

## AValiação de Classificadores para o Diagnóstico de Lesões de Mama em Ultrassonografia

M. A. Massote\*, W. Gómez\*\*, W. C. A. Pereira\*, M. R. Petraglia\*\*\*

\*Programa de Engenharia Biomédica, COPPE/UFRJ, Rio de Janeiro, Brasil

\*\*Laboratório de Tecnologías de Información/CINVESTAV-IPN, Ciudad Victoria, México

\*\*\* Programa de Engenharia Elétrica, COPPE/UFRJ, Rio de Janeiro, Brasil

e-mail: massotemariana@gmail.com

**Resumo:** A fim de avaliar o desempenho de dois classificadores, linear e quadrático, foi desenvolvido um método de análise de imagens de ultrassom de lesões de mama para diagnóstico assistido por computador (CAD). Foi realizada a segmentação da parte da imagem relativa à lesão, em seguida foram extraídas cinco características morfológicas da imagem resultante. Os dois algoritmos para os classificadores foram treinados e avaliados usando o método Bootstrap .632+. A área sob a curva ROC (AUC) foi o índice usado na avaliação do desempenho dos classificadores. Os resultados mostraram que o método linear obteve melhor performance com  $AUC=0,95$ , enquanto o classificador quadrático atingiu  $AUC=0,88$ .

**Palavras-chave:** Ultrassom, lesões de mama, processamento de imagens, segmentação, classificadores

**Abstract:** To evaluate the performance of the linear and quadratic classifiers, an ultrasound image analysis method was developed for computer-aided diagnosis (CAD). First, the lesion segmentation was performed to extract the region of interest. Thereafter, five morphological features were computed. Both classifiers were trained and evaluated by using the .632+ bootstrap method. The area under the ROC curve (AUC) was used to evaluate the classification performance. The results pointed out that the linear method reached the best performance with  $AUC=0.95$ , whereas the quadric classifier attained  $AUC=0.88$ .

**Keywords:** Ultrasound, breast lesions, image processing, segmentation, classifiers

### Introdução

Muitos métodos foram testados como coadjuvantes da mamografia para diagnósticos, mas somente o ultrassom foi considerado eficaz [1]. Ele permite visualizar a estrutura interna dos tecidos, mesmo que o tecido seja denso, e distingue claramente cistos de lesões sólidas, casos nos quais a mamografia não é eficaz [2] [3]. Além disso, o ultrassom emite radiações não ionizantes (som de alta frequência), é barato e portátil.

O ultrassom é usado para examinar um ponto específico do órgão. A imagem gerada e armazenada não compreende uma visão geral devido à variação entre as imagens geradas por diferentes radiologistas. Por isso é um exame que depende da experiência do médico radiologista e é subjetivo [4].

Trabalhos anteriores avaliaram alguns classificadores para imagens de ultrassom de lesões de mama. Ainda que alguns autores avaliem classificações para características morfológicas, esses estudos diferem quanto às características extraídas, quantidade de características usadas e métodos de classificação. No trabalho de Huang Y-L et al. [5], foram usadas 19 características morfológicas e o classificador por máquina de vetores de suporte, apresentou  $AUC = 0,909$ ,  $SEN = 0,941$  e  $ESP = 0,774$ . No trabalho de Alvarenga AV et al. [6], foram extraídas 7 características morfológicas e foi usada a Análise Discriminante Linear de Fisher (LDA) para classificar os dados. Obteve-se  $AUC = 0,86$ ,  $SEN = 0,83$  e  $ESP = 0,85$ .

Esse trabalho teve como objetivo avaliar dois classificadores um linear e outro quadrático para classificar as características extraídas da lesão.

### Materiais e métodos

Foi desenvolvido um sistema CAD (Diagnóstico Assistido por Computador) de lesões de mama em imagens de ultrassom a fim de diminuir a subjetividade ou melhorar a acurácia dos diagnósticos e contribuir com uma segunda opinião para o radiologista[7].

O CAD desenvolvido constituiu em: pré-processamento, segmentação, extração de características e classificação.

O banco de dados utilizado consiste de 641 imagens de lesões de mama em ultrassom, uma para cada paciente, adquiridas durante a rotina de diagnóstico no Instituto Nacional do Câncer(INCA), Rio de Janeiro, Brasil, conforme aprovado pelo Comitê de ética do mesmo (38/2001). As imagens foram obtidas no equipamento Sonoline Sienna® (Siemens, Erlangen, Germany) usando transdutor ultrassônico B-mode 40mm com feixe linear de 7,5 MHz. A resolução do feixe é de 0,5mm e o tamanho do pixel é 0,33 mm/pixel. As imagens foram capturadas diretamente de um sinal de vídeo de 8 bits e salvas em formato TIFF. Todos os

228 carcinomas e as 413 lesões benignas possuem biópsias realizadas pelo INCA.

O CAD se baseia nas características numéricas extraídas para classificar as lesões em benignas ou malignas. O modelo computacional usa programação em Matlab e consiste dos seguintes passos:

**Região de interesse** – É feito um corte na imagem para separar a região de interesse e diminuir o tamanho da figura para agilizar o processamento posterior. É pedido que o radiologista marque quatro pontos para os limites da lesão que serão usados para localizar o centroide e cortar a imagem em, no máximo, duas vezes o tamanho da lesão (Figura 1).

**Filtro de Mediana Híbrido** – É um filtro de estatística de ordem e faz parte do pré-processamento da imagem tendo a função de diminuir o *speckle* da imagem ao mesmo tempo que deve preservar detalhes das bordas (Figura 1). Tem a vantagem de borrar menos a imagem que os filtros lineares e preservar melhor as bordas que o filtro de mediana "tradicional". Ele consiste em substituir o pixel central pelo valor da mediana do pixel central, da mediana das diagonais (com exceção do centro) e da mediana dos valores verticais e horizontais ao redor do pixel central. Portanto, o tamanho do filtro deve ser ímpar para que haja sempre um pixel central.

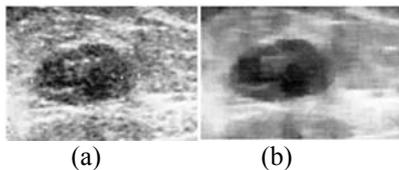


Figura 1: (a) Região de interesse cortada. (b) Redução de ruído com filtro de mediana híbrido com *kernel* de tamanho  $15 \times 15$  pixels.

**Função Gaussiana Constante** – Usamos a seguinte equação para calcular uma função Gaussiana com mesmo centro da lesão que foi multiplicada pelo negativo da imagem para atenuar os pixels distantes (Figura 2):

$$G_c(x, y) = \frac{\exp\left\{-\frac{(x-\mu_w)^2}{2\sigma_w^2} - \frac{(y-\mu_h)^2}{2\sigma_h^2}\right\}}{2\pi(\sigma_w^2\sigma_h^2)^{1/2}} \quad (1)$$

onde  $\sigma_w^2$  e  $\sigma_h^2$  são as variâncias definidas pela marcação do radiologista para delimitação do tumor dos pontos nos eixos  $x$  e  $y$ , e o par  $(\mu_w, \mu_h)$  é o centroide da lesão.

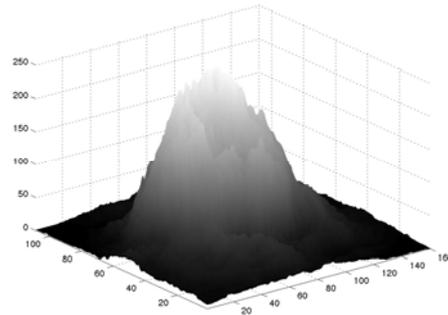


Figura 2: Função gaussiana multiplicada pelo negativo da imagem.

**Segmentação** – Gera uma imagem binária (Figura 3) do contorno da lesão a fim de permitir a extração de características da lesão quantitativamente. Essa etapa é importante, pois as características da borda da lesão devem ser preservadas, apesar de todos os filtros aplicados antes da segmentação, já que a extração de características se baseia nessa imagem para gerar parâmetros numéricos que permitam comparar as lesões e classificá-las em benignas ou malignas [8].

Após o cálculo da Gaussiana a imagem resultante é redimensionada para a faixa de 0-255 para extrair as possíveis margens. A Média Radial Derivativa (ARD) é calculada em cada uma das possíveis margens, sendo definida como:

$$ARD(\Gamma) = \frac{1}{N} \sum_{x,y \in \Gamma} \nabla I(x, y) \cdot u(x, y) \quad (2)$$

onde  $\Gamma$  é a possível margem,  $N$  é o número de pixels em  $\Gamma$ ,  $\nabla I(x, y)$  é o gradiente da imagem filtrada e  $u(x, y)$  é um vetor unitário na direção do gradiente.

O valor máximo da ARD fornece uma medida da máxima mudança de intensidade entre a lesão e o fundo [9]. Portanto, a margem associada ao máximo da ARD é selecionada como borda final da lesão (Figura 3).

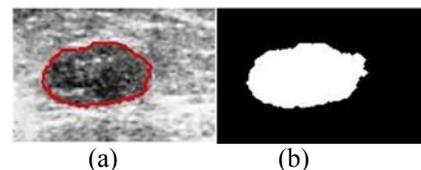


Figura 3: (a) Borda final da lesão em vermelho (b) Imagem binária extraída.

**Extração de Características** – As características mais comuns extraídas de imagens de ultrassom são as características morfológicas e as características da textura. Como já mostrado, as características morfológicas descrevem parâmetros da forma das lesões para posterior classificação delas [10]. Nesse trabalho foram extraídos cinco parâmetros morfológicos que foram previamente selecionadas pelo nosso grupo de pesquisa [11]:

1. Elipse Equivalente – Esse algoritmo calcula o perímetro e o ângulo da elipse equivalente, ou seja, o

perímetro de uma lesão de mesma área depois de calcular os coeficientes da equação geral da elipse.

2. Orientação – Calculando o ângulo  $d$  do eixo maior da lesão, podemos identificar se essa lesão está mais espalhada no mesmo tecido (horizontal) ou se estaria infiltrando nos tecidos adjacentes (vertical).

3. Número de Protuberâncias e Depressões – Esse número é dado pela quantidade de máximos e mínimos em relação ao polígono convexo relativo a uma lesão. Quanto mais lobulada a lesão maior o número de protuberâncias e depressões.

4. Razão Largura-Profundidade – Lesões benignas tendem a ser mais largas do que profundas, resultando em um valor de razão largura-profundidade mais baixo que lesões malignas que em geral, infiltram tecidos vizinhos.

5. Razão de Superposição – A razão de superposição é a relação entre a interseção entre a área do polígono convexo da lesão e sua área, ou seja, quanto maior esse valor significa que o polígono convexo coincide mais com a lesão e que portanto, é menos lobulado.

**Classificação** – Foram analisadas as respostas para dois classificadores: Análise Discriminante Linear de Fisher (LDA) e Análise Discriminante Quadrática (QDA). O LDA classifica uma amostra  $\mathbf{x}$  como:

$$D_{LDA}(\mathbf{x}) = \left[ \mathbf{x} - \frac{1}{2}(\mu_1 + \mu_2) \right]^T \mathbf{S}^{-1}(\mu_1 - \mu_2) \quad (3)$$

onde  $\mu_1$  e  $\mu_2$  são as médias das classes benigno e maligno a matriz de covariância conjunta é:

$$\mathbf{S} = \frac{1}{n_1 + n_2 - 2} \left[ (n_1 - 1)\Sigma_1 + (n_2 - 1)\Sigma_2 \right] \quad (4)$$

onde  $\Sigma_1$  e  $\Sigma_2$  são matrizes de covariância das classes benigno e maligno respectivamente, e  $n_1$  e  $n_2$  são os números de amostras em cada uma dessas classes.

O classificador QDA se define como:

$$D_{QDA}(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + b^T \mathbf{x} + c \quad (5)$$

onde os coeficientes são:

$$\mathbf{A} = \frac{1}{2}(\Sigma_2^{-1} - \Sigma_1^{-1}) \quad (6)$$

$$\mathbf{b} = \Sigma_1^{-1} \mu_1 - \Sigma_2^{-1} \mu_2 \quad (7)$$

$$\mathbf{c} = \frac{1}{2} \log \left( \frac{|\Sigma_2|}{|\Sigma_1|} \right) + \frac{1}{2} (\mu_2^T \Sigma_2^{-1} \mu_2 - \mu_1^T \Sigma_1^{-1} \mu_1) \quad (8)$$

**Avaliação** – O algoritmo Bootstrap .632+ [12] foi usado para selecionar aleatoriamente amostras e criar 50 conjuntos para treinar e testar os classificadores.

A área sob a curva ROC (AUC), a sensibilidade (SEN) e a especificidade (ESP) foram usadas para avaliar o desempenho dos classificadores. Um valor de “1” para as três medidas significa classificação perfeita.

Para testar a significância estatística entre os valores das médias dos classificadores foi usado o método Welch's t-teste. Foi usado este teste pois as amostras possuem distribuição normal, conferidos pelo Shapiro-Wilk para avaliar normalidade dos dados e variâncias diferentes conferidas pelo F-test. Todos os testes estatísticos usaram um Intervalo de Confiança de 95% ( $\alpha=0,05$ ).

## Resultados

Cada conjunto Bootstrap é usado no treino do LDA e do QDA, e testado com as imagens que não foram utilizadas no conjunto de treino gerando assim dados para a avaliação estatística.

Os valores das médias ( $\mu$ ) e desvio padrão ( $\sigma$ ) para cada classificador estão na Tabela 1.

Tabela 1: Resultados de desempenho dos classificadores para 50 conjuntos Bootstrap.

Medida	LDA		QDA	
	$\mu$	$\sigma$	$\mu$	$\sigma$
AUC	0,951	0,008	0,887	0,014
SEN	0,892	0,027	0,844	0,036
ESP	0,900	0,022	0,843	0,029

O classificador LDA foi melhor pois apresentou um desempenho maior para as três medidas avaliadas. Também a variação ao redor das médias no classificador LDA foi menor do que no QDA indicando que a resposta é mais estável.

## Discussão e Conclusão

A classificação usando as cinco características extraídas da imagem, foi diferente. Usando esses parâmetros, o LDA apresentou resultado melhor que o QDA. O QDA se ajusta mais às amostras de treino que o LDA, e portanto, o LDA funciona melhor para dados com muito ruído, já que funciona de uma forma mais generalista.

O LDA mostrou AUC = 0,951 e ESP = 0,900 maior que o trabalho de Huang Y-L et al. [5], que usou máquina de vetores de suporte, e obteve AUC = 0,909 e ESP = 0,774, porém com SEN = 0,892 menor que SEN = 0,941. Obteve-se AUC, SEN e ESP maiores que AUC = 0,86, SEN = 0,83 e ESP = 0,85 de Alvarenga AV et al. [6] que também usou LDA.

As cinco características que foram previamente selecionadas pelo nosso grupo de pesquisa [11], já garantem uma boa performance de classificação.

Como trabalho futuro, pretende-se repetir os testes usando outros parâmetros, como exemplo, extração de características da textura.

## Agradecimentos

Agradecemos o Departamento de Radiologia do Instituto Nacional do Câncer pelas imagens e resultados de biópsias utilizadas neste trabalho e ao CNPQ pelo apoio financeiro.

## Referências

- [1] Skaane P. Ultrasonography as Adjunct to Mammography in the Evaluation of Breast Tumors [dissertação]. Oslo: Ullevaal University Hospital; 1999.

- [2] Leconte I, Feger C, Galant C, Berlière M, Berg BV, D'Hoore W, Maldague B. Mammography and Subsequent Whole-Breast Sonography of Nonpalpable Breast Cancers: The Importance of Radiologic Breast Density. *American Journal of Roentgenology*. 2003;180:1675-9.
- [3] Crystal P, Strano SD, Shcharynski S, Koretz MJ. Using Sonography to Screen Women with Mammographically Dense Breasts. *American Journal of Roentgenology*. 2003;181:177-82.
- [4] Baker JA, Kornguth PJ, Soo MS, Walsh R, Mengoni P. Sonography of solid breast lesions: observer variability of lesion description and assessment. *American Journal of Roentgenology*. 1999;172:1621-5.
- [5] Huang Y-L, Chen D-R, Jiang Y-R, Kuo S-J, Wu H-K, Moon WK. Computer-aided diagnosis using morphological features for classifying breast lesions on ultrasound. *Ultrasound Obstet Gynecol*. 2008;32:565-572.
- [6] Pereira WCdA, Alvarenga AV, Infantosi AF, Macrini L, Pedreira CE. A non-linear morphometric feature selection approach for breast tumor contour from ultrasonic images. *Computers in Biology and Medicine*. 2010;40:912-918.
- [7] Ayer T, Ayyaci MUS, Liu ZX, Alagoz O, Burnside ES. Computer-aided diagnostic models in breast cancer screening. *Imaging Med* 2010.
- [8] Horsch K, Giger ML, Venta LA, Vyborny CJ. Automatic segmentation of breast lesions on ultrasound. *Med Phys*. 2001;28:1652-9.
- [9] Gomez W, Leija L, Alvarenga AV, Infantosi AF, Pereira WC. Computerized lesion segmentation of breast ultrasound based on marker-controlled watershed transformation. *Med Phys*. 2010;37:82-95.
- [10] Gómez W. Desarrollo de una Metodología Computacional para la Clasificación de Lesiones de Mama en Imágenes Ultrasonicas[dissertação]. México, D. F.: Centro de Investigación y de Estudios Avanzados del Intituto Politécnico Nacional; 2009.
- [11] Gómez Flores W, Pereira WCdA, Infantosi AFC. Improving classification performance of breast lesions on ultrasonography. [in press] *Pattern Recognition*. 2014. DOI:10.1016/j.patcog.2014.06.006.
- [12] Sahiner B, Chan H-P, Hadjiiski L. Classifier performance prediction for computer-aided diagnosis using a limited dataset. *Med Phys*. 2008;35:1559-1570.