

## DETECÇÃO E DIAGNÓSTICO DE MASSAS EM MAMOGRAFIA: revisão bibliográfica\*

MAMMOGRAPHY MASS DETECTION AND DIAGNOSIS: a survey

DETECCIÓN Y DIAGNÓSTICO DE MASAS EN UNA MAMOGRAFÍA:  
una revisión de la literatura

*Simara Vieira da Rocha  
Geraldo Braz Júnior  
Aristóphanes Correa Silva  
Anselmo Cardoso Paiva*

**Resumo:** O câncer de mama tem se tornado cada dia mais freqüente entre a população feminina acima dos 40 anos. Somente para o ano de 2011 são estimados, no Brasil, 49 mil novos casos. Uma das maneiras para detectar os tumores não palpáveis que causam câncer de mama é realizar uma radiografia (mamografia) das mamas. A mamografia é atualmente a melhor técnica de detecção precoce de lesões não palpáveis na mama com altas chances de ser um câncer curável. Sabe-se que as chances de cura do câncer de mama são, relativamente altas, se detectado nos estágios iniciais. Entretanto, a sensibilidade desse exame pode variar bastante, em decorrência de fatores como qualidade do exame ou experiência do especialista. Dessa forma, a utilização de sistemas CAD e CADx tem contribuído para aumentar as chances de uma detecção e diagnósticos corretos, ou seja, uma segunda opinião, auxiliando os especialistas na tomada de decisões em um tratamento do câncer de mama. Este artigo faz uma revisão bibliográfica de trabalhos voltados para detecção e diagnóstico de massas.

**Palavras-chave:** Massa. Mamografia. Detecção. Diagnóstico. Câncer de mama.

**Abstract:** Breast cancer has become increasingly common among the female population over 40 years old. Only for the year 2011 are estimated, in Brazil, 49 000 new cases. One way to detect non-palpable tumors that cause breast cancer is to perform an X-ray (mammogram) of the breasts. Mammography is currently the best technique for early detection of non-palpable breast lesions with high chances of being a curable cancer. It is known that the chances of a cure for breast cancer are relatively high if detected in early stages. However, the sensitivity of this exam can vary greatly due to factors such as quality of examination or experience of the specialist. Thus, the use of CAD systems and CADX has contributed to increase the chances of detection and correct diagnosis, working as a second opinion in treatment of breast cancer. This article is a literature review of studies focused on detection and diagnosis of masses.

**Keywords:** Mass. Mammography. Detection. Diagnosis. Breast cancer.

**Resumen:** El cáncer de mama se ha tornado cada vez más común entre la población femenina de más de 40 años. Sólo para el año 2011 se estima que en Brasil habrán 49 000 nuevos casos. Una forma de detectar tumores no palpables que causan el cáncer de mama es realizar una radiografía (mamografía) de los senos. La mamografía es actualmente la mejor técnica para la detección precoz de lesiones mamarias no palpables, con altas posibilidades de ser un cáncer curable. Se sabe que las posibilidades de una cura para el cáncer de mama son relativamente altas si se detecta en etapas tempranas. Sin embargo, la sensibilidad de esta prueba pueden variar considerablemente debido a factores como la calidad de los exámenes o la experiencia del especialista. Por lo tanto, el uso de sistemas CAD y CADX ha contribuido a aumentar las posibilidades de detección y el diagnóstico correcto, o una segunda opinión, ayudando a los expertos en la toma de decisiones en el tratamiento del cáncer de mama. Este artículo es una revisión de la literatura de trabajos sobre detección y diagnóstico de masas.

**Palabras clave:** Masa. Mamografía. Detección. Diagnóstico de cáncer de mama.

### 1 INTRODUÇÃO

De acordo com Instituto Nacional de Câncer (2010), o termo câncer é utilizado genericamente para representar um conjunto de mais de 100 doenças, incluindo tumores malignos de diferentes localizações. Importante causa de doença e morte

no Brasil, desde 2003, as neoplasias malignas constituem-se na segunda causa de morte na população, representando quase 17% dos óbitos de causa conhecida, notificados em 2007, no Sistema de Informações sobre Mortalidade.

\*Artigo recebido em outubro 2011  
Aprovado em dezembro 2011

Em 2008, a Agência Internacional para Pesquisa em Câncer (IARC)/OMS estimou que ocorreriam 12,4 milhões de casos novos e 7,6 milhões de óbitos por câncer no mundo. Destes, os mais incidentes foram o câncer de pulmão (1,52 milhões de casos novos), mama (1,29 milhões) e cólon e reto (1,15 milhões). Devido ao mau prognóstico, o câncer de pulmão foi a principal causa de morte (1,31 milhões), seguido pelo câncer de estômago (780 mil óbitos) e pelo câncer de fígado (699 mil óbitos). Para América do Sul, Central e Caribe, estimou-se em 2008 cerca de um milhão de casos novos de câncer e 589 mil óbitos. Nas mulheres, o mais frequente foi o câncer de mama, seguido do colo do útero, cólon e reto, estômago e pulmão (INTERNATIONAL AGENCY FOR RESEARCH ON CANCER, 2008).

No Brasil, as estimativas, para o ano de 2010, serão válidas também para o ano de 2011, e apontam para a ocorrência de 489.270 casos novos de câncer. Sendo 236.240 casos novos para o sexo masculino e 253.030 para sexo feminino. Estima-se que o câncer de pele do tipo não melanoma (114 mil casos novos) será o mais incidente na população brasileira, seguido pelos tumores de próstata (52 mil), mama feminina (49 mil), cólon e reto (28 mil), pulmão (28 mil), estômago (21 mil) e colo do útero (18 mil), acompanhando o mesmo perfil da magnitude observada para a América Latina (INSTITUTO NACIONAL DE CÂNCER, 2010).

Na região Sudeste, o câncer de mama é o mais incidente entre as mulheres, com um risco estimado de 65 casos novos por 100 mil. Sem considerar os tumores de pele não melanoma, este tipo de câncer também é o mais frequente nas mulheres das regiões Sul (64/100.000), Centro-Oeste (38/100.000) e Nordeste (30/100.000). Na região Norte é o segundo tumor mais incidente (17/100.000).

Segundo o Instituto Nacional de Câncer (2010) para o ano de 2010, no Maranhão, foram estimados 390 novos casos de câncer de mama feminino. Só na capital do estado, São Luís, foram estimados 150 novos casos de câncer de mama feminino, representando uma taxa bruta de 28,91% dos novos casos de câncer em mulheres na capital e 11,91% dos novos casos de câncer no estado.

Nas últimas duas décadas, a taxa de mortalidade por câncer de mama no Brasil apresentou elevação de 68%, passando de 5,77 em 1979 para 9,70 mortes por 100 mil mulheres em 1998 (INSTITUTO NACIONAL DE CÂNCER, 2010). Dentre as explicações para esse aumento estão: o envelhecimento da

população, a mudança do perfil reprodutivo, a exposição a poluentes, o sedentarismo, a obesidade, dentre outros. Como mudanças no perfil reprodutivo feminino podem ser citadas a gravidez tardia, a nuliparidade, a diminuição no número de gestações e o tempo de amamentação. O aparecimento do câncer de mama pode ser explicado, entre outros fatores, pela perda do controle normal da proliferação celular (ciclo celular) onde as células são submetidas a um processo contínuo de síntese, mitose (divisão celular) e apoptose (morte celular fisiológica). Em determinadas ocasiões certas células reproduzem-se com uma velocidade maior e de maneira anômala, desencadeando o aparecimento de massas celulares denominadas neoplasias.

### 1.1 O problema

Uma das maneiras para detectar os tumores não palpáveis que causam câncer de mama é realizar uma radiografia (mamografia) das mamas. A mamografia é atualmente a melhor técnica de detecção precoce de lesões não palpáveis na mama com altas chances de ser um câncer curável. Sabe-se que as chances de cura do câncer de mama são relativamente altas se detectado nos estágios iniciais. A partir do início da utilização da mamografia foi observada uma redução da taxa de mortalidade associada a essa patologia (AMERICAN CANCER SOCIETY, 2007).

Os exames de rastreamento são apontados como o principal método diagnóstico do câncer de mama em estágios iniciais, sendo capaz de detectar alterações ainda não palpáveis. Entretanto, a sensibilidade desse exame pode variar bastante, em decorrência de fatores como qualidade do exame, experiência do especialista ou idade do paciente.

A mamografia, nos programas de rastreamento, é proposta apenas para mulheres acima dos quarenta anos de idade (AZEVEDO; PEIXOTO, 1993). Este exame é bastante efetivo para tal faixa etária, o que fornece uma boa relação custo/benefício. Os tipos mais comuns de anormalidades visíveis em imagens de mamografia são: calcificações (benignas e malignas); massas circulares e bem definidas; massas espiculadas; massas mal definidas e distorção de arquitetura (HEATH et al., 2001).

Uma massa é um aglomerado de células que se unem de forma mais densa em relação ao tecido que a envolve. Este aglomerado pode ser causado por câncer de mama, assim como

também por condições benignas. Algumas características das massas são determinantes para estabelecer suas probabilidades de malignidade: tamanho, forma e disposição de suas margens.

Já as calcificações são depósitos de cálcio que aparecem como pontos brancos no mamograma. São de dois tipos: microcalcificações e macrocalcificações. As microcalcificações são depósitos pequenos e indicam, dependendo de sua forma, uma possível presença cancerígena. As macrocalcificações são grandes depósitos de cálcio e normalmente estão associadas com condições benignas, causadas, por exemplo, por inflamações ou envelhecimento das artérias.

Técnicas de processamento de imagens vêm adquirindo uma importância ímpar para o diagnóstico e auxílio na intervenção médica. O tempo gasto para trabalhar com essas imagens, a subjetividade dos atributos extraídos e a necessidade contínua de investigação para o progresso na área, tem contribuído para o surgimento de novas técnicas de processamento e análise das imagens médicas que melhoram a qualidade do diagnóstico médico.

O processamento de imagens na medicina representa um conjunto de técnicas computacionais, que aplicados, podem prover auxílio ao diagnóstico, planejamento de tratamentos, simulação de cirurgias, compressão de imagens em bancos de exames, recuperação de exames por conteúdo de imagens, auxílio à pesquisa em medicina, educação médica, dentre outros.

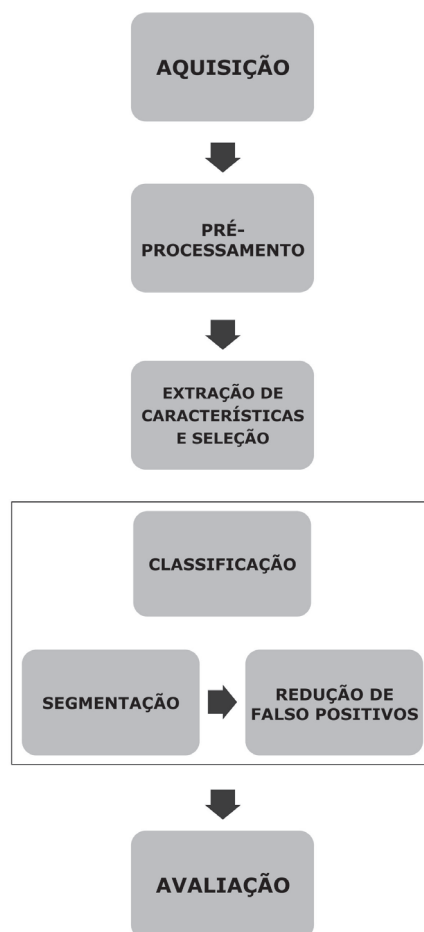
O objetivo do uso do processamento digital de imagens consiste em melhorar o aspecto visual de certas feições estruturais para o analista humano e fornecer subsídios para a sua interpretação, inclusive gerando produtos que possam ser posteriormente submetidos a outros processamentos. A evolução da tecnologia de computação digital, bem como o desenvolvimento de novos algoritmos para lidar com sinais bidimensionais está permitindo uma gama de aplicações cada vez maior.

Sistemas de detecção e diagnóstico auxiliado por computador (respectivamente CAD - Computer-Aided Detection e CADx - Computer-Aided Diagnosis) têm sido propostos com o objetivo de auxiliar o radiologista, indicando áreas suspeitas, bem como anormalidades mascaradas. Esses sistemas têm sido desenvolvidos por vários grupos de pesquisa, visando auxiliar na detecção e diagnóstico precoce do câncer de mama (GUPTA, 1995; KINOSHITA et al., 2004; MEERSMAN; SCHEUNDERS; DYCK, 1998).

Estudos indicam que, não importando o nível de habilidade, todos os especialistas falham em detectar uma anormalidade em alguns momentos. Estudos mostram que o índice de detecção da presença de câncer de mama poderia ser aumentado de 5% a 15% se ferramentas CAD fossem utilizadas (FREER; ULISSEY, 2001).

As ferramentas de detecção auxiliam os especialistas no planejamento de procedimentos invasivos e as ferramentas de diagnóstico na decisão a respeito da realização de certos procedimentos, os quais, tomados em um espaço de tempo curto, podem ser fundamentais em um tratamento adequado e com grandes chances de sucesso. Juntas, as ferramentas de detecção e diagnóstico constituem uma importante ferramenta de auxílio ao especialista a promover o desenvolvimento de tratamentos mais adequados aos pacientes. Segundo Cheng et al., (2006), pode-se afirmar que metodologias para detecção ou diagnóstico de massas, presentes em sistemas CAD/CADx, envolvem as fases descritas na Figura 1.

Figura 1 – Fluxograma de atividades para metodologias de detecção e diagnóstico



Fonte: Adaptado de (CHENG et al., 2006)

As etapas de aquisição, pré-processamento e extração de características são compartilhadas entre as duas abordagens. A aquisição se remete à digitalização de filmes de raio-x, no caso de metodologias de detecção, ou obtenção de regiões de interesse, no caso de metodologias de diagnóstico. Após a aquisição é feito o pré-processamento da imagem ou região, cujo objetivo é suprimir os ruídos e melhorar o contraste da mesma. Em seguida, são extraídas características que serão usadas na etapa seguinte.

No caso de uma abordagem de detecção a etapa que se segue é de segmentação, que consiste na localização das regiões suspeitas de conterem as massas. Normalmente são geradas muitas regiões suspeitas, que podem incluir de fato uma massa, chamada de caso positivo. Todavia, um grande número de regiões suspeitas, que não são efetivamente positivas, são geradas em conjunto com as massas. Daí a necessidade de mais uma etapa, redução de falsos positivos, que tem como premissa reavaliar as regiões segmentadas a partir de características extraídas.

Numa abordagem de diagnóstico, a próxima etapa é de classificação. Esta utiliza o conjunto de características previamente extraídos para informar se a região possui características de massa, podendo ainda informar quanto ao comportamento maligno.

A necessidade de criação de novas técnicas computacionais mais eficientes, que possam, por sua vez, aumentar as chances de uma detecção e diagnósticos corretos, provendo uma segunda opinião, motiva a realização desse trabalho, que no atual estágio, encontra-se na revisão bibliográfica de trabalhos correlatos.

O restante deste trabalho está organizado em três seções. A Seção 2 trata de metodologias estudadas a respeito de detecção de massas em mamografias. A Seção 3 apresenta trabalhos relacionados com o diagnóstico de massas. Por fim, a Seção 4 apresenta discussões e conclusões a respeito do estudo desenvolvido.

## 2 DETECÇÃO

As abordagens que propõem metodologias de detecção de massas tentam resolver um problema ainda em aberto na área. As imagens de mamografia representam as estruturas internas da mama com pouca informação. Normalmente, massas são estruturas muito semelhantes a outras ao seu redor. Isso torna as metodologias de detecção muito específicas e

pouco parametrizáveis. Também faz com que um grande número de abordagens seja proposto. A seguir, faremos a revisão de trabalhos que propõem soluções de detecção, organizados em grandes grupos relacionados com a forma que representam e tratam o problema.

### 2.1 Multiresolução

Liu, Babbs e Delp et al. (2001) utilizam transformada de wavelet para gerar  $N$  imagens em resoluções diferentes com o objetivo de detectar tumores espiculados em mamografias da base MIAS (*Mammographic Image Analysis Society Digital Mammogram Database*). Para cada imagem em resolução diferente (partindo da imagem de menor resolução em direção a que possui a melhor resolução) é calculado, para cada pixel, um vetor de cinco características. Essas características são utilizadas por uma árvore de decisão no processo de classificação de regiões e suas vizinhanças levando em consideração que uma classificação positiva feita em uma resolução menor será continuada para a maior resolução. A análise de características leva em consideração a média e desvio padrão das intensidades dos vizinhos, o desvio padrão de um histograma direcional em relação aos vizinhos e o desvio padrão das médias do gradiente repartida em  $K$  seções (sendo  $K$  a quantidade de pixels da vizinhança). Segundo estudo dos autores, o padrão comportamental de uma região normal e de tumor, em relação ao estudo dos gradientes, pode ser identificado e diferenciado por uma ferramenta de análise de padrões.

Na mesma linha do trabalho anterior, Sampat e Bovik (2003) utilizam Radon Transform como rotina de melhoramento, para codificar regiões lineares de uma imagem de mamografia (obtida da base DDSM: *Digital Database for Screening Mammography*) para um determinado limiar e ângulo. A partir da imagem transformada, os autores determinam filtros intitulados *Radial Spiculated Filters*, desenvolvidos especificamente para detectar massas espiculares. Os filtros levam em consideração os ângulos tomados na transformação anterior. Basicamente é feita uma varredura pela imagem, buscando ponto de máximo na aplicação dos filtros propostos. Estes pontos serão marcados como possíveis massas.

Campanini et al. (2004) utilizam uma abordagem sem extração de características para detecção de massas em mamografias da base DDSM. Primeiramente uma máquina de vetores de suporte (MVS) é treinada com



regiões de massa e não massa. As características passadas para a MVS são a decomposição completa de cada região de interesse previamente obtida através da transformada de wavelet. Logo após, a imagem original é representada em múltiplas escalas. Para cada escala, uma janela de 64x64 é transladada sobre a imagem recortando vários pedaços, em passos de 10% do tamanho da máscara com a intenção de obter regiões entrelaçadas. Cada janela é caracterizada por uma wavelet completa é submetida ao primeiro classificador. Como terceiro passo, é feita uma redução de falsos positivos, criando novamente um classificador MVS com as amostras utilizadas na primeira etapa somadas as amostras classificadas como positivas e falsas positivas na segunda etapa. Em seguida é feita a junção das classificações em múltiplas escalas da imagem original. A metodologia encontra resultados de 80% de acertos em massas com uma taxa de 1.1 falsos positivos por imagem.

## 2.2 Baseada em corte

Os trabalhos de Matsubara et al. (1997) e Brzakovic, Luo e Brzakovic (1990) trazem metodologias de utilização de um corte global para separação de regiões de massa de demais regiões de fundo. Por esta proposta gerar muitos falsos positivos, ambos em uma etapa posterior, caracterizam as regiões extraídas no intuito de utilizar um classificador que possa responder se são verdadeiras massas e em seguida quanto ao comportamento maligno. No primeiro trabalho, a imagem de mamografia passa por um pré-processamento onde é extraído o músculo peitoral e retirado o fundo (identificação da mamografia). Em seguida as mamografias são categorizadas conforme o grau de densidade (*granular, fatty, dense e high dense*). Um limiar específico é calculado para cada categoria com intuito de separar regiões mais densas. Essas regiões são posteriormente classificadas por área, circularidade, desvio padrão e contraste. Por último cada região encontrada é classificada analisando-se a dimensão fractal. Os autores encontram 97% de sensibilidade com a geração de 3,5 falsos positivos por imagem e 100% de acerto no diagnóstico.

O segundo trabalho calcula cortes globais utilizando multiresolução baseada em pirâmide fuzzy. De posse das regiões, é feita uma classificação usando árvore de decisão com regras bayesianas. A metodologia encontra 85% de acurácia e não gera falsos positivos.

Cortes globais normalmente assumem uma característica única para todas as imagens utilizadas na base. Esse fator pode influenciar negativamente na detecção. Por isso, novas técnicas baseadas em cortes adaptativos têm sido propostas. Kom et. al. (2007) utilizam métodos de corte adaptativo aliando um método linear de melhoramento baseado no padrão encontrado em massas. Os parâmetros para ajuste do método de melhoramento são empiricamente encontrados. A idéia é separar ao máximo a região da massa do fundo da imagem que comprovadamente varia de paciente para paciente. Este trabalho utiliza uma base de imagens oriundas do Yaounde Gynaeco-Obstetric and Pediatric Hospital. A metodologia consegue 95.91% de sensibilidade e 0.033% de falsos positivos por imagem.

## 2.3 Template matching

Tourass e Vargas-Voracek (2003) utilizam técnicas de template matching para reconhecimento de regiões de interesse. As regiões foram extraídas da base DDSM. A metodologia utiliza uma base de regiões previamente obtida. A partir de uma nova imagem que deve ser analisada, o *template matching* buscará na base por imagens semelhantes. A metodologia como um todo encontra 87% de acerto. Já Lai, Li e Bischof (1989) propõem uma metodologia para detecção de massas circunscritas utilizando técnicas de *template*. Antes da segmentação ser realizada com um *template*, a imagem é pré-processada com o filtro da mediana. Esta consegue diminuir o ruído ao mesmo tempo em que preserva as bordas da região de interesse. Com a imagem melhorada, é determinado um template circular e a métrica de similaridade. A metodologia consegue encontrar em todas as imagens testadas a região de massa, gerando 1,1 falsos positivos por imagem.

Hatanaka et al. (2001) utilizam template matching para melhorar o desempenho do CAD na detecção de massas em regiões de fronteira da captura da imagem – normalmente junto ao músculo peitoral. Esta metodologia é o complemento de um trabalho anterior do mesmo grupo. Os principais pontos da metodologia são a extração do músculo peitoral, extração de candidatos a massa utilizando o template matching, em seguida a eliminação de falsos positivos através de características de segunda ordem (como medidas de haralick). A detecção é realizada apenas na região delimitada nos limites do sensor de captura da imagem.

Nunes, Silva e Paiva (2009) utilizam *template matching* como principal técnica de detecção auxiliada por um classificador MVS treinado com características calculadas pelo índice de Simpson para redução de falsos positivos. A metodologia inicia com um *k-means* para gerar uma classificação inicial. Posteriormente o *template matching* é usado para detectar regiões circulares. Essas regiões são submetidas à descrição e classificação. A metodologia atinge 79,12% de acurácia sobre imagens selecionadas do DDSM. Haja vista a necessidade de utilização de características de geometria, Nunes, Silva e Paiva (2009) estendem o trabalho anterior, utilizando como características excentricidade, circularidade, compacidade, densidade circular e desproporção circular de cada região segmentada. Este incremento no trabalho aumenta a acurácia da metodologia para 83,94%.

## 2.4 Baseada em morfologia e estatística

Li et al. (2001) propõem o melhoramento morfológico da imagem, na etapa de pré-processamento, para realizar uma melhor segmentação (utilizando técnicas de segmentação baseada em modelagem estocástica). O filtro consegue separar de maneira clara a região de massa do tecido de glanular ao redor, mesmo em imagens densas. Todavia, os elementos estruturantes (tamanho) são dependentes da base de imagens a ser analisada.

Brake e Karssemeijer (2001) utilizam dois métodos de segmentação, um baseado em *discrete dynamic countour* e outro baseado em *crescimento de regiões*, sendo o crescimento medido na forma de uma distribuição estatística. Os dois métodos geraram boa sensibilidade, todavia baixa especificidade. Foi verificado que o crescimento de regiões se adequa melhor a formas genéricas enquanto o primeiro método é melhor aplicado para formas esféricas. Gurcan (2001) utiliza *kmeans* e contornos ativos para segmentar massas em regiões delimitadas de mamografias. Em seu trabalho, estas massas são posteriormente descritas através de medidas morfológicas e textura para serem classificadas quanto ao seu comportamento.

Yuan et al. (2007) propõem uma melhoria ao método de segmentação proposto em um trabalho anterior do mesmo grupo. Neste novo trabalho, o grupo trabalha com uma segmentação de massas baseado em duas etapas: a

primeira etapa e também de inicialização utiliza *Radial Gradient Index* (RGI) e tem a função de delimitar regiões de segmentação iniciais a serem utilizadas pela segunda etapa que usa um contorno ativo com *Level Sets*. Os dois métodos são dependentes de uma correção de fundo para separação de massas. Para tanto, os autores utilizam uma técnica intitulada *two-dimensional background trend correction* com uma função polinomial de primeira ordem. Os resultados encontrados demonstram mais estabilidade e melhor desempenho do que os resultados encontrados apenas utilizando RGI.

Eltonsy, Tourassi e Elmaghraby (2007) propõem uma metodologia de múltiplas camadas concêntricas (MCL) para detecção de massas malignas. As camadas são concêntricas a um ponto focal e não necessariamente precisam sobrepor completamente a região interna. Para os ajustes de parâmetros da metodologia, metade das imagens da base é usada para treinamento e a outra metade para teste. A metodologia encontra em seus resultados 81% de sensibilidade com a geração de 0,6 falsos positivos por imagem.

Timp e Karssemeijer (2004) propõem a utilização de técnicas de programação dinâmica para gerar segmentações mais precisas em métodos baseados em contorno. Nos experimentos está a utilização dos seguintes algoritmos de segmentação adaptados para programação dinâmica: *path finding*, crescimento de regiões, contornos ativos discretos. Os resultados encontrados na metodologia obtiveram melhorias nos resultados de sobreposição de massas em relação ao contorno indicado pelo especialista. Com um melhor contorno encontrado, foram verificados também melhores resultados na classificação das regiões quanto ao comportamento maligno ou benigno.

Com a idéia de segmentar automaticamente massas em mamografias, Braz Junior et al. (2006) aplicam o algoritmo *k-means* juntamente com características obtidas pela matriz de co-ocorrência e geometria. O trabalho apresenta uma idéia inicial de aplicação com limitações presentes principalmente na escolha do número *k* de grupos. Com o intuito de melhorar os resultados obtidos no trabalho anterior, Martins et al. (2009) propõe a utilização conjunta de *k-means* com MVS para segmentação de regiões em mamografias. A metodologia atinge 85% de acerto, e gera 1.2 falsos positivos por imagem.

Martins et al. (2008) utilizam a rede Growing Neural Gas para detecção de massas

em mamografias do DDSM. As regiões extraídas pelo algoritmo são descritas usando a função de padrão pontual K de Ripley e geometria e posteriormente classificadas usando MVS. A metodologia atinge 90,73% de sensibilidade e 0,93 falso positivos por imagem. Borges et al. (2011) utilizam cadeia de Markov escondida junto com a função K de Ripley e características geométricas para descrever regiões de interesse segmentadas por uma rede Neural Celular. A metodologia é aplicada a base de imagens DDSM e atinge sensibilidade de 80% para encontrar corretamente as massas, com 5,3 falsos positivos por imagem que são reduzidos numa etapa de redução de falso positivos para 0,84 casos por imagem.

## 2.5 Múltiplas visões

Yin et al. (1991) trabalharam, mesmo com alinhamento manual, com os pares de visões Cranio Caudal (CC) e Médio-Lateral Obliqua (MLO) das mamas esquerda e direita e concluem que esta abordagem reduz significativamente a quantidade de falso positivos. Zheng e Chang (1995) comparam a utilização de uma visão da mamografia com a utilização da mamografia das mamas esquerda e direita. Em seu trabalho, é constatado que independente do ganho de informação obtido com a mamografia das duas mamas, as melhorias de desempenho são mais dependentes dos algoritmos de segmentação utilizados.

Mendez et al. (1998) utilizam a subtração bilateral dos pares de imagem a fim de identificar assimetrias. Inicialmente, o alinhamento é feito levando em consideração a borda e bico da mama. Em seguida, são retiradas diferenças de iluminação o que poderia influenciar na subtração das imagens. Após a normalização, as imagens são subtraídas entre si, e através de limiarização binarizadas para que uma rotina de crescimento de regiões possa identificar regiões diferentes entre as mamas. Filtro de área e geometria foram utilizados para diminuir a quantidade de falso positivos. O desempenho geral do sistema em capturar regiões positivas foi de 71%, sendo que, em média, são geradas 0,67 falsos positivos por imagem.

Ericeira, Silva e Paiva (2010) apresentam uma metodologia para detecção de regiões suspeitas em mamografias utilizando a informação presente em assimetrias das mamografias das duas mamas. Inicialmente, as imagens

são registradas e todas as áreas assimétricas são divididas em janelas descritas através da função de semivariograma cruzado. As descrições são submetidas a um classificador MVS que obtém 84% de sensibilidade com um alto índice de falsos positivos por imagem.

## 3 DIAGNÓSTICO

Como explicado anteriormente, a etapa de diagnóstico é feita a partir dos conjuntos de características previamente extraídas, as quais são submetidas à avaliação de um classificador, o qual, a partir de um treinamento prévio, informa se a região possui características de massa, podendo ainda informar quanto ao comportamento maligno.

Existem diversos trabalhos publicados que discutem esse assunto. Dessa forma, abordaremos alguns dos mais relevantes, organizados em grandes grupos relacionados com a forma que representam e tratam o problema.

### 3.1 Redes neurais

Em Sahiner et al. (2000) é feita uma avaliação da robustez de um método computado-rizado que foi desenvolvido para a classificação das massas benignas e malignas no que diz respeito a variações tanto da sutileza das lesões quanto na digitalização dos mamogramas. O método inclui, além da segmentação de regiões de massa e extração de características de maneira automática, a caracterização da lesão também de forma automática. Para tanto, utilizou uma Rede Neural Artificial na etapa de caracterização das massas a partir das características extraídas na etapa antecedente. O desempenho do método foi avaliado com a análise do *Receptor Operating Characteristic* (ROC), obtendo um índice Az de 0.90 de sensibilidade.

Mann e Shanmugam (2000) analisam o desempenho de um sistema automatizado de classificação das formas de massa em mamografias. Uma Rede Neural de duas camadas foi desenvolvida para automatizar a classificação das massas de mamografia em duas classes de forma: redonda ou irregular. A rede é treinada para dados segmentados tanto manualmente quanto automaticamente, e os resultados foram comparados. Observou-se que independentemente da forma de segmentação o classificador consegue um índice Az na curva ROC superior a 0.95 nos dois casos.

O objetivo do trabalho de Bovis e Singh (2000) é detectar massas em mamogramas,

com base em características de textura, em que as regiões suspeitas são identificadas pela subtração das imagens bilaterais de pares (esquerda e direita) da mama. Assim, o mamilo do seio é utilizado como um ponto comum de rotação, cuja finalidade é facilitar um alinhamento com a maior correlação antes da etapa de subtração. Cinco matrizes co-ocorrência são construídas em quatro distâncias diferentes para cada região suspeita em cada uma das duas imagens resultantes da subtração, produzindo cada matriz 14 características de textura (totalizando 70 características por região), resultando em 4 vetores de características (um para cada distância). Porém, somente aquele vetor com melhor taxa de verdadeiro positivo (determinado pela técnica de análise discriminante linear) é usado pela Rede Neural Artificial, produzindo uma taxa média de reconhecimento de 77%.

Cheng et al. (2006), Kumar, Zhang e Verma (2006), Mazurowski et al. (2008) e, ainda, Verma e Panchal (2006) demonstram que classificadores baseados em redes Neurais mostraram ser mais adequados para o dilema de classificar condições médicas, como câncer de mama. Na verdade, as redes Neurais têm demonstrado sua superioridade nestes domínios segundo Kumar, Zhang e Verma (2006) e Mazurowski et al. (2008), devido a sua capacidade para classificar anormalidades (como benigno ou maligno), com base em características de entrada quando estas são fracamente correlacionadas com o diagnóstico, do que utilizando outras técnicas para este fim. No entanto, trabalhos como os de Cheng et al. (2006), Manrique e Rios (2006) e também Verma e Panchal (2006) mostram que as redes neurais podem sofrer de vários problemas que limitam sua eficácia.

### 3.2 Redes bayesianas

Em Oliver et al. (2008) é apresentado um novo classificador Bayesiano, o qual é baseado na combinação de dois classificadores: *K-means* e Árvore de Decisão. Essa combinação é conseguida através de uma abordagem de atribuição "soft", onde os resultados binários (ou discretos) da classificação são transformados em valores contínuos que retratam membros da classe. Assim, a fim de alcançar uma adesão para cada classe, ao invés de considerar os critérios de votação, é levado em conta o resultado de cada classificador. Adicionando todos os resultados para a mesma classe e normalizando-os, os membros de cada classe são obtidos.

O resultado obtido de classificação correta foi cerca de 78% das mamografias.

Zheng, B. et al. (1999), compararam o desempenho de uma rede Neural Artificial e uma Bayesiana, usadas no mesmo esquema CAD. Para tanto, utilizou o mesmo banco de dados e algoritmo genético para otimizar ambas. Os resultados experimentais demonstraram que, com o uso do algoritmo genético, o desempenho das duas redes convergiu para o mesmo nível na detecção de massas de mamografias digitalizadas. Portanto, o estudo concluiu que a melhoria do desempenho dos sistemas de CAD pode ser mais dependente da seleção da característica de otimização e diversidade da base de dados de treinamento do que de qualquer paradigma particular da máquina de classificação.

Finalmente, Martins et al. (2006) fazem uso de características de segunda ordem extraídas de matrizes de co-ocorrência sobre regiões de interesse do MIAS passadas como entrada para uma rede neural Bayesiana, alcançando 86,84% de acerto total na tarefa de distinção de regiões em benignas e malignas.

### 3.3 Árvore de decisão binária

Zhen e Chan (2001) utilizam uma Árvore de Decisão Binária para determinar se as regiões encontradas nas etapas anteriores do método proposto pelo estudo são suspeitas de câncer. Assim, cada nó nesta árvore contém um limiar (*threshold*) para a característica que ele representa e os dados de entrada irão fluir ao longo da direção das setas com base no resultado em cada nó. Quando o sinal chegar ao fim significa que um resultado de classificação foi feito. Os resultados da verificação mostram que o algoritmo proposto tem uma sensibilidade de 97,3% e o número de falsos positivos por imagem é 3,92%. Seguindo a mesma linha, ou seja, utilizando Árvores de Decisão Binária na etapa de classificação Zheng, A. K. et al. (1999) também produziram os mesmos resultados.

Li et al. (1995) mostram a utilização da Árvore de Decisão Binária a partir de regiões de interesse segmentadas utilizado um método com uma modificação da cadeia de Markov aleatória. Assim, foram obtidos para os 16 casos que constituíam a base de dados, uma sensibilidade de 94% foi observada com 1,5 falsos positivos por imagem.



### 3.4 Máquina de vetor de suporte (MVS)

Martins et al. (2010) comparam o desempenho dos classificadores MVS e Rede Neural Bayesiana na tarefa de classificação regiões extraídas de mamografias em tecidos normais e anormais baseados em características calculadas através do semivariograma, levando em consideração a distribuição espacial de informações introduzida, inicialmente, neste contexto em Silva, Carvalho e Gattas (2005). As variáveis são inicialmente selecionadas usando *Stepwise*. O estudo observa como resultado o melhor desempenho de classificadores MVS, com 86,11% de acerto geral nos testes contra 76,85% das redes neurais Bayesianas.

Com o intuito de analisar o desempenho dos classificadores utilizando multiresolução, Braz Junior et al. (2007) utilizam a transformada de *Wavelet*, juntamente com geoestatística, para criar vetores descritivos da textura de regiões extraídas de mamografias e classificá-las, usando MVS, em massa ou normais. A metodologia alcança 98,24% de acerto total e comprova o desempenho da utilização de geoestatística para descrição de regiões mamográficas.

Martins et al. (2007) introduzem a utilização de descrição pontual de padrões, usando a função K de *Ripley* como descritor de textura de regiões extraídas de mamografias do DDSM. Os vetores de características são submetidos a um classificador MVS e atingem 94,25% de acurácia na classificação das regiões em massa e não massa.

Haja vista o melhor desempenho dos classificadores MVS, esta categoria de máquinas de reconhecimento de padrões é utilizada com sucesso nos trabalhos de Braz Junior et al. (2009a) e também Braz Junior et al. (2009b) para classificação de massas quanto ao caráter anormal. Nos seus trabalhos, são utilizados os classificadores *One-Class SVM* e *Two-Class MVS* juntamente com estatísticas geoespaciais obtidas através do semivariograma parametrizado. As imagens base são obtidas do DDSM. O primeiro trabalho, usando o índice de *Getis* para descrever a textura de regiões extraídas alcança 95,19% de acerto total. O segundo trabalho que utiliza Índice de *Moran* e Coeficiente de *Geary* alcança, respectivamente, para cada métrica, 99,39% e 96,04% na tarefa de classificação de massas e não massas.

Este último trabalho ainda realiza a classificação de massas em benignas ou malignas utilizando os mesmos índices e classificador

atingindo 88% de acerto. O trabalho ainda apresenta o comportamento estatístico das métricas e comprovam ser distinguíveis, fatos que influenciam positivamente no momento do reconhecimento dos padrões, anteriormente notado no trabalho de Silva Junior et al. (2006). Neste são usadas, com sucesso, medidas extraídas do semivariograma juntamente com redes neurais MLP para distinguir massas obtidas da base MIAS, quanto ao comportamento de malignidade, atingindo 88% de acerto total.

## 4 DISCUSSÃO E CONCLUSÃO

O processamento de imagens juntamente com técnicas de diferentes áreas como estatística, espacial, reconhecimento de padrões e inteligência artificial têm proposto a aplicação de uma vasta lista de técnicas. O objetivo principal, para estudos na área de câncer de mama, é criar uma metodologia capaz de prover um mecanismo que auxilie o especialista a acompanhar, delimitar e diagnosticar, através de imagens radiográficas, qualquer aparecimento de tecido que represente uma neoplasia numa mama. Dentre as prioridades está a diminuição da quantidade de exames invasivos, como a biopsia. Exames que seriam realizados em situações de desconfiança pelo especialista, poderiam ser descartados pela aplicação que estaria baseando sua recomendação em muitas outras bases de informação. Evidente que a aplicação não tem o intuito de substituir o especialista, mas sim prover uma segunda opinião de maneira clara e rápida.

Os desafios que compõe esta tarefa iniciam-se no momento da captura da imagem radiográfica, tipicamente mamografias. Este processo inclui ruídos ou mesmo falhas que podem mascarar um tecido anormal. Logo se faz necessário que uma grande parte da metodologia leve em consideração um tratamento inicial das imagens de maneira a torná-las normalizadas sem desfigurar quaisquer feições presentes internamente. E esta foi uma característica percebida em todos os trabalhos relacionados. Mesmo que em alguns deles sejam utilizados mamografias digitais, ou seja, não teriam o processo de captura de filme que normalmente introduz erros.

O desafio continua para metodologias que propõem a detecção de massas em mamografia quando se analisa o grau de riqueza de informações presentes nesta imagem concluindo ser escasso. Pode-se concluir que a

massa visualmente será uma região semelhante a outras regiões em uma imagem de mamografia. Todavia, se analisarmos a forma que os pontos que compõem essa imagem se distribuem e se associam espacialmente podemos perceber que a massa possui um comportamento homogêneo e em camadas, o que permite criar técnicas que, empiricamente parametrizadas, são capazes de detectar essas regiões em detrimento as vizinhas semelhantes que não são massas.

Partindo desse pressuposto, pesquisadores relatados neste trabalho utilizaram métodos de corte, transformadas de wavelet, crescimento de regiões, contornos ativos, dentre outras técnicas, com o intuito de obter os parâmetros específicos capazes de realizar a segmentação da mamografia levando em consideração que apenas regiões pertencentes a massas deveriam estar presentes no final da etapa. Todavia, é bastante comum que estes métodos não consigam realizar esse trabalho com maestria e quase todas as metodologias revisadas possuem um etapa posterior de redução de regiões que foram identificadas como regiões positivas (que contenham massas).

Essa última etapa, redução de falsos positivos, geralmente é auxiliada por algum conhecimento prévio ou observado sobre massas. Uma observação pertinente sobre massas é que possuem forma circular e achatada. Auxiliado a geometria está a análise de textura aplicada nos trabalhos revisados utilizando desde índices estatísticos, mas também espaciais. Unindo estas características obtidas e aplicando a uma máquina de reconhecimento de padrões é possível diminuir drasticamente o número de falso positivos gerados em uma imagem. A importância da redução de falsos positivos se resume a necessidade de preservar o paciente de procedimentos invasivos como biopsia.

Mesmo se tendo uma receita a ser seguida de etapas, é muito comum as metodologias de detecção prover novas etapas, ou incluir subetapas. Isso acontece principalmente pela dificuldade de se analisar a imagem mamográfica e faz com que os estudos de uma forma geral sejam bastante diferentes entre si. Por isso, a grande quantidade de técnicas diferentes empregadas mesmo levando em consideração a mesma base de imagens como teste. Conclui-se, portanto, que a tarefa de detecção de massas em mamografias está intimamente relacionada com a imagem mamográfica, a

base de onde foi obtida e com o ajuste e especialização de técnicas de segmentação.

A partir de regiões identificadas como suspeitas ou regiões de interesse, o novo desafio para esta área passa a ser distinguir o comportamento das mesmas quanto à malignidade. As técnicas então, contando com informações empiricamente observadas como a forma espiculada de massas malignas ou mesmo através do comportamento estatístico dos pontos que compõem a região estudada, fazem o processo de classificação ou categorização de padrões, neste caso, quanto ao padrão benigno ou maligno.

Muitas técnicas, antes da definição dos diagnósticos, possuem inúmeras regiões selecionadas, prováveis resultados de um método de detecção e precisam descobrir dentre este conjunto, quais regiões realmente representam massas. A literatura apresenta trabalhos que propõem soluções para os dois problemas de maneira individualizada e alguns em conjunto. Todavia, verifica-se o consenso ao afirmar que metodologias que classificam regiões em massas e não massas têm tido um desempenho superior a metodologias que tratam o problema de classificar massas em malignas ou benignas.

A observação nos leva a concluir que nesta circunstância o padrão gerado por uma massa é bem diferente que o padrão de um tecido normal. Todavia, o padrão gerado por uma massa maligna não possui uma quantidade de diferenças suficientemente grande para distingui-las computacionalmente em grupos diferentes. Este fato tem instigado alguns trabalhos a criar uma etapa anterior na qual é realizada uma nova representação da informação através de multiresolução ou decomposição de fatores primordiais dessas imagens. O objetivo destes últimos trabalhos é diminuir a correlação existente entre massas benignas e malignas e aumentar o desempenho das ferramentas.

Por fim, os estudos revisados trazem como resultados propostas de metodologias que poderiam ser aplicadas diretamente em sistemas CAD/CADx e distribuídas para aumentar a eficiência no tratamento de câncer de mama em geral. Trazem também o conhecimento acumulado e aplicado pela área nos últimos anos para servir de referencial para novos estudos que estão iniciando. Pode-se concluir que ainda é uma área em aberto, com muitos objetivos a serem trabalhos até chegar a situação ideal de se conseguir uma metodologia de precisão máxima no reconhecimento de uma neoplasia de mama.

## REFERÊNCIAS

- AMERICAN CANCER SOCIETY. Learn about breast cancer. 2007. Disponível em <<http://www.cancer.org>>. Acesso em: 20 set. 2011.
- AZEVEDO, C. M.; PEIXOTO, J. E. *Falando sobre Mamografia*. Rio de Janeiro: INCA, 1993.
- BORGES, S. W.; et al. Detection of masses in mammogram images using CNN, geostatistic functions and SVM. *Computers in Biology and Medicine*, p. 653-664, 2011.
- BOVIS, K.; SINGH, S. Detection of masses in mammograms using texture features. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 15., 2000. *Proceedings...* [S.l.: s.n.], 2000. p. 267-269.
- BRAKE, G. M.; KARSSEMEIJER, A. Segmentation of suspicious densities in digital mammograms. *Med. Phys.*, p. 259-266, 2001.
- BRAZ JUNIOR, G. et al. Classification of breast tissues using getis-ord statistics and support vector machine. *Journal Intelligent Decision Technologies*, v. 3, p. 197-205, 2009a.
- \_\_\_\_\_. Classification of breast tissues using Moran's index and Geary's coefficient as texture signatures and SVM. *Computers in Biology and Medicine*, v. 39, p. 1063-1072, 2009b.
- BRAZ JUNIOR, G. et al. Identificação de massas em mamografias usando textura, geometria e algoritmos de agrupamento e classificação. In: WORKSHOP DE INFORMÁTICA MÉDICA, 2006, Vila Velha. *Workshop...* Vila Velha: [s.n.], 2006.
- BRAZ JUNIOR, G. et al. Breast tissues classification based on the application of geostatistical features and wavelet transform. In: CONFERENCE ON 6TH INTERNATIONAL SPECIAL TOPIC INFORMATION TECHNOLOGY APPLICATIONS IN BIOMEDICINE, 2007, Tokyo. *Anais...* [S.l.:s.n.], 2007. p. 227-230.
- BRZAKOVIC, D.; LUO, X. M.; BRZAKOVIC, P. An approach to automated detection of tumors in mammograms. *IEEE Trans. Med. Imaging*, v. 9, n. 3, p. 233-241, 1990.
- CAMPANINI, R. et al. A novel featureless approach to mass detection in digital mammograms based on support vector machines. *Phys. Med. Biol.*, v. 49, n. 6, p. 961-975, 2004.
- CHENG, H. D. et al. Approaches for automated detection and classification of masses in mammograms. *The Journal of Pattern Recognition*. v. 39, p. 646-668, 2006.
- ELTONSY, N. H.; TOURASSI, G. D.; ELMAGHRABY, A. S. A concentric morphology model for the detection of masses in mammography. *IEEE Trans. Med. Imag.*, v. 26, n. 6, p. 880-889, jun. 2007.
- ERICEIRA, R. D.; SILVA, C. A.; PAIVA, C. A. Detecção de regiões suspeitas em mamografias digitais utilizando descrição espacial com variograma cruzado. In: CONGRESSO BRASILEIRO DE INFORMÁTICA EM SAÚDE, 12., 2010, Porto de Galinhas. *Anais...* Porto de Galinhas: [s.n.], 2010.
- FREER, T. W.; ULISSEY, M. J. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology*, v. 22, n. 3, p.781-786, 2001.
- GURCAN, M. N. Computer-aided characterization of mammographic masses: accuracy of mass segmentation and its effects on characterization. *IEEE Trans. Med. Imaging*, v. 20, n. 12, p. 1275-1284, 2001.
- GUPTA, R.; UNDRILL, P. E. The use of texture analysis to delineate suspicious masses in mammography. *Phys. Med. Biol.*, v.40, p.835-855, 1995.
- HATANAKA, Y. et al. Development of an automated method for detecting mammographic masses with a partial loss of region. *IEEE Trans. Med. Imaging*, v. 20, n. 12, p. 1209-1214, 2001.
- HEATH, M. et al. *Proceedings of the fifth international workshop on digital mammography*. [S.l.]: M.J. Yaffe; Medical Physics Publishing, 2001. p. 212-218.
- HUO, Z. et al. Analysis of spiculation in the computerized classification of mammographic masses. *Med. Phys.*, v. 22, n. 10, p. 1569-1579, 1995.
- INSTITUTO NACIONAL DE CÂNCER. *Estimativa 2010*. Incidência do câncer no Brasil. Rio de Janeiro: INCA, 2010. Disponível em: <[http://www1.inca.gov.br/estimativa/2010/index.asp?link=conteudo\\_view.asp&ID=5](http://www1.inca.gov.br/estimativa/2010/index.asp?link=conteudo_view.asp&ID=5)>. Acesso em: 2 out. 2011.
- KINOSHITA, S. K. et al. An automatic method for detection of the nipple and pectoral muscle in digitized mammograms. In: CONGRESSO LATINO AMERICANO DE ENGENHARIA BIOMÉDICA, João Pessoa. *Proceeding of the International Federation for Medical and Biological Engineering*, v. 5. p. 1303-1306. 2004.
- KOM, G.; TIEDEU, A.; KOM, M. Automated detection of masses in mammograms by local adaptive thresholding. *Comput. Biol. Med.*, v. 37, n. 1, p. 37-48, 2007.

- KUMAR, K.; ZHANG, P.; VERMA, B. Application of decision trees for mass classification in mammography. In: INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS AND KNOWLEDGE DISCOVERY, 2006, Pequim. *Anais...* Pequim: [s.n.], 2006. p. 366-376.
- KUPINSKI, M. A.; GIGER, M. L. Automated seeded lesion segmentation on digital mammograms. *IEEE Trans. Med. Imaging*, v. 17, n. 4, p. 510-517, 1998.
- LAI, S. M.; LI, X.; BISCOF, W. F. On techniques for detecting circumscribed masses in mammograms. *IEEE Trans. Med. Imag.*, v. 18, n. 4, p. 377-386, Dec. 1989.
- LI, H. D. et al. Computerized radiographic mass detection C part I: lesion site selection by morphological enhancement and contextual segmentation. *IEEE Trans. Med. Imaging*, v. 20, n. 4, p. 289-301, 2001.
- LI, H.D. et al. Markov random field for tumor detection in digital mammography. *IEEE Trans. Med. Imaging*, v. 14, n. 3, p. 565-576, 1995.
- LIU, S.; BABBS, C. F.; DELP, E. J. Multiresolution detection of speculated lesions in digital mammograms. *IEEE Trans. Image Process.*, v. 10, n. 6, p. 874-884, jun. 2001.
- MANN, B. L; SHANMUGAM, N. Using neural networks with wavelet transforms for an automated mammographic mass classifier. In: ANNUAL EMBS INTERNATIONAL CONFERENCE, 22., 2000. *Proceedings...* v. 2, 2000, p. 985-987.
- MANRIQUE, D.; RIOS, J.; Rodriguez-Paton, A. Evolutionary system for automatically constructing and adapting radial basis function networks. *Neurocomputing*, v. 69, p. 2268-2283, Oct. 2006.
- MARTINS, O. L. et al. Comparison of support vector machines and bayesian neural networks performance for breast tissues using geostatistical functions in mammographic images. *International Journal on Computational Intelligence and Applications*, v. 9, p. 271-288, 2010.
- MARTINS, O. L. et al. Detection of masses in digital mammograms using K-means and support vector machine. *Electronic Letters on Computer Vision and Image Analysis*, v. 8, p. 39-50, 2009.
- MARTINS, O. L. et al. Classification of normal, benign and malignant tissues using co-occurrence matrix and bayesian neural network in mammographic images. In: BRAZILIAN NEURAL NETWORKS SYMPOSIUM, 9., 2006, Ribeirão Preto. *Ninth Brazilian Symposium on Neural Networks (SBRN'06)*. Los Alamitos, CA. USA: IEEE Computer Society, 2006. p. 5.
- MARTINS, O. L. et al. Detection of breast masses in mammogram images using growing neural gas algorithm and ripley's. *Journal of Signal Processing Systems*, v. 10, p. 1-14, 2008.
- MARTINS, O. L. et al. Classification of breast masses in mammogram images using ripley's K function and support vector machine. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND DATA MINING, 5., 2007. *Lecture Notes in Artificial Intelligence*, Leipzig, v. 4571. p. 784-794, 2007.
- MATSUBARA, T. et al. Development of new schemes for detection and analysis of mammographic masses. *Intel. Inform. Systems*, p. 63-66, 1997.
- MAZUROWSKI, M. et al. Training neural network classifiers for medical decision making: the effects of imbalanced datasets on classification performance. *Neural Networks*, 21, p. 427-436, 2008.
- MEERSMAN, D.; SCHEUNDERS, P.; DYCK, V. D. Detection of microcalcifications using non-linear filtering. In: EUROPEAN SIGNAL PROCESSING CONFERENCE, 4., 1998. *Anais...* [S.l.:s.n.], 1998. p. 2465-2468.
- MENDEZ, A. J. et al. Computer-aided diagnosis: automatic detection of malignant masses in digitized mammograms. *Med. Phys.*, v. 25, n. 6, p. 957-964, 1998.
- NUNES, P. A.; SILVA, C. A.; PAIVA, C. A. Detection of masses in mammographic images using geometry, simpson's diversity index and SVM. *International Journal of Signal and Imaging Systems Engineering*, v. 3, p. 40-51, 2010.
- \_\_\_\_\_. Detection of masses in mammographic images using simpsons diversity index in circular regions and SVM. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND DATA MINING, 2009. Leipzig. *Machine Learning and Data Mining in Pattern Recognition*, Berlin Heidelberg: Springer Verlag in LNAI series, 2009. v. 5632, p. 540-553.
- OLIVER, A. et al. A novel breast tissue density classification methodology: information technology in biomedicine. *IEEE Transactions On*, v. 12, n. 1, p. 55, 2008.
- PETRICK, N. et al. Combined adaptive enhancement and region-growing segmentation of breast masses on digitized mammograms, *Med. Phys.* v. 26, n. 8, p. 1642-1654, 1999.



- SAHINER, B. et al. Computerized classification of benign and malignant masses on digitized mammograms: a study of robustness. *Acad. Radiol.*, v. 7, p. 1077-1084, 2000.
- SAMPAT, M. P.; BOVIK, A. C. Detection of spiculated lesions in mammograms, in Proc. 25th Annu. Int. Conf. *IEEE Eng. Med. Biol. Soc.*, v. 1, p. 810-813, 2003.
- SILVA JUNIOR, R. V. et al. Semivariogram applied for classification of benign and malignant tissues in mammography. In: INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND RECOGNITION, 2006, Póvoa de Varzim. *Lecture Notes in Computer Science...* Berlin Heidelberg: Springer-Verlag, 2006. v. 4142. p. 570-579.
- SILVA, C. A.; CARVALHO, P. C. P.; GATTASS, M. Diagnosis of lung nodule using semivariogram and geometric measures in computerized tomography images. *Computer Methods and Programs in Biomedicine*, v. 79, p. 31-38, 2005.
- TIMP; KARSSEMEIJER, N. A new 2d segmentation method based on dynamic programming applied to computer aided detection in mammography, *Med. Phys.*, v. 31, n. 5, p. 958-971, 2004.
- TOURASS, G. D.; VARGAS-VORACEK, G. D. Computer-assisted detection of mammographic masses: A template matching scheme based on mutual information. *Med. Phys.*, v. 30, n. 8, p. 2123-2130, 2003.
- VERMA, B.; PANCHAL, R. Neural networks for the classification of benign and malignant patterns in digital mammograms. In: FULCHER, John (Ed.). *Advances in applied artificial intelligence*. USA: Idea Group, Inc., 2006.
- INTERNATIONAL AGENCY FOR RESEARCH ON CANCER. *World Cancer Report*. 2008. Disponível em: <<http://www.iarc.fr/en/Publications/PDFs-online/World-Cancer-Report/World-Cancer-Report-2008>>. Acesso em: 30 set. 2011.
- YIN, F. F. et al. Computerized detection of masses in digital mammograms: analysis of bilateral subtraction images. *Med. Phys.*, v. 18, n. 5, p. 955-963, 1991.
- YUAN, Y. et al. A dual-stage method for lesion segmentation on digital mammograms. *Med. Phys.*, v. 34, n. 11, p. 4180-4193, 2007.
- ZHEN, L.; CHAN, A. K. An artificial intelligent algorithm for tumor detection in screening mammogram. *IEEE Trans. Med. Imaging*, v. 20, n. 7, p. 559-567, 2001.
- ZHENG, A. K. et al. Detection of cancerous masses for screening mammography using discrete wavelet transform-based multiresolution Markov random field, *J. Digital Imaging*, v. 12, n. 2, p. 18-23, 1999.
- ZHENG, B. et al. Comparison of artificial neural network and bayesian belief network in a computerassisted diagnosis scheme for mammography. *IEEE International Conference on Neural Networks*, p. 4181-4185, 1999.
- ZHENG, Y .H.; CHANG, D. G. Computerized detection of masses from digitized mammograms: comparison of singleimage segmentation and bilateral-image subtraction, *Assoc. Univ. Radiologists*, v. 2, n. 12, p. 1056-1061, 1995.