

ALIMENTAÇÃO, MANUTENÇÃO E SEGURANÇA DO BANCO DE DADOS DE IMAGENS TERMOGRÁFICAS DA UFPE

José Reginaldo Alves de Queiroz Júnior¹; Rita de Cássia Fernandes Lima²

¹Estudante do Curso de Engenharia de Controle e Automação – CTG – UFPE; E-mail: reginaldoqueirozjr3@gmail.com,

²Docente/pesquisador do Depto de Engenharia Mecânica – CTG – UFPE. E-mail: ritalima@ufpe.br.

Sumário: A termografia é uma técnica que vem sendo proposta como um exame complementar na detecção do câncer de mama. O tratamento das imagens termográficas visa segmentar a região da mama, diferenciando-a do ambiente externo e das regiões corporais que não correspondem à área de interesse. Para armazenar prontuários e imagens das pacientes, foi desenvolvido um banco de dados (BD), que permite consulta por parte de profissionais ligados à pesquisa, mediante permissão prévia. Neste trabalho, foi estudada uma forma de garantir a manutenção e alimentação deste BD, visando aumentar a segurança do sistema. Além disso, foi analisada a influência da umidade do ar, sobre a segmentação e a classificação das patologias presentes nas imagens termográficas armazenadas.

Palavras-chave: banco de dados; câncer de mama; classificação; imagens termográficas; segmentação; umidade do ar;

INTRODUÇÃO

A detecção precoce do câncer de mama deve ser priorizada em todas as regiões do Brasil. No país, a estimativa para o ano de 2014, que será válida também para o ano de 2015, aponta para a ocorrência de aproximadamente 57 mil casos novos no país [1]. A detecção precoce é um fator de diminuição da mortalidade do câncer de mama, pois abre possibilidade para alternativas de tratamento, incluindo cirurgias menos agressivas.

A termografia pode se tornar importante na detecção precoce de câncer de mama como uma análise qualitativa e quantitativa, pois pode ser usada na validação de suspeitas médicas não-confirmadas por outras formas de diagnóstico, e até mesmo ajudar a confirmar um problema fisiológico existente. Para o caso do câncer da mama, a termografia, associada à mamografia e ao exame clínico prévio, pode elevar a sensibilidade da detecção deste tipo de câncer para aproximadamente 98% [2].

A proposta desse trabalho é o melhoramento do banco de dados (BD) de imagens termográficas da UFPE através da sua manutenção e alimentação. Adicionalmente, foi realizado um estudo para verificar a relação entre a umidade relativa do ar e a segmentação e a classificação de uma imagem.

MATERIAIS E MÉTODOS

As imagens termográficas utilizadas foram obtidas com uma câmera de infravermelho FLIR S45, adquirida pelo Departamento de Engenharia Mecânica da Universidade Federal de Pernambuco (DEMEC/UFPE). A análise térmica das imagens foi realizada através do software *FLIR QuickReport* e processadas no MATLAB, através de uma rotina desenvolvida por Dourado Neto [4].

A confiabilidade na aquisição e o processamento adequado das imagens dependem da padronização conveniente dos processos. Com este objetivo, as imagens são armazenadas em três locais diferentes, incluindo o servidor da UFPE, onde está sediado o BD.

Aquisição de imagens e manutenção do banco de dados: O processo de aquisição das imagens térmicas das mamas é simples e indolor. O procedimento é realizado em uma sala do HC/UFPE, onde foi instalado um aparato mecânico para auxiliar no posicionamento da

paciente. Durante a aquisição das imagens são realizadas duas séries de imagens, a primeira (D1) a uma distância fixa, com a câmera posicionada mais longe da paciente. A segunda série (D2) é realizada com a câmera mais próxima da paciente, com distância que pode variar entre as imagens.

O uso do BD visa o armazenamento e a transferência de dados do paciente no intuito de reduzir erros clínicos e aumentar a segurança dos resultados dos exames. Os dados adquiridos foram armazenados em um sistema desenvolvido em PHP, utilizando gerenciamento baseado no padrão *Oracle*, que foi usado como linguagem de manipulação de dados. Após finalizar o cadastro da paciente, uma pasta é criada no servidor e nomeada com o número do prontuário da mesma. A imagem anexada é então movida para esta pasta. Deste modo, cada paciente mantém uma pasta única, identificada a partir do seu número de prontuário e contendo suas imagens termográficas. Paciente que por algum motivo não possua número de prontuário, recebe no momento do cadastro um número fictício dado pelo sistema. Este número fictício nomeará a pasta acima referida [3].

Durante o processo de estudo e da alimentação do BD, houve um problema técnico nele ou no servidor, que ocasionou a sua inatividade durante um longo período em 2014. Após algumas análises e tentativas, foi possível restaurar o BD, mesmo que com algumas falhas quanto ao número de pacientes cadastradas.

Seleção e correção das umidades das imagens: De acordo com a teoria de espalhamento da radiação na atmosfera, quanto maior for a quantidade de água presente no meio, maior será a atenuação da radiação por espalhamento, influenciando nas temperaturas registradas nos termogramas. Dourado Neto [4] identificou uma grande variação entre as umidades na sala de exames, medidas antes de 2010 e os valores mais recentes. Nesta oportunidade, houve a troca do termo-higromêtro utilizado após a quebra do equipamento mais antigo (Fig. 1).

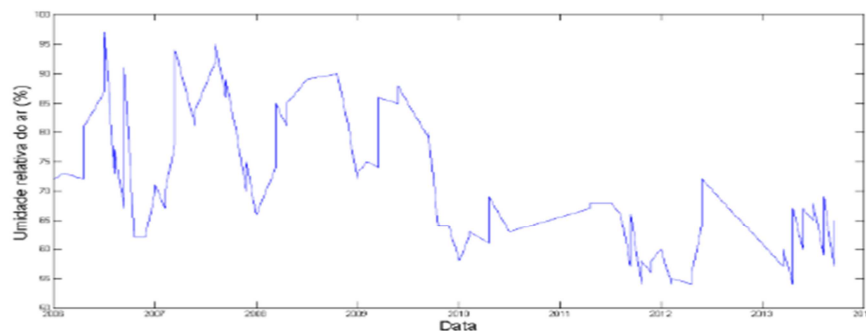


Figura 1 – Umidade relativa do ar para cada data de aquisição (Fonte: DOURADO, 2014).

Baseado nessa figura foram calculadas as médias trimestrais a partir de 2010, usando-se uma rotina em linguagem C. Esses valores passaram a ser usados nas imagens termográficas anteriores a 2010, para correção das temperaturas. As umidades relativas ajustadas trimestrais obtidas para um ano foram: 1º trimestre – 60,9, 2º trimestre – 63,9, 3º trimestre – 64,2 e 4º trimestre – 58,2. Vale ressaltar que a umidade medida no momento da aquisição das imagens é também usada como parâmetro interno da câmera que, a partir desse dado e de outros, realiza cálculos para apresentar os valores medidos para a temperatura do objeto.

Uso do segmentador e do classificador automático: O sistema de classificação estatístico, usado para classificação das patologias mamárias, foi desenvolvido por Dourado Neto [4]. Inicialmente identificam-se e isolam-se os objetos de interesse, no processo de segmentação. Da imagem segmentada, foram extraídas as características estatísticas relevantes que permitirão a identificação de cada patologia mamária. Dourado Neto [4] desenvolveu um programa computacional em MATLAB que gerencia a subrotina de segmentação e da extração de características; os testes estatísticos para a seleção de características; e a classificação das patologias utilizando diferentes classificadores. Com uso do sistema descrito, fo-

ram segmentadas e classificadas as imagens que tiveram sua umidade ajustada (Bloco 1). Em seguida, foram também classificadas e segmentadas as imagens com as umidades anteriores (Bloco 2). Esse procedimento visa comparar os resultados obtidos a fim de analisar a influência da umidade sobre a classificação dos termogramas.

RESULTADOS

Após a segmentação das imagens dos Blocos 1 e 2, os resultados foram divididos em três grupos. O primeiro contém as imagens que apresentaram uma melhoria na segmentação, definindo melhor as regiões de interesse após a mudança da umidade (Fig. 2). O segundo grupo possui as imagens que, após o reajuste da umidade, regrediram quanto à qualidade da segmentação, isto é, foram perdidas áreas de interesse, antes presentes na segmentação (Fig. 3). E por fim, o terceiro grupo apresentava as imagens que tiveram poucas mudanças quanto à segmentação pós-ajuste de umidade.

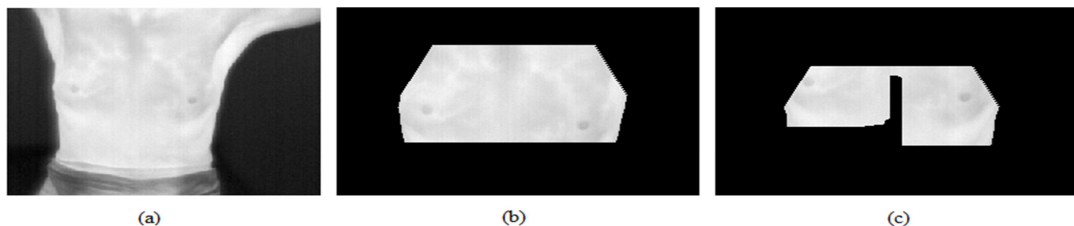


Figura 2 – Representante do Grupo 1, onde (a) Termograma em escala de cinza, (b) Segmentação original e (c) Segmentação após o ajuste de umidade.

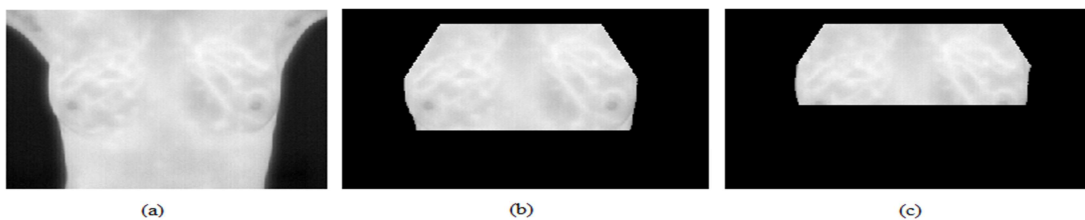


Figura 3 – Representante do Grupo 2, onde (a) Termograma em escala de cinza, (b) Segmentação original e (c) Segmentação após o ajuste de umidade.

Com uso do classificador estatístico citado, os Blocos 1 e 2 foram classificados com as características definidas como c_6 e c_5 , que melhor separavam as classes Maligno e Não-Maligno, segundo [4]. Depois, esses mesmos blocos, foram classificados com características que melhor se adaptaram à nova amostra analisada. As Tabelas 1 e 2 apresentam os resultados obtidos após a classificação das amostras com a umidade ajustada e a anterior, respectivamente.

Tabela 1 – Resultados da classificação do Bloco 1

Características	Número de imagens	Acertos	Sensibilidade	Especificidade	Acerto	Acerto	Acerto
					Benigno	Cisto	Normal
[6 5]	80 (27, 33, 13, 7)	56,25 %	51,85 %	58,49 %	69,70 %	61,54 %	00,00 %
[6 10]	80 (27, 33, 13, 7)	67,50 %	70,37 %	66,04 %	75,76 %	61,54 %	28,57 %

Tabela 2 – Resultados da classificação do Bloco 2

Características	Número de imagens	Acertos	Sensibilidade	Especificidade	Acerto	Acerto	Acerto
					Benigno	Cisto	Normal
[6 5]	80 (27, 33, 13, 7)	53,75 %	51,85 %	54,72 %	69,70 %	38,46 %	14,29 %
[6 18]	80 (27, 33, 13, 7)	70,00 %	59,26 %	75,47 %	84,85 %	61,54 %	57,14 %

As imagens foram alocadas em classes Maligno, Benigno, Cisto e Normal. Foi usado o classificador *Support Vector Machine* (SVM), que apresentou os melhores resultados [4]. Finalmente, foi considerada toda a amostra usada por Dourado Neto, com a substituição

das 80 imagens antigas pelas que tiveram sua umidade ajustada. Usando as características c_6 e c_5 , obtiveram-se os dados contidos na Tabela 3.

Tabela 3 – Comparação dos resultados com amostra completa

Situação	Número de imagens	Acertos	Sensibilidade	Especificidade	Acerto	Acerto	Acerto
					Benigno	Cisto	Normal
Bloco 1.	234 (43,79,42,70)	78,21 %	62,79 %	81,68 %	79,75 %	76,19 %	87,14 %
Bloco 2	234 (43,79,42,70)	79,49%	67,44%	82,20%	81,01%	76,19%	87,14%

DISCUSSÃO

Quando o BD parou de funcionar, havia cerca de 320 pacientes cadastradas, muitas com diagnóstico concluído. Após a restauração do BD, apenas 120 pacientes do cadastro anterior foram restabelecidas. Um problema que surgiu após a restauração do BD está relacionado aos usuários administradores e está sendo resolvido.

Para o estudo da umidade, foram usadas as amostras de 2007, que apresentou a maior discrepância nos valores da variável, comparados com os anos cujos índices foram usados para a normalização (2010-2013). Devido à diminuição da amostra, pode ter ocorrido algum desbalanceamento na base de dados, ou seja, a distribuição de imagens nas quatro patologias pode ter ficado irregular, interferindo na taxa de acerto do classificador.

CONCLUSÕES

Este trabalho propôs o estudo do BD de imagens termográficas de mama da UFPE visando: o controle da alimentação dos dados, uma forma de melhorar a manutenção; e maneiras de suprir a sua segurança. Paralelo a isso, foi visto como a umidade relativa do ar influenciou na qualidade da segmentação e nos resultados da classificação das imagens contidas no BD.

Antes de manipular o BD, este foi alimentado com novas pacientes, para aumentar a base de dados. Contudo, em 2014, o BD ficou fora ar devido a algum problema nele ou na máquina onde o servidor está instalado. A restauração só foi possível graças a um módulo desse mesmo BD, que estava instalada em outra máquina do laboratório. Simultaneamente à restauração do BD, foi realizado um estudo a fim de encontrar a relação entre a umidade e o processamento das imagens termográficas de mama. A metodologia adotada apresentou bons resultados preliminares, tendo-se verificado que a área segmentada automaticamente tornou-se mais precisa. A classificação da nova amostra apresentou maiores taxas de acerto e sensibilidade (proporção de indivíduos com o problema e que são classificados corretamente pelo teste), se comparados às obtidas pela mesma amostra, com umidade não-coriçada.

AGRADECIMENTOS

O aluno agradece ao CNPq pelo apoio financeiro para a pesquisa e a Hugo Dourado de Moraes Neto.

REFERÊNCIAS

- [1] INCA. Estimativas 2014: Incidência de câncer no Brasil. Brasil: Ministério da Saúde. Instituto Nacional de Câncer. Rio de Janeiro, 2014.
- [2] NG, E. Y-K., SUDHARSAN, N. M., "Computer simulation in conjunction with medical thermography as an adjunct tool for early detection of breast cancer", *BMC Cancer*, V. 4, 2004.
- [3] Araújo, M.C., Lima, R.C.F., Santos, F.G.S. 2009. Uso de um banco de dados como ferramenta auxiliar da detecção precoce de câncer de mama. XXX Iberian Latin American Congress on Computational Methods in Engineering.
- [4] Dourado Neto, H.M. 2014 Segmentação e análise automáticas de termogramas: um método auxiliar na detecção de câncer de mama. *Dissertação de mestrado*. Departamento de Engenharia Mecânica, Universidade Federal de Pernambuco.
- [5] Araújo, M.C. de, Lima, R.C.F. de, Souza, R.M.C.R. Interval symbolic feature extraction for thermography breast cancer detection, *Expert Systems with Applications*, ISSN 0957-4174, v 41, pp 6728-6737, 2014.