

DIAGNÓSTICO AUXILIADO POR COMPUTADOR NA RADIOLOGIA*

Paulo Mazzoncini de Azevedo-Marques¹

Resumo Vários desenvolvimentos tecnológicos estão convergindo de forma a aumentar a influência da área de imagens nas pesquisas biomédicas e na medicina clínica. Muitos pesquisadores têm trabalhado no desenvolvimento de sistemas computadorizados para detecção automatizada e quantificação de anormalidades em imagens radiológicas. Estes sistemas são dedicados ao diagnóstico auxiliado por computador. Este artigo discute os conceitos básicos relacionados ao diagnóstico auxiliado por computador e apresenta uma revisão bibliográfica sobre o assunto.

Unitermos: Diagnóstico auxiliado por computador. Radiografia digital. Processamento de imagem. Visão computacional. Inteligência artificial.

Abstract *Computer-aided diagnosis in radiology.*

Several technological developments are contributing to increase the use of imaging techniques in biomedical research and clinical medicine. Many investigators have attempted to develop computerized routines for automated detection and quantitative analysis of abnormalities in radiological images. These routines have been referred to as computer-aided diagnosis. In this paper we discuss the basic concepts of computer-aided diagnosis and review the literature.

Key words: Computer-aided diagnosis. Digital radiography. Image processing. Computer vision. Artificial intelligence.

INTRODUÇÃO

Diagnóstico auxiliado por computador (“computer-aided diagnosis” – CAD) pode ser definido como um diagnóstico feito por um radiologista que utiliza o resultado de análises quantitativas automatizadas de imagens radiográficas como uma “segunda opinião” para a tomada de decisões diagnósticas^(1,2). É importante ressaltar que o computador é utilizado somente como uma ferramenta para obtenção de informação adicional, sendo o diagnóstico final sempre feito pelo radiologista, o que diferencia claramente o conceito básico de diagnóstico auxiliado por computador do conceito de “diagnóstico automatizado”, que foi um conceito proposto e estudado nas décadas de 60 e 70⁽¹⁾.

A finalidade do CAD é melhorar a acurácia do diagnóstico, assim como a consistência da interpretação da imagem radiológica, mediante o uso da resposta

do computador como referência. A resposta do computador pode ser útil, uma vez que o diagnóstico do radiologista é baseado em avaliação subjetiva, estando sujeito a variações intra e interpessoais, bem como perda de informação devido à natureza sutil do achado radiológico, baixa qualidade da imagem, sobreposição de estruturas, fadiga visual ou distração. Além disso, foi demonstrado que uma dupla leitura (por dois radiologistas) pode aumentar a sensibilidade do diagnóstico⁽³⁾. A proposta do CAD é funcionar como um segundo especialista.

Basicamente, existem dois tipos de aplicações de sistemas CAD. Um é o *auxílio à detecção* de lesões, a partir da localização de padrões anormais através da varredura da imagem pelo computador (por exemplo, agrupamentos de microcalcificações em imagens mamográficas ou nódulos pulmonares em imagens de tórax). O outro é o *auxílio ao diagnóstico*, através da quantificação de características da imagem e sua classificação como correspondendo a padrões normais ou anormais (por exemplo, a associação da quantidade e forma das microcalcificações presentes em um agrupamento com a malignidade ou não do tumor, ou a associação da textura dos pulmões com lesões intersticiais em imagens de tórax). Em geral, os sistemas CAD utilizam-se

de técnicas provenientes de duas áreas do conhecimento: *visão computacional*, que envolve o processamento de imagem para realce, segmentação e extração de atributos, e *inteligência artificial*, que inclui métodos para seleção de atributos e reconhecimento de padrões⁽⁴⁾.

Por ter base conceitual genérica e ampla, a idéia do CAD pode ser aplicada a todas as modalidades de obtenção de imagem, incluindo radiografia convencional, tomografia computadorizada, ressonância magnética, ultra-sonografia e medicina nuclear. Pode-se, também, desenvolver esquemas de CAD para todos os tipos de exame de todas as partes do corpo, como crânio, tórax, abdome, osso e sistema vascular, entre outros. Porém, os principais objetos de pesquisa para o desenvolvimento de sistemas CAD têm sido as áreas de mamografia, para a detecção precoce do câncer de mama; tórax, para a detecção de nódulos pulmonares, lesões intersticiais e pneumotórax; e angiografia, para a análise quantitativa de estenoses e de fluxo sanguíneo⁽¹⁾.

VISÃO COMPUTACIONAL E INTELIGÊNCIA ARTIFICIAL

Visão computacional envolve a utilização de um computador para extração de atributos de uma imagem em formato

* Trabalho realizado no Serviço de Radiodiagnóstico do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (FMRP-USP), Ribeirão Preto, SP. Apoio financeiro da Fapesp.

1. Engenheiro Eletrônico, Doutor em Física Aplicada, Docente do Centro de Ciências das Imagens e Física Médica do Departamento de Clínica Médica da FMRP-USP.

Endereço para correspondência: Prof. Dr. Paulo M. de Azevedo Marques. Rua Ceará, 3005, apto. 22. Ribeirão Preto, SP, 14090-300. E-mail: pmarques@fmrp.usp.br

Aceito para publicação em 19/6/2001.

digital. De modo geral, os conceitos de visão computacional utilizados em sistemas CAD podem ser divididos em duas etapas principais.

A primeira é o processamento da imagem para realce e segmentação das lesões. A segmentação é o processo que subdivide a imagem em suas partes ou objetos constituintes e utiliza as propriedades básicas de descontinuidade dos níveis de cinza para segmentação através de bordas, fronteiras e linhas, ou de similaridade destes para separação de regiões que apresentem determinada característica em comum⁽⁴⁾. Como resultado da segmentação, obtém-se um conjunto de objetos que permitirá uma descrição da imagem digitalizada por meio da quantificação de determinadas propriedades.

Várias técnicas de processamento de imagens têm sido utilizadas para diferentes tipos de lesões. Algumas das técnicas mais empregadas incluem filtragem baseada na análise de Fourier, transformada "wavelet", filtragem morfológica e técnicas de imagem-diferença, entre outras. A Figura 1 apresenta um exemplo de realce de microcalcificações mamográficas utilizando a transformada "wavelet"⁽⁵⁾, e a Figura 2 mostra a segmentação de uma região de interesse de uma mamografia digitalizada contendo um agrupamento de microcalcificações suspeitas⁽⁶⁾. O método de segmentação exemplificado é composto pela combinação de um método de "stretching", uma técnica de imagem-diferença e um método de "thresholding" local adaptativo. O método de "stretching", ou linearização do histograma, é um método simples mas muito utilizado em processamento de imagens e tem por objetivo aumentar o contraste da imagem. Este método pode ser aplicado à imagem antes da etapa de segmentação e extração de atributos, para aumentar o contraste e normalizar os níveis de cinza entre os limites máximo e mínimo da escala de tons de cinza. A técnica imagem-diferença é geralmente utilizada para o realce de microcalcificações e nódulos mediante supressão das estruturas de fundo da imagem, causadas pela anatomia normal da mama. Este procedimento é realizado por dois filtros espaciais gerando uma imagem

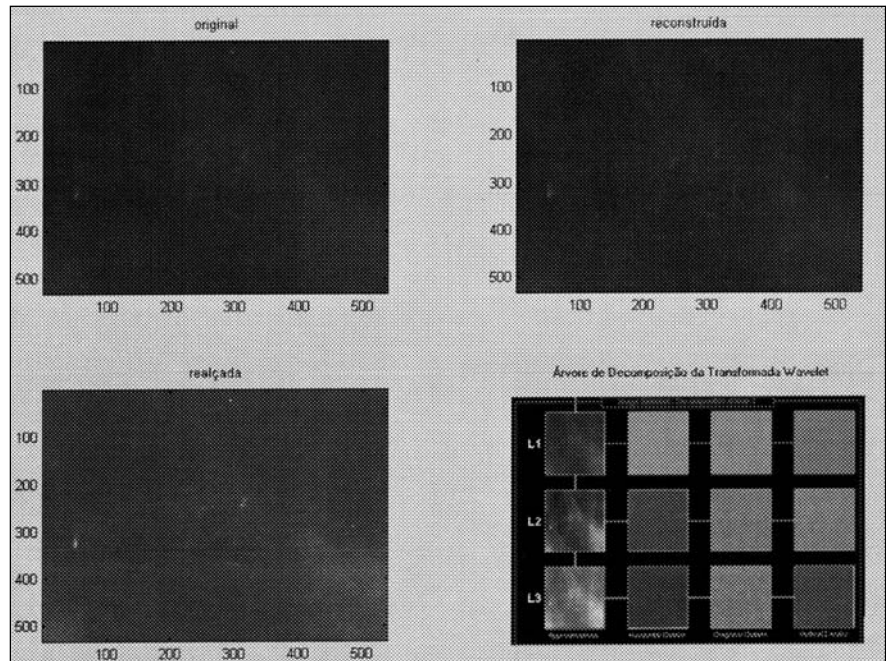


Figura 1. Realce de microcalcificações através da aplicação da transformada "wavelet".

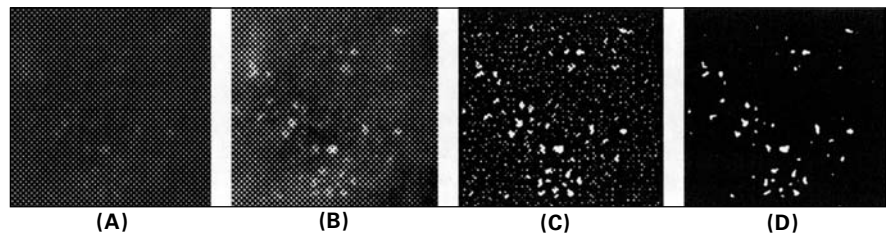


Figura 2. Exemplo de segmentação de imagem. **A:** Região de interesse contendo um agrupamento suspeito de microcalcificações. **B:** Imagem após "stretch". **C:** Imagem-diferença. **D:** Imagem segmentada.

com as microcalcificações realçadas e outra com as microcalcificações suprimidas. Nesta última permanecem apenas as estruturas de fundo do mamograma, uma vez que os sinais são suprimidos. Subtraindo as duas imagens, tem-se uma terceira imagem em que as altas variações de densidade de fundo são removidas. A técnica de "thresholding" local adaptativo é baseada em estatísticas locais (média e desvio padrão) dos valores de cinza dos pixels contidos em uma região quadrada de $m \times m$ pixels, tendo por finalidade a binarização da imagem. Para cada região da imagem é calculado o valor do limiar ("threshold"), sendo o valor do seu pixel central ajustado para o nível de cinza máximo (255) ou mínimo (0), dependendo se o seu valor inicial é, respectivamente, maior ou menor

do que o valor calculado. No exemplo da Figura 2 o tamanho da região foi fixado em 7×7 pixels após um ensaio com diferentes tamanhos de máscaras. Após a binarização, a imagem é varrida para a separação das microcalcificações, que são os objetos constituintes da imagem.

Após o realce e a segmentação, a etapa seguinte do processamento envolve a quantificação de atributos da imagem, como tamanho, contraste e forma dos seus objetos constituintes. Dependendo da sua natureza, estes atributos podem ou não ser compatíveis com a visão humana. É possível a definição de um grande número de atributos baseados em fórmulas matemáticas que podem não ser facilmente compreendidos por um observador humano. Porém, geralmente é mais útil definir, particularmente nos estádios

iniciais de desenvolvimento de sistemas CAD, atributos de imagem que já tenham sido anteriormente reconhecidos e descritos subjetivamente pelos radiologistas, uma vez que estes são baseados na observação de um número muito grande de casos e possibilitam diagnósticos com alta acurácia e confiabilidade⁽¹⁾. De modo geral, os atributos são quantificados a partir de propriedades métricas, topológicas e de textura dos objetos.

Conceitualmente, propriedades métricas utilizam generalizações da noção de distância euclidiana, sendo pois sensíveis a qualquer distorção “elástica” do plano imagem⁽⁴⁾. As propriedades métricas mais diretas que podem ser utilizadas na descrição de formas são a área, o perímetro, a razão de “delgadeza” e a razão de aspecto. A razão de delgadeza é uma relação entre a área da figura e o perímetro ao quadrado e fornece uma noção de quanto a figura é arredondada (gorda) ou alongada (magra). Figuras mais arredondadas terão razão de delgadeza próxima de 1 (1 é o valor máximo e será obtido quando a figura for um círculo), e figuras com formas próximas a linhas terão razão de delgadeza próxima de zero. Além disso, a razão de delgadeza é adimensional, dependendo somente da forma da figura e não da escala. A razão de aspecto é também uma medida do alongamento de uma figura. A razão de aspecto de um retângulo é a razão do seu comprimento por sua largura, de modo que uma razão de aspecto próxima de 1 indica um retângulo “gordo”. Para estender este conceito a figuras arbitrárias, deve-se envolver a figura em um retângulo e considerar que a razão de aspecto da figura será a mesma que a do retângulo que a envolve. Pode-se notar que muitos resultados são possíveis, dependendo das várias maneiras que se coloca o retângulo ao redor da figura. A definição mais simples especifica que os lados do retângulo devem ser paralelos aos eixos coordenados do plano da imagem.

Outra maneira possível de se caracterizar um subconjunto de uma imagem é especificar algumas de suas propriedades topológicas. Uma propriedade topológica é uma propriedade invariante (insensível) às distorções “elásticas” do

plano imagem, ou seja, é uma propriedade que não será modificada por deformações do tipo alongamento ou contração do plano imagem. São propriedades que não envolvem noções de distância e nem propriedades que dependam dessas noções. As propriedades topológicas geralmente utilizadas são o número de elementos conectados e o número de “buracos” existentes na figura. Pode-se utilizar também combinações dessas propriedades, como, por exemplo, o número de Euler (definido como o número de elementos conectados menos o número de buracos). De modo geral, as propriedades topológicas são utilizadas como auxílio para outras formas de descrição. Normalmente, um problema para o qual uma descrição topológica sozinha seja suficiente é bastante raro⁽⁴⁾.

Uma definição genérica de textura poderia ser “a disposição ou características dos elementos constituintes de alguma coisa, especialmente no que se refere à aparência superficial ou à qualidade tátil”. Porém, no caso de imagens, uma definição mais adequada poderia ser “uma característica representativa da distribuição espacial dos níveis de cinza dos pixels de uma região”⁽⁷⁾. Ou seja, um atributo de textura é um valor, calculado a partir da imagem de um objeto, que quantifica algumas características da variação dos níveis de cinza desse objeto. Normalmente, um atributo de textura é independente da posição, orientação, tamanho, forma e brilho do objeto.

Estatisticamente, os atributos de textura de uma imagem podem ser obtidos a partir de cálculos efetuados sobre os valores do seu histograma de segunda ordem. O histograma de segunda ordem $p(i,j)$, também chamado de matriz de dependência espacial de nível de cinza ou matriz de co-ocorrência, é uma representação da distribuição de probabilidade de ocorrência de um par de valores semelhantes de nível de cinza, separados por uma distância “d”, na orientação “ θ ”. Em outras palavras, $p(i,j)$ indica a frequência de ocorrência de um particular par de nível de cinza i e j , medido a partir de uma distância “d” e de um ângulo “ θ ”. As estatísticas do histograma de segunda ordem, como, por exemplo, entropia, inér-

cia e energia, correlacionam-se muito bem com as estruturas da imagem e têm sido largamente utilizadas na literatura para a análise de características como homogeneidade, contraste, presença de estruturas organizadas, transição de nível de cinza e complexidade da imagem⁽⁸⁾.

Após a extração de informação, os atributos quantificados são utilizados no processo de classificação das estruturas presentes na imagem. Inteligência artificial envolve o uso do computador para o processamento de dados, visando à distinção entre padrões normais e anormais, a partir dos atributos extraídos das imagens. As técnicas relacionadas a essa área do conhecimento incluem métodos para a seleção de atributos, como os baseados na separabilidade entre as distribuições de probabilidades das classes e algoritmos genéticos, e classificadores, como os baseados em técnicas de análise discriminante, sistemas especialistas baseados em regras específicas, métodos estatísticos e redes neurais artificiais^(2,4,9). A Figura 3 apresenta um esquema de uma rede neural artificial multicamadas do tipo “feedforward”.

AUXÍLIO À DETECÇÃO

A detecção automatizada de lesões envolve a localização pelo computador de regiões contendo padrões radiológicos suspeitos, porém com a classificação da lesão sendo feita exclusivamente pelo radiologista. Sistemas para auxílio à detecção têm sido desenvolvidos principalmente para imagens de tórax e de mama.

Segundo a literatura especializada, de 30% a 50% dos casos de câncer de mama detectados por meio de mamografia apresentam agrupamentos de microcalcificações associados⁽²⁾. Além disso, estudos demonstraram que 26% dos casos de câncer não-palpável de mama apresentam nódulos associados na mamografia e 18% apresentam nódulos e microcalcificações⁽¹⁰⁾. Por isso, a maioria dos sistemas de auxílio ao diagnóstico em mamografia é voltada para a detecção de nódulos e microcalcificações⁽²⁾.

Desde 1987, pesquisadores da Universidade de Chicago, EUA, têm apresentado uma metodologia para a detecção

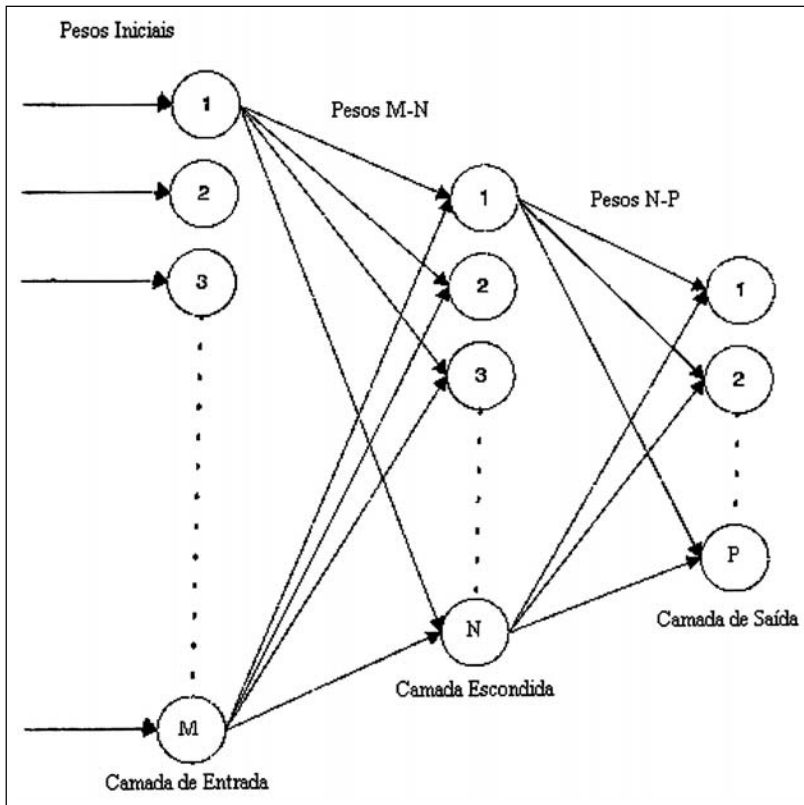


Figura 3. Esquema de uma rede neural artificial multicamadas do tipo "feedforward".

automatizada de microcalcificações em mamogramas. Esta metodologia, descrita inicialmente por Chan *et al.*⁽¹¹⁾, inclui etapas de pré-processamento para realce das imagens, filtragem para remoção de estruturas de fundo, segmentação das microcalcificações para extração de atributos e o uso de redes neurais artificiais para redução do número de detecções falsas-positivas.

A detecção automatizada de nódulos em mamogramas digitalizados é mais difícil do que a detecção de microcalcificações, devido à similaridade de muitos nódulos com o tecido parenquimatoso normal que os envolve. Segundo Giger⁽²⁾, provavelmente um dos primeiros estudos sobre a análise computadorizada de imagens mamográficas foi apresentado em 1967 por Winsberg *et al.*⁽¹²⁾, que descrevem um método computadorizado para comparação de padrões de densidade em várias áreas de uma mesma mama e entre as mamas direita e esquerda. O método foi aplicado em um número pequeno de casos selecionados. Atualmente existem muitos trabalhos voltados para a detec-

ção de lesões espiculadas, assimetrias, aumento de densidade, lesões circunscritas e distorções de parênquima. Pesquisadores da Universidade de Chicago têm trabalhado, também, no desenvolvimento de uma metodologia para a detecção de nódulos em mamogramas, a qual é baseada nas diferenças de simetria de arquitetura entre as mamas esquerda e direita, com as assimetrias indicando possíveis nódulos⁽¹³⁾. Em um estudo prospectivo de 10.000 mamogramas, realizado pelos pesquisadores da Universidade de Chicago, empregando um sistema CAD como segunda opinião, o computador indicou a presença de câncer cerca de um ano antes de este ser clinicamente diagnosticado em aproximadamente 15% de todos os casos de câncer, e em 56% dos casos foi detectado câncer em um estudo retrospectivo de mamogramas considerados clinicamente negativos. A taxa de falsos-positivos foi de aproximadamente 1,3 agrupamento falso por imagem e 2,1 nódulos falsos por imagem⁽¹⁴⁾.

Em junho de 1998, o Food and Drug Administration (FDA), dos Estados Uni-

dos, aprovou o ImageChecker M1000 (R2 Technology, Los Altos, Califórnia) como um sistema de auxílio à detecção, para ser utilizado por radiologistas na rotina de leitura de mamografias. Um amplo estudo, envolvendo a análise retrospectiva de 1.083 casos de câncer e o acompanhamento prospectivo de 14.817 casos, foi realizado como parte dos requisitos para a sua aprovação pelo FDA. Neste estudo, a sensibilidade alcançada foi de 98,3% para a detecção de agrupamentos de microcalcificações e 72% para a detecção de nódulos, com uma taxa de falso-positivo média de um (agrupamento ou nódulo) por imagem⁽¹⁵⁾. Segundo Giger⁽²⁾, em um estudo realizado por Hendricks, da Universidade de Nijimen (Holanda), utilizando o ImageChecker, não foi observada diferença entre uma dupla leitura com dois radiologistas e uma dupla leitura utilizando o sistema de CAD.

Ao contrário da análise da mamografia, na qual existem basicamente duas respostas possíveis, câncer ou não, a radiografia de tórax pode apresentar uma variedade de patologias possíveis, cuja detecção é o objetivo de alguns sistemas de CAD. Estes sistemas incluem métodos de processamento para realce de imagens, como subtração de dupla energia e temporal, e esquemas para detecção automatizada de padrões suspeitos, como nódulos, pneumotórax, cardiomegalia, assimetrias torácicas e lesões intersticiais. A detecção radiográfica de pequenos nódulos é um aspecto importante no diagnóstico e condução do tratamento de pacientes com câncer de pulmão. Porém, a não detecção pode ocorrer devido à sobreposição com costelas, brônquios, vasos sanguíneos e outras estruturas anatómicas normais. A taxa média de perda na detecção radiográfica de nódulos precoces é estimada em cerca de 30%⁽¹⁶⁾. Segundo MacMahon⁽¹⁷⁾, um sistema para detecção de nódulos pulmonares vem sendo desenvolvido pelos Laboratórios Rossmann, da Universidade de Chicago, há mais de dez anos, tendo recebido numerosos refinamentos neste período. A finalidade básica deste sistema é a distinção de opacidades focais anormais do fundo anatômico complexo da radiografia de tórax, tendo por objetivo chamar a

atenção do radiologista para a região suspeita. Atualmente, o programa de computador possui sensibilidade de aproximadamente 70% para nódulos pulmonares sutis, com uma média de um a dois falsos-positivos por imagem. Segundo o autor, embora essa acurácia seja menor do que a apresentada pela maioria dos radiologistas, os erros cometidos pelo computador tendem a ser diferentes daqueles cometidos pelos radiologistas, o que faz com que o uso do programa possa auxiliar mesmo radiologistas experientes na tarefa de detecção de nódulos.

A detecção e quantificação de lesões pulmonares intersticiais pelos radiologistas está reconhecidamente sujeita a uma grande variação inter e intrapessoal, sendo, por isso, uma área particularmente interessante para aplicação de análises computadorizadas. Katsuragawa *et al.*⁽¹⁸⁾ e Monnier-Cholley *et al.*⁽¹⁹⁾ apresentaram esquemas computadorizados para avaliação de lesões intersticiais a partir de atributos de textura extraídos de imagens radiográficas. Inicialmente, regiões de interesse são automaticamente selecionadas em radiografias de tórax e, então, processadas para eliminação de estruturas anatômicas e realce de detalhes finos da textura dos pulmões. Um processamento da imagem baseado na transformação de Fourier é utilizado para a descrição da magnitude e rugosidade da textura. Os autores descrevem valores de sensibilidade e especificidade na ordem de 89% para uma base de dados contendo 100 casos patológicos.

Outros programas de CAD têm sido desenvolvidos, dirigidos para a detecção de pneumotórax⁽²⁰⁾, cardiomegalia⁽²¹⁾ e assimetria anormal de tórax⁽²²⁾. Artigos e detalhes podem ser encontrados nos anais do International Workshop on Computer-Aided Diagnosis⁽²³⁾.

AUXÍLIO AO DIAGNÓSTICO

Uma vez que uma lesão foi detectada, cabe ao radiologista decidir o encaminhamento do caso, ou seja, se é necessária a realização de algum outro exame ou de uma biópsia, por exemplo. Ou então, se um simples acompanhamento é suficiente e, neste caso, qual deve ser o in-

tervalo até o próximo exame. Sistemas de análise computadorizada estão sendo desenvolvidos para auxiliar a distinção entre lesões malignas e benignas e aumentar a sensibilidade e especificidade do diagnóstico. Estes sistemas utilizam atributos extraídos e quantificados de forma automatizada e também por radiologistas. A vantagem da extração automatizada (pelo computador) é a objetividade e reprodutibilidade da medida dos atributos escolhidos. Por outro lado, os radiologistas utilizam uma grande quantidade de atributos, os quais são extraídos e interpretados simultânea e instantaneamente. É importante ressaltar que um dos objetivos dos sistemas de auxílio à classificação é a redução do número de casos benignos encaminhados para biópsia. Porém, como o "custo" da perda de uma lesão maligna é muito maior que o de uma classificação errônea de um caso benigno, os sistemas de auxílio ao diagnóstico devem ser desenvolvidos com o propósito de aumentar a especificidade, porém sem reduzir a sensibilidade. No que se refere ao câncer de mama, métodos de auxílio à classificação estão sendo desenvolvidos para a caracterização de lesões a partir de informações extraídas de imagens de mamografia, ultrassonografia e de ressonância magnética.

Vários pesquisadores têm utilizado a análise computadorizada de atributos extraídos por inspeção visual para auxílio à tomada de decisão quanto ao encaminhamento do tratamento do paciente⁽²⁾. De modo geral, este tipo de análise utiliza uma lista de atributos que devem ser avaliados e quantificados pelo radiologista em uma escala preestabelecida, em conjunto com um método para composição destes valores utilizando coeficientes otimizados. As respostas obtidas, em geral, apresentam a probabilidade da presença de determinada condição (por exemplo, de uma lesão previamente detectada ser maligna). Gale *et al.*⁽²⁴⁾ apresentam um aumento de especificidade de 67% para 88% entre o diagnóstico de câncer de mama feito sem auxílio e com auxílio de um sistema computadorizado. Porém, houve redução na sensibilidade de 87% para 79% com o sistema computadorizado. Cook e Fox⁽²⁵⁾ relatam um

aumento na acurácia diagnóstica de médicos residentes, de 40% para 73%, e de médicos radiologistas, de 70% para 73%, com a utilização de um sistema especialista voltado para a diferenciação de 16 tipos de lesões de mama circunscritas, a partir da identificação de achados radiológicos por inspeção visual. Pesquisadores da Duke University, EUA⁽²⁶⁾, utilizaram as categorias do padrão BI-RADS (American College of Radiology Breast Imaging Reporting and Data System)⁽²⁷⁾ como entradas para redes neurais artificiais, visando à distinção entre lesões benignas e malignas e à previsão da potencialidade de invasão do câncer de mama. Os resultados mostraram o potencial do uso dos padrões do BI-RADS no diagnóstico auxiliado por computador.

Ashizawa *et al.*⁽²⁸⁾ apresentaram os resultados da utilização de uma rede neural artificial para auxílio ao diagnóstico diferencial de lesões intersticiais de pulmão. A rede neural foi projetada para distinguir entre 11 lesões intersticiais a partir de dez parâmetros clínicos e 16 achados radiológicos obtidos por radiologistas especializados em tórax. Foram utilizados 150 casos reais, 110 casos publicados e 110 casos hipotéticos. O desempenho da rede foi avaliado por meio de curva ROC ("receiver operating characteristic"), tendo sido obtida uma área sob a curva igual a 0,947 para os casos clínicos. Foram obtidos valores de sensibilidade e especificidade em torno de 90%, considerando-se a resposta correta contida dentro dos dois maiores valores de saída para as 11 patologias.

O desenvolvimento de métodos de auxílio à classificação, baseados na extração e quantificação automatizada de atributos, exige que inicialmente sejam definidos quais atributos são clinicamente significativos e depois estabelecidos os métodos para sua extração e quantificação. Segundo Giger⁽²⁾, nos últimos cinco anos houve expressivo aumento no número de trabalhos sobre desenvolvimento de métodos de classificação automatizada de lesões em imagens mamográficas. De modo geral, um método automatizado para classificação de microcalcificações agrupadas envolve a extração de atributos que se correlacionam qualitativa-

vamente com a experiência dos radiologistas, estimando a probabilidade de malignidade a partir da aparência das microcalcificações existentes no agrupamento. O computador extrai automaticamente atributos relacionados ao tamanho, forma, quantidade e distribuição espacial das microcalcificações, e um classificador (uma rede neural artificial, por exemplo) é utilizado para combinar estes atributos e fornecer uma estimativa da probabilidade de malignidade. Esta estimativa pode então ser utilizada pelos radiologistas como auxílio na decisão se a lesão é maligna ou benigna. Resultados de estudos aplicados têm mostrado melhorias estatisticamente significativas de desempenho quando os radiologistas utilizam sistemas de auxílio computadorizado, em comparação com situações em que estes sistemas não são empregados. São relatados aumentos de sensibilidade, de 73% para 87%, e de especificidade, de 32% para 40%⁽²⁴⁾. Resultados como estes indicam que o computador pode potencialmente ser utilizado para auxiliar os radiologistas a decidirem quais mulheres devem ser encaminhadas para biópsia, podendo reduzir o número de biópsias desnecessárias.

No caso de nódulos, uma lesão maligna pode ser diferenciada de uma lesão benigna, em virtude da sua aparência mais espiculada na imagem mamográfica⁽²⁾ e, de modo geral, para a classificação deste tipo de lesão, o nódulo é primeiramente extraído do fundo anatômico do mamograma utilizando-se alguma técnica de processamento de imagem (como crescimento de região, por exemplo). Então, atributos de forma, contorno, níveis de cinza e textura são extraídos e utilizados por um classificador para o cálculo da probabilidade de malignidade associada. A Figura 4 mostra um exemplo de segmentação de nódulo através do método de crescimento de região⁽²⁹⁾.

Em trabalho realizado por pesquisadores da Universidade de Chicago, com um sistema desenvolvido para a classificação de nódulos mamográficos a partir do seu grau de espicularidade⁽²⁾, foram obtidos, para a base de dados utilizada, sensibilidade de 100% e valor preditivo positivo de 83% (12% acima do valor

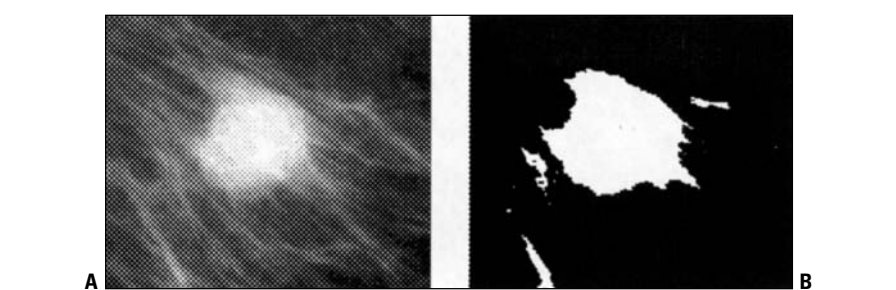


Figura 4. Exemplo de segmentação de nódulo mamográfico. **A:** Região de interesse original. **B:** Nódulo segmentado.

obtido por radiologistas especializados em mamografia e 21% acima da média de outros radiologistas menos experientes).

O ultra-som de mama também é utilizado como importante auxiliar no diagnóstico mamográfico, sendo tipicamente aplicado para avaliar lesões palpáveis e mamograficamente detectadas quanto à sua natureza sólida ou cística. A acurácia do ultra-som tem sido apontada como sendo de 96% a 100% no diagnóstico de lesões císticas benignas. Porém, o ultra-som de mama não é utilizado na rotina para distinguir lesões sólidas malignas de benignas, devido à considerável sobreposição das suas aparências ultra-sonográficas⁽²⁾. Chen *et al.*⁽³⁰⁾ relataram um método baseado em rede neural artificial para auxílio ao diagnóstico diferencial de tumores sólidos de mama em imagens de ultra-som. Foram avaliadas 140 imagens de nódulos sólidos de mama, sendo 52 correspondentes a tumores malignos e 88 a tumores benignos, todos confirmados por biópsia. Atributos de textura foram extraídos, após a seleção manual da região de interesse contendo a lesão. A acurácia da rede neural foi de 95% para a classificação das lesões malignas, com sensibilidade de 98%, especificidade de 93%, valor preditivo positivo de 89% e valor preditivo negativo de 99%. Segundo esses autores, o sistema foi capaz de diferenciar os nódulos sólidos de mama com uma acurácia relativamente alta e auxiliou operadores inexperientes a evitar diagnósticos errados.

A ressonância magnética pode também auxiliar no diagnóstico do câncer de mama, como técnica complementar à mamografia, com a vantagem de propiciar a obtenção de informação tridimensional da mama. O realce decorrente da presen-

ça de contraste é utilizado para a diferenciação entre lesões e tecidos normais, devido ao aumento de vascularização e permeabilidade capilar do tumor. Porém, algumas lesões benignas também são realçadas na ressonância e, por isso, sua especificidade é limitada. Pode-se utilizar atributos temporais e espaciais das lesões para a sua caracterização. Atributos temporais incluem taxa máxima de realce da lesão, o tempo em que isto ocorre, a velocidade máxima de realce, e mudança na não-homogeneidade da lesão no processo de realce. Atributos espaciais incluem relação de contraste entre a lesão e o fundo, não-homogeneidade dos valores dos voxels na lesão, características das margens da lesão, textura, circularidade e irregularidade da lesão. Gilhuijs *et al.*⁽³¹⁾ apresentaram um sistema de classificação automatizada de lesões de mama a partir de imagens de ressonância magnética, utilizando atributos temporais e espaciais calculados nas três dimensões. Em um estudo inicial para distinção entre lesões benignas e malignas, realizado com 27 casos, foi obtida uma área sob a curva ROC igual a 0,96. Nunes *et al.*⁽³²⁾ concluíram, como resultado de um estudo feito com atributos de arquitetura da mama, que a caracterização da margem da lesão é um determinante primário da sua malignidade.

Armato III *et al.*⁽³³⁾ desenvolveram um método automatizado para detecção de assimetria anormal em radiografias pósterio-anteriores digitais de tórax. O método é baseado em um "thresholding" de nível de cinza para segmentação e extração de contorno de regiões do pulmão contendo ar. Foram realizados testes em uma base de dados contendo 600 imagens clínicas de tórax, e a habilidade do

método em distinguir entre casos normais e casos com assimetria anormal foi avaliada por intermédio da curva ROC. O desempenho do método alcançou uma área sob a curva de 0,84. Segundo os autores, a sensibilidade do método torna-o potencialmente útil como componente para integração de um sistema de CAD para radiografia digital de tórax. Vargas-Voracek *et al.*⁽³⁴⁾ apresentaram um estudo realizado com atributos de textura extraídos de radiografias de tórax. O experimento foi realizado utilizando-se uma base de dados contendo 200 regiões de interesse (ROIs) extraídas de radiografias digitalizadas de pacientes com lesões intersticiais de pulmão e pacientes com radiografias normais. Nenhum pré-processamento foi realizado nas imagens. Os atributos calculados para cada ROI foram utilizados como entrada em uma rede neural artificial. A “performance” diagnóstica da rede foi bastante boa, alcançando sensibilidade de 90% e especificidade de 84%, com uma área sob a curva ROC igual a 0,85. McNitt-Gray *et al.*⁽³⁵⁾ desenvolveram uma metodologia para caracterização de nódulos pulmonares solitários, obtidos por exame de tomografia computadorizada de alta resolução, como sendo benignos ou malignos, por meio da quantificação de atributos de tamanho, forma, atenuação e textura. Um estudo inicial, realizado com 31 pacientes com diagnósticos definitivos (14 benignos e 17 malignos) e utilizando dois atributos de textura considerados os mais discriminantes (entropia e correlação), resultou em uma taxa de acerto na classificação de 90,3%.

MEDIDAS DE DESEMPENHO E CARACTERÍSTICAS

Existem vários índices de desempenho que podem ser utilizados na avaliação de sistemas de auxílio ao diagnóstico. Uma medida possível e muito utilizada é a análise da característica de resposta do observador (ROC), introduzida no contexto das imagens médicas por Swets⁽³⁶⁾ e Metz⁽³⁷⁾. A análise por curva ROC é fundamentada na teoria de detecção de sinal, tendo como base a idéia de que para qualquer sinal sempre existirá um fundo rui-

do que varia aleatoriamente sobre um valor médio. Quando um estímulo está presente, a atividade que ele cria no sistema de obtenção de imagem é adicionada ao ruído existente naquele momento. Este ruído pode estar dentro do próprio sistema ou fazer parte do padrão de entrada. A tarefa do observador (ou sistema automático) é determinar se o nível de atividade no sistema é devido apenas ao ruído ou resultado de um estímulo adicionado ao ruído. Em sua forma mais simples, a tarefa de diagnosticar consiste na apresentação de imagens contendo ou não uma anormalidade associada, a um observador que deve responder se existe ou não uma anormalidade presente. Para a análise da “performance” em um teste de diagnóstico, a curva ROC plota a fração de verdadeiro-positivo (sensibilidade) contra a fração de falso-positivo (1 – especificidade) para toda a faixa de resultados possíveis para o teste.

A análise por curva ROC baseia-se na premissa de que o processo de decisão do observador pode ser modelado por uma variável randômica que se ajusta a um modelo bi-normal (Figura 5). As funções de densidade de probabilidade desta variável para as duas hipóteses (presença de anormalidade e ausência de anormalidade) são consideradas normalmente distribuídas. A menos que as condições do teste sejam perfeitas, estas distribui-

ções se sobrepõem e um limiar de decisão (“threshold”) irá sempre envolver um compromisso entre os resultados falsos-positivos (não existe anormalidade, mas o observador indica que existe) e falsos-negativos (existe anormalidade, mas o observador indica que não existe). Uma mudança no limiar irá causar mudanças na especificidade e sensibilidade do teste. Os observadores humanos, assim como os sistemas de auxílio ao diagnóstico, operam em pontos ao longo da curva ROC. Os pontos de operação sobre a curva (sensibilidade vs. especificidade) mudam de acordo com os valores de custo atribuídos aos resultados falsos-positivos ou falsos-negativos.

A medida mais comum extraída da curva ROC é a área sob a curva. A área sob a curva ROC varia de 0,5 (comportamento aleatório) a 1,0 (discriminação perfeita), sendo que quanto mais a curva se aproxima do canto esquerdo superior do gráfico melhor a “performance” do teste. As Figuras 6 e 7 mostram as curvas ROC obtidas para duas configurações diferentes de um classificador estatístico não-paramétrico, variando-se os limiares de separação entre as classes maligna e benigna, utilizado em um sistema de classificação automatizada de microcalcificações presentes em mamogramas digitalizados⁽⁶⁾.

Os parâmetros de aquisição da imagem digital, as características da base de

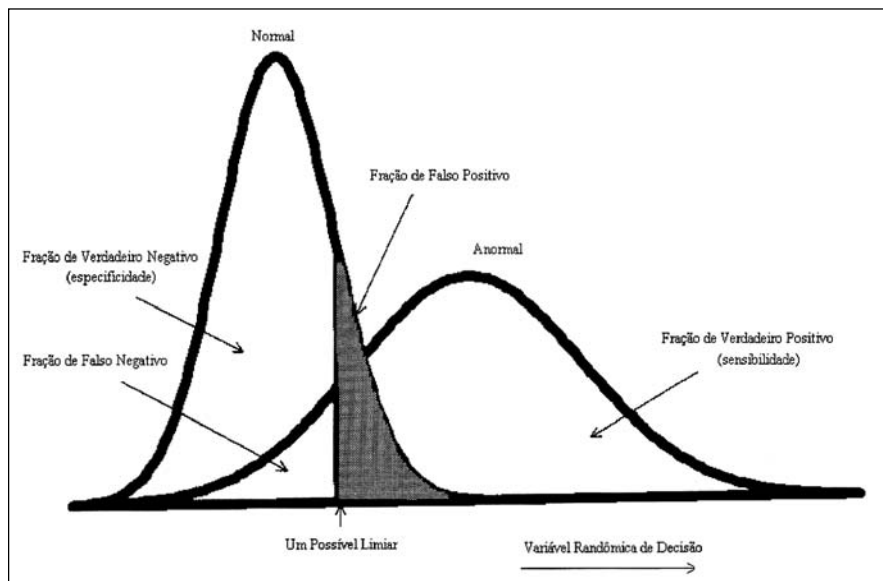


Figura 5. Modelo bi-normal de distribuição de probabilidade para análise por curva ROC.

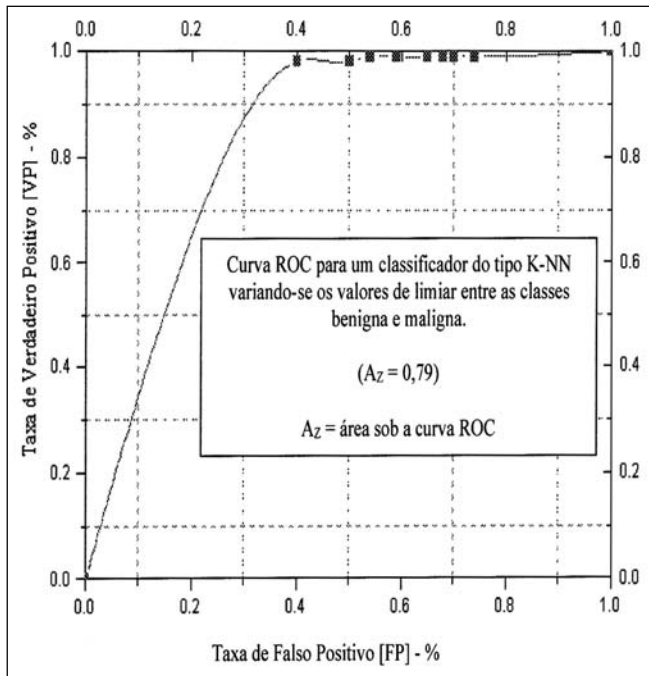


Figura 6. Curva ROC para um classificador do tipo K-NN.

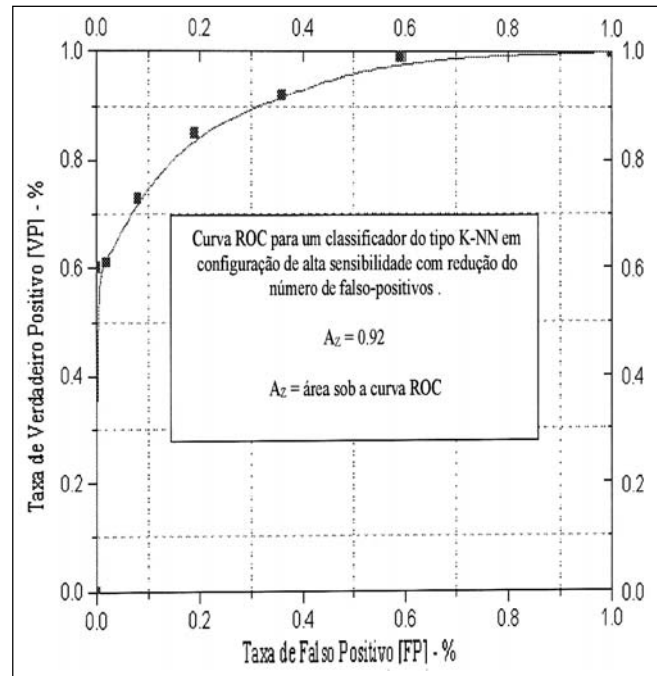


Figura 7. Curva ROC para um classificador do tipo K-NN em configuração de alta sensibilidade.

dados utilizada para treinamento e teste dos algoritmos, a comparação entre métodos e a determinação da viabilidade de aplicação clínica são também características fundamentais no processo de avaliação de sistemas de CAD. As imagens podem ser obtidas diretamente em formato digital ou a partir da digitalização de filmes. Para a digitalização de filmes, a qualidade da imagem original é fundamental, uma vez que imagens de baixa qualidade irão afetar tanto a inspeção visual como a “performance” da análise automatizada. O tamanho do pixel e o número de níveis de quantização também irão afetar a resposta do sistema, uma vez que modificações nos valores de entrada irão produzir saídas diferentes. É bastante difícil a comparação entre os vários métodos de diagnóstico auxiliado por computador que estão sendo desenvolvidos, devido às diferentes bases de dados utilizadas e os variados métodos de descrição dos resultados. Algumas bases de dados estão sendo estruturadas para desenvolvimento e teste de sistemas de CAD^(38,39). A forma como a base de dados é utilizada irá também influenciar o desenvolvimento e a apresentação de resultados do sistema. Por exemplo, um

sistema computadorizado para auxílio na detecção de lesões de mama, treinado com projeções médio-laterais oblíquas, irá fornecer resultados tendenciosos (terá um “bias”) se forem utilizadas projeções crânio-caudais nos testes. Os resultados também serão tendenciosos se, com uma base de dados limitada, os parâmetros do classificador forem determinados usando-se os mesmos dados que foram empregados para a seleção dos atributos mais discriminantes. Além disso, diferentes formas de medida irão causar variações nos resultados obtidos. Por exemplo, se a sensibilidade do sistema é dada em função da porcentagem de lesões detectadas por imagem ou por caso, ou então, se a porcentagem de sobreposição entre a lesão real e a lesão detectada pelo computador é calculada levando-se em consideração sua intersecção ou a intersecção dividida pela sua união.

DISCUSSÃO

A dificuldade na realização do diagnóstico visual de lesões em estágio precoce, juntamente com o grande volume de dados gerados num sistema de diagnóstico hospitalar, têm induzido vários

grupos de pesquisa a buscarem soluções computacionais para auxiliar os médicos nesse problema.

Em uma visão global dos trabalhos apresentados sobre sistemas de CAD, pode-se constatar que, de modo geral, estes sistemas podem detectar lesões perdidas pelos radiologistas, são suficientemente robustos em relação a mudanças na aquisição e digitalização das imagens, podem ser integrados com relativa facilidade nas rotinas existentes e podem classificar lesões com acurácia comparável à dos especialistas. Porém, o número de testes dos sistemas em ambientes clínicos ainda é bastante limitado e uma avaliação cuidadosa é necessária antes de sua inserção na rotina de serviço.

Os bons resultados conseguidos em testes de desempenho, em conjunto com a difusão da tecnologia de imagem digital, incentivam a integração de métodos automatizados de detecção e classificação em um número cada vez maior de protótipos clínicos. Porém, é necessária uma introdução sistemática e gradual dos sistemas CAD junto aos serviços de radiologia, de forma a minimizar a ruptura dos hábitos atuais de leitura, garantindo a sua aceitabilidade e a otimização do

desempenho do radiologista. Desse modo, um papel específico para cada sistema CAD deverá ser definido para cada radiologista, de acordo com o seu treinamento e habilidade individuais, reduzindo as variações intra-observador e melhorando o desempenho diagnóstico.

CONCLUSÕES

De modo geral, o uso de computadores para auxílio à análise de imagens radiológicas tem-se mostrado eficiente na melhoria da detecção e classificação de lesões, especialmente em imagens mamográficas e de tórax. Muitos artigos descrevendo essa tecnologia têm enfatizado que a análise automatizada é um auxílio, e não um substituto, para o radiologista, não necessitando ser perfeita e sim possuir um desempenho próximo ao do especialista. Nesse contexto, as pesquisas que estão sendo desenvolvidas indicam ser bastante promissor o uso de sistemas de diagnóstico auxiliado por computador como uma segunda opinião no processo de diagnóstico por imagem.

REFERÊNCIAS

- Doi K. Computer-aided diagnosis and its potential impact on diagnostic radiology. *In: Doi K, MacMahon H, Giger ML, Hoffmann KR, eds. Computer-aided diagnosis in medical imaging.* Amsterdam: Elsevier Science, 1999:11–20.
- Giger ML. Computer-aided diagnosis. *RSNA Categorical Course in Breast Imaging 1999:249–72.*
- Thurfjell EL, Lernevall KA, Taube AA. Benefit of independent double reading in a population-based mammography screening program. *Radiology 1994;191:241–4.*
- Castleman KN. *Digital image processing.* New Jersey: Prentice-Hall, 1996.
- Pereira Jr RR, Azevedo-Marques PM, Frère AF. Wavelet transform improves contrast enhancement of digital mammographic images. *Proceedings of the 6th International Workshop on Digital Mammography, Toronto, Canadá, 2000 (in press).*
- Ferrari RJ, Azevedo-Marques PM, Frère AF, Kinoshita SK, Spina LAR. Characterization of breast cancer using statistical approaches. *In: Doi K, MacMahon H, Giger ML, Hoffmann KR, eds. Computer-aided diagnosis in medical imaging.* Amsterdam: Elsevier Science, 1999:281–6.
- IEEE standard 610.4-1990, IEEE standard glossary of image processing and pattern recognition terminology. New York: IEEE Press, 1990.
- Haralick RM, Shanmugan K, Dinstein I. Textural features of images classification. *IEEE Trans Systems Man Cybernetics SMC-3 1973:610–21.*
- Duda RO, Hart PE. *Pattern classification and scene analysis.* New York: John Wiley & Sons, 1973.
- Sickles EA. Mammographic features of 300 consecutive nonpalpable breast cancers. *AJR 1986; 146:661–3.*
- Chan HP, Doi K, Galhota S, Vyborny CJ, MacMahon H, Jokich PM. Image feature analysis and computer-aided diagnosis in digital radiography. I. Automated detection of microcalcifications in mammography. *Med Phys 1987;14:538–48.*
- Winsberg F, Elkin M, Macy J, Bordaz V, Weymouth W. Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis. *Radiology 1967;89:211–5.*
- Yin FF, Giger ML, Doi K, Metz CE, Vyborny CJ, Schmidt RA. Computerized detection of masses in digital mammograms: analysis of bilateral-subtraction images. *Med Phys 1991;18:955–63.*
- Nishikawa RM, Giger ML, Wolverton DE, *et al.* Prospective testing of clinical mammography workstation for CAD: analysis of the first 10,000 cases. *In: Karssemeijer N, Thijssen M, Hendriks J, van Erming L, eds. Digital mammography.* Dordrecht: Kluwer Academic Publishers, 1998:401–6.
- Doi T, Hasegawa A, Hunt B, *et al.* Clinical results with the R2 ImageChecker™ mammographic CAD system. *In: Doi K, MacMahon H, Giger ML, Hoffmann KR, eds. Computer-aided diagnosis in medical imaging.* Amsterdam: Elsevier Science, 1999:201–7.
- Forrest JV, Friedman PJ. Radiologic errors in patients with lung cancer. *West J Med 1981;134: 485–90.*
- MacMahon H. Clinical application of CAD in the chest. *In: Doi K, MacMahon H, Giger ML, Hoffmann KR, eds. Computer-aided diagnosis in medical imaging.* Amsterdam: Elsevier Science, 1999:23–33.
- Katsuragawa S, Doi K, MacMahon H. Image feature analysis and computer-aided diagnosis in digital radiography: detection and characterization of interstitial lung disease in digital chest radiographs. *Med Phys 1988;15:311–9.*
- Monnier-Cholley L, MacMahon H, Katsuragawa S, Morishita J, Doi K. Computerized analysis of interstitial infiltrates on chest radiographs: a new scheme based on geometric pattern features and Fourier analysis. *Acad Radiol 1995;2:455–62.*
- Sanada S, Doi K, MacMahon H. Image feature analysis and computer aided diagnosis in digital radiography: automated detection of pneumothorax in chest images. *Med Phys 1992;19:1153–60.*
- Nakamori N, Doi K, Sabeti V, MacMahon H. Image feature analysis and computer-aided diagnosis in digital radiography: automated analysis of sizes of heart and lung in chest images. *Med Phys 1990;17: 342–50.*
- Armato SG III, Giger ML, MacMahon H. Computerized detection of abnormal asymmetry in digital chest radiographs. *Med Phys 1994;21:1761–8.*
- Doi K, MacMahon H, Giger ML, Hoffmann KR. Computer-aided diagnosis in medical imaging. *Excerpta Medica, International Congress Series II82, Radiology.* Amsterdam: Elsevier, 1999.
- Gale AG, Roebuck EJ, Riley P, Worthington BS. Computer aids to mammographic diagnosis. *Br J Radiol 1987;60:887–91.*
- Cook HM, Fox MD. Application of expert systems to mammographic image analysis. *Am J Physiol Imaging 1989;4:16–22.*
- Baker JA, Kornguth PJ, Lo JY, Floyd CE Jr. Artificial neural network: improving the quality of breast biopsy recommendations. *Radiology 1996; 198:131–5.*
- D'Orsi CJ, Basset LW, Feig AS, *et al.* *Breast Imaging Reporting and Data System (BI-RADS).* Reston, VA: American College of Radiology, 1998.
- Ashizawa K, Ishida T, MacMahon H, Vyborny CJ, Katsuragawa S, Doi K. Artificial neural networks in chest radiography: application to differential diagnosis of interstitial lung disease. *Acad Radiol 1999;6:2–9.*
- Kinoshita SK, Azevedo-Marques PM, Slaets AFF, Marana HRC, Ferrari RJ. Characterization of breast masses using texture and shape features. *In: Doi K, MacMahon H, Giger ML, Hoffmann KR, eds. Computer-aided diagnosis in medical imaging.* Amsterdam: Elsevier Science, 1999:265–70.
- Chen DR, Chang RF, Huang YL. Computer-aided diagnosis applied to US of solid breast nodules by using neural networks. *Radiology 1999;213:407–12.*
- Gilhuijs KGA, Giger ML, Bick U. Computerized analysis of breast lesions in three dimensions using dynamic magnetic resonance imaging. *Med Phys 1998;25:1647–54.*
- Nunes LW, Schnall MD, Orel SG, *et al.* *Breast MR imaging: interpretation model.* *Radiology 1997; 202:833–41.*
- Armato SG III, Giger ML, MacMahon H. Automated abnormal asymmetry detection in digital posteroanterior chest radiographs. *In: Doi K, MacMahon H, Giger ML, Hoffmann KR, eds. Computer-aided diagnosis in medical imaging.* Amsterdam: Elsevier Science, 1999:89–94.
- Vargas-Voracek R, Floyd CE Jr. Stochastic and deterministic texture modeling and analysis for image processing in chest radiography. *In: Doi K, MacMahon H, Giger ML, Hoffmann KR, eds. Computer-aided diagnosis in medical imaging.* Amsterdam: Elsevier Science, 1999:95–9.
- McNitt-Gray MF, Wyckoff N, Hart EM, Sayre JW, Goldin JG, Aberle DR. Computer-aided techniques to characterize solitary pulmonary nodules imaged on CT. *In: Doi K, MacMahon H, Giger ML, Hoffmann KR, eds. Computer-aided diagnosis in medical imaging.* Amsterdam: Elsevier Science, 1999: 101–6.
- Swets JA. ROC analysis applied to the evaluation of medical imaging techniques. *Invest Radiol 1979;14:109–21.*
- Metz CE. ROC methodology in radiologic imaging. *Invest Radiol 1986;21:720–33.*
- Mascio LN, Frankel SD, Hernandez JM, Logan CM. Building the LLNL/UCSF digital mammogram library with image groundtruth. *In: Doi K, Giger ML, Nishikawa RM, Schmidt RA, eds. Digital mammography'96. Excerpta Medica, International Congress Series 1119.* Amsterdam: Elsevier, 1996:427–30.
- Bowyer K, Kopans D, Kegelmeyer WP Jr, *et al.* The digital database for screening mammography. *In: Doi K, Giger ML, Nishikawa RM, Schmidt RA, eds. Digital mammography'96. Excerpta Medica, International Congress Series 1119.* Amsterdam: Elsevier, 1996:431–4.