



Aplicação de conceitos da lógica nebulosa à classificação BI-RADS de nódulos de mama

Applying concepts of fuzzy logic classification of BI-RADS breast nodules

Gisele Helena Barboni Miranda¹
Paulo Mazzoncini de Azevedo Marques²
Joaquim Cezar Felipe¹

Descritores: : Lógica Fuzzy

RESUMO

Este trabalho apresenta um método para avaliação e classificação de nódulos de mama que utiliza a Lógica Nebulosa no processo de inferência da categoria BI-RADS relativa à avaliação. Um dos principais obstáculos que se contrapõem ao processo de representação e processamento automático de imagens médicas nos ambientes de estudo e diagnóstico auxiliado por computador é o chamado “hiato semântico”, que se caracteriza pela divergência dos resultados obtidos pelos algoritmos de reconhecimento de padrões e apoio à decisão com os resultados esperados pelo radiologista. Esta divergência se deve principalmente às dificuldades encontradas pelos sistemas computacionais em representar e simular o raciocínio e o estilo adotado pelo radiologista no processo de análise da imagem. A proposta deste trabalho consiste na aplicação dos conceitos da Lógica Nebulosa à definição dos atributos descritores da terminologia empregada pelos radiologistas para a caracterização de achados, de forma que a adaptação de definições e regras linguísticas aos modelos numéricos, proporcionada pela Lógica Nebulosa, forneça uma aproximação entre o sistema computacional e o usuário especialista, gerando resultados mais efetivos e fomentando, assim, a redução do hiato semântico. Entrevistas com especialistas do domínio de aplicação do trabalho auxiliaram as etapas de modelagem e implementação do sistema nebuloso proposto. Questionários contendo itens para avaliação dos atributos padronizados pelo BI-RADS foram aplicados a radiologistas possibilitando o estudo da relação entre as características de análise e as categorias BI-RADS de nódulos de mama. Uma análise comparativa realizada entre diagnósticos gerados pelos radiologistas e a saída automática gerada pelo sistema aponta na direção da precisão e da efetividade do mesmo.

Descriptors: Fuzzy Logic

ABSTRACT

This work presents a method for evaluation and classification of breast nodules that utilizes the Fuzzy Logic to infer the BI-RADS assessment related to the evaluation. One of the main obstacles that opposes to the process of representation and automatic processing of medical images in study environments and computer aided diagnosis is the “semantic gap”, that is characterized for the divergence of the results obtained through pattern recognition and decision support algorithms with the results expected by the radiologist. This divergence is due, mainly to the difficulties found by the computational systems in representing and simulating the reasoning and the style adopted by the radiologist in the process of image analysis. The proposal of this work consists in the application of the concepts of Fuzzy Logic in the definition of the describing attributes of the terminology used by the radiologists to characterize findings. This way the adaptation of definitions and linguistic rules to the numerical models, achieved through the Fuzzy Logic, can support an approach between the computational system and the specialist user, leading to more effective results and fomenting, thus, the reduction of the semantic gap. Interviews with specialists of the application domain of this work assisted the stages of modeling and implementation of the proposed fuzzy system. Questionnaires containing items for evaluation of the BI-RADS standardized attributes were applied to radiologists making possible the study of the relation between the characteristics of analysis and the BI-RADS categories of breast nodules. A comparative analysis made between the diagnostics proposed by the radiologists and the automatic system output indicates its precision and effectiveness.

Autor Correspondente:
**Gisele Helena Barboni
Miranda**
e-mail: gimiranda@usp.br

¹ Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo - USP - Ribeirão Preto (SP), Brasil.

² Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo - USP - Ribeirão Preto (SP), Brasil.

Artigo recebido: 16/03/2009
Aprovado: 10/07/2009

INTRODUÇÃO

O objetivo geral dos sistemas de informação na medicina têm sido disponibilizar a informação necessária ao médico de maneira rápida e precisa a fim de melhorar a qualidade e a eficiência dos processos no cuidado à saúde. Paralelamente, os registros clínicos do paciente e as anotações relativas aos laudos têm seguido o caminho da digitalização e da informatização. Bases de dados contendo informações sobre o paciente, sua evolução clínica, assim como exames realizados têm se tornado comum em grandes hospitais. A possibilidade de se tratar computacionalmente imagens e dados permite o desenvolvimento de ferramentas para o diagnóstico auxiliado por computador (*Computer-Aided Diagnosis* – CAD)⁽¹⁻²⁾, as quais têm se mostrado coadjuvantes capazes de promover um aumento da qualidade do trabalho dos radiologistas⁽³⁾.

Os sistemas CAD para radiologia têm como objetivo auxiliar o diagnóstico do radiologista, proporcionando um aumento de precisão, por meio do uso dos resultados gerados pelo computador como referência, como, por exemplo, a indicação de áreas suspeitas da imagem ou de anormalidades mascaradas. Este auxílio do computador é importante, já que o diagnóstico do radiologista é baseado em avaliações subjetivas, sujeitas a variações pessoais (tais como perda de informações devido à fadiga visual e distração), qualidade da imagem e sobreposição de estruturas, entre outros fatores⁽⁴⁾. Desta forma, os sistemas CAD têm mostrado que podem melhorar o desempenho dos diagnósticos radiológicos.

Ao longo da sua evolução, os projetos e pesquisas abordando métodos para CAD depararam-se com um forte obstáculo: o chamado “hiato semântico” (*semantic gap*). O hiato semântico consiste no distanciamento observado entre os resultados de métodos automatizados e os resultados esperados pelos usuários, que surge devido às dificuldades desses métodos em representar as situações reais. Um médico, diante de certa imagem, poderia estar buscando variadas abordagens e diversos objetivos diagnósticos, os quais não estariam necessariamente cobertos pelos métodos disponíveis.

Para minimizar este problema, um procedimento que tem sido frequentemente adotado é a delimitação de contextos bastante específicos para a definição dos métodos de CAD. Entretanto, independentemente de se abordar um contexto específico ou amplo, a representação de uma imagem ou de um achado por meio de valores numéricos exatos ou de intervalos disjuntos, por si só, já caracteriza um hiato semântico significativo, uma vez que esta representação está relativamente distante da forma de análise e interpretação do analista humano. Ao analisar o formato de um nódulo, o laudo será gerado em linguagem natural com uma frase do tipo “formato

ligeiramente irregular”. Para tomar um descritor numérico relativo ao formato e torná-lo adequado à semântica do usuário, precisamos fazer uso de uma técnica numérica um pouco mais sofisticada. Os conceitos da Lógica Nebulosa destinam-se a tratar este tipo de situação.

A Lógica Nebulosa utiliza conceitos intuitivos e métodos de raciocínio aproximado para representar um modelo matemático clássico⁽⁵⁾. Atualmente, pode ser considerada uma técnica de Inteligência Artificial que busca simular o raciocínio humano em sistemas de apoio à tomada de decisão. Os conjuntos nebulosos tratam a separação entre classes caracterizadas por faixas de valores numéricos por medidas de pertinência sobrepostas, substituindo transições bruscas. Esta abordagem aplica-se a situações nas quais a fronteira entre conjuntos não é bem delineada – a maioria dos casos de avaliação realizada por um ser humano – aproximando conhecimento objetivo (representado por dados numéricos) e conhecimento subjetivo (representado por termos lingüísticos).

No auxílio ao diagnóstico de nódulos de mama, para guiar radiologistas e médicos no processo de tomada de decisão, há o padrão de classificação BI-RADS (*Breast Imaging Reporting and Data System*), desenvolvido em colaboração por membros e comitês do ACR (*American College of Radiology*), que tem por objetivo unificar e padronizar a terminologia utilizada em exames de mamografias, as categorias de avaliação dos achados e a ação resultante para cada uma delas, facilitando o monitoramento dos resultados⁽⁶⁾. A utilização do léxico BI-RADS é altamente recomendada para guiar radiologistas e médicos no processo de tomada de decisão.

Em trabalhos correlatos, algumas abordagens foram propostas utilizando a lógica nebulosa na etapa de processamento de imagens de lesões de mama e em ambientes de auxílio ao diagnóstico. Em um deles é apresentado um método para diferenciação de contornos lobulados e microlobulados. Funções de pertinência são criadas para representar características, tais como número de ondulações do contorno e tamanho da ondulação⁽⁷⁾. Em outro, três regiões distintas em mamografias são comparadas: regiões que contêm tumores, regiões que contêm microcalcificações e regiões que não contêm lesões⁽⁸⁾. Um sistema de inferência nebuloso é proposto baseado na intensidade dos pixels dessas três regiões, podendo-se inferir o tipo de lesão (tumor ou microcalcificação) ou a ausência da mesma. Em contrapartida, a metodologia utilizada neste projeto, além de tratar os atributos de análise como conjuntos nebulosos, também fornece uma saída nebulosa, gerando uma resposta quantitativa e padronizada pelo BI-RADS, o que não foi apresentado em trabalhos anteriores.

O BI-RADS obedece a uma estrutura de linguagem natural. Por exemplo: “se o nódulo tem formato arredondado, margens bem definidas e densidade

mediana, então sua classificação será: achado com alta probabilidade de benignidade”. Este tipo de estrutura se ajusta perfeitamente à forma de tratamento dos Conjuntos Nebulosos e da lógica de utilização dos mesmos.

O presente trabalho está focado no estudo e no desenvolvimento de métodos que apliquem os conceitos da Lógica Nebulosa à representação de imagens de radiografias de mama por características visuais às mesmas, visando à redução do hiato semântico que ocorre nos ambientes de análise, recuperação por similaridade e geração de diagnóstico. Com isto, pretende-se investigar a eficácia desses métodos no sentido de tornar as aplicações de auxílio ao diagnóstico mais precisas e efetivas, possibilitando que os resultados dos sistemas automatizados sejam mais próximos dos resultados gerados e esperados pelos médicos usuários. Além disso, a utilização de parâmetros e regras definidos em termos mais próximos da linguagem natural deverá servir como estímulo para que os médicos enquanto potenciais usuários passem a fazer uso efetivo dos sistemas CAD.

BASE CONCEITUAL

A seguir serão apresentados os conceitos e padrões utilizados neste trabalho como base para o método proposto.

Lógica Nebulosa

A Lógica Nebulosa foi apresentada inicialmente por Lofti Zadeh, professor da Universidade de Berkeley, na década de 60. A utilização dos conceitos da Lógica Nebulosa aplicada aos diversos domínios do conhecimento vem crescendo ao longo dos anos. Sistemas baseados em controladores nebulosos têm possibilitado a automação de diversos processos⁽⁹⁾, incluindo o controle de sofisticados procedimentos industriais. Na área médica, essa abordagem é bastante adequada ao processo de determinação do diagnóstico, sujeito à subjetividade do médico.

A teoria de conjuntos nebulosos distingue-se da teoria clássica de conjuntos por tratar do aspecto vago da informação. Esta teoria é menos restritiva e mais adequada ao tratamento de informações fornecidas por seres humanos⁽¹⁰⁾. Ela estende a lógica clássica binária (0,1) para o domínio contínuo, o intervalo (0,1), permitindo uma transição gradual entre a não pertinência e a pertinência de um elemento a um conjunto⁽¹¹⁾.

Segundo Zadeh⁽¹²⁾, com a lógica nebulosa é possível a utilização de subconjuntos representativos de conceitos intermediários como: “verdadeiro”, “muito verdadeiro”, “pouco verdadeiro”, “falso”. Esses valores poderão ser estruturados em regras semânticas, que juntamente com os conectivos *if*, *and*, *or*, *then* formam a base do processo de inferência nebuloso, o qual é aproximado e não exato.

Seja A um conjunto nebuloso do universo de discurso U e “ x ” um elemento a ser caracterizado, a proposição “ x pertence a A ”, não necessariamente será verdadeira ou falsa, como na lógica binária. O elemento “ x ” possuirá um grau de pertinência ao conjunto A , que define o quão relacionado está “ x ” ao conjunto A (Figura 1).

É a função de pertinência que caracteriza um conjunto nebuloso⁽¹³⁾. Como exemplo, consideremos o conjunto universo “Temperatura”. Podemos dividi-lo em subconjuntos nebulosos representativos dos conceitos “temperatura baixa”, “temperatura média” e “temperatura alta”. Para um valor igual a 5°C, poderíamos obter um grau de pertinência μ_1 deste elemento ao subconjunto “temperatura baixa”, um grau de pertinência μ_2 ao subconjunto “temperatura média” e um grau de pertinência μ_3 ao subconjunto “temperatura alta”.

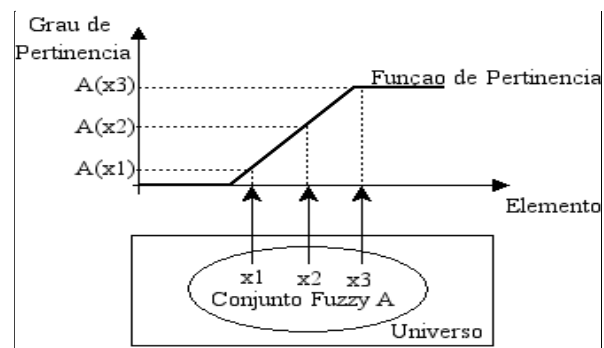


Figura 1 - Determinação do grau de pertinência.

Geralmente, a representação de conjuntos nebulosos é feita através de funções dos tipos triangulares, trapezoidais, gaussianas, entre outras. Essas funções são conhecidas por famílias de funções parametrizáveis. Na Figura 2 é possível visualizar algumas dessas funções típicas.

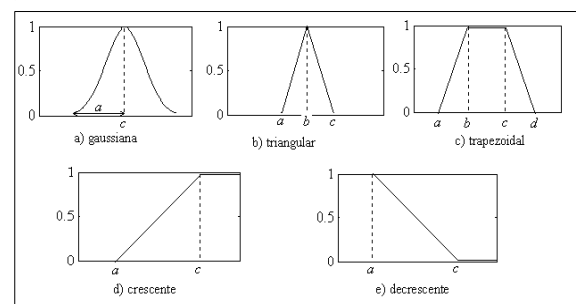


Figura 2 – Exemplos de funções de pertinência.

Inferência Nebulosa

A entrada de um sistema nebuloso assume valores precisos, ainda não nebulosos, que passarão por uma etapa inicial chamada “nebulização”, na qual, serão mapeados em valores lingüísticos pelas funções de pertinência características de cada subconjunto nebuloso de entrada. A criação de uma base de regras

composta por uma coleção de proposições nebulosas caracteriza a estratégia de controle e juntamente com os dados de entrada será processada pelo modelo de inferência escolhido para fornecer a saída nebulosa do sistema. A etapa denominada “denebulização” é utilizada quando a saída do sistema necessita de uma representação discreta. Esta última etapa transforma a saída nebulosa em ação de controle não nebulosa⁽¹⁰⁾ (Figura 3).

Uma base de regras nebulosas possui proposições cujos antecedentes consequentes ou ambos representam conjuntos nebulosos. Geralmente, as regras encontram-se no formato “se <premissa> então <conclusão>”, descrevendo uma relação entre elementos de conjuntos diferentes. Essa abordagem facilita o processo de interpolação, ou seja, uma regra pode ser definida pelo produto cartesiano entre conjuntos nebulosos⁽¹⁴⁾. Especialistas da área de interesse, geralmente, auxiliam as etapas de criação da base de regras e dos conjuntos nebulosos. Porém, também é possível a utilização de métodos matemáticos e algoritmos computacionais nestas etapas.

O modelo de inferência nebulosa adotado foi o modelo de Mamdani, proposto na década de 70, e utilizado como padrão para utilização da lógica nebulosa por muito tempo. As regras semânticas definidas neste modelo utilizam operadores de união e intersecção (*Max-Min*) e tanto os antecedentes como os consequentes das regras possuem relações nebulosas. Para cada regra da base, é definido o produto cartesiano *fuzzy* (intersecção *fuzzy*) dos componentes da parte antecedente da regra, gerando um coeficiente de disparo para cada uma. Quando este coeficiente é maior que zero, pode-se dizer que a regra foi ativada pela entrada atual e contribuirá para o cálculo da saída fornecida pelo sistema⁽¹⁵⁾. Entre as regras ativadas, a operação união *fuzzy*, será aplicada sobre as regras com o mesmo subconjunto de saída, gerando o seguinte universo de discurso para cada subconjunto como descrito em (1):

$$\mu(C) = \max_{K=1..n} [\min(D_K, \mu C_K)] \quad (1)$$

onde, $D(k)$ representa o coeficiente de ativação de cada regra que possui a saída C_i . O conjunto $\mu(C)$ pode ser utilizado no processo de tomada de decisão

em uma avaliação qualitativa. Para um resultado quantitativo, ou seja, uma variável escalar em sistemas de controle, é necessário realizar a etapa de denebulização, no qual $\mu(C)$ assumirá um valor não nebuloso e passará a ser escalar. O método do centróide e o método da média dos máximos estão entre os mais utilizados⁽¹⁵⁾. No primeiro, é calculada a abscissa do ponto do centro de massa para cada subconjunto ativado. No segundo, o valor numérico é dado pela média dos pontos de máximo das funções de pertinência dos conjuntos de saída ativados no processo de inferência.

Descritores BI-RADS

Os principais atributos para classificação de nódulos de mama utilizados neste trabalho e padronizados pelo BI-RADS são: formato, contorno e densidade. O formato pode ser classificado como redondo, oval, lobular ou irregular. O contorno assume as características: circunscrito, microlobulado, obscurecido, indistinto e espiculado. A densidade do nódulo quando comparada ao tecido subjacente pode ser alta, isodensa, baixa ou de aspecto gorduroso. Formato bastante irregular, contorno muito espiculado e densidade alta geralmente implicam em um alto grau de malignidade do tumor. A combinação desses três descritores permite uma transição entre a caracterização de uma lesão como benigna ou maligna. A Figura 4 apresenta alguns exemplos de formato, contorno e densidade em lesões de mama.

Categorias de Avaliação BI-RADS

O BI-RADS possui seis categorias de avaliação, numeradas de 0 a 5, para a caracterização de achados. A categoria 0 indica que é necessário uma avaliação adicional da imagem, o que ocorre em situações de acompanhamento. Exames como ultra-som e ressonância magnética poderão ser recomendados nesta categoria e sempre que possível a mamografia deve ser comparada a exames anteriores. Na categoria 1, não há presença de massas ou calcificações suspeitas, caracterizando uma mamografia negativa. A categoria 2 apresenta achados benignos, sem indicações de malignidade. Um achado na categoria 3 apresenta alto grau de benignidade, porém menor do que o grau da categoria 2 e uma pequena probabilidade de

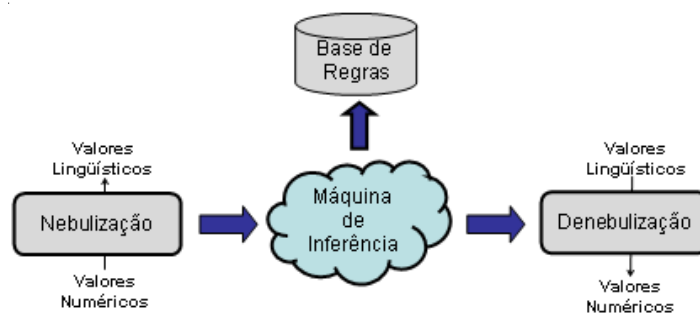


Figura 3 - Máquina de Inferência Nebulosa.

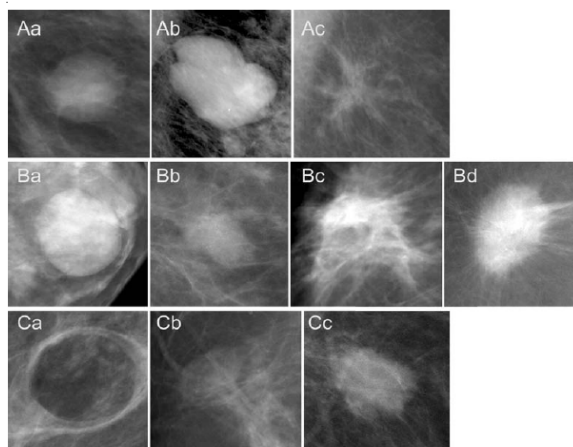


Figura 4 - Formato: Aa redondo, Ab lobular, Ac irregular. Contorno: Ba bem definido, Bb mal definido, Bc obscurecido, Bd espiculado. Densidade: Ca contém gordura, Cb isodensa, Cc alta⁽⁶⁾.

malignidade (menor que 2%) está associada a esta categoria⁽⁶⁾, recomendando-se o pedido de uma nova mamografia após seis meses. As categorias 4 e 5 são consideradas malignas, ou seja, positivas. Na categoria 4, há lesões com alta probabilidade de malignidade, porém elas não apresentam características morfológicas de câncer de mama. Para uma lesão classificada como BI-RADS 4 deverá ser realizado um estudo histológico da lesão. Por fim, a categoria 5 apresenta lesões com alta probabilidade de serem câncer. A Figura 5 apresenta exemplos de lesões classificadas como BI-RADS 0, 2, 3, 4 e 5.

MATERIAIS E MÉTODOS

A seguir será apresentada a metodologia empregada na modelagem do sistema de inferência nebuloso desenvolvido

Participação de Especialistas do Domínio

A fim de conhecer os parâmetros perceptuais utilizados por médicos radiologistas na avaliação de nódulos de mama com a aplicação prática do léxico BI-RADS, foram realizadas reuniões com radiologistas do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (HC-FMRP) da Universidade de São Paulo (USP). As reuniões embasaram a escolha da metodologia a ser utilizada no projeto, permitindo a modelagem do sistema na criação de escalas representativas da malignidade da lesão e na criação dos subconjuntos nebulosos, a partir do conhecimento desses especialistas.

Base de Imagens

Foi utilizada neste trabalho a base de imagens de mamografia digital DDSM (*Digital Database for Screening Mammography*)⁽¹⁶⁾ construída por pesquisadores do Hospital Geral de Massachusetts e da Universidade do Sul da Flórida, com lesões recortadas por

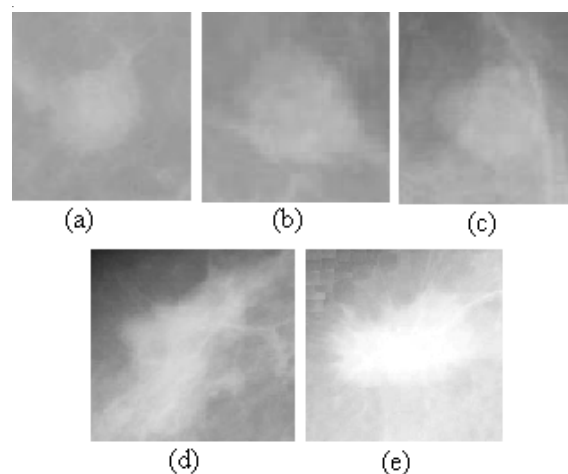


Figura 5 - Exemplo de lesões classificadas como (a) BI-RADS 0, (b) BI-RADS 2, (c) BI-RADS 3, (d) BI-RADS 4 e (e) BI-RADS 5⁽¹⁶⁾.

pesquisadores da Universidade de Chicago. Para este trabalho foram utilizadas 250 imagens compostas por regiões de interesse contendo lesões benignas ou malignas, diagnosticadas por radiologistas, com auxílio de exames anátomo-patológicos.

Modelagem do Sistema

A modelagem do sistema nebuloso proposto foi dividida em duas etapas.

Na **primeira etapa** foram utilizados os atributos visuais para classificação de nódulos de mama padronizados pelo BI-RADS, apresentados na Introdução. Para avaliar a contribuição dos atributos e suas características na determinação da classificação de um achado, a malignidade do tumor foi tratada de maneira escalar. Foram criadas escalas representativas do conceito malignidade, variando no intervalo (0,1), que indicassem a relação entre este conceito e uma determinada característica de um atributo em análise.

Essas escalas foram apresentadas na forma de um questionário a cinco radiologistas seniores, experientes em diagnóstico de mama, sendo que três deles são docentes da Faculdade de Medicina de Ribeirão Preto e os outros dois profissionais atuantes em clínicas especializadas no diagnóstico de imagens de mama. Os questionários continham campos para a avaliação da forma, do contorno e da densidade do nódulo. Para a seleção das imagens foi utilizado o BI-RADS associado a cada uma delas, armazenado na base IARL⁽¹⁶⁾. Foram selecionadas 35 imagens de diferentes categorias BI-RADS a fim de se obter uma amostra diversificada para ser analisada por cada radiologista. Durante a avaliação, para cada atributo era adotada uma característica e apontado um valor na escala que indicasse a influência da característica na malignidade da lesão. O valor escolhido atuou como peso na determinação da malignidade.

Para o atributo *formato*, foi utilizada uma escala para relacionar as características “redonda”, “oval”,

“lobular” e “irregular” aos seus respectivos graus de malignidade, já que usualmente apenas uma delas é adotada durante a avaliação do radiologista. Como exemplo, poderíamos obter a seguinte avaliação: “formato lobular com grau de malignidade 0,6”. Similar ao atributo *formato*, o atributo *densidade* possui também apenas uma escala que relaciona as características “alta”, “isodensa”, “baixa” e “de aspecto gorduroso” à escala de malignidade.

Quanto ao atributo *contorno*, a fim de se manter a fidelidade ao padrão BI-RADS, foi permitida a escolha de mais de uma característica para um mesmo achado. Uma escala foi adotada para cada característica: “circunscrito”, “microlobulado”, “obscuro”, “indistinto” e “espiculado”. Também relativo ao contorno, foram determinadas quatro variáveis lingüísticas, que atuam como pesos na avaliação final do contorno: “totalmente”, “muito”, “parcialmente” e “pouco”. No questionário permitiu-se a escolha de uma das variáveis acima descritas para cada característica escolhida associada a um grau de malignidade, por exemplo, “contorno parcialmente circunscrito com grau de malignidade 0,2 e parcialmente obscurecido com grau de malignidade 0,5”. É importante notar que a variável “totalmente” exclui a escolha de mais de uma característica e a quantidade de características selecionadas dependerá da avaliação do radiologista. Por fim, a categoria BI-RADS foi determinada pelo radiologista para cada exemplo analisado.

Após a avaliação das lesões pelos radiologistas, um banco de dados foi modelado e implementado para armazenar as informações obtidas e facilitar a extração dos parâmetros. As informações armazenadas representam o conhecimento dos especialistas do domínio, importantes em um sistema nebuloso, e serão a referência para a criação dos conjuntos nebulosos e da base de proposições. O próximo passo consistiu na determinação de três coeficientes: *formato*, *contorno* e *densidade*, que representarão a entrada nebulosa do sistema proposto. Os coeficientes de formato e de densidade assumem os valores atribuídos como grau de malignidade na entrada do sistema pelo radiologista. Para o cálculo do coeficiente de contorno, realizou-se uma média ponderada das avaliações obtidas, caso mais de uma característica fosse escolhida. O peso utilizado foi a média aritmética da característica obtida das avaliações. Seja $R_{i,j}$ o grau de malignidade definido pela combinação qualificador “i” com a característica “j”, na qual i poderá assumir os seguintes valores: totalmente, muito, parcialmente ou pouco; e j representado pela característica do contorno do nódulo. O coeficiente de contorno é dado pela seguinte relação:

$$\text{coef}_{\text{contorno}} = \frac{\sum_{k=1}^n R_{i,j} \cdot P_j}{\sum_{k=1}^n P_j} \quad (2)$$

na qual, n representa o número de combinações entre variável e característica escolhidas pelo radiologista e P_j , representa o peso atribuído à característica j. A Tabela 1 apresenta os pesos obtidos (P_j) para cada característica do contorno e $R_{i,j}$ o grau de malignidade da combinação variável / característica.

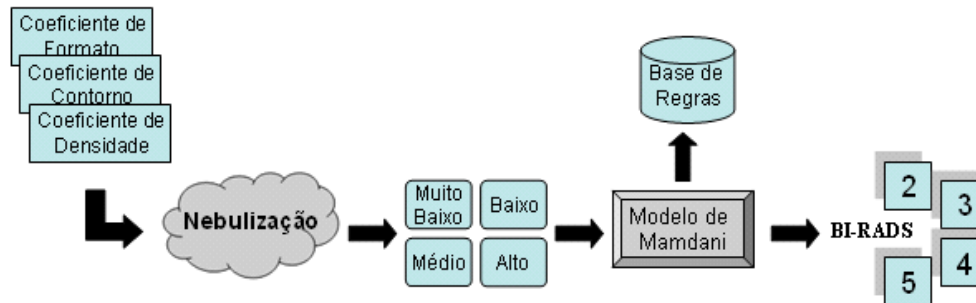
A **segunda etapa** consistiu na criação do sistema de inferência nebuloso e na escolha de seus parâmetros estruturais, tais como o número de variáveis de entrada e saída e o tipo de função de pertinência utilizada.

A entrada do sistema é composta pelos três coeficientes determinados na etapa anterior: coeficiente de formato, coeficiente de contorno e coeficiente de densidade. Eles representam a entrada numérica do sistema e serão mapeados em variáveis lingüísticas representativas dos conceitos “alto”, “médio”, “baixo” e “muito baixo”. Quanto mais alto o coeficiente, maior influência sobre a malignidade da lesão. Optou-se pela representação dos coeficientes por meio de conjuntos nebulosos, utilizando funções Sino para representar as funções de pertinência. Este tipo de função apresenta uma transição suave das fronteiras entre os subconjuntos, adequando-se ao problema em análise. Elas diferem das funções de pertinência gaussianas por possuírem um intervalo de valores com grau de pertinência máxima. O modelo de inferência escolhido foi o método de Mamdani. As regras nebulosas serão formadas pela combinação das variáveis correspondentes aos três coeficientes de entrada. As categorias BI-RADS 2, 3, 4 e 5 formarão a parte conseqüente das regras. A categoria 0 não será considerada na modelagem, pois como descrito anteriormente, ela é atribuída quando há necessidade de exame complementar e somente após realização do exame é que, uma das outras categorias poderá ser atribuída à lesão. Uma modelagem esquemática pode ser visualizada na Figura 6.

Foi utilizado o Fuzzy Toolbox® do Matlab® para implementação do sistema nebuloso e simulação do procedimento de inferência. A parte conseqüente das regras foi determinada utilizando-se as avaliações obtidas a partir de questionários aplicados aos radiologistas, representando o conhecimento de especialistas da área. Foram considerados os subconjuntos dos coeficientes de formato, contorno e densidade como antecedentes das regras. Os subconjuntos são representados por meio das funções de pertinência geradas. Para cada avaliação, foram determinadas as combinações de subconjuntos ativados, a partir dos graus definidos pelos radiologistas durante as entrevistas. As combinações obtidas foram armazenadas na base de dados de avaliações. Após cadastradas todas as combinações ativadas para todos os exemplos da base, foram selecionadas as relações que representassem as mesmas combinações de entrada do sistema e a classe majoritária do conjunto selecionado foi utilizada para determinar a parte conseqüente da regra.

Tabela 1 - Pesos obtidos para as características do atributo contorno.

Característica	Peso característica	Totalmente	Muito	Parcialmente	Pouco
Circunscrito	0.353	0.243	0.350	0.466	0.500
Microlobulado	0.696	0.821	0.750	0.450	0.533
Obscurecido	0.572	0.500	0.400	0.585	0.525
Indistinto	0.672	0.800	0.637	0.648	0.600
Espiculado	0.830	0.965	0.816	0.760	0.766

**Figura 6** - Determinação de Coeficientes a partir das Avaliações.

RESULTADOS

A variável de entrada *coeficiente de formato* foi dividida em três subconjuntos nebulosos, assumindo os seguintes valores lingüísticos: baixo, médio e alto, variando no intervalo (0,1) (Figura 7).

A variável de entrada *coeficiente de contorno* foi dividida em quatro subconjuntos nebulosos, assumindo os seguintes valores lingüísticos: muito baixo, baixo, médio e alto, variando no intervalo (0,1) (Figura 8).

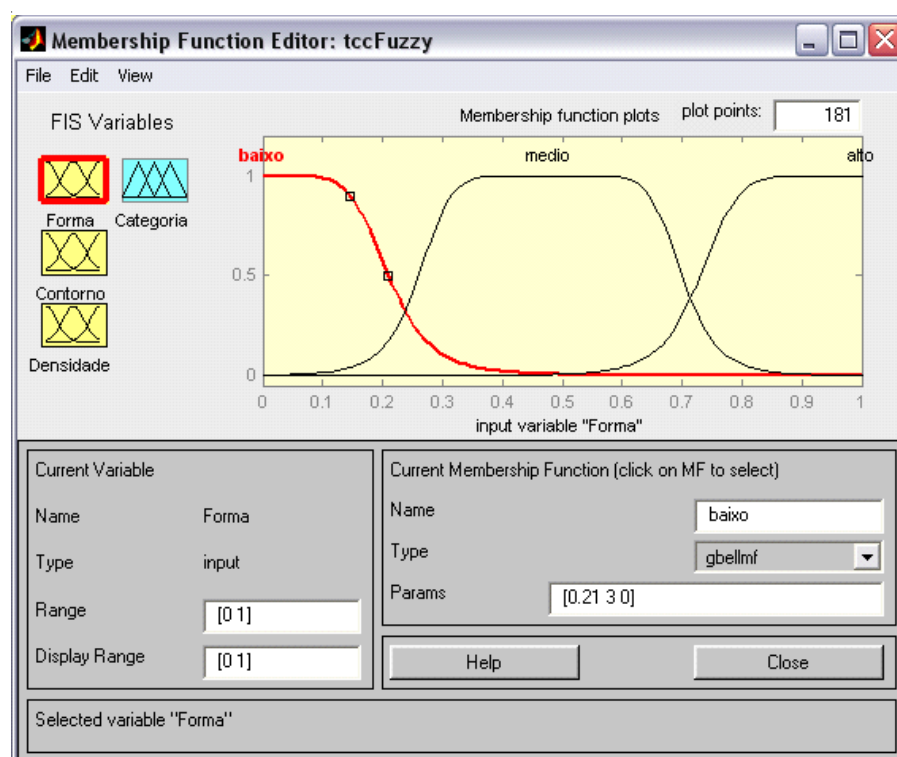
A variável de entrada *coeficiente de densidade* foi dividida em três subconjuntos nebulosos, assumindo os

seguintes valores lingüísticos: baixo, médio e alto, também variando no intervalo (0,1) (Figura 9)

Foram consideradas 36 possibilidades para a base de regras, combinando-se os subconjuntos do *coeficiente de formato*, os subconjuntos do *coeficiente de contorno* e os subconjuntos do *coeficiente de densidade*.

Avaliação do Sistema

Testes foram realizados comparando-se as avaliações dos radiologistas com a saída do sistema nebuloso. Na Tabela 2 são apresentados alguns resultados

**Figura 7** - Funções de Pertinência para o coeficiente de formato.

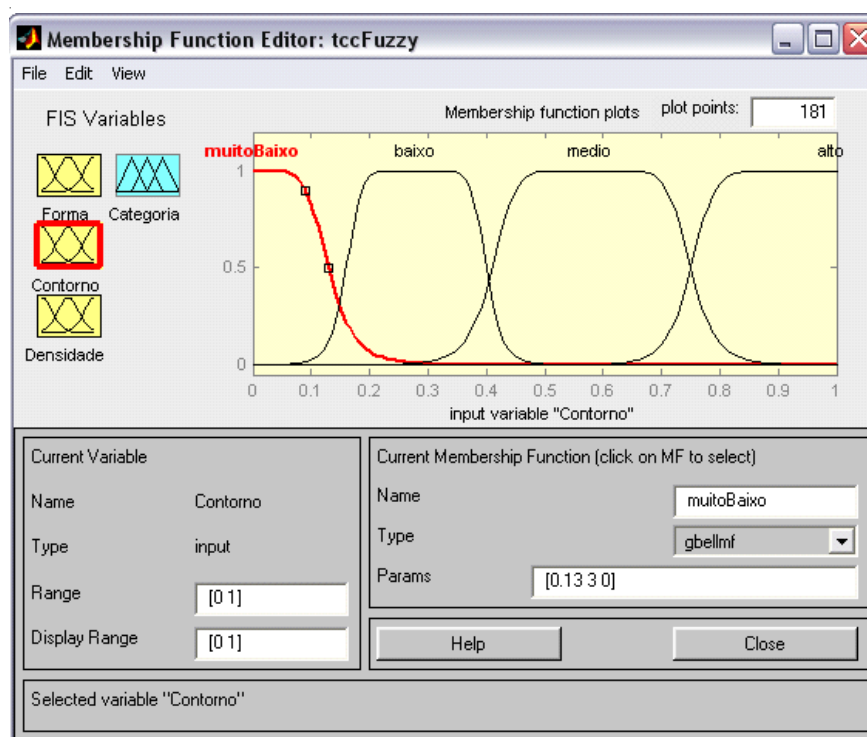


Figura 8 - Funções de Pertinência para o coeficiente de contorno.

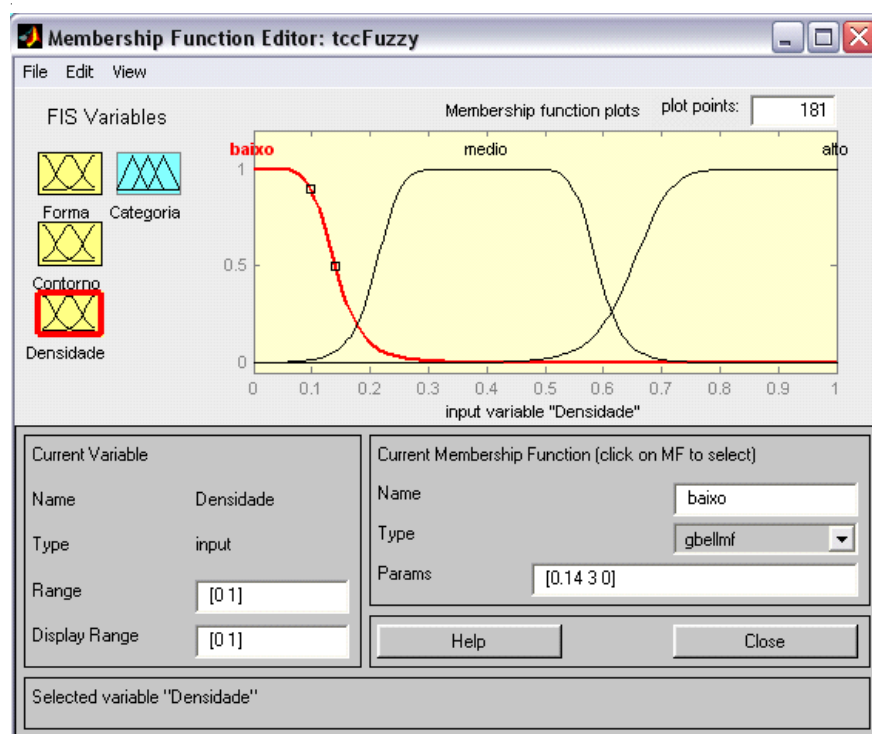


Figura 9 - Funções de Pertinência para o coeficiente de densidade.

Para o exemplo 1 (E1), temos a seguinte avaliação: formato irregular, contorno muito espiculado e densidade alta. A categoria 4 foi atribuída ao exemplo pelos radiologistas R1 e R2 e a categoria 5, pelo radiologista R3. O sistema gerou a seguinte saída: categoria 4 com pertinência de 67% e categoria 5 com pertinência de 20%.

Para o exemplo 2 (E2), consideramos a seguinte avaliação: formato lobular, contorno parcialmente

obscuro e densidade alta. Similar ao exemplo anterior, a categoria 3 foi atribuída ao exemplo pelos radiologistas R1 e R3 e a categoria 4, pelo radiologista R2. A saída gerada pelo sistema foi: categoria 3 com pertinência de 61% e categoria 4 com pertinência de 12%.

Por fim, para o exemplo 3 (E3), foi utilizado como entrada do sistema a seguinte avaliação: formato irregular, contorno totalmente espiculado e densidade

Tabela 2 - Exemplos de resultados obtidos.

Parâmetros de Entrada				Avaliação BI-RADS	
Formato	Contorno	Densidade			
E1	Irregular	Muito espiculado	Alta	R1	4
				R2	4
				R3	5
				Sistema	4 (pert. 67%)
				Sistema	5 (pert. 20%)
E2	Lobular	Parcialmente obscuro	Alta	R1	3
				R2	4
				R3	3
				Sistema	3 (pert. 61%)
				Sistema	4 (pert. 12%)
E3	Irregular	Totalmente espiculado	Alta	R1	5
				R2	5
				R3	5
				Sistema	5 (pert. 100%)

alta. A categoria 5 foi atribuída ao exemplo pelos 3 radiologistas e resultou em pertinência máxima como saída do sistema.

Os exemplos 1 e 2 caracterizam regiões de fronteira entre conjuntos, representadas pelos diferentes graus de pertinência atribuídos a cada conjunto. O exemplo 3 é um elemento com pertinência total à categoria 5. Nota-se que o qualificador “totalmente” junto a um contorno espiculado, geram um coeficiente de formato alto, elevando o grau de pertinência do exemplo à categoria 5.

DISCUSSÕES E CONCLUSÕES

Este trabalho apresenta um método para caracterização nebulosa dos principais descritores da terminologia lingüística utilizada na avaliação de nódulos de mama, para a determinação da categoria BI-RADS associada. O sistema desenvolvido contou com a participação de especialistas do domínio de aplicação e métodos computacionais que auxiliaram toda a etapa de modelagem e implementação do sistema.

A partir dos exemplos armazenados na base de dados de avaliações, foi possível observar que, dada uma mesma imagem de referência, diferentes avaliações entre os radiologistas foram obtidas. Os exemplos analisados mostram que há uma intersecção entre as categorias BI-RADS adotadas pelos radiologistas, evidenciando a ausência de uma fronteira bem definida entre as diferentes categorias de avaliação e até mesmo entre os atributos de análise dos nódulos.

Nas avaliações realizadas com os radiologistas, houve a preocupação em fornecer um questionário intuitivo, com uma análise representativa do processo de categorização de nódulos em exames mamográficos, a fim de facilitar a obtenção de coeficientes a partir dos atributos de análise. A criação de uma escala associada à malignidade da lesão, variando com as características dos atributos, foi um

ponto crucial na determinação das funções de pertinência, visto que apenas variáveis lingüísticas são descritas no BI-RADS. As escalas possibilitaram todo o processo de inferência nebuloso e atingiram o objetivo de atuar como um peso na determinação da categoria de avaliação do nódulo, uma vez que, durante as entrevistas realizadas, uma mesma característica foi “pesada” de maneiras diferentes. Por exemplo, “formato lobular com grau 0,7” e “formato lobular com grau 0,4”. Evidencia-se que um grau menor para o formato pesará menos na determinação da malignidade da lesão.

A utilização da Lógica Nebulosa permitiu a representação da subjetividade observada nas avaliações dos achados mamográficos, durante as entrevistas realizadas com radiologistas. Ao tratar os atributos *forma*, *contorno* e *densidade* como conjuntos nebulosos, foi possível determinar um grau de pertinência de um nódulo a cada uma das categorias de avaliação BI-RADS. Além disso, o método proposto possibilitará a criação de um sistema CAD mais amigável, no qual os parâmetros de entrada serão bastante próximos à forma natural como o radiologista avalia e qualifica os achados da imagem, caracterizando-se, portanto, como um sistema computacional não-intrusivo capaz de promover a redução do hiato semântico.

Os resultados discutidos permitem concluir que a metodologia proposta poderá ser utilizada em sistemas de auxílio a diagnóstico, fornecendo uma segunda opinião consistente, embasada em características quantitativas de análise.

AGRADECIMENTOS

Os autores agradecem à FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) pelo apoio financeiro e aos radiologistas Drs. Clóvis Simão Trad, Angélica Godinho, Valdair Francisco Muglia, Jorge Elias Junior, e André Luis Caliento Paiva pela avaliação das lesões selecionadas.

REFERÊNCIAS

1. Doi K. Current status and future potential of computer-aided diagnosis in medical imaging. *Br J Radiol.* 2005;78:S3-s19.
2. Yin T-K, Chiu N-T. A computer-aided diagnosis for locating abnormalities in bone scintigraphy by a fuzzy system with a three-step minimization approach. *IEEE Trans Med Imaging* 2004;23(5):639-54.
3. Doi K. Diagnostic Imaging over the last 50 years: research and development in medical imaging science and technology. *Phys Med Biol* 2006; 51(13):5-27.
4. Marques PMA. Computer-aided diagnosis in radiology. *Radiol Bras* 2001;34(5):285-93.
5. Oliveira-JR, HIME A. Lógica difusa: aspectos práticos e aplicações. Rio de Janeiro: Interciência; 1999.
6. Obenauer S, Hermann KP, Grabbe E. Applications and literature review of the BI-RADS classification. *Eur Radiol.* 2005; 15(5):1027-36.
7. Kovalerchuk B, Triantaphyllou E, Ruiz JF, Clayton J. Fuzzy logic in computer-aided breast cancer diagnosis: analysis of lobulation. *Artif Intell Med.* 1997;11(1):75-85.
8. Bayram B, Acar U. An approach to the detection of lesions in mammograms using fuzzy image processing. *J Int Med Res.* 2007;35(6):790-5.
9. Amendola M, Souza A. Manual do uso da teoria dos conjuntos Fuzzy no Matlab 6.5. Campinas: CPG/FEAGRI/UNICAMP; 2004.
10. Sandri S, Correa C. Lógica Nebulosa. V Escola de Redes Neurais, Promoção: Conselho Nacional de Redes Neurais, 19 de julho, 1999 – ITA, São José dos Campos – SP. c073-c090
11. Zadeh LA. Fuzzy logic and approximate reasoning (In memory of Grigore Moisil). *Synthese* 1975; 30:407-28.
12. Zadeh LA, Bellman RE. Local and Fuzzy logics. In: Dunn JM, Epstein G. Modern uses of multiple-valued logic. Netherland: D.Reidel Publishing Co; 1977. p.103-65.
13. Klir GJ, Yuan B. Fuzzy Sets And Fuzzy logic: theory and applications. Rio de Janeiro: Prentice-Hall; 1995.
14. Dubois D, Prade H. What are fuzzy rules and how to use them. *Fuzzy sets and systems.* 1996;84(2):169-85.
15. Almeida PEM, Evsukoff AG. Sistemas Fuzzy. In: Rezende, S. O. (org). *Sistemas inteligentes: fundamentos e aplicações.* São Paulo: Manole; 2003. p.169-200.
16. IARL. [homepage on the Internet]. Digital Database for Screening Mammography. [citado em 13 jul 2008]. Disponível em: <http://marathon.csee.usf.edu/Mammography/>.