

Implementação de recursos inéditos para gerenciamento e utilização do *BancoWeb* - uma base de imagens para auxílio a pesquisas na área de mamografia digital

Pâmela S. Rodrigues, Bruno R. N. Matheus, Homero Schiabel

*Universidade de São Paulo, São Carlos – SP, Brasil

e-mail: pamelarodrigues@usp.br

Abstract – *The aim of this work is to implement improvements in a mammographic images database with public access, which may aid in the researches intended to diagnosis of breast cancer. The database contains images of several patients with reports and important information to classify the images, providing their consultation, visualization, clipping and classification.*

Palavras-chave: *images database, mammography, classification.*

Introdução

Os dois métodos eficazes para diagnóstico de câncer de mama são através do exame clínico de mama (ECM) e mamografia [1]. Estes exames passam por uma análise médica resultando em um diagnóstico. Estudos relatam que entre 10 a 15% das mamografias são erroneamente avaliadas [2], portanto, pesquisas e métodos são elaborados para auxílio no rastreamento e diagnóstico preciso.

Uma base de dados é um repositório de informação relacionado com determinado assunto ou finalidade, ou seja, uma coleção de dados [3]. Esse tipo de repositório na área médica é fundamental para possibilitar testes e pesquisas variadas; particularmente no atual estado da arte dos processamentos computacionais, são fonte para validação e verificação da eficácia de diversos tipos de técnicas e ferramentas para auxílio, por exemplo, ao diagnóstico. Na área de pesquisas relacionadas a *computer-aided diagnosis* (CAD), bases de imagens/dados são fundamentais para testes tanto de desenvolvimento como de desempenho de esquemas computacionais que visam ao auxílio na identificação de diversos tipos de patologias em variadas áreas da Medicina [4].

Assim, em função dessa necessidade, foi elaborada ao longo de muitos anos em nosso grupo uma base de dados composta essencialmente de imagens digitalizadas de filmes mamográficos obtidos em diferentes hospitais,

com as respectivas informações correspondentes aos laudos médicos dos exames. [5]. Essa base tem sido constantemente desenvolvida em termos tanto de cadastramento de novos casos, como de ferramentas de gerenciamento, tendo resultado num projeto que disponibiliza *on line* parte do conjunto de mais de 6000 imagens digitalizadas em formato .TIFF, com resolução de contraste de 12 bits e espacial de 0,075mm na maioria dos casos. O sistema de gerenciamento previamente desenvolvido – projeto *BancoWeb*¹ – permite que se faça também a busca de casos para o usuário baixar e montar seu próprio conjunto de imagens de interesse para testes de eventuais ferramentas de processamento de imagens digitais na área de mamografia [6].

Nossa base de imagens mamográficas tem armazenadas imagens obtidas em alguns hospitais com seus respectivos laudos e informações das pacientes (não confidenciais) [6]. Estas destinam-se ao uso em pesquisas, sobretudo para testes de desenvolvimento de técnicas de processamento.

A base possui cadastro de paciente, exame, *scanner*, aparelho de raios X, hospital, tipo de achado, patologia, assim como permite busca de imagens e recorte. O presente trabalho descreve a implementação de novos recursos para classificação de tipos de achado, melhorias na busca de imagens, inserção de imagens de novas fontes – imagens digitais diretas –, *download* do exame completo e o principal recurso, a integração com um esquema CADx (Computer-Aided Diagnosis), que corresponde a um programa de auxílio ao diagnóstico [7].

Estes recursos visam tornar a utilização da base mais ampla e versátil no seu sistema de gerenciamento – como, por exemplo, informar e distinguir a fonte de aquisição da imagem (CR, DR ou filme digitalizado) –, na busca de imagens, em usabilidade e disposição das informações. Também, existe a possibilidade de efetuar o *download* de todos os exames da paciente,

¹ Disponível em <http://lapimo.sel.eesc.usp.br/bancoweb/>.

categorizar os tipos de achado e classificar as imagens com o CADx.

Materiais e métodos

Usuários cadastrados na base têm permissão para visualizar as imagens, laudos, consultar, classificar, recortar e baixar todos os exames de uma determinada paciente. Nesse trabalho, houve quatro implementações importantes no sistema, que são descritas a seguir.

A primeira é a possibilidade da inserção de imagens na base considerando como dado-chave a sua fonte de aquisição, ou seja, indicar se é uma imagem do tipo CR (de *Computed Radiography*), ou DR (de *Digital Radiography*), ou se se trata de filme digitalizado. Existem atualmente na base cerca de 200 imagens digitais provenientes de sistemas DR e 1700 digitalizadas.

O segundo recurso introduzido é no resultado da busca. Quando se realiza uma busca são listadas todas as imagens encontradas com suas respectivas informações, como por exemplo: posição, mama direita ou esquerda, tamanho do pixel, tipo de sistema digitalizador, número de bits de contraste, entre outras informações. Também possui um totalizador da quantidade de imagens encontradas e quantidade de imagens na base, que atualmente está em 1900.

Na tela de busca de imagens é possível aplicar filtros e visualizar todas as informações da imagem, conforme *Figura 1*.

Segundo Matheus (2011), a partir de pesquisas na literatura, foram encontradas algumas bases de imagens com características que, porém, não atendem adequadamente as exigências para o desenvolvimento de esquemas CADx [7]. Por isso, o *BancoWeb*, nesta nova versão, traz o

recurso inédito que é a integração com um programa CADx.

Ao se buscar uma determinada imagem, existe a opção de “Recorte” que permite selecionar parte dela e baixar apenas essa região, em vez de baixar o arquivo da imagem inteira. Além disso, é possível “Classificar” uma determinada imagem a partir desse recorte: ao selecionar uma determinada área suspeita, o sistema automaticamente envia o recorte ao CADx que analisa a região e retorna se se trata de achado “Suspeito” ou “Normal”, conforme *Figura 2*. Atualmente, essa ferramenta permite enviar apenas um recorte para o CADx, mas pretende-se futuramente recortar até duas áreas de imagens para essa análise.



Figura 2. Recorte da imagem e classificação do CADx.

Estes sistemas têm o intuito de analisar as informações fornecidas para apresentar, com base nos dados da imagem processada, uma informação diagnóstica. Na mamografia as marcações das regiões de interesse podem indicar ao pesquisador informações de um possível achado, além de servir como auxílio na prática médica, por exemplo, em um ambiente de ensino.

A quarta implementação realizada é a categorização dos tipos de achado. Quando inserimos um novo exame com os dados da paciente e a imagem, podemos informar as características pertinentes a ela, como posição,

Total de Imagem(s) encontrada(s) : 1900 imagem(s)		Total de Imagem(s) cadastrada(s) no banco : 1900 imagem(s)	
Próximo >> 1475		NOVA PESQUISA MENU	
	Posição: Crânio-caudal Mama: direita Tamanho do Pixel: 0.075 Tempo de lactação: não consta Reposição Hormonal: não consta Scanner: Lumiscan 75 (Lumiasy)	Bits de Contraste: 12 BI-RADS: 2 Achado: sim Câncer na Família: não consta Nódulo Palpável: não consta Scanner: Lumiscan 75 (Lumiasy)	Sexo: feminino Ano de Nascimento: 01/01/1958 Idade Telarca: não consta Uso de Anticoncepcional: não consta Raio-X: Senographe 500t
	Posição: Crânio-caudal Mama: direita Tamanho do Pixel: 0.075 Tempo de lactação: não consta Reposição Hormonal: não consta Scanner: Lumiscan 75 (Lumiasy)	Bits de Contraste: 12 BI-RADS: 3 Achado: sim Câncer na Família: não consta Nódulo Palpável: não consta Scanner: Lumiscan 75 (Lumiasy)	Sexo: feminino Ano de Nascimento: 01/01/1958 Idade Telarca: não consta Uso de Anticoncepcional: não consta Raio-X: Senographe 500t
	Posição: Crânio-caudal Mama: esquerda Tamanho do Pixel: 0.075 Tempo de lactação: não consta Reposição Hormonal: não consta Scanner: Lumiscan 75 (Lumiasy)	Bits de Contraste: 12 BI-RADS: 3 Achado: sim Câncer na Família: não consta Nódulo Palpável: não consta Scanner: Lumiscan 75 (Lumiasy)	Sexo: feminino Ano de Nascimento: 01/01/1958 Idade Telarca: não consta Uso de Anticoncepcional: não consta Raio-X: Senographe 500t
	Posição: Médio-lateral Mama: esquerda Tamanho do Pixel: 0.075 Tempo de lactação: não consta Reposição Hormonal: não consta Scanner: Lumiscan 75 (Lumiasy)	Bits de Contraste: 12 BI-RADS: 3 Achado: sim Câncer na Família: não consta Nódulo Palpável: não consta Scanner: Lumiscan 75 (Lumiasy)	Sexo: feminino Ano de Nascimento: 01/01/1958 Idade Telarca: não consta Uso de Anticoncepcional: não consta Raio-X: Senographe 500t

Figura 1. Busca de imagens mamográficas

bits de contraste, tipo de digitalizador, e tipo de achado, entre outras informações. Também pode-se indicar se há alguma suspeita de anormalidade na mama e, em caso positivo, é possível descrever o tipo de achado, como: nódulo, calcificações, densidades, linfonodos, entre outros.

Segundo Moreira (2012), existem vários bancos de dados de imagens, alguns públicos e outros restritos [8]. No entanto, geralmente não atendem a todos os requisitos necessários para estudo na área do câncer de mama. Foi realizada anteriormente [6] uma comparação com essas bases e a maioria não apresenta uma busca funcional por essas imagens. Portanto, recursos desenvolvidos nessa base buscam superar algumas limitações de bases de dados existentes.

Para desenvolver os recursos citados acima, foram utilizadas as linguagens de programação PHP e Java Script, HTML como linguagem de marcação para interação com usuário e o banco de dados MySQL. Os testes foram realizados na maioria dos browsers como: Google Chrome, Internet Explorer 7 ou superior, Mozilla Firefox e Safari, mostrando-se funcional e atingindo todos os objetivos descritos anteriormente.

Resultados e Discussão

Com o objetivo de testar a viabilidade do uso prático das novas implementações no sistema, realizaram-se diversos testes com inclusão de exames com imagens de parâmetros diferentes e recortes de diversas áreas da mesma. Os resultados dos testes indicaram que a busca e classificação das imagens permitiram uma visualização de forma clara das informações e dados da mesma. O programa CADx associado é de fácil utilização e também oferece instruções ao usuário sobre sua usabilidade, exibindo um resultado evidente.

A utilização da fonte da imagem tornou os dados do exame ainda mais completo, pois indica se se trata de um sistema convencional (filme digitalizado), CR ou DR, além de permitir busca pelo tipo de fonte de aquisição. Atualmente, no *BancoWeb* existem cadastradas 196 imagens DR, 1704 por filme digitalizado, e até o momento não há nenhuma CR, conforme *Figura 3*.

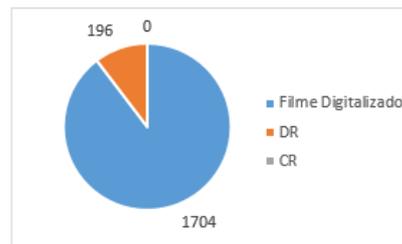


Figura 3. Gráfico ilustrativo da quantidade de imagens por fonte de aquisição na base *on line*.

Já a busca e lista de resultados apresentadas através de um recurso de grade horizontal pode exibir os dados com melhor visualização e clareza. Além disso, oferece um recurso de *download* do exame completo, o que antes não existia.

A integração com o esquema CADx é uma funcionalidade totalmente nova no sistema e de extrema importância para o auxílio ao rastreamento e diagnóstico do médico. Trata-se de um recurso inédito em termos de sistema gerenciador de base de dados, pois, além da possibilidade de levantamento das imagens de mamografia, pode-se agora também utilizar uma ferramenta computacional (o CADx associado [7]) para auxiliar na procura por um achado em uma determinada área selecionada na imagem.

A classificação do tipo de achado está atualmente em desenvolvimento e tem como intuito indicar para cada arquivo de imagem inserido na base se há ou não achado de interesse clínico, qual o tipo caso exista, e todas as características adicionais pertencentes a este tipo de achado, conforme *Figura 4*.

De fato, todas as implementações realizadas têm sido bem difundidas, de modo que, ao realizar uma consulta via comandos SQL no banco de dados do *BancoWeb*, observa-se que a quantidade de usuários cadastrados no sistema tem aumentado significativamente (em média têm sido registrados pedidos de cadastramento de cerca de 20 novos usuários por mês). Atualmente a estrutura da base que está disponível *on line* contém 671 usuários ativos, sendo de acesso livre e gratuito.

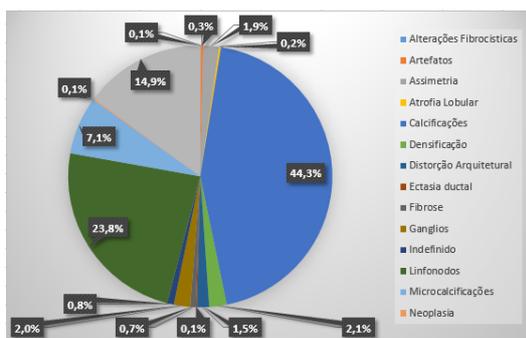


Figura 4. Desenvolvimento da classificação de tipos de achado.

Conclusões

Neste artigo apresentamos uma aplicação para pesquisa e rastreamento a partir de nossa base de imagens mamográficas. As implementações realizadas propiciaram a fácil utilização do sistema de gerenciamento, permitindo uma rápida classificação e busca das imagens. Além disso, este trabalho aperfeiçoou o sistema do *BancoWeb* para o cadastro e armazenamento dos dados, auxiliando também os pesquisadores a melhorar a qualidade e processamento de imagens para alcançar soluções que possam beneficiar o diagnóstico preciso.

Além de concluir a implementação do tipo de achado para classificação dos casos, pretende-se, futuramente, agregar a inserção de imagens de *phantom*, que simulam as estruturas presentes em uma mamografia [9].

Agradecimentos

Agradecimentos à FAPESP pelo apoio financeiro.

Referências

- [1] OSHIRO, M. L., et al. **Câncer de mama avançado como evento sentinela para avaliação do programa de detecção precoce do câncer de mama no Centro-Oeste do Brasil**. Revista Brasileira de Cancerologia 60.1 (2014): 15-23.
- [2] PEIXOTO, J. E.; CANELLA, E.; AZEVEDO, A. C. **Mamografia: da prática ao controle**. Rio de Janeiro: Gráfica Esdeva, 2007.
- [3] CALDEIRA, C. P. **Introdução ao Modelo de Dados Relacional**. Universidade de Évora. Disponível em: <<http://www.portalwebmarketing.com/Tecnologia/Introdu%C3%A7%C3%A3oaoModelodeDadosRelacional/MDRDefini%C3%A7%C3%A3oBas eDadosRelacional/tabid/654/Default.aspx>> Acesso em: 11 de set. 2017.

[4] SCHIABEL, H. Esquemas CAD: uma análise dos seus aspectos e aplicações como ferramenta de auxílio ao diagnóstico em mamografia – Cap. do livro **“Tecnologias, Técnicas e Tendências em Engenharia Biomédica”** (ISBN 978-85-7917-289-2), Canal 6 Editora (Ed. Adriano O. Andrade; Alcimar B. Soares; Alexandre Cardoso; Edgard A. Lamounier), p. 34-54, 2014.

[5] BENATTI, R. H. **Elaboração de um banco de imagens mamográficas digitalizadas**. (Dissertação de Mestrado). São Carlos: Escola de Engenharia de São Carlos da Universidade de São Paulo, 2003.

[6] MATHEUS, B. R. N.; SCHIABEL, H. – **Online mammographic images database for development and comparison of CAD schemes** – Journal of Digital Imaging (Publicado on-line em 18/05/2010 – doi: 10.1007/s10278-010-9297-2), v. 24, n. 3, p. 500-506, 2011.

[7] SCHIABEL, H.; MATHEUS, B. R. N.; ANGELO, M. F.; PATROCINIO, A. C.; VENTURA, L. – **A prototype of mammography CADx scheme integrated to imaging quality evaluation techniques** – SPIE Medical Imaging 2011: Computer-aided Diagnosis, v. 7963, p. 796323-1 - 796323-12 – Orlando, FL (USA) – Feb, 12-17, 2011.

[8] MOREIRA, I. C. et al. **INbreast: Toward a Full-field Digital Mammographic Database**. Academic Radiology, Vol 19, Nº 2, Feb 2012.

[9] SOUSA, M. A. Z.; SIQUEIRA, P. N.; SCHIABEL, H. **Evaluation of a simulation procedure designed to recognize shape and contour of suspicious masses in mammography** – SPIE Medical Imaging 2015: Image Perception, Observer Performance and Technology Assessment, v. 9416, p. 941618-1 – 941618-6 – Orlando, FL, USA, Feb, 22-26, 2015.