técnicas fica inviabilizada.

Dessas afirmativas decorrem os esforços de muitos grupos de pesquisa na criação de bancos de imagens robustos, constituídos de um número estatisticamente significativo de amostras de imagens e cujos achados presentes nas mesmas podem ser confrontados com um padrão ouro. Por exemplo, no diagnóstico de câncer de mama, quando um pesquisador quer aferir a robustez de um algoritmo e, para, além disso, quer comparar diretamente os seus resultados com os de outros pesquisadores, ele tem a disposição vários bancos de dados e imagens públicos. Os mais acessados e usados são o *Mammographic Image Analysis Society* (MIAS) database e o *Digital Database for Screening Mammography* (DDSM) [1], [2].

A técnica mais utilizada para o diagnóstico da Tuberculose (TB) é a baciloscopia convencional (de campo claro). O exame de baciloscopia - pesquisa de bacilos álcool-ácido resistentes (BAAR) - consiste na análise direta de lâmina microscópica (contendo pequena fração de escarro do paciente) preparada com soluções corantes específicas. O diagnóstico é realizado a partir da visualização e contagem das estruturas bacilares (coradas em vermelho, em meio a um fundo azulado) presentes nos campos microscópicos. Dependendo da quantidade média de bacilos por campo, essa análise se estende de 20 até 100 campos microscópicos.

Do exposto, tem-se que a leitura baciloscópic, é, portanto, dependente da *expertise* do técnico responsável pela realização do exame e, em face da alta incidência de TB (“em 2012, cerca de 8,6 milhões de pessoas desenvolveram tuberculose e 1,3 milhões morreram da doença” [3]), o número elevado de baciloscopias realizado pelos profissionais induz a fadiga visual e, em consequência, a erros diagnósticos.

Assim, a análise automatizada da baciloscopia, apresenta-se como uma demanda que pode minimizar os erros decorrentes da análise visual humana. No entanto,

essa tarefa típica da área de reconhecimento de padrões, não é uma tarefa trivial em face das características das imagens baciloscópias.

Costa e colaboradores publicaram o primeiro trabalho sobre detecção automática de bacilos da Tuberculose (*Micobacterium tuberculosis*) em imagens de microscopia de campo claro [4]. A partir de então novas propostas de métodos de segmentação do bacilo da TB tem surgido [5] [6] [7] [8]. No entanto, a avaliação de desempenho de cada um dos métodos propostos pelos autores mencionados, foi realizada em bases de dados proprietárias e, portanto, sem a possibilidade de comparação de desempenho entre as mesmas.

Dentro desse contexto, o presente trabalho apresenta o *TbImages* – o primeiro banco de imagens de baciloscopia de campo claro com o fim de viabilizar a análise comparativa das técnicas de detecção do *Micobacterium tuberculosis*. Objetiva-se com isso, potencializar o desenvolvimento no menor tempo possível de uma ferramenta de auxílio ao diagnóstico automático da tuberculose a partir na análise automática de imagens de campos baciloscópicos.

Materiais e métodos

Materiais – Lâminas contendo secreção pulmonar de 12 pacientes com suspeita de tuberculose foram utilizadas. A preparação das lâminas seguiu o método de coloração já estabelecido. Neste caso, utilizou-se o método de Kinyoun. A utilização das amostras de secreção dos pacientes foi previamente aprovada, sem restrições, pelo Comitê de Ética em Pesquisas com Seres Humanos do INPA (protocolo 186/08).

Aquisição de imagens – Com vistas à formação de um banco de imagens representativo das condições de infraestrutura e de equipamentos disponíveis nos locais onde os exames de baciloscopia são rotineiramente realizados, foi utilizado um microscópio convencional (modelo Zeiss Axioskop 4 – com ampliação de 100x e abertura numérica de 1,25), ao qual foi acoplada uma câmara fotográfica digital de 10 megapixels. De cada lamina foram adquiridas imagens de 10 campos.

A utilização do microscópio referido (sem foco automático) em confronto com o objetivo de automatização do processo de aquisição de imagens evidenciou a necessidade de se estabelecer uma metodologia de foco passivo. Esse tipo de foco é assim denominado por ser baseado somente na análise do conteúdo das imagens capturadas. A posição do foco é determinada pela procura do máximo de uma função de autofoco usando a medida de uma série de imagens capturadas em diferentes planos focais. Para a realização dessa atividade foram adquiridas 10 imagens por campo, diferenciadas pela profundidade focal. O processo de aquisição relatado é ilustrado na Figura 1. A Tabela 1 sumaria a composição do banco de imagens. A avaliação da melhor métrica de foco foi objeto de outro trabalho científico [9].

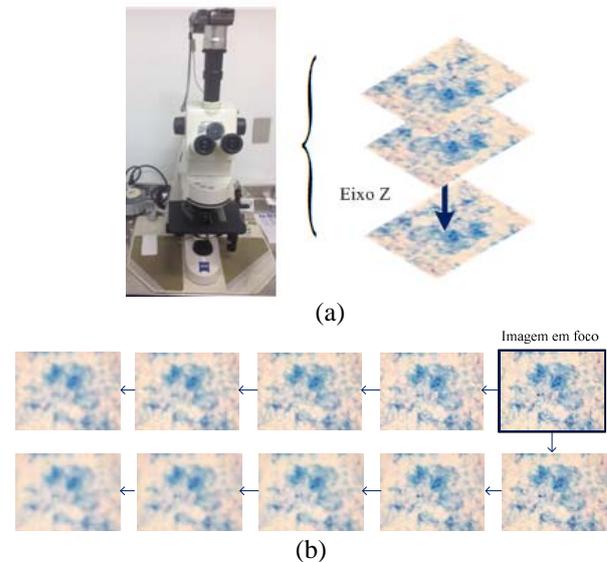


Figura 1: O foco passivo é definido em uma faixa profundidade focal que engloba a posição da imagem em foco (a) ilustração do processo; (b) amostra de imagens de um campo microscópico que diferem pela profundidade focal.

Tabela 1: Constituição do banco de imagens

Parâmetro	Quantidade
Pacientes	12
Lâminas por paciente	1
Campos por lamina	10
Imagens por campo	10
Total de Imagens adquiridas	1200
Total de Imagens em foco*	120

*uma por campo

Classificação das imagens em foco – As 120 imagens em foco, avaliadas pela métrica de foco implementada em [9], foram, posteriormente, submetidas a um processo de agrupamento, de acordo com a densidade de artefatos presentes no *background* e a densidade de bacilos por campo. Tal classificação visou, tão somente, possibilitar a avaliação das técnicas de detecção de bacilos por grupo de imagens.

A classificação das imagens, segundo a densidade de artefatos no fundo foi realizada a partir da análise da quantidade de tonalidade azul ($0,5 \leq \text{tonalidade} \leq 0,7$), característica do contra corante presente na imagem. Para cada uma das 120 imagens identificou-se o percentual de pixels na faixa de tonalidade do azul. Em seguida, através de um processo de agrupamento, método do vizinho mais próximo, separou-se essas imagens em dois grupos. De forma experimental, verificou-se que esses dois grupos poderiam ser separados por uma operação de limiar, conforme descrita a seguir:

$$\%Azul_{imagem} \begin{cases} \geq 13,56 \rightarrow \text{Alta densidade} \\ < 13,56 \rightarrow \text{Baixa densidade} \end{cases} \quad (1)$$

Esses dois grupos foram depois subdivididos em função da densidade de bacilos presentes. A Figura 2

apresenta a classificação realizada e mostra uma imagem exemplo de cada grupo. A Tabela 2 mostra a divisão das 120 imagens em foco nos diversos grupos.

Tabela 2 – Distribuição das 120 imagens em foco nos grupos

Grupo	#Paciente	Lâmina/Paciente	#Imagens
A-A*	1	1	10
	2	1	10
	3	1	10
			60
A-B	4	1	10
	5	1	10
	6	1	10
			30
			120
B-A	7	1	10
	8	1	10
	9	1	10
			60
B-B	10	1	10
	11	1	10
	12	1	10
			30

*A-A: Alta Densidade de artefato/fundo – Alta Densidade de bacilos

Observa-se na Figura 2 que as imagens com alta densidade de artefatos no fundo são caracterizadas pela presença marcante do contra corante azul de metileno no

fundo, enquanto que as de baixa densidade tem presença mais discreta do mesmo. Em geral esses tipos de variações são decorrentes das características da amostra (secreção pulmonar) e do processo de coloração.

Leitura das imagens pelo especialista – As 120 imagens em foco foram submetidas à análise de um especialista com o fim de identificação dos bacilos. Essa análise resultou em três classes de objetos: bacilos verdadeiros, aglomerados de bacilos e bacilos duvidosos. Esta última classe deveu-se ao fato de que, como a superfície do esfregaço presente na lâmina não é, na maioria das vezes, uniforme, o foco não está presente em todo o campo capturado, gerando dúvidas se o objeto é realmente um bacilo. Objetos pertencentes a cada uma das classes referidas foram circundados por uma forma geométrica distinta:

- Um bacilo verdadeiro → um círculo;
- Um aglomerado de bacilos → um retângulo
- Um bacilo duvidoso → um polígono

Adicionalmente, foi realizada a contagem dos objetos por imagem.

Resultados

Na Figura 3, são apresentados exemplos dos objetos, identificados pelo especialista, nas imagens de baciloscopia de campo claro e suas respectivas formas de identificação. A Figura 4 ilustra amostras do *Tbimages*.

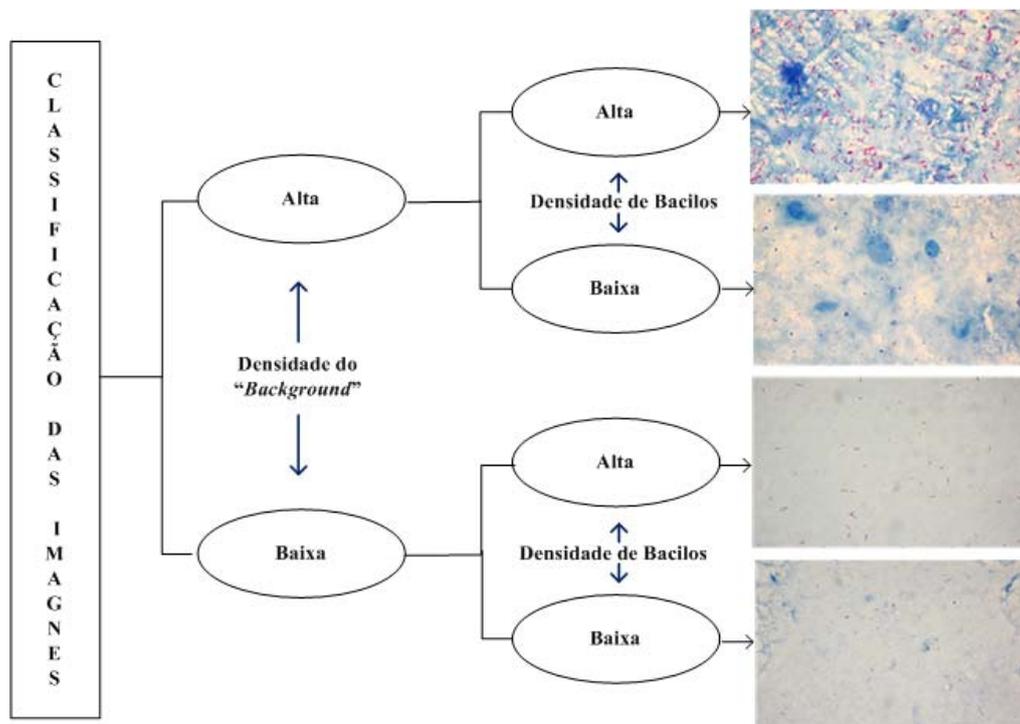


Figura 2: Classificação das imagens segundo a densidade de artefatos presentes no fundo e a densidade de bacilos

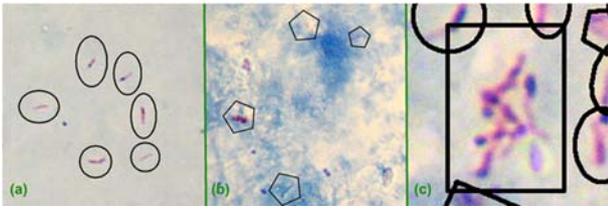


Figura 3- Exemplos de marcação dos objetos: (a) bacilos verdadeiros; (b) bacilos duvidosos; (c) aglomerados de bacilos.

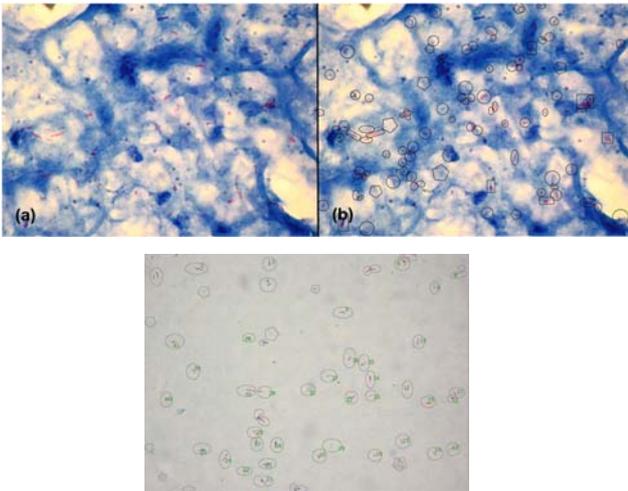


Figura 4 – Imagens baciloscópicas: (a) Imagem original. (b) Imagem com marcação dos bacilos, bacilos duvidosos e aglomerados (c) exemplo de imagem com a contagem dos bacilos marcados.

Discussão

O projeto, a construção e a disponibilização do primeiro banco de imagens de microscopia de campo claro de esfregaços baciloscópicos que viabiliza o *benchmark* de algoritmos de detecção de bacilos em bacilosopia de campo claro foi o objeto desse trabalho.

O referido banco é constituído de um conjunto de imagens baciloscópicas de lâminas cujas secreções são oriundas de pacientes anônimos. A seleção das lâminas e dos campos microscópicos capturados foi realizada de tal forma a redundar em um conjunto de imagens representativo dos diferentes índices bacilares, bem como das diferentes apresentações do *background*. Tais imagens foram adquiridas em diferentes profundidades focais, com a imagem em foco sendo validada por métrica já estabelecida em pesquisas anteriores. Adicionalmente, todas as imagens em foco foram submetidas a um especialista que, através de uma interface gráfica própria, identificou todos os bacilos, as globias (aglomerados de bacilos) e os bacilos duvidosos presentes nas mesmas. O referido banco foi publicado na Web (www.tbimages.ufam.edu.br) para o conhecimento da comunidade acadêmica. O mesmo já foi disponibilizado para pesquisadores interessados.

Os autores estão trabalhando em uma segunda versão do banco, desta feita com o objetivo de possibilitar o

benchmark da bacilosopia automática, ou seja, de técnicas de diagnóstico automático da tuberculose através da análise de imagens baciloscópicas.

Agradecimentos

Este trabalho contou com o suporte financeiro da FAPEAM, SUFRAMA e FINEP.

Referências

- [1] J Suckling et al: The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica. International Congress Series 1069 pp375-378, 1994.
- [2] Michael Heath et al: The Digital Database for Screening Mammography, in Proceedings of the Fifth International Workshop on Digital Mammography, M.J. Yaffe, ed., 212-218, Medical Physics Publishing, 2001.
- [3] WHO, Global TB Control report; 2013. [accessed in 2014, jan 24]. Available from: <http://www.who.int/tb/publications/global_report/en>.
- [4] Costa, M. G. F.; Costa Filho, C. F. F.; Sena, J. F.; Salem, J.; Lima, M. O. de. Automatic identification of mycobacterium tuberculosis with conventional light microscopy. In the 30th Annual International IEEE EMBS Conference Vancouver, British Columbia, Agosto, p. 382-385, 2008.
- [5] Raof RAA, Salleh Z, Sahidan SI, Mashor MY, Md Noor SS, Idris FM and Hasan H. Color thresholding method for image segmentation algorithm of Ziehl-Neelsen sputum slide images. Proceedings of 5th International Conference on Electrical Engineering, Computing Science and Automatic Control. 2008: 212-17.
- [6] Sadaphal P, Rao J, Comstock GW and Beg MF. Image processing techniques for identifying mycobacterium tuberculosis in Ziehl-Neelsen stains. International Journal of Tuberculosis Lung Disease. 2008; 12(5): 579-82.
- [7] Makkapati V, Agrawal R and Acharya R. Segmentation and classification of tuberculosis bacilli from ZN-stained sputum smear images. Proceedings of 5th Annual IEEE Conference on Automation Science and Engineering. 2009: 217-20.
- [8] Khutlang R, Krishnan S, Dendere R, Whitelaw A, Veropoulos K, Learmonth G and Douglas TS. Classification of mycobacterium tuberculosis in images of ZN-Stained sputum smears. IEEE Transactions on Information Technology in Biomedicine. 2010; 14(4): 949-57.
- [9] Kimura Junior, A.; Costa, M. G. F.; Costa Filho, C.F.F. Evaluation of autofocus functions of conventional sputum smear microscopy for tuberculosis. In: 32nd Annual International IEEE EMBS Conference, 2010, Buenos Aires. Proceedings of 32nd Annual International IEEE EMBS Conference, 2010. v.1. p.3041-3044.