

XV Congresso Brasileiro de Informática em Saúde 27 a 30 de novembro - Goiânia - Brasil

DETECÇÃO DE ARTEFATOS ESTRANHOS EM RADIOGRAFIAS DE TÓRAX

Afonso Ueslei da Fonseca^{1*}, BSc; Fabrízzio Alphonsus A. M. N. Soares, PhD; Leandro Luís Galdino de Oliveira, PhD.

> Instituto de Informática, Universidade Federal de Goiás. Goiânia, Goiás – Brasil.

Resumo: A radiografia de tórax é um dos exames preconizados para detecção de pneumonia na infância pela Organização Mundial de Saúde. Durante a execução do exame é comum encontrar artefatos nessas imagens radiográficas como: rótulos de identificação, marcas de dedos, botões de camisa, dentre outros, principalmente quando digitalizadas por meio de câmeras fotográficas. Esses artefatos além de tirarem o anonimato da radiografia, afetam significativamente sua análise por sistemas informatizados de detecção e suporte a identificação de doenças. Este trabalho apresenta um método eficiente para identificação dos artefatos, composto de 3 etapas principais: filtragem de *pixels* baseada em histograma, detecção de bordas com algoritmo de Roberts e filtragem espacial por filtro de desvio padrão. Este método foi experimentado em uma base de 200 imagens e inspecionado visualmente para identificação de erros. Resultados experimentais como, eficiência (tempo processamento/radiografia) \approx 7ms e precisão de 0,98 demonstram que o método é bastante promissor.

Palavras-chave: Radiografia Infantil, Processamento de Imagem Assistida por Computador, Detecção de objetos.

Abstract: Chest radiography is one of recommended imaging test by World Health Organization for childhood pneumonia diagnosis. However, during patient examination is very usual finding artifacts in these images, such as identification labels, fingerprints, shirt buttons, and so forth. Moreover, when these images are digitally scanned, other problems raise such as noise, brightness control and so on. Artifacts can reveal private data and expose patient identification. Furthermore, these artifacts can significantly damage automatic analysis by computer diagnosis aided systems. This works presents an efficient method for artifact identification spacial filtering. This method has been experimented upon 200 images database and presented about 7ms of time processing per image. Visually inspection was used to error measuring and we achieve 0,98 of precision. As a result of this, the method demonstrate a very promising preprocessing tool.

Keywords: Chest X-ray; Computer aided; Image processing; Object detection; Child

Introdução

Imagens digitalizadas de radiografias de tórax compõem uma preciosa fonte de informação para estudo, treinamento, investigação e auxilio a diagnóstico principalmente em sistemas informatizados^{1,2}. Inúmeros bancos de dados como JSRT³₃ SCR⁴, ©HARRY⁵, Chandrasekhar⁶, JAEGER⁷ estão disponíveis com os mais variados tipos de radiografia, adquiridas com diferentes configurações e aparelhos o que permite uma boa amostragem para análises automatizadas por computador.

A organização mundial de saúda (OMS) recomenda e tem como preconizado para detecção de pneumonia na infância o exame de radiografia de tórax⁸, dada à sua simplicidade e o seu baixo custo.

1^{*} Autor correspondente: Afonso Ueslei da Fonseca. e-mail: afonso@inf.ufg.br

Entretanto radiografias pediátricas – crianças menores de 5 anos – comumente apresentam grande heterogeneidade. Por exemplo, além das diferenças de configurações e aparelhos, são ainda afetadas pela questão técnica relativa ao profissional e ao posicionamento e manobra dos pacientes nessa faixa etária e que reduz ainda mais a qualidade da radiografia⁹. A Figura 1 mostra exemplos de radiografias pediátricas com presença de artefatos.

Além dos problemas técnicos apresentados, por questões de ética, privacidade e segurança é também desejável que tais imagens preservem o anonimato^{10,11} do paciente. Da mesma foram, é necessária uma mínima qualidade para que o processamento por algoritmos de análise automática tenham eficiência¹². Assim, a remoção de artefatos é de substancial contribuição visto que favorecem o anonimato e contribuem com a performance do processamento^{12,13,14,15}.



Figura 1 – Exemplos de imagens com presença de artefatos estranhos: a) radiografia com identificação textual à direita b) radiografia com marcas de dedos na região inferior, c) radiografia com etiquetas na parte superior esquerda.

Sistemas informatizados de detecção e apoio a identificação de doenças vem se tornando a cada dia mais comuns em ambientes hospitalares, quer seja auxiliando no treinamento de profissionais, ou servindo de suporte a identificação de doenças^{2,16,17,18,19}. Isso reforça ainda mais a importância de tais sistemas e a necessidade de que tenham eficiência naquilo que propõem.

No que tange a imagens radiografias estudos recentes têm endereçado a detecção e remoção de objetos estranhos em radiografias de tórax. Hogeweg¹³, trabalhou a detecção usando classificação de *pixels* supervisionada combinada com um classificador KNN, e uma segmentação pós-processamento dos *pixels* por regiões sobre certo limiar. Xue¹² usou dois métodos, um baseado na transformada circular de Hough e outro no algoritmo de Viola-Jones para detecção de botões. Em ambos os trabalhos fez-se uso de métodos supervisionados que exigem etapa de treinamento. Em Hogeweg¹⁴ trabalhou a detecção de estruturas translucidas alongadas por meio de filtragem subespacial por *Blind Source Separation* (BSS) realizada por meio da análise de componentes principais (PCA). Em todos os trabalhos apenas a região dos pulmões foi considerada para aplicação dos métodos.

Este trabalho propõe a detecção de artefatos em toda a área da radiografia a partir da combinação de filtragem de *pixels* baseada em histograma, detecção de bordas com algoritmo de *Roberts e* filtragem espacial por filtro de desvio padrão.

O artigo é organizado nas seguintes seções: A base de dados usada nesse trabalho; detalhes dos passos do método proposto para detecção automática de artefatos; discussão dos resultados e conclusão.

Base de dados de radiografia de tórax infantil

Para a detecção dos objetos contamos com uma base de dados de radiografias posterior de tórax obtida do resultado do Programa de Vigilância de Pneumonia Radiológica em Crianças do Município de Goiânia, cidade com população estimada em 2015 de 1.430.697 habitantes²⁰. A Secretaria de Saúde

iniciou em julho de 1999 a vacinação contra o Hib nos serviços locais de saúde, e concomitantemente foi implementado um sistema de vigilância populacional de pneumonias adquiridas na comunidade admitidas em hospitais pediátricos da cidade. Vinte hospitais pediátricos participaram do estudo, que contemplam 85% das internações por pneumonia no município.

O sistema de vigilância teve início em maio de 2000. Crianças menores de cinco anos, com infecção respiratória aguda atendida pelos pediatras nos ambulatórios dos serviços de saúde com diagnóstico clínico sugestivo de pneumonia eram encaminhadas para radiografia de tórax.

Na admissão hospitalar, uma ficha clínico-epidemiológica era preenchida e as radiografias de todos os casos potenciais eram fotografadas com câmeras digitais (Mavica FD90) de alta resolução, de acordo com especificações e configurações recomendadas pela Organização Mundial da Saúde/OMS.

As fotos digitais eram imediatamente enviadas para os dois radiologistas treinados para interpretar as radiografias de acordo com procedimentos padronizados durante workshop realizado em Santiago - Chile, patrocinado pela OMS/PAHO. Assim, "caso de pneumonia radiológica" foi determinado nas seguintes situações: 1) presença de apenas opacidade alveolar; 2) presença de qualquer tipo de opacidade (alveolar, intersticial ou mista) associada a derrame pleural; 3) apenas derrame pleural.

A base de dados contém cerca de 23.000 imagens (padrão técnico adequado em 85% delas) provenientes das radiografias de tórax de crianças com suspeita clínica de pneumonia. A base constitui um excelente acervo digital para construção de um sistema especialista de visão computacional que possa gerar evidências e contribuir para tomada de decisão na prática clínica. Esta base de dados encontra--se armazenada no Departamento de Saúde Coletiva da Universidade Federal de Goiás (IPTSP/UFG).

A aprovação ética para esta investigação foi concedida pelo Comitê de Ética da Universidade Federal de Goiás, Goiânia, Brasil (protocolo nº. 100/11).

Método proposto

O método consiste de 3 etapas principais executados em toda a imagem: filtragem de *pixels* baseada em histograma, detecção de bordas por *Roberts e* filtragem espacial por filtro de desvio padrão. Complementarmente, em todas as etapas, operações morfológicas são realizadas para remover ruídos. Ao final, os resultados dos estágios são então combinados para obtenção de uma máscara final que cubra os *pixels* correspondentes aos objetos identificados. A Figura 2 mostra um esquema geral do método proposto.



Figura 2 – Visão geral do método proposto, uma imagem de entrada é submetidas a etapas de processamento e a saída é uma máscara dos informações estranhas identificadas.

Limiarização baseada em histograma

Suponha que uma imagem *I* tenha *N pixels* com *L níveis de cinza L* = {0,1, ..., *L*-1} o histograma da imagem é indicado por h(i), i=0, ..., *L*-1 onde h(i) representa o número de *pixels* com o nível de cinza *i*, o desvio padrão do histograma é indicado desvp(h) e dado *t* o valor de limiar correspondente a 1/6 de desvp(h). Seja f(x,y) as coordenadas de um *pixel* na imagem bidimensional, a filtragem g(x,y) de *I* é obtida pela seguinte limiarização:

$$g(x,y) = \begin{cases} 1 & \text{para todo } f(x,y) \text{ de intensidade i tal que } h(i) < t \\ 0 & \text{para todo } f(x,y) \text{ de intensidade i tal que } h(i) > t \end{cases}$$

A filtragem resulta em uma imagem indicada por I_{hist} que representa uma máscara dos *pixels* da imagem com menor taxa de amostragem, em seguida I_{hist} é submetida a duas operações morfológicas, quais sejam:

<i>Majority</i> ²	seta um pixel para 1 se 5 ou mais pixel. rio seta para 0.	s de	e um	ia vi	izin	hanç	ca [.	3x3]	' tive	erem valor 1, caso contrá
		0	0	0	0		0	0	0	c
		0	0	0	0		0	0	0	C
Spur ¹	remove <i>pixels</i> esporões, por exemplo:	0	0	1	0	-	0	0	0	C
		0	1	0	0		0	1	0	C
		1	1	0	0		1	1	0	C

Detecção de bordas por 'Roberts'

Essa detecção de bordas é obtida usando a aproximação das derivadas de $Roberts^{21}$ e retorna as extremidades nos pontos de uma imagem onde o gradiente de I é máximo, isso resulta em uma borda em regiões da imagem onde o limites são bem definidos. Seja I_{rob} a imagem resultante da detecção de bordas de I pelo operador de *Roberts*, sobre ela a fim de promover ligação de pontos de descontinuidade, remover pequenos ruídos, e preencher buracos são realizadas as seguintes operações morfológicas:

Bwareaopen ²	remoção de objetos com menos de <i>S pixels</i> . O tamanho de S foi definido empiricamente levando em conta o tamanho médio dos objetos presentes nas radiografias de nossa base de dados.
Imdilate ²	ligação dos pontos descontínuos através de um elemento estruturante definido como [(1 1 1 1 1) (1; 1; 1; 1; 1)].
Imfill ²	preenchimento de buracos.
Imerode ²	erosão por um elemento estruturante em forma de disco com raio igual a 3.

² *Majority, Spur, bwareaopen, imdilate, imfill, imerode* são funções de processamento digital de imagens presentes no programa *Matlab*^[21] para execução de operações morfológicas^[20].

Filtragem desvio padrão

Sobre a imagem I é aplicado uma filtragem pelo desvio padrão²¹ local gerando uma nova imagem I_{desvp} que é submetida a uma limiarização que seta em '1' todos os valores de intensidade se estes são maiores que 3 vezes o seu desvio padrão e em '0' caso contrário. Em seguida, sobre o resultado são aplicadas as mesmas operações morfológicas realizadas na imagem anterior I_{rob} .

Combinação dos resultados das etapas principais

Para se obter o melhor de cada etapa são realizadas sobre as imagens I_{hist} , $I_{rob} \in I_{desvp}$ as seguintes operações:

 $I_{hisbr} = I_{hist} \cup I_{rob}$, que representa a união das duas imagens.

 $I_{mask} = I_{hisbr} \cup' I_{desup}$, aqui união representada por U' por se tratar de operação adapta.

A operação U' que gera I_{mask} é obtida pela união de dos elementos de I_{desvp} com aqueles elementos de I_{hisbr} que tenha ao menos um ponto de interseção com I_{desvp} . A Figura 3 mostra um exemplo das duas operação de união.



Figura 3 - Em I_{hisbr} temos a união simples entre I_{hist} e I_{rob} . Em I_{hisbr} o objeto de cor verde representa aquele que tem pelo menos um ponto de interseção com I_{desvp} e os de cor vermelha representam o contrário. Em I_{desvp} objetos de cor azul representam elementos sem interseção com I_{hisbr} e o de amarelo o contrário. Em I_{mask} observa-se que todos os objetos de I_{desvp} foram preservados enquanto que os de I_{hisbr} preservou apenas aquele que tinham alguma interseção com I_{desvp} (objeto de cor verde). Em I_{mask} o objeto de cor marrom representa a união dos elementos com interseção.

Após a operação de U' a imagem resultante I_{mask} é submetida a uma operação que a compara com sua respectiva esqueletização, então para cada objeto cujo esqueleto represente mais de 1/3 da sua área este então é removido de I_{mask} . Dessa forma objetos como linhas alongadas e pequenos pontos são removidas de I_{mask} . O resultado final pode ser observado no exemplo da Figura 4.



Figura 4 - Imagem I_{mask} final após remoção de artefatos com área do esqueleto superior a 1/3 da área do artefato que o originou. Observe artefatos em forma de traços (cor lilás) foram removidos.

Resultados e discussão

Para a realização dos experimentos foram extraídas 200 imagens da base descrita anteriormente, todas contendo artefatos estranhos. Para assegurar uma maior homogeneidade das amostras, das imagens selecionadas, 100 possuem presença de infiltrações pulmonares (pneumonia), 50 são de parênquima normal, e outras 50 não possuem diagnóstico conclusivo (imagens de baixa qualidade visual). As imagens estão no padrão RGB, com 768x1024 *pixels* e 255 níveis de intensidade em cada canal e armazenadas em formato JPEG²³. Nenhum pré-processamento foi realizado antes da submissão destas ao método proposto.

Para a condução dos testes foi utilizado o software Matlab R2014a²² em um computador com processador Core™ i5-2500 CPU @3.30Ghz, memória RAM 8 GB, com Windows 7 - 64 bits.

O tempo de execução total sobre as 200 radiografias foi de 147.7408 segundos dando uma média de 0.7387 ms, mostrando a eficiência do método proposto e sua habilitação para uso em sistemas de tempo real.

Para avaliação do método, as radiografias foram inspecionadas manualmente uma a uma e então foram contabilizados 3.422 artefatos a serem confrontados com aqueles detectados pelo método. Cada artefato detectado (região da sua máscara) foi definido como um *blob*^{3 24} seguindo os critérios de tamanho mínimo de 128 *pixels*, detecção válida somente se a máscara cobre todo o artefato da imagem original e contabilização única de artefatos sobrepostos.

A Figura 5 mostra exemplos de artefatos encontrados nas radiografias e os *blob's* correspondentes, cada *blob* tem uma única cor e corresponde a máscara do artefato identificado.

³ Blob – binary large object – do inglês grande objeto binário.









d) h) Figura 5 – Amostras de imagens com *blob's*: de a) a d) exemplos de radiografias com artefatos, de e) a h) imagens resultantes com *blob's* correspondentes aos artefatos identificados. Para avaliar a eficiência do método foram geradas matrizes de confusão para cada categoria a partir dos resultados da detecção, e calculadas as medidas de sensibilidade, precisão, taxa de erro e *FlScore*²⁵. As Tabelas 1.a), 1.b) e 1.c) apresentam o resumo das matrizes de confusão.

Tabela 1. - Resultados da detecção de artefatos, e suas medidas estatísticas nos seus respectivos conjuntos: a) amostra com 100 radiografias com infiltrações pulmonares (pneumonia), b) amostra com 50 radiografias parênquima normal (sem pneumonia) e c) amostras com 50 radiografias com diagnóstico indeterminado.

População	Verdadeiro Positivo	Falso Positivo	Falso Negativo				
1800	1759	28	41				
Sensibilidade	Precisão	Taxa de Erro	F1 Score				
0,98	0,98	0,02	0,98				
a)							
População	Verdadeiro Positivo	Falso Positivo	Falso Negativo				
827	812	11	15				
Sensibilidade	Precisão	Taxa de Erro	F1 Score				
0,98	0,99	0,02	0,98				
b)							
População	Verdadeiro Positivo	Falso Positivo	Falso Negativo				
795	727	27	68				
Sensibilidade	Precisão	Taxa de Erro	F1 Score				
0,91	0,96	0,08	0,94				
	C						

Conforme pode ser visto na Tabela 1, os menores valores de sensibilidade, precisão e F1 Score foram 0,91, 0,96 e 0,94, respectivamente, sendo estes valores significativamente altos. As taxas de erro foram bastante pequenas sendo 0,08 o maior de seus valores. Ressalta-se que a categoria com os piores resultados é exatamente o da Tabela 1.c) que refere-se às Radiografias Indeterminadas e que possuem imagens de pior qualidade.

Tabela 2 – Resultados da detecção de artefatos, e suas medidas estatísticas no conjunto de 200 radiografias sem separação por categorias.

População	Verdadeiro Positivo	Falso Positivo	Falso Negativo		
3422	3298	66	12		
Sensibilidade	Precisão	Taxa de Erro	F1 Score		
0,96	0,98	0,04	0,97		

A Tabela 2 traz um consolidado e representa as mesmas imagens sem separação por categorias. É possível observar que as variáveis Sensibilidade, Precisão e F1 Score mantêm-se significativamente altos e a Taxa de erro permanece bastante baixa

Conclusão

Neste trabalho foi apresentado um método para identificação de artefatos estranhos em radiografias de tórax, contendo como fases principais uma filtragem de *pixels* baseada em histograma, detecção de bordas com algoritmo de *Roberts e* filtragem espacial por filtro de desvio padrão. Tais algoritmos foram aplicados a uma base de dados de radiografias de tórax usadas para diagnóstico de pneumonia infantil. Para testes foram selecionadas 200 imagens distribuídas nas categorias, parênquima normal, infiltradas, e com diagnóstico inconclusivo todas contendo artefatos estranhos.

O método foi aplicado nas imagens selecionadas e conforme os resultados, observou-se que a taxa de acerto foi bastante elevada com aproximadamente 96% de sensibilidade e 98% de precisão. Ressalta-se que em 69% das imagens foram detectados corretamente todos os artefatos, e nas demais a quantidade de artefatos não detectados ou detectados incorretamente é significativamente pequena e desta forma espera-se que em processamentos adicionais possa reduzir a interferência causada por tais artefatos.

Destaca-se que a construção desse método foi projetada para ser bastante simples, e desta forma o processamento de cada imagem foi de aproximadamente 7*ms*, demonstrando que o método pode ser usado numa maior cadeia de processamento sem grandes impactos no desempenho.

Por fim, os resultados demostraram que o método é bastante satisfatório e pode ser recomendado para utilização futura como etapa de pré-processamento de imagens de radiografias de tórax.

Referências

- [1] Oliveira LLG, e Silva SA, Ribeiro LHV, de Oliveira RM, Coelho CJ, Andrade ALSS. Computeraided diagnosis in chest radiography for detection of childhood pneumonia. 2008 I. J. Medical Informatics, 77(8):555–564.
- [2] MACEDO SOD. Desenvolvimento de um sistema de auxílio ao diagnóstico de pneumonia na infância utilizando visão computacional. 2012 In Workshop de Visão Computacional.
- [3] Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K, et al. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. American Journal of Roentgenology, 2000, 174, pp. 71–74
- [4] Van Ginneken B, Stegmann MB, Loog M. "Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database". Medical Image Analysis, 2006, vol. 10, pp. 19-40.
- [5] ©Harry Shulman, Harry's Chest Radiology Atlas, use allowed for educational purposes. URL http://chestatlas.com/cover.htm (acessado em 28-jun-2016).
- [6] Chandrasekhar J. Chest X-ray Atlas. 2002. URL http://goo.gl/3vzIhv (acessado em 28-jun-2016).
- [7] Jaeger S, Candemir S, Antani S, Wáng Y-XJ, Lu P-X, G. Thoma. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. Quantitative Imaging in Medicine and Surgery. 2014; 4(6):475-477. doi:10.3978/j.issn.2223-4292.2014.11.20.
- [8] World Health Organization. Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children. 2001
- [9] Erondu OF. Challenges and peculiarities of paediatric imaging. 2013 INTECH Open Access Publisher.
- [10] Andrade M A, Prestes DP, Varani ML, Holsbach LR. TI Móvel na Saúde: Privacidade, Confidencialidade e Segurança. 2014 XXIV Congresso Brasileiro de Engenharia Biomédica - CBEB

- [11] Macedo HAS, Rodrigues VMCP. Programa de controle de qualidade: a visão do técnico de radiologia. 2009 Radiologia Brasileira, 42(1).
- [12] Xue Z, Candemir S, Antani S, Long LR, Jaeger S, Demner-Fushman D, et al (2015, November). Foreign object detection in chest X-rays. In Bioinformatics and Biomedicine (BIBM), 2015 IEEE International Conference on (pp. 956-961). IEEE.
- [13] Hogeweg L, Sánchez CI, Melendez J, Maduskar P, Story A, Hayward A, et al. Foreign object detection and removal to improve automated analysis of chest radiographs. 2013 Medical Physics, 40(7), 071901.
- [14] Hogeweg L, Sanchez CI, Van Ginneken B. Suppression of translucent elongated structures: Applications in chest radiography. 2013 IEEE Transactions on Medical Imaging, 32(11), 2099– 2113. http://doi.org/10.1109/TMI.2013.2274212
- [15] Iakovidis DK, Tsevas S, Savelonas MA, Papamichalis G. Image analysis framework for infection monitoring. 2012 IEEE Transactions on Biomedical Engineering, 59(4), 1135–1144. http://doi. org/10.1109/TBME.2012.2185049
- [16] Pinheiro MG, Canalli HL, de Figueiredo LR, Junior JE. Ambiente computacional para ensino de radiologia e diagnóstico por imagem: uma proposta para arquivo didático. 2009 Journal of Health Informatics, 1(2)
- [17] van Gemert-Pijnen L, Karreman J, Vonderhorst S, Verhoeven F, Wentzel J. Participatory Development via User-Involvement-A Case Study about the development of a Web-based Patientcommunication system about Methicillin-resistant Staphylococcus aureus. 2011 electronic Journal of Health Informatics, 6(4).
- [18] Marin HF. Sistemas de informação em saúde: considerações gerais. 2010 Journal of Health Informatics, 2(1).
- [19] Caritá EC, Nini RA, de Souza Melo A. Sistema de auxílio aos diagnósticos de enfermagem para vítimas de trauma no atendimento avançado pré-hospitalar móvel utilizando as Taxonomias NANDA e NIC. 2010 Journal of Health Informatics, 2(4).
- [20] Brasil, I. B. G. E. Instituto Brasileiro de geografia e Estatística. 2016
- [21] Gonzalez RC, Woods RC. Processamento digital de imagens 3.ed. São Paulo: 2010 Pearson Prentice Hall.
- [22] Guide, Matlab User's. The MathWorks Inc. 1998 Natick, MA, 4, 382.
- [23] Adams M D. The JPEG-2000 still image compression standard. 2001
- [24] Danker AJ, Rosenfeld A. Blob detection by relaxation. 1981 IEEE Transactions on Pattern Analysis and Machine Intelligence, (1), 79-92.
- [25] Powers DMW. Evaluation: from Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation. 2011 Journal of Machine Learning Technologies, 2(1), 37-63.

Contato

Afonso Ueslei da Fonseca Instituto de Informática – UFG. Sala 242. Alameda Palmeiras, Quadra D, Câmpus Samambaia, CEP 74690-900 -Goiânia – GO E-mail: afonso@inf.ufg.br Fone: (62) 3521-1245 / Fax: (62) 3521-1182