

## EMPREGANDO ALGORITMOS GENÉTICOS NA SEGMENTAÇÃO AUTOMÁTICA DE NODULO MAMARIO

**Danilo César Pereira**

UFABC – Universidade Federal do ABC  
Rua Santa Adélia, 166, Bangu, Santo André, São Paulo – Brasil  
danilo.pereira@ufabc.edu.br

**Marcelo Zanchetta do Nascimento**

UFABC – Universidade Federal do ABC  
Rua Santa Adélia, 166, Bangu, Santo André, São Paulo – Brasil  
marcelo.nascimento@ufabc.edu.br

### RESUMO

O Instituto Nacional de Câncer relata mais de 9 mil óbitos somente no Brasil devido ao câncer de mama. Atualmente, a mamografia vem sendo utilizada pelos especialistas para detectar possíveis anormalidades em seu estágio inicial. Porém, entre 10% e 30% das anormalidades do tipo malignas visíveis em mamogramas, em estudos retrospectivos, não foram detectadas pelos especialistas durante exames de rotina. Para minimizar essa taxa de erro, sistemas computacionais de apoio à detecção (Computer-Aided Detection - CADe) e ao diagnóstico (Computer-Aided Diagnosis – CADx) vêm sendo desenvolvidos para auxiliar os especialistas na detecção e classificação de possíveis anormalidades. Nesse trabalho, filtros de realce de brilho foram aplicados na etapa inicial do processamento e quantificados. Após a etapa de pré-processamento, um método automático de segmentação baseado em limiarização multi-modal foi desenvolvido pela combinação entre algoritmo genético e transformada wavelet. Um conjunto de casos selecionados da base de dados pública Digital Database for Screening Mammography (DDSM) nas visões Crânio-Caudal (CC) e Médio-Lateral Oblíqua (MLO) foram utilizadas para avaliar os métodos propostos. De acordo com os resultados, o método de detecção foi capaz de identificar corretamente 88% dos casos analisados.

**PALAVRAS CHAVE.** Segmentação de imagens mamográficas, Algoritmo Genético, Processamento de Imagens Biomédicas.

### ABSTRACT

The National Cancer Institute reports more than 9 thousand deaths only in Brazil due to breast cancer. Currently, mammography has been used by radiologists to detect possible abnormalities in early stage. So, between 10% and 30% of malignant abnormalities of the type seen in mammograms in retrospective studies were not detected by the radiologists during routine examinations. To minimize this error rate, Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx) have been developed to assist the radiologists in the detection and classification of possible abnormalities. This work, filters for brightness enhancement was applied in the stage of improvement and quantified. Afterwards the preprocessing stage, a method to determine automatic threshold values based in multilevel threshold was developed by the combination of genetic algorithm and wavelet transform. A set of selected cases from the public database Digital Database for Screening Mammography (DDSM) in CranioCaudal (CC) and MedioLateral Oblique (MLO) views was used to assess the proposed methods. According to results, the detection method was able correctly identify 88% of the analyzed cases.

**KEYWORDS.** Mammographic image segmentation, Genetic Algorithm, Biomedical image processing.

## 1. Introdução

O câncer de mama é um dos principais responsáveis pelo alto número de óbitos entre mulheres. No Brasil, o Instituto Nacional de Câncer relata que são diagnosticados mais de 49 mil novos casos de câncer de mama todos os anos, e são contabilizados mais de 9 mil óbitos devido sua recorrência (INCA, 2008). Estudos realizados em mulheres americanas mostram que há um aumento de 93% na taxa de sobrevivência quando o diagnóstico é realizado em estágio inicial, esse valor é reduzido para 72% no estágio intermediário da doença e somente 13% dos pacientes sobrevivem quando a doença encontra-se no estágio avançado (TANG *et. al*, 2009).

Atualmente, a mamografia vem sendo um dos exames mais utilizados pelos especialistas para detectar possíveis anormalidades em seu estágio inicial. Nesse exame obtêm-se quatro imagens, duas correspondentes à mama da direita e duas à mama da esquerda, tendo assim duas vistas de cada mama, a Cranio-Caudal (CC) e a Médio-Lateral Oblíqua (MLO). A obtenção das imagens CC e MLO melhoram a visualização do tecido mamário e aumentam as chances da detecção de sinais que caracterizam uma possível presença do câncer de mama não palpável. No entanto, estudos mostram que entre 10% e 30% das anormalidades do tipo maligna visíveis em mamogramas, em estudos retrospectivos, não foram detectadas pelos especialistas durante exames de rotina.

Recentemente, sistemas computacionais de apoio à detecção (Computer-Aided Detection - CADe) e ao diagnóstico (Computer-Aided Diagnosis - CADx) vêm sendo desenvolvidos para auxiliar os especialistas na detecção e classificação de possíveis anormalidades, fornecendo uma segunda opinião ou pré-leitura aos casos analisados e possibilitando um diagnóstico mais preciso sobre as regiões onde se encontram anormalidades, minimizando a necessidade da realização de novos exames (KINOSHITA *et. al*, 2008).

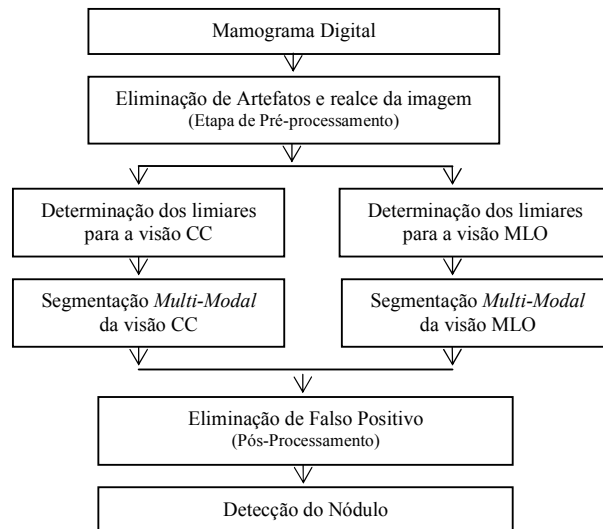
Apesar dos sistemas CAD terem proporcionado um grande número de pesquisas e aumentado a sensibilidade, a grande maioria analisa as visões *Médio-Lateral Oblíqua* (MLO) e *Crânio-Caudal* (CC) de forma independente. Em algumas situações, o sistema detecta anormalidades em apenas uma das visões (ENGELAND & KARSSMEIJER, 2007). Porém, especialistas consideram que há uma inconsistência se uma determinada lesão é semelhante em ambas as visões e o sistema não tem a capacidade de encontrá-la. Com isso, há uma redução na confiabilidade e nos benefícios oferecidos pelo sistema (NASCIMENTO & RAMOS, 2008). Estudos vêm mostrando que essas limitações têm provocado mudanças e por isso os especialistas estão ignorando os resultados proporcionados por esses sistemas (QIAN *et. al.*, 2001).

Esse trabalho apresenta um método computacional que emprega uma combinação entre algoritmo genético e transformada *wavelet*. A combinação desses métodos foi empregada para determinar automaticamente o número e os valores de limiares, sendo posteriormente utilizados em uma limiarização *multi-modal* para detectar lesões causadas por nódulo mamário nas imagens mamográficas. Também foram empregados, nesse estudo, diferentes filtros de realce baseado em transformada *wavelet* (Haar, Daubechies e Coiflets) para avaliar o desempenho na etapa de melhoramento da relação sinal/ruído antes da aplicação do método de detecção. O método foi aplicado sobre um conjunto de casos selecionados da base de dados pública *Digital Database for Screening Mammography* (DDSM).

Esse artigo foi organizado da seguinte forma: no item 2 descreve-se as etapas do sistema para realce e segmentação *multi-modal* de imagens mamográficas, o item 3 ilustra os resultados da técnica desenvolvida e no item 4 são apresentadas as discussões e conclusões do trabalho.

## 2. Materiais e Métodos

Este item descreve os procedimentos empregados nesse estudo para detecção automática de nódulo mamário. A Fig. 1 apresenta o diagrama em blocos dos processos aplicados para detecção.



**Fig. 1** Diagrama em blocos do algoritmo desenvolvido para detecção de nódulos em imagens mamográficas adquiridas nas visões CC e MLO.

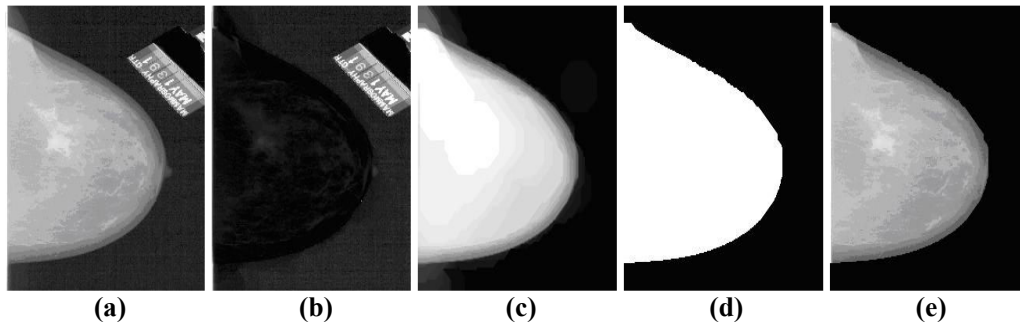
## 2.1. Banco de Imagens Mamográficas

A base *Digital Database for Screening Mammography* (DDSM), um projeto colaborativo envolvendo o Massachusetts General Hospital (D. Kopans, R. Moore), a Universidade do Sul da Flórida (K. Bowyer) e o Sandia National Laboratories (P. Kegelmeyer) foi utilizada para avaliação dos métodos propostos. Essa base dispõe de aproximadamente 2.500 casos de mamogramas, obtidos entre os anos de 1988 e 1999 de quatro instituições dos Estados Unidos: Massachusetts General Hospital (MGH), Wake Forest University School of Medicine (WFUSM), Sacred Heart Hospital (SHH) e Washington University e digitalizadas pelos seguintes equipamentos: Howtek (WU e MGH), Lumisys (WFUSH e SHH) e DBA (MGH).

Cada caso mamográfico contém duas imagens nas visões: *Médio-Lateral Obliqua* (MLO) e *Crânio-Caudal* (CC); além de informações associadas aos pacientes, como: idade na época do estudo, índice da densidade de mama ACR BI-RADS™, subgrupo de índice para anormalidades, descritores de anormalidade ACR, informações da imagem (equipamento de digitalização, resolução espacial, tipo de anormalidade e localização) e um arquivo com extensão “.overlay” que contém o código em cadeia da posição da lesão, em ambas as visões. Essas imagens foram compactadas no padrão LJPEG e digitalizadas nas resoluções de 12 ou 16 bits por pixel para a representação dos níveis de cinza. Para avaliar o desempenho do método proposto, selecionou-se um conjunto de 40 imagens mamográficas, sendo 20 na visão CC e 20 na visão MLO.

## 2.2. Algoritmo para Eliminação de Artefatos

O desempenho do algoritmo de segmentação pode ser prejudicado por diversos fatores, como por exemplo, os artefatos e informações pessoais dos pacientes demarcados na imagem por dispositivos de chumbo. Por isso, nessa etapa foi desenvolvido um algoritmo para eliminação desses artefatos. Nesse algoritmo, primeiramente uma operação morfológica denominada *Top-Hat*, com um elemento estruturante com raio de 60 pixels foi aplicada. Em seguida, uma subtração entre a imagem obtida e a imagem original foi realizada. Então, o método de limiarização de Otsu (OTSU, 1979) foi aplicado sobre a imagem resultante. Após obter a imagem limiarizada, uma operação de multiplicação de imagens foi aplicada entre a imagem limiarizada e a imagem original, permitindo obter uma imagem sem a presença de artefatos na região de fundo. A Fig. 2 apresenta as imagens obtidas em cada etapa desse algoritmo.



**Fig. 2** Imagens obtidas com o algoritmo para eliminação de artefatos: (a) imagem mamográfica na visão CC, obtida do caso 1134, (b) imagem modificada pela operação morfológica Top-Hat, (c) imagem resultante da operação de subtração, (d) imagem resultante da limiarização de Otsu e (e) imagem resultante sem a presença de artefatos.

### 2.3. Método de Realce e Melhoramento de Contraste dos Níveis de Brilho da Imagem

A etapa de realce e melhoramento têm como objetivo proporcionar um aumento de contraste entre as estruturas internas à mama e assim proporcionar melhores resultados para o algoritmo de segmentação. PEREIRA *et. al.*, (2008) apresentou um estudo mostrando que filtros de realce no domínio *wavelet* foram superiores aos filtros de realce no domínio de Fourier.

Existem várias possibilidades para a escolha da função  $f(t)$ , denominada *função mãe*, aplicada à transformada *wavelet*. Cada função *mãe* possui uma característica distinta que permite maior ou menor adaptação a um determinado contexto. Com isso, foram desenvolvidos algoritmos para filtragem utilizando as seguintes funções *mãe*: Haar, Daubechies e Coiflets para realçar o contraste da imagem.

Nesses algoritmos, as seguintes etapas foram aplicadas: (a) decomposição da imagem em três níveis de profundidade usando uma das funções *mãe* (REIGOTTA, 2007). Esse procedimento resultou quatro subimagens: uma de aproximação e três de detalhes (horizontal, vertical e diagonal). Então, utilizando as informações obtidas da aproximação aplicou-se o filtro de *Wiener*, dado por:

$$F_{res} = \left[ \frac{|H(u,v)|^2}{|H(u,v)|^2 + S_{\eta}(u,v)/S_f(u,v)} \right] G(u,v) \quad (1)$$

onde

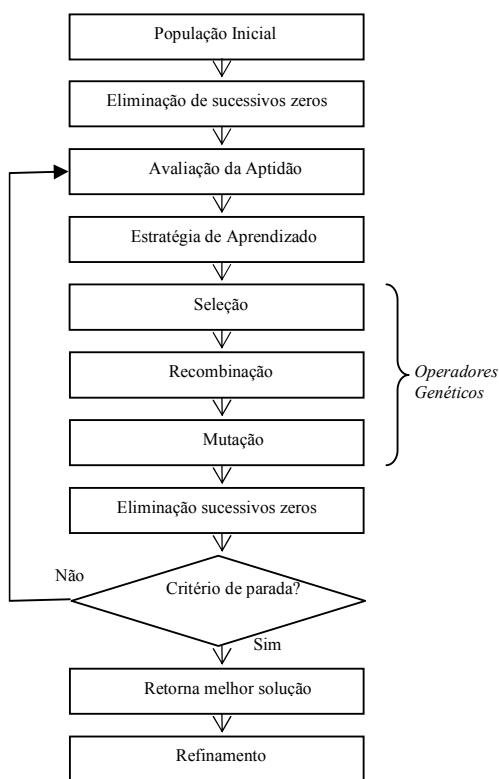
- $H(u,v)$  = função de degradação;
- $|H(u,v)|^2 = H(u,v) * H(u,v)$  ;
- $H(u,v)^*$  = conjugado complexo de  $H(u,v)$  ;
- $S_{\eta}(u,v) = |N(u,v)|^2$  = espectro de potência do ruído;
- $S_f(u,v) = |F(u,v)|^2$  = espectro de potência da imagem não-degradada pelo ruído.

Após a filtragem das subimagens, a transformada *wavelet* inversa foi aplicada com objetivo de retornar a imagem ao domínio espacial.

## 2.4. Algoritmo Genético Aplicado na Detecção Automática de Nódulos Mamários

Para detecção automática de nódulo mamário foi desenvolvido um método computacional que combina algoritmo genético e transformada wavelet para determinar a quantidade e os valores de limiar necessários para a segmentação multi-modal. Para isso, inicialmente, o histograma de níveis de cinza foi calculado para a imagem. Em seguida, a transformada *wavelet* Daubechies de ordem 1 (db1), com um nível de decomposição, foi empregada sobre o histograma de níveis de cinza com o objetivo de reduzir em 50% original. A aplicação da *wavelet* permitiu extrair duas informações: o sinal de aproximação e sinal de detalhe.

A partir do sinal de aproximação, que contém todas as informações originais do histograma, com tamanho reduzido, um Algoritmo Genético (AG) foi construído para determinação da quantidade e dos valores de limiar (Fig. 3). Nesse algoritmo, uma população randômica de soluções foi construída utilizando a representação binária. Nessa metodologia, os valores do vale do histograma foram representados com o *bit* 0 (zero) e as regiões de pico foram simbolizados pelo *bit* 1 (um). Esse procedimento de codificação também produziu alguns cromossomos apenas com *bits* 0 (zero) consecutivos. Então, aplicou-se o método proposto por Hammouche (HAMMOUCHE *et. al.*, 2008) sobre esses cromossomos para alterar essa seqüência de *bits*, preservando apenas o primeiro deles com valor 0 e modificando os demais.



**Fig. 3** Diagrama de blocos ilustrando as etapas do AG construído para determinação dos valores limiares para segmentação *multi-modal* de imagens mamográficas.

Então, a aptidão de cada indivíduo da população foi calculada usando a função de aptidão (YEN *et. al.*, 1995) definida por:

$$F(k) = \rho(Disk(k))^{1/2} + (\log_2(k))^2 \quad (2)$$

$$Disk(k) = \Omega_w^w(k) = \Omega_T^w - \Omega_B^w(k) \quad (3)$$

$$\Omega_T^2 = \sum_{j=0}^{L-1} (j-m)^2 p_j, \quad \Omega_T^2(k) = \sum_{i=1}^k p_i (m_i - m)^2 \quad (4)$$

$$P_i = \sum_{j=t_{i-1}}^{t_i-1} p_j, \quad m_i = \frac{\sigma_i}{P_i}, \quad \sigma_i = \sum_{j=t_{i-1}}^{t_i-1} j p_j, \quad m = \sum_{j=0}^{L-1} j p_j \quad (5)$$

onde  $\rho$ ,  $P_i$ ,  $m_i$ ,  $m$ ,  $L'$  e  $t_i$  correspondem, respectivamente, a uma constante com valor positivo, a probabilidade da classe  $C_i$ , a média de níveis de cinza da classe, a média de níveis de cinza da imagem, o histograma reduzido e o  $i$ -ésimo valor de limiar. O parâmetro  $k$  representa o número otimizado de classes.

Para selecionar os melhores indivíduos da população com o AG, empregou-se o operador de seleção por torneio com tamanho quatro. Esse processo escolheu, aleatoriamente, quatro indivíduos da população, sendo que apenas o indivíduo com o maior valor de aptidão foi copiado para um local isolado, chamado de depósito genético. A seleção foi repetida até que o depósito genético formado estivesse do mesmo tamanho da população inicial.

Para aumentar o valor de aptidão dos indivíduos e, conseqüentemente, reduzir o tempo de convergência do AG, uma estratégia de aprendizado proposta por Yin (Yin *et. al.* 1999) foi aplicada. Esse método copiou o valor de um gene escolhido aleatoriamente para a posição anterior e posterior do cromossomo, ou seja,  $t_1$  para  $t_1 - 1$  e  $t_1 + 1$ .

As operações genéticas de recombinação e mutação, respectivamente, foram aplicadas pelo operador de combinação em um ponto, com uma probabilidade de 0,9, e mutação flip, com probabilidade de 0,01. Em conseqüência do emprego desses métodos, novas sequências de *bits* 0 foram novamente estabelecidas. Então, o método para eliminação dos sucessivos valores de limiar foi aplicado sobre os dados para eliminação dessas combinações.

O AG foi iterado por 50 vezes e após seu término, o cromossomo do indivíduo com valor de aptidão mais elevada foi selecionado para fornecer o número e os valores limiares. Como os valores de limiares foram obtidos sobre o histograma reduzido, multiplicou-se cada valor de limiar por um fator de  $2^r$ , sendo que  $r$  representou o nível de decomposição empregado na transformada *wavelet*. Com o objetivo de proporcionar estabilidade e maior exatidão nos valores de limiar produzidos pelo AG, empregou-se um procedimento de refinamento proposto por Hammouche (HAMMOUCHE *et. al.*, 2008), onde os valores de aptidão foram atualizados de acordo com a média dos níveis de cinza ( $m_i$ ) da classe  $C_i$ , dado pela equação 6. Esse processo foi aplicado recursivamente até que a estabilidade dos valores limiar foi obtida, ou seja,  $t_i(s+1) = t_i(s), i = 1, 2, \dots, k-1$ .

$$t_i(s+1) = \frac{m_i(s) + m_{i+1}(s)}{2} \quad (6)$$

A partir desses valores, aplicou-se a Equação (7) para segmentar a imagem com o objetivo de detectar as possíveis lesões.

$$B(x, y) = k \text{ se } t_{k-1} \leq f(x, y) \leq t_k \text{ para } k = 1, 2, \dots, n \quad (7)$$

onde  $f(x, y)$  é a imagem original,  $B(x, y)$  é a imagem segmentada,  $n$  é o número de regiões e  $t_{k-1}, \dots, t_k$  são os valores de limiar que definem o intervalo das  $n$  regiões (ou classes).

Com o propósito de reduzir possíveis regiões com falso positivo, uma etapa de pós-processamento foi aplicada após a segmentação da imagem em *multi-modais*. Nessa etapa, uma comparação entre as áreas obtidas pela limiarização nas duas visões da mama (CC e MLO) foi

empregada para eliminar as regiões que não apresentaram correspondência e consequentemente reduzir os falsos positivos.

## 2.5. Avaliação dos Filtros de Realce Aplicados nos Mamogramas

Para quantificar o desempenho dos filtros de realce foi empregado o método proposto por Singh e Bovis (SINGH & BOVIS, 2005), que combina três diferentes métricas: uma nomeada por Medida de Separação de Distribuição (MSD) e duas medidas de realce de contraste entre fundo e objeto, uma baseada em desvio padrão dos níveis de cinza (Medida do Melhoramento de Contraste baseado em desvio padrão dos níveis de cinza - MMCd) e outra baseada em entropia (Medida de Melhoramento de Contraste baseado em entropia – MMCe). Essas métricas avaliam as influências das técnicas de processamento na área do fundo (área vizinha à massa) e do objeto (massa presente na região de interesse) das regiões de interesse (ROIs).

A medida MSD foi empregada para avaliar a redução de sobreposição entre a distribuição da intensidade dos níveis de cinza na área de objeto e região de fundo, dada por:

$$MSD = \left( M_2 - m_{fun}^P \right) + \left( M_2 - m_{obj}^P \right) - \left( M_1 - m_{fun}^O \right) + \left( M_1 - m_{obj}^O \right) \quad (8)$$

onde o valor médio de níveis de cinza do objeto (massa) e do fundo da ROI original são determinados por  $m_{obj}^O$  e  $m_{fun}^O$ , respectivamente. Os termos  $m_{obj}^P$  e  $m_{fun}^P$  consideram, consecutivamente, os valores médios dos níveis de cinza do objeto e do fundo após aplicação do filtro de realce na ROI. Os parâmetros  $M_1$  e  $M_2$  são definidos pela Equação (9).

$$M_1 = \frac{(m_{fun}^O \cdot d_{obj}^O) + (m_{obj}^O \cdot d_{fun}^O)}{(d_{obj}^O \cdot d_{fun}^O) + c}, \quad M_2 = \frac{(m_{fun}^P \cdot d_{obj}^P) + (m_{obj}^P \cdot d_{fun}^P)}{(d_{obj}^P \cdot d_{fun}^P) + c} \quad (9)$$

onde  $d_{obj}^O$  e  $d_{obj}^P$  são os valores de desvio padrão do objeto na ROI original e realçada, respectivamente. Os parâmetros  $d_{fun}^O$  e  $d_{fun}^P$  são os valores de desvio padrão do fundo na ROI original e realçada, respectivamente. A constante  $C$  impede a divisão por zero. Para essa métrica, quanto maior o valor de MSD, melhor a separação entre as distribuições de níveis de cinza de objeto e fundo.

Também foi avaliado o realce de contraste dos níveis de cinza do fundo e do objeto por meio das métricas MMC<sub>d</sub> e MMC<sub>e</sub> (Equação (10)).

$$MMC_d = \left[ \frac{\left( \frac{m_{obj}^P}{m_{fun}^P} \right) - \left( \frac{m_{obj}^O}{m_{fun}^O} \right)}{\frac{d_{obj}^P}{d_{obj}^O}} \right], \quad MMC_e = \left[ \frac{\left( \frac{m_{obj}^P}{m_{fun}^P} \right) - \left( \frac{m_{obj}^O}{m_{fun}^O} \right)}{\frac{Ent_{obj}^P}{Ent_{obj}^O}} \right] \quad (10)$$

onde  $Ent_{obj}^P$  e  $Ent_{obj}^O$  são definidos pela entropia da ROI realçada e original, respectivamente. Nessas métricas, uma boa capacidade de melhoramento está associada a um valor positivo.

Com o objetivo de determinar um índice de classificação dos filtros, calculou-se a distância euclidiana entre pontos correspondentes ao espaço 3D e os pontos  $(I, I, I)$ . Essa métrica recebe o nome de Melhoramento Combinado (MC) e é definida pela Equação (11). Analisando essa métrica, quanto menor o valor de MC, maior a capacidade do filtro para realce.

$$MC = \sqrt{(1 - MSD)^2 + (1 - MMC_d)^2 + (1 - MMC_e)^2} \quad (11)$$

Para avaliar o realce dos filtros também aplicou-se a métrica proposta por Pereira (PEREIRA *et. al.*, 2007), no qual as imagens foram avaliadas em relação à diferença das intensidades médias dos níveis de cinza entre fundo e objeto (Equação (12)).

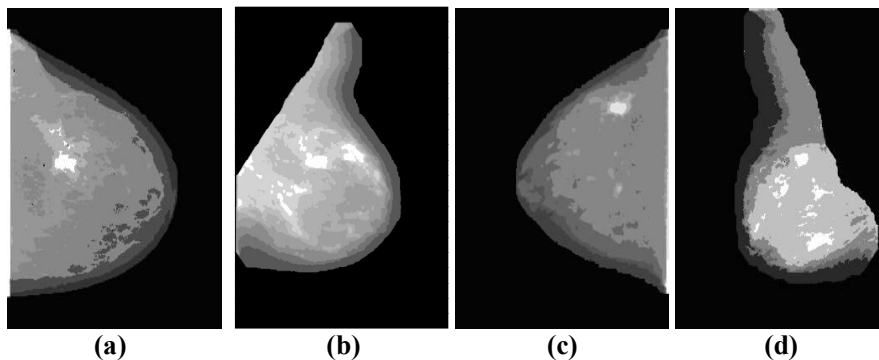
$$CNR = \frac{(m_{obj}^P - m_{fun}^P)}{\sqrt{\frac{((d_{obj}^P)^2 + (d_{fun}^P)^2)}{2}}} \quad (12)$$

## 2.6. Avaliação do Algoritmo de Segmentação Automática de Nódulos Mamários

A taxa de acerto foi estimada pela comparação do tamanho da área da região segmentada pelo algoritmo proposto e pelo código em cadeia (*code-chain*) contido no arquivo “.overlay” e disponibilizado pela base DDSM.

## 3. Resultados

Os resultados obtidos pelo método proposto de segmentação automática nas visões CC e MLO dos casos 1134 e 1211 são apresentados nas Figs. 4(a)-(d). Esses casos foram escolhidos aleatoriamente para demonstrar a capacidade dos algoritmos desenvolvidos. A região de músculo peitoral das imagens na visão MLO foi retirada de forma manual por um especialista, antes da aplicação do algoritmo de segmentação. Os resultados obtidos após o emprego do algoritmo de segmentação automática *multi-modal* para o caso 1134 nas visões CC e MLO são mostrados, respectivamente, nas Figs. 4(a) e 4(b). Os resultados para o caso 1211 para as visões CC e MLO são apresentados nas Figs 4(c) e 4(d), respectivamente.

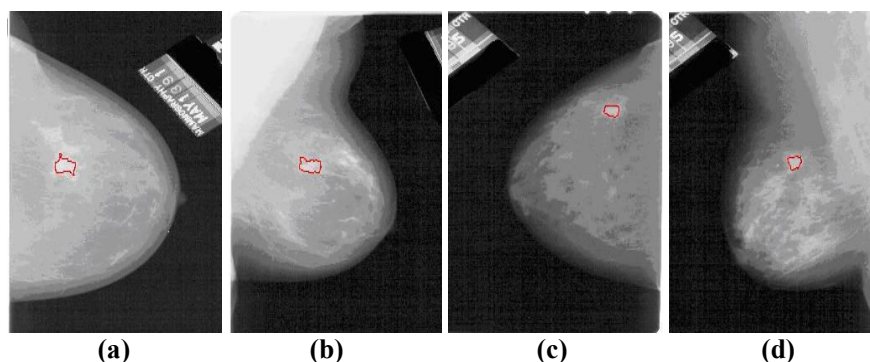


**Fig. 4** Imagens obtidas após a etapa de segmentação *multi-modal*: (a) imagem na visão CC do caso 1134, (b) imagem na visão MLO do caso 1134, (c) imagem na visão CC do caso 1211 e (d) imagem na visão MLO do caso 1211.

As Figs. 5(a) e 5(b) exibem, respectivamente, os resultados obtidos após a etapa de pós-processamento para as imagens CC e MLO do caso 1134. Os resultados obtidos após a etapa de pós-processamento para as imagens CC e MLO do caso 1211 são mostradas nas Figs. 5(c) e 5(d), respectivamente. Nessas imagens, as linhas sólidas correspondem aos limites do nódulo identificado pelo método proposto.

Nas Tabelas 1, 2 e 3 são apresentados, respectivamente, os resultados obtidos com os filtros de realce baseados em transformada *wavelet* Haar, Daubechies e Coiflets aplicados nas duas visões das imagens dos casos 1134 e 1211.





**Fig. 5** Imagens obtidas após a etapa de pós-processamento: (a) imagem na visão CC do caso 1134, (b) imagem na visão MLO do caso 1134, (c) imagem na visão CC do caso 1211 e (d) imagem na visão MLO do caso 1211.

**Tabela 1** Resultados obtidos com a filtragem baseada em transformada *wavelet* Haar.

Imagem	MSD	MMC <sub>d</sub>	MMC <sub>e</sub>	MC	CNR
1134_CC	0,0366	0,0335	0,0282	1,6754	6,4618
1134_MLO	0,0029	0,0096	0,2133	1,8735	4,4659
1211_CC	0,0071	0,0060	0,0059	1,7211	7,8748
1211_MLO	0,0075	0,0065	0,0065	1,5202	6,5497

**Tabela 2** Resultados obtidos com a filtragem baseada em transformada *wavelet* Daubechies.

Imagem	DSM	MMC <sub>d</sub>	MMC <sub>e</sub>	MC	CNR
1134_CC	0,0915	0,0919	0,0758	1,5825	7,0466
1134_MLO	0,0081	0,0084	0,0108	1,7513	6,2944
1211_CC	0,0233	0,0316	0,0249	1,6860	8,9566
1211_MLO	0,0183	0,0227	0,0188	1,6975	7,9719

**Tabela 3** Resultados obtidos com a filtragem baseada em transformada *wavelet* Coiflets.

Imagem	DSM	MMC <sub>d</sub>	MMC <sub>e</sub>	MC	CNR
1134_CC	0,1230	0,1297	0,1066	1,5247	7,6960
1134_MLO	0,0724	0,1034	0,0751	1,5874	10,5141
1211_CC	0,0600	0,0610	0,0474	1,6465	9,0691
1211_MLO	0,0331	0,0789	0,0791	1,3622	8,0283

Para estimar a taxa de acerto, foi realizada uma comparação entre a área delimitada pelo código em cadeia e a área demarcada pelo método proposto. Nessa análise foi obtido um erro médio de 11,78% e um desvio padrão de 18,94%.

## 5. Discussões e Conclusão

Analisando as Tabelas 1, 2 e 3 observa-se que a transformada *wavelet* Coiflets proporcionou resultados superiores para a métrica DSM, pois quanto maior o valor numérico, maior a diferença entre a função de distribuição de probabilidade dos níveis de cinza do objeto e fundo. Isso resulta em uma maior distinção de estruturas com anormalidade.

Na Tabela 3, verifica-se que os resultados obtidos com a transformada *wavelet* Coiflets foram superiores para a medida MMC<sub>d</sub>, o que significa que a homogeneidade dos níveis de cinza entre a região de objeto e do fundo foi maior após a aplicação dessa transformada *wavelet*. De forma semelhante, a *wavelet* Coiflets também foi superior na medida MMC<sub>e</sub>. Através da métrica MC, que busca classificar o desempenho das técnicas, observou-se que a transformada *wavelet* Coiflets também proporciona os melhores resultados, pois quanto menor esse valor, maior é a capacidade de realce do filtro.

De forma semelhante às métricas anteriores, a técnica de filtragem baseada em transformada *wavelet* Coiflets também apresentou resultados superiores para a medida de melhoramento da intensidade dos níveis de cinza (CNR).

A segmentação automática de lesões baseado na combinação da transformada *wavelet* e algoritmo genético foi capaz de detectar as anormalidades com resultados relevantes para os casos analisados. Os resultados foram satisfatórios, uma vez que o erro médio da diferença entre as técnicas foi inferior a 12% (Tabela 4). Porém, em alguns casos, o nódulo identificado pelo código em cadeia não contorna a anormalidade semelhante ao diagnóstico de um especialista como demonstrado no trabalho de Mencattini (MENCATTINI *et al.*, 2008). Com isso, uma diferença entre a região demarcada pelo método proposto e a disponibilizada pelo código em cadeia (arquivo “.overlay”) pode ocorrer para alguns dos casos analisados.

## Referências

- Engeland V. S., Karssemeijer N., Combining two mammographic projections in a computer aided mass detection method. *Medical Physics*, Mar – 34(3), p. 898-905, 2007.
- Hammouche, K., Diaf, M.; Siarry, P., A multilevel automatic thresholding method based on a genetic algorithm for a fast image segmentation. *Computer Vision and Image Understanding*, n. 109, p. 163-175, 2008.
- Huo Z., Giger M. L., Vyborny C. J., Breast cancer: effectiveness of computer aided diagnosis - observer study with independent database of mammograms. *Academic Radiology*, p. 224–256, 2002.
- Inca – Instituto Nacional do Câncer. Ministério da Saúde. Falando sobre Doenças de Mama, <http://www.inca.gov.br>, 2008, 27/03/2008.
- Tang J., Rangayyan R. M., Xu J., Naqa I. E., Yang Y., Computer-Aided Detection and Diagnosis of Breast Cancer with Mammography: Recent Advances. *IEEE Trans on Information Technology in Biomedicine*, v. 13, n. 2, p. 236-251, 2009.
- Mencattini, A.; Salmeri, M.; Lojacono, R.; Frigerio, M.; Caselli, F., Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing. *IEEE Trans. on Instrumentation and Measurement*, n. 57, p. 1422-1430, 2008.
- Nascimento M. Z., Ramos R. P., Combinando duas visões mamográficas em extração de características com Ridgelet. In: *XI Congresso Brasileiro de Informática em Saúde, Campos do Jordão*. XI Congresso Brasileiro de Informática em Saúde, v. 1, 2008.
- OTSU, N. A threshold selection method from grey level histograms. *IEEE Trans. on Systems*, v. 9, p. 62-66, 1979.
- Pereira, C. S., Mendonça, A. M., Campilho, A., Evaluation of contrast enhancement filters for lung nodule detection, *LNCS 4633*, p. 878, 2007.
- Pereira, D. C., Nascimento, M. Z., Scott, L. P., Avaliação de filtros para pré-processamento aplicados à mamogramas, *21º CBEB – Congresso Brasileiro de Engenharia Biomédica.*, p. 1-4, Nov. 2008.

**Qian W., Sun X., Song D., Clark, R. A.,** Digital mammography: wavelet transform and Kalman-filtering neural network in mass segmentation and detection. *Academic Radiology*, n. 8, p. 1074-1082, 2001.

**Reigotta, N. S.,** Comparação da transformada wavelet discreta e da transformada cosseno para a compressão de imagens de impressão digital. 143 f. Dissertação (Doutorado). USP: Escola de Engenharia de São Carlos, São Carlos, 2007.

**Singh, S., Bovis, K.,** An evaluation of contrast enhancement techniques for mammographic breast masses, *IEEE Trans. on Informatic Technology in Biomedicine* , vol. 9, n. 1, p. 109-119, Mar. 2005.

**Tourassi G. D., Catarious D. M., Floyd C.,** Computer-assisted detection of mammographic masses: A template matching schema based on mutual information. *Medical Physics*, vol. 30, n. 6, p. 2123-2130, 2003.

**Yin, P. Y.,** A fast scheme for optimal thresholding using genetic algorithms. *Signal Processing*, n. 72, p. 85-95, 1999.