

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

Gisele Helena Barboni Miranda

Lógica Nebulosa Aplicada à Avaliação de Nódulos de
Mama Visando a Redução do Hiato Semântico no Ambiente
de Diagnóstico Auxiliado por Computador

Ribeirão Preto, 2008

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

Gisele Helena Barboni Miranda

Lógica Nebulosa Aplicada à Avaliação de Nódulos de
Mama Visando a Redução do Hiato Semântico no Ambiente
de Diagnóstico Auxiliado por Computador

Monografia de conclusão de curso de
graduação apresentada à Faculdade de
Filosofia, Ciências e Letras de Ribeirão Preto
e à Faculdade de Medicina de Ribeirão Preto,
Universidade de São Paulo, como requisito da
disciplina Desenvolvimento de Projeto II
(IBM1064).

Prof. Dr. Joaquim Cezar Felipe
Orientador

Prof. Dr. Paulo Mazzoncini de A. Marques
Co-Orientador

Ribeirão Preto, 2008

Gisele Helena Barboni Miranda

Lógica Nebulosa Aplicada à Avaliação de Nódulos de
Mama Visando a Redução do Hiato Semântico no Ambiente
de Diagnóstico Auxiliado por Computador

Monografia de conclusão de curso de graduação
apresentada à Faculdade de Filosofia, Ciências e Letras
de Ribeirão Preto e à Faculdade de Medicina de Ribeirão
Preto, Universidade de São Paulo.

Aprovado em: _____

Conceito: _____

Banca Examinadora

Prof. Dr. Joaquim Cezar Felipe
Faculdade de Filosofia Ciências e Letras de Ribeirão Preto
Universidade de São Paulo

Prof. Dr. Paulo Mazzoncini de Azevedo Marques
Faculdade de Medicina de Ribeirão Preto
Universidade de São Paulo

Prof. Dr. Renato Tinós
Faculdade de Filosofia Ciências e Letras de Ribeirão Preto
Universidade de São Paulo

*Aos meus familiares e amigos,
Pelo apoio, preocupação e incentivo
durante esses quatro anos.*

AGRADECIMENTOS

À Deus, por mais uma conquista.

Aos meus pais, pela oportunidade única que me proporcionaram, pelo apoio incondicional, pelo conforto nos momentos difíceis, pelo exemplo de perseverança e dedicação que pude observar em diversos momentos da minha vida.

Ao meu orientador, o Prof. Dr. Joaquim Cezar Felipe, pela credibilidade e confiança dispensadas a mim no desenvolvimento deste trabalho, pelo apoio acadêmico e profissional que obtive durante as atividades realizadas e à amizade adquirida ao longo deste ano.

Ao meu co-orientador, o Prof. Dr. Paulo Mazzoncini de Azevedo Marques pela colaboração e apoio durante o desenvolvimento deste projeto.

Aos radiologistas Dr. André Luis Caliento Paiva, Dra Angélica Maria Reis Godinho, Dr. Clóvis Simão Trad, Dr. Jorge Elias Júnior e Dr. Valdair Francisco Muglia pela contribuição nas avaliações das lesões selecionadas utilizadas na fase de modelagem e avaliação do sistema desenvolvido.

Aos meus amigos e colegas da 3ª Turma do curso de Informática Biomédica, pelo carinho, apoio, amizade e convivência que me fizeram crescer e aprimorar os meus conhecimentos e que me auxiliaram na escolha de uma área de atuação como Informata Biomédica. Pelos ótimos momentos e recordações que me proporcionaram.

À Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP, pelo auxílio financeiro.

RESUMO

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, a fim de se gerar automaticamente a classificação BI-RADS para o diagnóstico de lesões de mama, de forma que a adaptação de definições e regras lingüísticas aos modelos numéricos, proporcionada pela Lógica Nebulosa, possa propiciar uma aproximação entre o sistema computacional e o usuário especialista, gerando resultados mais efetivos e fomentando, assim, a redução do hiato semântico.

Palavras-chave: Lógica Nebulosa, Diagnóstico Auxiliado por Computador, BI-RADS.

ABSTRACT

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 by the divergence between the results obtained through pattern recognition and decision support algorithms and the results expected by the radiologist. This divergence occurs, mainly, due 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 related to Fuzzy Logic in the definition of the describing attributes of the terminology used by the radiologists to characterize findings, to generate, automatically, the BI-RADS classification of a breast lesion. This way, the adaptation of definitions and linguistic rules to the numerical models, achieved through 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.

Key-words: Fuzzy Logic, Computer-aided Diagnosis, BI-RADS

SUMÁRIO

LISTA DE FIGURAS	V
LISTA DE TABELAS	VI
LISTA DE SIGLAS	VII
CAPÍTULO 1 INTRODUÇÃO	1
1.1. CONTEXTUALIZAÇÃO	1
1.2. MOTIVAÇÃO	2
1.3. OBJETIVOS	3
1.4. ORGANIZAÇÃO DA MONOGRAFIA	4
CAPÍTULO 2 FUNDAMENTOS TEÓRICOS	5
2.1. CONSIDERAÇÕES INICIAIS	5
2.2. DIAGNÓSTICO AUXILIADO POR COMPUTADOR	5
2.3. CLASSIFICAÇÃO DE NÓDULOS DE MAMA	6
2.3.1. DESCRITORES DE MASSAS	7
2.3.2. CATEGORIAS DE AVALIAÇÃO	8
2.4. LÓGICA NEBULOSA	9
2.4.1. CONJUNTOS NEBULOSOS	9
2.4.2. OPERAÇÕES COM CONJUNTOS NEBULOSOS	11
2.4.3. RELAÇÕES NEBULOSAS	11
2.4.4. MÁQUINA DE INFERÊNCIA NEBULOSA	12
2.4.5. MODELOS DE INFERÊNCIA	13
2.4.5.1. <i>Modelo de Mamdani</i>	13
2.4.5.2. <i>Modelo de Takagi-Sugeno-Kang</i>	14
CAPÍTULO 3 METODOLOGIA	15
3.1. CONSIDERAÇÕES INICIAIS	15
3.2. REPRESENTAÇÃO NEBULOSA DA TERMINOLOGIA LINGÜÍSTICA	15
3.2.1. <i>PRIMEIRA ETAPA</i> : PARTICIPAÇÃO DE ESPECIALISTAS DO DOMÍNIO	15
3.2.1.1. BASE DE DADOS DE AVALIAÇÕES	17
3.2.2. <i>SEGUNDA ETAPA</i> : DETERMINAÇÃO DOS PARÂMETROS DE ENTRADA	18
3.3. DETERMINAÇÃO DAS FUNÇÕES DE PERTINÊNCIA	19
3.3.1. <i>OMEGA</i> : DISCRETIZAÇÃO DE DADOS E SELEÇÃO DE ATRIBUTOS	20
3.3.2. <i>FUZZY OMEGA</i> : DETERMINAÇÃO DAS FUNÇÕES DE PERTINÊNCIA	21
3.4. BASE DE PROPOSIÇÕES NEBULOSAS	24
CAPÍTULO 4 RESULTADOS	25
4.1. CONSIDERAÇÕES INICIAIS	25
4.2. FUNÇÕES DE PERTINÊNCIA OBTIDAS	25
4.3. BASE DE PROPOSIÇÕES NEBULOSAS	30
4.4. SISTEMAS NEBULOSOS EM AMBIENTES DE AUXÍLIO AO DIAGNÓSTICO	31
4.4.1. PROTÓTIPO IMPLEMENTADO: <i>FUZZY BI-RADS</i>	31

4.5. VALIDAÇÃO DOS RESULTADOS	35
<u>CAPÍTULO 5 DISCUSSÃO E CONCLUSÕES</u>	<u>38</u>
5.1. DISCUSSÃO	38
5.2. POSSIBILIDADE DE TRABALHOS FUTUROS	40
5.3. CONCLUSÕES	40
<u>REFERÊNCIAS BIBLIOGRÁFICAS</u>	<u>42</u>
<u>APÊNDICE A – AVALIAÇÃO REALIZADA COM OS RADIOLOGISTAS</u>	<u>47</u>
<u>APÊNDICE B – ATRIBUTOS DE ENTRADA DO ALGORITMO FUZZY-OMEGA</u>	<u>48</u>
<u>APÊNDICE C – EQUAÇÕES POLINOMIAIS DAS FUNÇÕES DE PERTINÊNCIA</u>	<u>51</u>
<u>APÊNDICE D – DIAGRAMA DE CLASSES DO SISTEMA FUZZY-BI-RADS</u>	<u>55</u>

LISTA DE FIGURAS

Figura 2.1. Forma: Aa redonda, Ab lobular, Ac irregular. Contorno: Ba bem definido, Bb mal definido, Bc obscurecido, Bd espiculado. Densidade: Ca contém gordura, Cb isodensa, Cc alta. [Obenauer, S. '05].	7
Figura 2.2. Exemplo de lesões classificadas como (a) BI-RADS 0, (b) BI-RADS 2,.....	8
Figura 2.3. Determinação do grau de pertinência por meio da função de pertinência.	10
Figura 2.4. Exemplos de funções de pertinência.	11
Figura 2.5. Máquina de Inferência Nebulosa.	12
Figura 3.1. Diagrama Entidade - Relacionamento da Base de Dados de avaliações.....	17
Figura 3.2. Modelagem do Sistema Nebuloso.....	19
Figura 3.3. Passo 1 do algoritmo Omega.	20
Figura 3.4. Passo 2 do algoritmo Omega.	20
Figura 3.5. Passo 3 do algoritmo Omega.	21
Figura 3.6. Criação de subintervalos para o cálculo da taxa de inconsistência por elemento..	23
Figura 3.7. Exemplo de intervalo gerado pelo algoritmo Fuzzy-Omega.	23
Figura 4.1. Funções de Pertinência para o atributo coeficiente de contorno. (a) Função de pertinência com ajuste polinomial não ótimo. (b) Função de pertinência não ordinária.	26
Figura 4.2. Funções de pertinência para o atributo Coeficiente de Formato.....	27
Figura 4.3. Funções de pertinência para o atributo Coeficiente de Contorno.	28
Figura 4.4. Funções de pertinência para o atributo Coeficiente de Densidade.	28
Figura 4.5. Funções finais de pertinência para o atributo Coeficiente de Formato.	29
Figura 4.6. Interface gráfica do sistema desenvolvido.....	32
Figura 4.7. Visualização da resposta gerada pelo sistema.....	33
Figura 4.8. Subconjuntos nebulosos ativados.....	34
Figura 4.9. Imagem de referência (a) usf3424_LCC_L1_MS_B, (b) usf4077_LCC_L1_MS_M, (c) usf1187_RMLO_L1_MS_M, (d) usf0305_RMLO_L1_MS_B, (e) usf1820_LCC_L1_MS_M,	35

LISTA DE TABELAS

Tabela 1. Algumas operações com conjuntos nebulosos.....	11
Tabela 2. Atributos utilizados na classificação de nódulos de mama.	16
Tabela 3. Pesos obtidos para as características do atributo contorno.	18
Tabela 4. Regras ativadas para o exemplo em análise.	35
Tabela 5. Comparação dos resultados obtidos das avaliações com os resultados gerados pelo sistema.....	36
Tabela 6. Exemplos de avaliações realizadas.	38

LISTA DE SIGLAS

ACR	<i>American College of Radiology</i>
BI-RADS	<i>Breast Imaging Reporting and Data System</i>
CAD	<i>Computer-aided Diagnosis</i> (Diagnóstico Auxiliado por Computador)
DDSM	<i>Digital Database for Screening Mammography</i>
GBDI	Grupo de Bases de Dados e Imagens
HC-FMRP	Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto
ICMC	Instituto de Ciências Matemáticas e de Computação
NMD	<i>National Mammography Database</i>
PACS	<i>Picture Archiving and Communication System</i>

CAPÍTULO 1

Introdução

1.1. Contextualização

Nas últimas décadas, o volume de imagens médicas digitais produzidas em hospitais e centros de diagnóstico por imagens vem crescendo em escala exponencial. As atividades relacionadas à aquisição, ao gerenciamento e à disponibilização destas imagens têm exigido esforços de pesquisadores e profissionais envolvidos na informatização dos sistemas hospitalares, no sentido de estudar e propor métodos computacionais para dar suporte efetivo às tarefas cotidianas dos núcleos de análise de imagens.

Para a integração dos processos de aquisição, gerenciamento e disponibilização de imagens no ambiente médico são propostos os Sistemas de Comunicação e Arquivamento de Imagens (*Picture Archiving and Communication System - PACS*) [MARSH, A. '97]. Estes sistemas de arquivamento de imagens possibilitam a criação de grandes bases de dados de diferentes modalidades de imagens, englobando as diversas patologias registradas no hospital. As bases de imagens são consideradas uma fonte preciosa para o desenvolvimento de ferramentas de suporte à prática da medicina baseada em evidências, possibilitando o uso de casos similares como auxílio à tomada de decisão.

Paralelamente à evolução das imagens digitais, 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 do paciente, sua evolução clínica, assim como exames realizados têm se tornado comuns 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*) [DOI, K. '05] [YIN, T.-K. '04], as quais têm se mostrado coadjuvantes capazes de promover um aumento da qualidade do trabalho dos radiologistas [DOI, K. '06].

Os sistemas CAD têm como objetivo auxiliar e aumentar a precisão do diagnóstico do radiologista, através do uso dos resultados do computador como referência, como, por exemplo, a indicação de áreas suspeitas da imagem assim como anormalidades mascaradas. Este auxílio do computador é importante, já que o diagnóstico do radiologista é baseado em avaliações subjetivas, sujeitas as variações pessoais (tais como perda de informações devido à

fadiga visual e distração), qualidade da imagem e sobreposição de estruturas, entre outros fatores [MARQUES, P. M. A. '01]. Desta forma, os sistemas CAD têm mostrado que podem melhorar o desempenho dos diagnósticos radiológicos.

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

A Lógica Nebulosa ou Lógica *Fuzzy* foi desenvolvida por Lofti Zadeh, professor da Universidade de Berkeley, em 1965. Ela utiliza conceitos intuitivos para representar um modelo matemático clássico [OLIVEIRA-JR, H. A. '99]. 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 as transições bruscas. Esta abordagem aplica-se a situações, nas quais, a fronteira entre os 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).

Dessa forma, a proposta deste projeto consiste no estudo e no desenvolvimento de métodos que apliquem os conceitos da Lógica Nebulosa à representação de imagens médicas por características visuais utilizadas na avaliação das mesmas, visando à redução do hiato semântico que ocorre nos ambientes de análise, recuperação por similaridade e geração de diagnóstico de radiografias de mama baseado na classificação BI-RADS. 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.

1.2. Motivação

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 do cuidado à saúde.

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 estar abordando um contexto específico ou amplo, a representação puramente numérica de uma imagem por meio de um vetor de características, por si só, já caracteriza um hiato semântico significativo, uma vez que esta representação está relativamente distante da forma como o analista humano interpreta a imagem. 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.

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. 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.

1.3. Objetivos

Este projeto tem por objetivo investigar a efetividade do uso da representação dos parâmetros de análise dos nódulos de mama por meio da Lógica Nebulosa. Para isto, pretende-se desenvolver um sistema nebuloso que forneça uma análise quantitativa automatizada (computadorizada) de nódulos de mama, baseado em parâmetros de observação

da imagem fornecidos pelo usuário, funcionando como uma segunda opinião para auxiliar o radiologista (e também outros médicos) na tomada de decisão do diagnóstico e na classificação do exame mamográfico da paciente.

Para essa análise computadorizada, serão considerados os principais atributos utilizados nos laudos mamográficos para achados do tipo “massas”, padronizados pelo sistema BI-RADS de classificação, que são: “forma”, “margem” e “densidade”. Também serão consideradas as categorias de avaliação resultantes dos laudos dos exames e o seu prognóstico correspondente.

O sistema deverá caracterizar os descritores escolhidos, assim como as categorias de avaliação, através de variáveis nebulosas e seus respectivos subconjuntos. O processo de inferência da categoria de acesso será determinado pela “máquina de inferência nebulosa”, que fornecerá a saída nebulosa do sistema para cada entrada nebulosa, a partir das relações definidas na base de regras, também descritas em formato nebuloso.

1.4. Organização da Monografia

O documento está organizado da seguinte forma: a Seção 2 apresenta os fundamentos teóricos sobre lógica nebulosa, diagnóstico auxiliado por computador e avaliação de nódulos de mama; a Seção 3 descreve a metodologia utilizada neste trabalho; a Seção 4 apresenta os resultados obtidos e, finalmente, a Seção 5 apresenta as conclusões e as possibilidades de trabalhos futuros.

CAPÍTULO 2

Fundamentos Teóricos

2.1. Considerações Iniciais

Neste capítulo serão abordados os fundamentos teóricos do projeto. Os temas incluem a teoria de Lógica Nebulosa, sistemas de Diagnóstico Auxiliado por Computador, e o processo de classificação de nódulos de mama em exames radiográficos.

2.2. Diagnóstico Auxiliado por Computador

Diagnóstico auxiliado por computador pode ser definido como o processo em que um radiologista usa resultados de uma análise quantitativa automatizada (computadorizada) de imagens radiológicas como uma “segunda opinião” para a detecção de lesões ou para tomar uma decisão de diagnóstico [DOI, K. ‘99] [GIGER, M. L. ‘99]. Vale lembrar que o diagnóstico final é sempre efetuado por um médico radiologista.

Os sistemas CAD podem ser basicamente divididos em dois tipos. O primeiro é o auxílio à detecção de lesões, localizando-se padrões anormais a partir da varredura da imagem, como, por exemplo, a análise computadorizada para detecção de nódulos de pulmão e também de mamas. O segundo é o auxílio ao diagnóstico, o qual pode ser descrito como a classificação das características e padrões da imagem como anormal ou normal, como, por exemplo, a classificação de nódulos de mamas em grupos que apresentam ou não malignidade do tumor. Estas análises em geral estão baseadas em detecção de bordas, realce de contraste e métodos de processamento hierárquico, como árvores de decisão, mas existem outros métodos que vêm sendo utilizados, tais como morfologia matemática, redes neurais, análise de fractais e transformadas de *wavelets* [GIGER, M. L. ‘00]. Essas técnicas podem ser divididas em duas áreas do conhecimento da Ciência da Computação: a Inteligência Artificial, que tem como função a detecção de padrões ou seleções de atributos que mais influenciam na decisão e a Visão Computacional, que, de forma geral, aplica as técnicas clássicas de processamento de imagens [CASTLEMAN, K. N. ‘96].

Por ser um conceito muito abrangente, a idéia de CAD pode ser aplicada em todas as áreas da radiologia, tanto na aquisição de imagens, como para diferentes tipos de exames

médicos. Assim, os sistemas CAD têm se tornado uma importante ferramenta de auxílio em várias áreas radiológicas, destacando-se sua eficiência nas análises de imagens de tórax, angiografia e, sobretudo mamografia [DOI, K. '99] [JOO, S. '04] [ANDRE, T. C. S. S. '02], onde é de suma importância a detecção precoce do câncer de mama, bem como a minimização dos casos falso-negativos e falso-positivos, com o intuito de se reduzir a quantidade de biópsias desnecessárias [BAKER, J. A. '96], já que apenas 10% a 20% dos casos enviados para a biópsia são comprovadamente malignos.

2.3. Classificação de Nódulos de Mama

O BI-RADS é uma ferramenta criada por membros e comitês do ACR (*American College of Radiology*) [ACR, '08], com o objetivo de padronizar a terminologia lingüística utilizada para a classificação de achados em mamografias, facilitando o monitoramento dos resultados [OBENAUER, S. '05].

O tutorial BI-RADS¹ está dividido em seções. A primeira apresenta um léxico da terminologia médica da área em questão, que inclui descrições de “massas”, distorções arquiteturais, calcificações, entre outros, juntamente com um conjunto de variáveis lingüísticas utilizadas para caracterizar cada item.

A segunda seção apresenta uma padronização para o processo de elaboração do diagnóstico do exame, no qual são recomendadas, primeiramente, a descrição da composição do tecido mamário (em mamografias mais densas torna-se mais difícil a detecção de pequenos nódulos); a descrição dos achados: se estes são massas ou calcificações, se possuem ou não distorções arquiteturais e a descrição de casos especiais e achados associados; por fim, há uma descrição das categorias de avaliação e a ação resultante para cada uma.

A terceira seção descreve a etapa de acompanhamento. Dados são coletados e utilizados pelo radiologista para uma avaliação estatística das mamografias. Esta seção também apresenta um glossário de termos estatísticos. A última seção provê informação técnica e instruções para a utilização da base de dados NMD (*National Mammography Database*), cujo objetivo é fornecer um meio de consulta e pesquisa para radiologistas e assegurar a qualidade do sistema BI-RADS.

¹ www.birads.at

2.3.1. Descritores de Massas

Nesta seção serão detalhados os descritores de massas, presentes no BI-RADS, utilizados neste trabalho. “Massas” são lesões observadas em duas projeções diferentes. Se vistas em apenas uma projeção, são denominadas “densidades” [ACR, ‘98]. Uma massa pode ser descrita em função dos seguintes descritores: forma, contorno (margem) e densidade.

A forma poderá ser classificada como: redonda, oval, lobular ou irregular. O contorno pode assumir as seguintes características: circunscrito (contorno bem definido), microlobulado, obscuro, indistinto e espiculado (contorno mal definido). Por fim, a densidade poderá 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 gradual entre a caracterização de uma lesão como benigna ou maligna. A Figura 2.1 apresenta alguns exemplos de forma, contorno e densidade em lesões de mama.

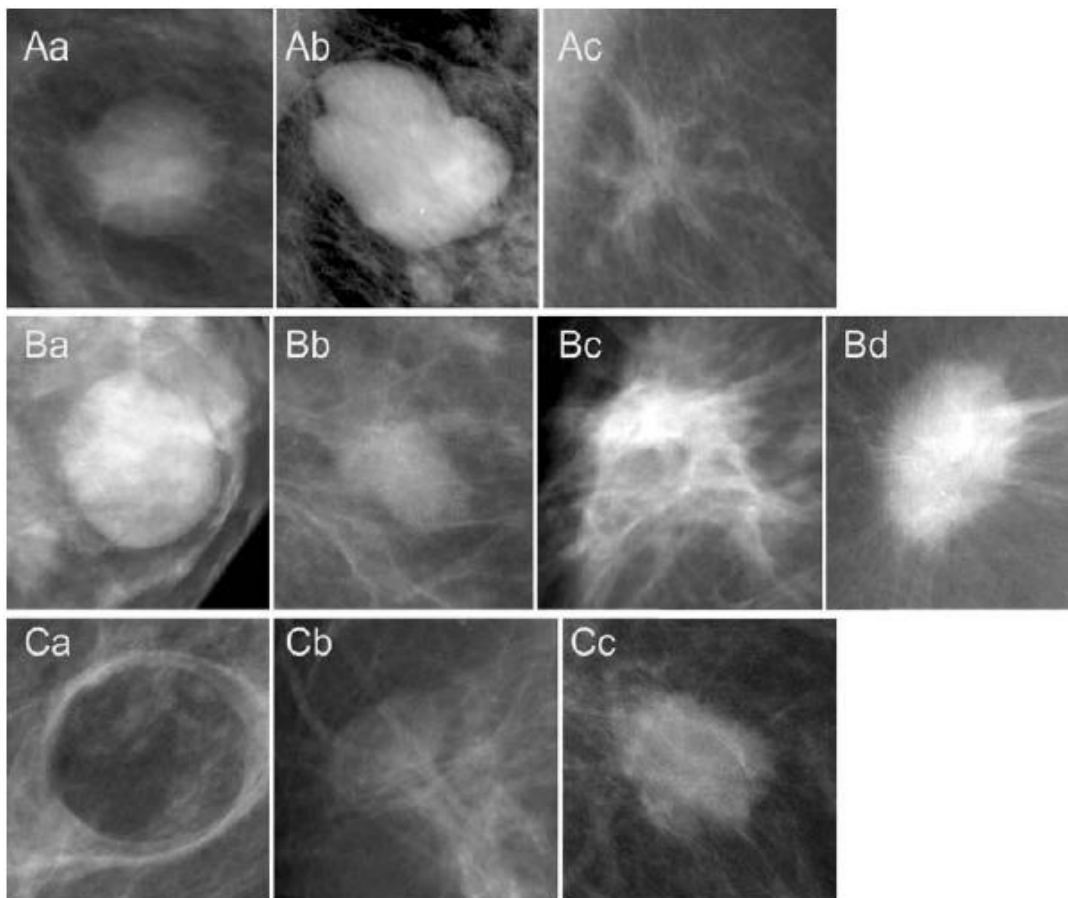


Figura 2.1. Forma: *Aa* redonda, *Ab* lobular, *Ac* irregular. Contorno: *Ba* bem definido, *Bb* mal definido, *Bc* obscurecido, *Bd* espiculado. Densidade: *Ca* contém gordura, *Cb* isodensa, *Cc* alta. [Obenauer, S. ‘05].

2.3.2. Categorias de Avaliação

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, é 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 malignidade (menor que 2%) está associada a esta categoria [OBENAUER, S. '05]. Uma nova mamografia é recomendada após seis meses. Durante a etapa de acompanhamento, não são esperadas mudanças na avaliação de uma lesão classificada como categoria 3.

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 2.2 apresenta exemplos de lesões classificadas como BI-RADS 0, 2, 3, 4 e 5.

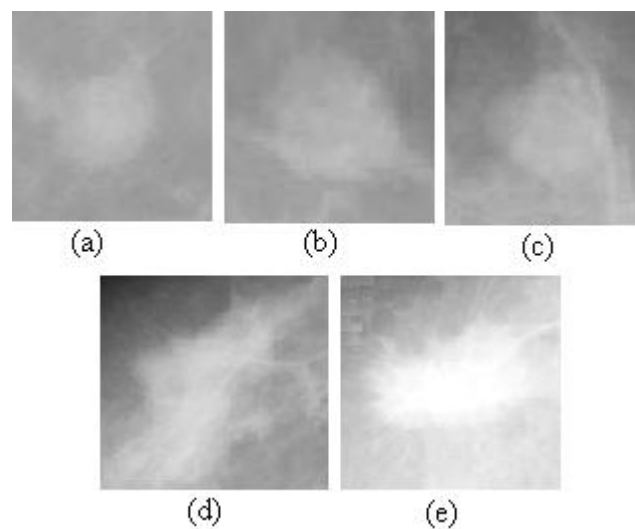


Figura 2.2. Exemplo de lesões classificadas como (a) BI-RADS 0, (b) BI-RADS 2, (c) BI-RADS 3, (d) BI-RADS 4 e (e) BI-RADS 5. [IARL, '08].

2.4. *Lógica Nebulosa*

O conceito de incerteza, para a Ciência, vem mudando ao longo dos anos. De acordo com a visão tradicional, a incerteza está relacionada aos conceitos de imprecisão e inconsistência. Não obstante, em uma visão moderna, a incerteza é considerada essencial e de grande utilidade para a Ciência [KLIR, G. J. '95]. A teoria de lógica nebulosa fornece mecanismos para a representação matemática de informações vagas e imprecisas, obtidas através de variáveis lingüísticas, simulando o raciocínio humano no processo de tomada de decisão e englobando o conceito de “incerteza” intrínseco a este.

A abordagem nebulosa facilita o tratamento de problemas complexos que não possuem uma fronteira bem definida entre conjuntos, pois incorpora tanto conhecimento objetivo como conhecimento subjetivo, obtidos através de dados numéricos e variáveis lingüísticas, respectivamente. Essas variáveis lingüísticas fazem parte do mecanismo *fuzzy* para manipulação de informações imprecisas.

Segundo Zadeh [ZADEH, L. A. '77], a lógica nebulosa permite 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.

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 [AMENDOLA, M. '04], incluindo o controle de sofisticados processos industriais. Na área médica, essa abordagem é bastante adequada ao processo de determinação do diagnóstico, já que este está sujeito à subjetividade do médico na caracterização de doenças.

2.4.1. **Conjuntos Nebulosos**

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 [SANDRI, S. '99]. Criada por Lofti Zadeh, professor da Universidade de Berkeley, em 1965, essa teoria 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 [ZADEH, L. A. '75].

Seja A um conjunto nebuloso e “ x ” um elemento a ser caracterizado, a proposição “ x

pertence à 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 2.3). Os graus de pertinência “zero” e “um” indicam, respectivamente a não pertinência e a total pertinência ao conjunto em questão.

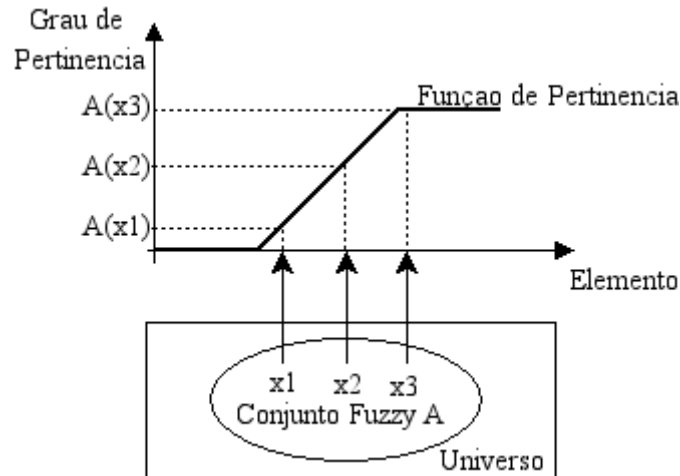


Figura 2.3. Determinação do grau de pertinência por meio da função de pertinência.

É a função de pertinência que caracteriza um conjunto nebuloso [KLIR, G. J. '95]. 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”.

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 2.4 é possível visualizar algumas dessas funções.

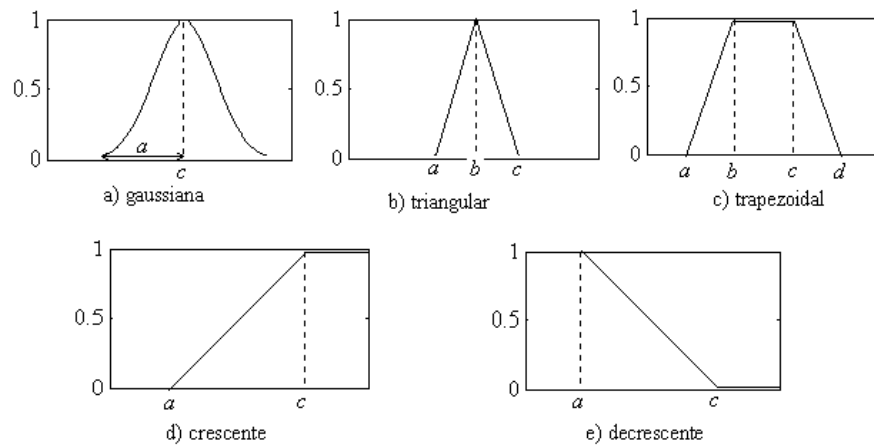


Figura 2.4. Exemplos de funções de pertinência.

2.4.2. Operações com Conjuntos Nebulosos

Conjuntos nebulosos também admitem as operações definidas sobre conjuntos clássicos. Porém, operações como complemento, união e intersecção, que na teoria clássica são únicas, em conjuntos nebulosos podem ser implementadas por uma classe de operações.

Sejam os conjuntos A e B dois conjuntos nebulosos, x um elemento qualquer e $A(x)$ e $B(x)$ os graus de pertinência do elemento x ao conjunto A e ao conjunto B, respectivamente, podemos ter algumas das seguintes operações, descritas na Tabela 1:

Tabela 1. Algumas operações com conjuntos nebulosos.

Complementos	Padrão: $\bar{A}(x) = 1 - A(x)$
	Yager: $\bar{A}_w(x) = (1 - x^w)^{1/w}$, onde w define um complemento nebuloso involutivo
Intersecções	Padrão: $(A \cap B)(x) = \min[A(x), B(x)]$
	Produto Algébrico: $(A * B)(x) = A(x) * B(x)$
	Produto Limitado: $(A \otimes B)(x) = \max[0, A(x) + B(x) - 1]$
Unões	Padrão: $(A \cup B)(x) = \max[A(x), B(x)]$
	Soma Algébrica: $(A + B)(x) = A(x) + B(x) - A(x) * B(x)$
	Soma Limitada: $(A \oplus B)(x) = \min[1, A(x) + B(x)]$

2.4.3. Relações Nebulosas

As relações nebulosas determinam o grau de associação entre dois ou mais conjuntos nebulosos. Elas representam associações entre elementos pertencentes a uma mesma relação. Supondo U e V dois conjuntos nebulosos, uma relação nebulosa $R(U, V)$ resultará no

subconjunto nebuloso $U \times V$, que também é caracterizado por uma função de pertinência $\mu_R(x,y)$ [MENDEL, J.M. '95]. Um valor no intervalo $[0,1]$ é atribuído a essa associação, indicando a intensidade com que os elementos se relacionam. Normalmente, essa associação é representada através de matrizes.

Os conceitos básicos e operações, anteriormente citados, com conjuntos nebulosos também são aplicados às relações nebulosas. Porém, algumas operações com relações nebulosas binárias não são aplicáveis a conjuntos nebulosos ordinários.

2.4.4. Máquina de 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. Opcionalmente, a entrada do sistema poderá conter também apenas valores lingüísticos. A partir desses valores é criada uma base de regras composta por uma coleção de proposições nebulosas. Essas regras são descritas em variáveis lingüísticas e poderão ser construídas com a ajuda de um especialista da área de interesse. Posteriormente, a base de regras será utilizada para fornecer a saída nebulosa do sistema para cada entrada nebulosa a partir das relações já definidas. Por fim, os valores de saída são “denebulizados”, pelo processador de saída, tornando-se valores precisos representativos do grau de pertinência do elemento ao conjunto [JAFELICE, R. S. M. '05]. Isso pode ser verificado na Figura 2.5.

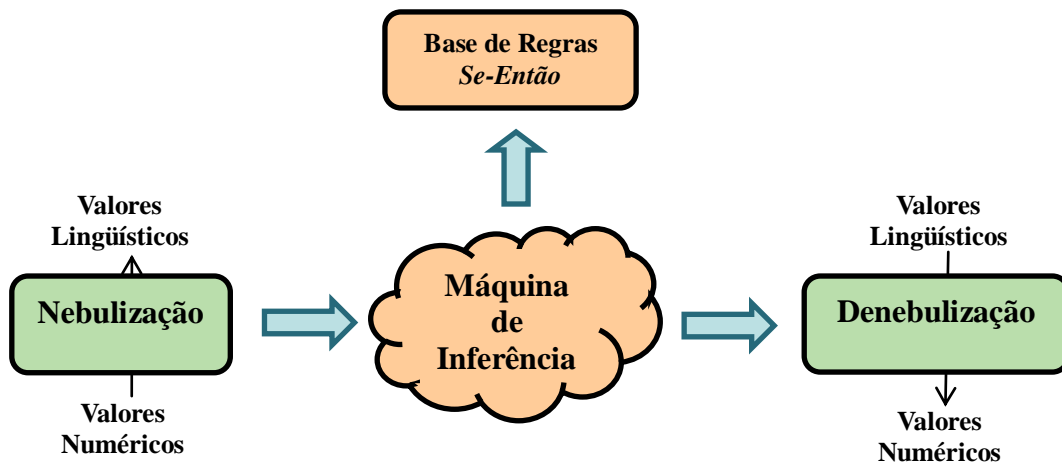


Figura 2.5. Máquina de Inferência Nebulosa.

Uma base de regras nebulosas possui proposições cujos antecedentes, conseqüentes ou ambos representam conjuntos nebulosos. Geralmente, as regras encontram-se no formato “se – então”, 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 [DUBOIS, D. '96].

2.4.5. Modelos de Inferência

Modelos de inferência nebulosos possuem algumas propriedades já definidas como os indicadores de disparo das regras e os operadores utilizados sobre os conjuntos nebulosos. O modelo utilizado depende da forma de armazenamento das informações escolhidas [ALMEIDA, P. E. M. '03]. Nas próximas seções serão apresentados os modelos de Mamdani e TSK.

2.4.5.1. Modelo de Mamdani

O modelo de Mamdani, proposto na década de 70, foi o modelo 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 conseqüentes das regras possuem relações nebulosas, como vemos a seguir:

- Se x é A_1 e y é B_1 então z é C_1
- Se x é A_2 e y é B_2 então z é C_2
- Se x é A_i e y é B_i então z é C_i

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 [ALMEIDA, P. E. M. '03]. 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 de saída:

$$\mu(C_i) = \max_{k=1..n} [\min(\mu A_k, \mu B_k)] \quad (1)$$

Onde, μA e μB representam os graus de pertinência à cada subconjunto nebuloso de entrada. 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 $\mu(C_i)$ 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. 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.

2.4.5.2. Modelo de Takagi-Sugeno-Kang

Este modelo é conhecido como modelo de inferência *fuzzy* paramétrico ou TSK e foi proposto na década de 80. O modelo também se baseia na utilização de uma base de regras como estrutura de controle nebuloso, porém a parte consequente de cada regra é determinada por uma função estritamente monotônica, formada por uma combinação linear das entradas [SANDRI, S. '99].

- Se x é A_1 e y é B_1 então $z_1 = f_1(x,y)$
- Se x é A_2 e y é B_2 então $z_2 = f_2(x,y)$
- Se x é A_i e y é B_i então $z_i = f_i(x,y)$

Um único valor z_i é obtido para cada regra, através de uma interpolação dos valores de entrada, representando a ação de controle da regra. Por fim, as regras são agregadas em uma única ação de controle (z), determinada pela média ponderada entre os valores de z_i , onde D_k é o coeficiente de ativação de cada regra.

$$z = \frac{\sum_{i=1}^k D_i \cdot f_i(x, y)}{\sum_{i=1}^k D_i} \quad (2)$$

Diferentemente do modelo de Mamdani, controladores baseados no modelo TSK não necessitam de uma interface de denebulização, pois a ação de controle resultante já é um valor não nebuloso.

CAPÍTULO 3

Metodologia

3.1. Considerações Iniciais

No presente capítulo será descrita a metodologia utilizada neste trabalho. A modelagem de um sistema nebuloso inclui a determinação dos parâmetros a serem utilizados como entrada e saída do sistema, os conjuntos e subconjuntos nebulosos, a criação da base de regras e o modelo de inferência escolhido. Especialistas da área de interesse, geralmente, auxiliam as etapas de criação da base de regras e geração dos conjuntos nebulosos, porém, nestas etapas, também é possível a utilização de métodos matemáticos e algoritmos computacionais.

Foi utilizada neste trabalho a base de imagens de mamografia digital DDSM (*Digital Database for Screening Mammography*) [IARL, '08] construída por pesquisadores do Hospital Geral de Massachusets e da Universidade do Sul da Flórida, com lesões recortadas por pesquisadores da Universidade de Chicago. Para este trabalho foram utilizadas 250 imagens compostas por regiões de interesse contendo lesões (nódulos) benignas e malignas, diagnosticadas por radiologistas e por exames anátomo-patológicos.

3.2. Representação Nebulosa da Terminologia Lingüística

A modelagem dos atributos de entrada do sistema proposto foi dividida em duas etapas. A primeira consiste na determinação de parâmetros numéricos de avaliação dos nódulos. A segunda etapa consiste na determinação das variáveis de entrada do sistema a partir dos parâmetros obtidos na primeira etapa.

3.2.1. Primeira Etapa: 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 auxiliaram a escolha da metodologia a ser utilizada no projeto, otimizando a modelagem do sistema e a criação dos

conjuntos nebulosos.

Como descrito na seção 2.3, os principais atributos para classificação de nódulos de mama utilizados neste trabalho e padronizados pelo BI-RADS são: forma, contorno e densidade. Na tabela abaixo é possível visualizar as características relativas a cada um desses atributos.

Tabela 2. Atributos utilizados na classificação de nódulos de mama.

Forma	Contorno	Densidade
<i>redonda</i>	<i>circunscrito</i>	<i>alta</i>
<i>oval</i>	<i>microlobulado</i>	<i>isodensa</i>
<i>lobular</i>	<i>obscuro</i>	<i>baixa</i>
<i>irregular</i>	<i>indistinto</i>	<i>contém gordura</i>
	<i>espiculado</i>	

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 a cinco radiologistas, na forma de um questionário, como apresentado no Apêndice A, contendo campos para a avaliação da forma, do contorno e da densidade do nódulo. Foram selecionadas 35 imagens, da base já descrita, a fim de serem analisadas 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 atua 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 quanto ao formato: “formato lobular com grau de malignidade 0,65” ou “formato irregular com grau de malignidade 0,9”. 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, 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 descritas para cada característica escolhida associada a um grau de malignidade, por exemplo, poderíamos obter a seguinte avaliação: “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 pessoal do radiologista. Por fim, a categoria de avaliação foi determinada para cada imagem analisada. As categorias BI-RADS atribuídas foram 0, 2, 3, 4 e 5.

3.2.1.1. Base de dados de avaliações

Após a fase de avaliação das lesões pelos radiologistas, um banco de dados foi modelado e implementado para armazenar as informações obtidas e facilitar a extração de parâmetros. As informações armazenadas representam o conhecimento dos especialistas do domínio de interesse, importantes em um sistema nebuloso, e representam a base para a criação dos conjuntos nebulosos e da base de proposições. Na Figura 3.1 é apresentado o Diagrama Entidade-Relacionamento do banco de dados das avaliações.

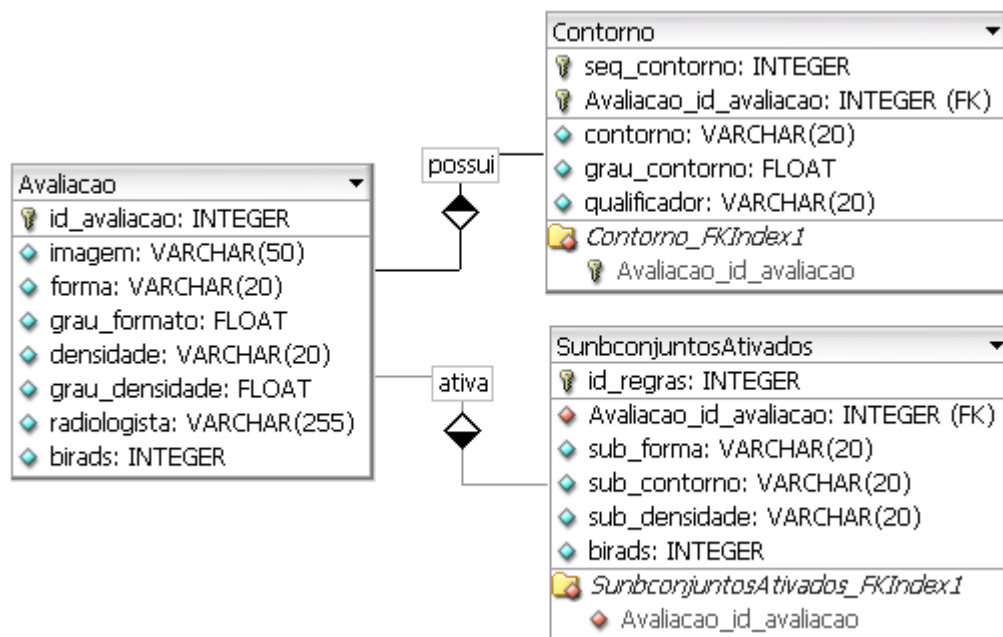


Figura 3.1. Diagrama Entidade - Relacionamento da Base de Dados de avaliações.

A tabela “Avaliação” armazena, para cada imagem, o formato, a densidade e seus respectivos graus de malignidade, assim como a categoria BI-RADS e o radiologista que realizou a avaliação. Devido à possibilidade de se obter mais de uma característica para o contorno, sendo um atributo multivalorado, foi criada a tabela “Contorno”, para armazenar o

qualificador (*totalmente, muito, parcialmente* ou *pouco*) e o coeficiente para cada característica. Inicialmente, com as avaliações obtidas na primeira etapa, foram preenchidas as tabelas “Avaliação” e “Contorno”. A terceira tabela, “SubconjuntosAtivados”, será utilizada na etapa de criação da base de proposições nebulosas.

3.2.2. Segunda Etapa: determinação dos parâmetros de entrada

A partir dos valores numéricos representativos do grau de malignidade associado às lesões, 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, refletindo a influência do atributo.

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 escolhida, 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:

$$coef_contorno = \frac{\sum_{k=1}^n R_{i,j} \cdot P_j}{\sum_{k=1}^n P_j} \quad (3)$$

Onde, n representa o número de combinações variável / característica escolhidas pelo radiologista e P_j , representa o peso atribuído à característica j. A Tabela 3 apresenta os pesos obtidos (P_j) para cada característica do contorno e $R_{i,j}$ o grau de malignidade da combinação variável / característica.

Tabela 3. Pesos obtidos para as características do atributo contorno.

<i>Característica</i>	<i>Peso Característica</i>	<i>totalmente</i>	<i>muito</i>	<i>parcialmente</i>	<i>pouco</i>
<i>circunscrito</i>	0.353	0.243	0.350	0.466	0.500
<i>microlobulado</i>	0.696	0.821	0.750	0.450	0.533
<i>obscurecido</i>	0.572	0.500	0.400	0.585	0.525
<i>indistinto</i>	0.672	0.800	0.637	0.648	0.600
<i>espiculado</i>	0.830	0.965	0.816	0.760	0.766

Com base nos coeficientes obtidos, o seguinte modelo nebuloso foi proposto para obter a categoria BI-RADS:

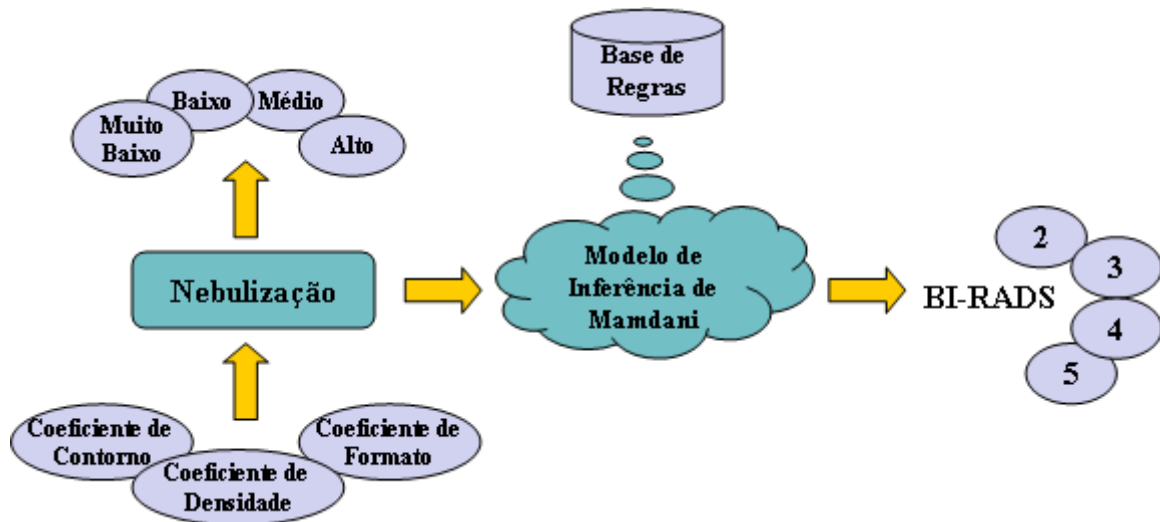


Figura 3.2. Modelagem do Sistema Nebuloso.

Os coeficientes 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. O método para obtenção das funções de pertinência dos subconjuntos nebulosos representativos desses conceitos será detalhado na seção 3.3. O modelo de inferência escolhido foi o método de *Mamdani*. A base de proposições nebulosas será formada pela combinação das variáveis correspondentes aos três coeficientes. 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 na seção 2.3, ela é atribuída quando há necessidade de exame adicional, como, por exemplo, a ultrasonografia. Após realizado o exame, qualquer uma das outras categorias poderá ser atribuída à lesão.

3.3. Determinação das Funções de Pertinência

Em sistemas nebulosos, a etapa de criação dos conjuntos, assim como a criação da base de proposições nebulosas, são auxiliadas pelo especialista do domínio de aplicação. No caso em estudo, o conhecimento do especialista foi representado a partir da análise direta da base de dados de avaliações, da qual foram obtidos os conjuntos nebulosos e as funções de pertinência. Para isso, foram utilizados os algoritmos *Omega*, para discretização de atributos e *Fuzzy Omega* para criação das funções de pertinência dos subconjuntos.

3.3.1. Omega: discretização de dados e seleção de atributos

O algoritmo Omega [RIBEIRO, M. X. '08], desenvolvido pelo Grupo de Bases de Dados e Imagens (ICMC - USP) promove simultaneamente a discretização de conjuntos de dados com valores contínuos e seleção de atributos. Ele possui complexidade linear em relação ao número de instâncias a serem processadas. A discretização de dados pelo Omega gera intervalos de valores caracterizados por uma classe majoritária, facilitando etapas posteriores de processamento dos dados.

A entrada do algoritmo é representada pelo conjunto de N instâncias associadas a uma classe $I = (f_i, c_i)$. O Omega processa cada atributo separadamente. As instâncias deverão estar ordenadas pelo valor do atributo em análise. Os limites dos intervalos (bins) representam os pontos de corte do conjunto de dados. Inicialmente, toda vez que ocorre uma mudança na classe associada à instância, um ponto de corte é criado. Neste passo, são produzidos bins puros, com entropia zero, ou seja, os intervalos gerados possuem somente instâncias de uma mesma classe. A Figura 3.3 apresenta elementos associados a três classes diferentes (A, B, C) e os pontos de corte dos intervalos gerados no primeiro passo do Omega (Figura 8).

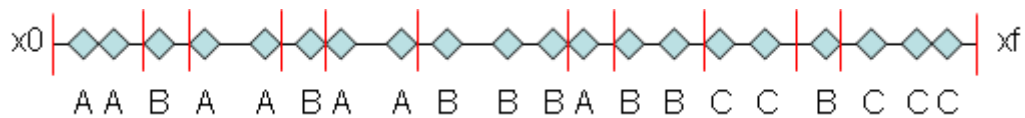


Figura 3.3. Passo 1 do algoritmo Omega.

O próximo passo do Omega restringe a quantidade de elementos por intervalo a uma frequência mínima (H_{Min}), evitando a utilização de um grande número de pontos de corte. O Omega remove os segundos pontos de corte dos intervalos que não satisfazem H_{Min} (Figura 3.4). Quanto maior o valor de H_{Min} , menor o número de intervalos obtidos no segundo passo. Porém, um valor muito alto de H_{Min} poderá introduzir inconsistências no processo de discretização dos dados.

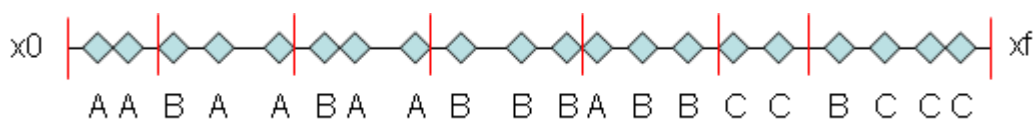


Figura 3.4. Passo 2 do algoritmo Omega.

No passo 3, o Omega agrega intervalos consecutivos utilizando uma medida de inconsistência para determinar quais intervalos devem ser agregados. O cálculo da taxa de

inconsistência para cada intervalo (ζ_{T_i}) é dado por:

$$\zeta_{T_i} = \frac{|T_i| - |M_{T_i}|}{|T_i|} \quad (4)$$

Onde $|T_i|$ é o número de instâncias do intervalo T_i e $|M_{T_i}|$ é o número de instâncias da classe majoritária de T_i . O Omega agrega intervalos consecutivos que possuam a mesma classe majoritária e suas taxas de inconsistência sejam menores ou iguais a um valor limite ζ_{Max} (Figura 3.5).

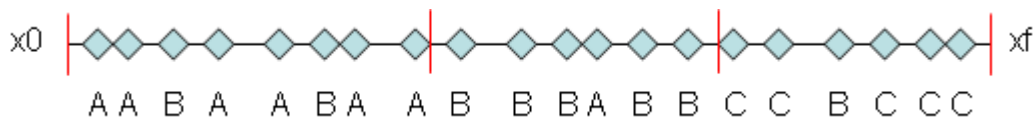


Figura 3.5. Passo 3 do algoritmo Omega.

Por fim, o último passo consiste na seleção de atributos. Seja T o conjunto de intervalos, no qual um atributo foi discretizado, o Omega calcula uma taxa de inconsistência global de acordo com a equação 5.

$$\zeta_G = \frac{\sum_{T_i \in T} |T_i| - |M_{T_i}|}{\sum_{T_i \in T} |T_i|} \quad (5)$$

Os atributos cujas taxas de inconsistência forem menores do que um valor limite ζ_{Gmax} são selecionados.

3.3.2. Fuzzy Omega: determinação das funções de pertinência

O algoritmo Fuzzy Omega é um método para geração de funções de pertinência baseado em uma análise estatística do conjunto de dados. Ele utiliza taxas de inconsistência, como as descritas em [RIBEIRO, M. X. '08], para mapear a distribuição das diferentes classes no conjunto de dados. Funções matemáticas podem ser geradas a partir de ajustes polinomiais. Essa abordagem auxilia a etapa de modelagem das funções, fornecendo uma análise menos subjetiva e gerando resultados mais próximos dos valores reais.

Para gerar automaticamente as funções de pertinência relacionadas a um conjunto nebuloso específico, o algoritmo Fuzzy-Omega realiza 3 passos (Quadro 1):

Quadro 1. Algoritmo Fuzzy – Omega.

O método Fuzzy-Omega

Entrada. Conjunto de N instâncias (I): $I = (A_i, C_i)$.

Saída. Taxas de inconsistência pontuais para cada instância I , com base na sua vizinhança.

Passo 1. Separar os valores do atributo desejado junto com a respectiva classe de cada instância e ordená-los segundo o valor de A_i ;

Passo 2. Aplicar o algoritmo Omega para determinar intervalos a partir do conjunto de dados e sua respectiva classe majoritária (C_M). Cada intervalo, de cada atributo, terá uma função de pertinência relacionada.

Passo 3. Para determinar a função de pertinência de cada intervalo gerado:

- Selecionar a classe majoritária M_{T_i} do intervalo;
 - Visitar cada valor do intervalo e , para cada um, criar um subintervalo consistindo em uma vizinhança de n valores, centrada no valor visitado;
 - Para cada subintervalo, calcular a taxa de inconsistência (ζ_{T_i}), considerando M_{T_i} como classe majoritária, de acordo com a equação (4);
 - Calcular o grau de pertinência $\mu(T_i)$ para cada valor visitado:
 $\mu(T_i) = 1 - \zeta_{T_i}$;
 - Mantendo a classe majoritária M_{T_i} , repetir os 3 últimos sub passos para o intervalo anterior e para o intervalo posterior (se eles existirem) em relação ao intervalo em análise.
 - Agrupar os pares (valor, grau de pertinência) obtidos e normalizar os valores dos graus de pertinência no intervalo $[0,1]$. Calcular uma função polinomial que forneça um ajuste ótimo ao conjunto de pares. Esta será a função de pertinência do intervalo.
-

A entrada do algoritmo Fuzzy-Omega consiste em um conjunto de N instâncias na forma $I(A_i, C_i)$, na qual, A_i é o valor do atributo para a instância e C_i é sua classe associada. No passo 1, esses valores deverão ser ordenados de acordo com os valores de A_i . O passo 2 consiste na aplicação do algoritmo Omega para obtenção dos intervalos resultantes do processo de discretização de dados realizado pelo algoritmo.

No passo 3, considerando os intervalos gerados na Figura 3.5, para o 1° intervalo a ser processado, serão considerados os elementos pertencentes ao 1° e 2° intervalos. Para o 2° intervalo a ser processado, serão considerados os elementos do intervalo anterior e posterior ao mesmo, assim como os elementos do próprio intervalo, ou seja, serão utilizados o 1°, o 2° e o 3° intervalos. Finalmente, para o 3° intervalo, serão considerados o 2° e o 3° intervalos.

Na Figura 3.6, podem ser visualizados exemplos de subintervalos considerados no cálculo da taxa de inconsistência por elemento. Foi adotada uma vizinhança $n = 5$. Supondo a função de pertinência a ser gerada para o 2° intervalo, a classe majoritária (C_M) a ser

considerada no cálculo de x_i é a classe B.

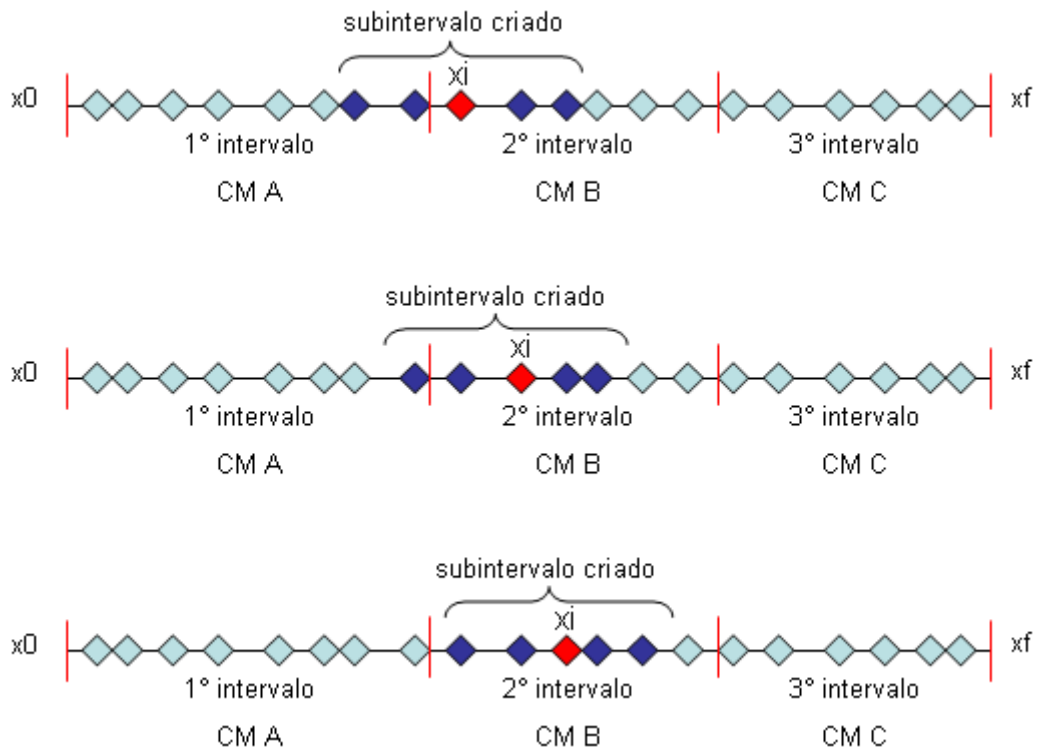


Figura 3.6. Criação de subintervalos para o cálculo da taxa de inconsistência por elemento.

As taxas de inconsistência calculadas são subtraídas de 1. Os pares (valor, grau de pertinência) obtidos para cada intervalo são plotados em um gráfico e um ajuste polinomial é realizado, encontrando-se uma curva ótima para a distribuição dos dados. Um exemplo de função gerada pelo algoritmo pode ser visualizada na Figura 3.7.

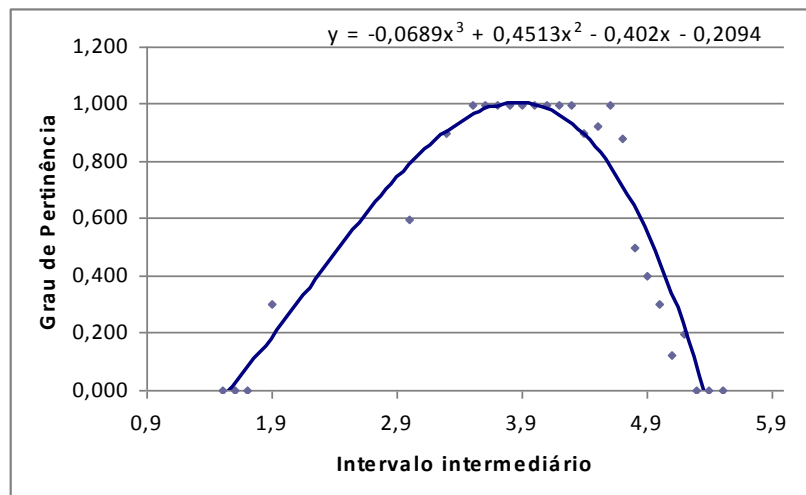


Figura 3.7. Exemplo de intervalo gerado pelo algoritmo Fuzzy-Omega.

Um estudo com a base de dados “iris”, disponível no repositório do UCI², e alguns descritores de Haralick das imagens da base de dados [IARL, '08] foi realizado a fim de se avaliar a efetividade do algoritmo Fuzzy-Omega³.

3.4. Base de Proposições Nebulosas

A base de regras em um sistema nebuloso, normalmente, é determinada pela combinação dos subconjuntos de entrada do sistema. Considerando um sistema com 3 atributos de entrada, cada um dividido em 3 subconjuntos nebulosos, a base de proposições possuirá 27 regras. A etapa de criação da base de regras também é caracterizada pelo auxílio de especialistas do domínio de aplicação na determinação da parte conseqüente da regra.

No sistema proposto neste trabalho, 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 na etapa anterior.

Para cada avaliação, foram determinadas as combinações dos subconjuntos ativados, a partir dos graus definidos pelos radiologistas durante as entrevistas. As combinações foram armazenadas na tabela “SubconjuntosAtivados” da base de dados de avaliações (Figura 3.1). 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.

² <http://archive.ics.uci.edu/ml/machine-learning-databases/iris/>

³ Este método é descrito em um artigo que se encontra em fase de submissão

CAPÍTULO 4

Resultados

4.1. Considerações Iniciais

Neste capítulo estão os resultados obtidos neste trabalho. São apresentadas as funções de pertinência para os coeficientes de entrada do sistema, a base de proposições nebulosas e a implementação de um protótipo de sistema de apoio ao diagnóstico que utiliza a lógica nebulosa para sugerir ao usuário a classificação BI-RADS a partir de parâmetros de entrada.

4.2. Funções de Pertinência Obtidas

A partir da base de dados de avaliações descrita na seção 3.2.1, foram selecionados os graus de malignidade associados ao formato do nódulo, à densidade e ao contorno, já cadastrados na base. Foram criados arquivos de entrada, para cada atributo, para serem processados pelo algoritmo Fuzzy-Omega. No Apêndice B é possível visualizar os valores dos atributos já ordenados e o respectivo BI-RADS associado ao valor. Cada atributo é processado separadamente pelo Fuzzy-Omega. Foram obtidos 126 valores para os graus de malignidade associados ao formato e à densidade e 180 valores para o grau de malignidade associado ao contorno devido ao fato do radiologista poder escolher mais de uma característica para o mesmo durante a análise da lesão. Avaliações que obtiveram classificação BI-RADS 0, não foram consideradas na geração das funções de pertinência.

Primeiramente, aplicando-se o Fuzzy-Omega, foram obtidas divisões em 2 e 3 intervalos no passo 2, o que levou a criação de 2 a 3 funções de pertinência por atributo no passo 3. Porém, testes realizados indicaram que a utilização de apenas duas funções de pertinência representativas dos coeficientes de formato e densidade não foram suficientes para gerar regras que distinguíssem as categorias BI-RADS 2 e 5, apesar da diferenciação entre as categorias 3 e 4. Quanto à utilização de três intervalos, não foram obtidas funções de pertinência características de conjuntos nebulosos ordinários (com grau pertinência entre 0 e 1) para todos os intervalos (Figura 4.1), o que também dificultou a utilização das mesmas. Essas funções não foram utilizadas no sistema nebuloso proposto.

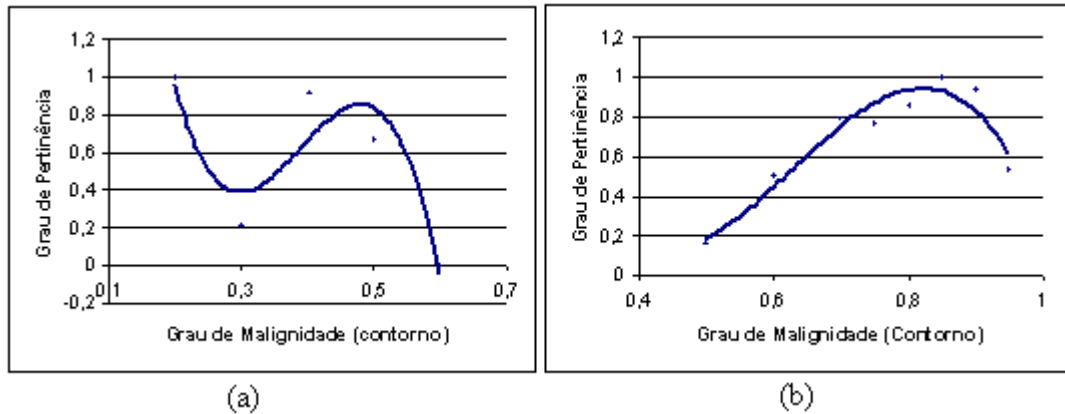


Figura 4.1. Funções de Pertinência para o atributo coeficiente de contorno. (a) Função de pertinência com ajuste polinomial não ótimo. (b) Função de pertinência não ordinária.

Para evitar que fossem geradas regras que não representassem as categorias 2 ou 5 e obter ajustes polinomiais para 3 ou mais intervalos, os valores de cada atributo de entrada foram considerados como um único intervalo. E as ações descritas no passo 3 do algoritmo Fuzzy-Omega (Seção 3.3.2) foram executadas adotando-se cada uma das classes BI-RADS como classe majoritária, ou seja, a taxa de inconsistência pontual para cada elemento foi calculada em relação a classe 2 para cada valor do atributo presente no conjunto de entrada, posteriormente, em relação às classes 3, 4 e 5. Deste modo, foi possível obter até 4 funções de pertinência para cada atributo. As variáveis lingüísticas utilizadas no processo de nebulização foram “muito baixo”, “baixo”, “médio” e “alto” conceituando os valores dos coeficientes.

A Figura 4.2 apresenta as funções de pertinência para o coeficiente de formato. A curva “coeficiente muito baixo” (alaranjada) é um ajuste polinomial de ordem 3, no qual foi adotada como classe majoritária do conjunto de dados a categoria 2. Devido à pequena quantidade de exemplos BI-RADS 2 obtidos, a curva possui um intervalo de variação menor do que as outras e reflete a baixa malignidade associada às lesões classificadas como BI-RADS 2.

A curva “coeficiente baixo” (azul) é um ajuste polinomial de ordem 4. Para esta curva foi adotada a categoria BI-RADS 3 como classe majoritária. A curva “coeficiente médio” (verde) é um ajuste polinomial de ordem 3 considerando-se o BI-RADS 4 como classe majoritária. As classes 3 e 4 são as mais frequentes no conjunto de dados e a variabilidade dos graus de malignidade associados à essas classes é alta. Porém, verificam-se regiões de abrangência diferentes para as duas curvas. A curva “coeficiente baixo” possui graus de pertinência maiores que 50% para graus de malignidade menores do que 0.4, a partir desse valor, verifica-se maior pertinência à curva “coeficiente médio”, evidenciando maiores graus de malignidade associados à categoria 4 na análise do formato do nódulo.

Por fim, a curva “coeficiente alto” (vermelha) é um polinômio de ordem 2 considerando-se o BI-RADS 5 como classe majoritária. Similar à curva “muito baixo”, o conjunto de dados possui poucos exemplos pertencentes à classe 5, também possuindo pequeno intervalo de variação dos dados. A curva 5 relaciona graus de malignidade altos, associados ao formato da lesão, à categoria BI-RADS 5.

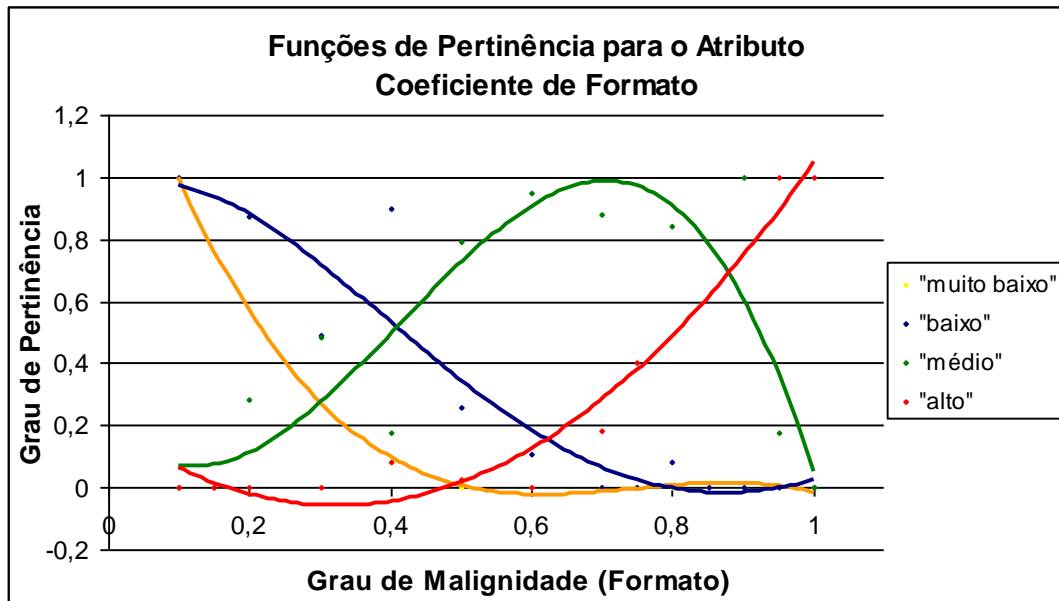


Figura 4.2. Funções de pertinência para o atributo Coeficiente de Formato.

As Figuras 4.3 e 4.4 apresentam as funções de pertinência para os coeficientes de contorno e densidade. Similar às funções do coeficiente de formato, verifica-se que quanto maior o grau de malignidade, maior pertinência às curvas “médio” e “alto”. O padrão de abrangência das curvas foi semelhante ao obtido para o coeficiente de formato. O coeficiente de contorno e de densidade também foram mapeados nos subconjuntos nebulosos “muito baixo”, “baixo”, “médio” e “alto”.

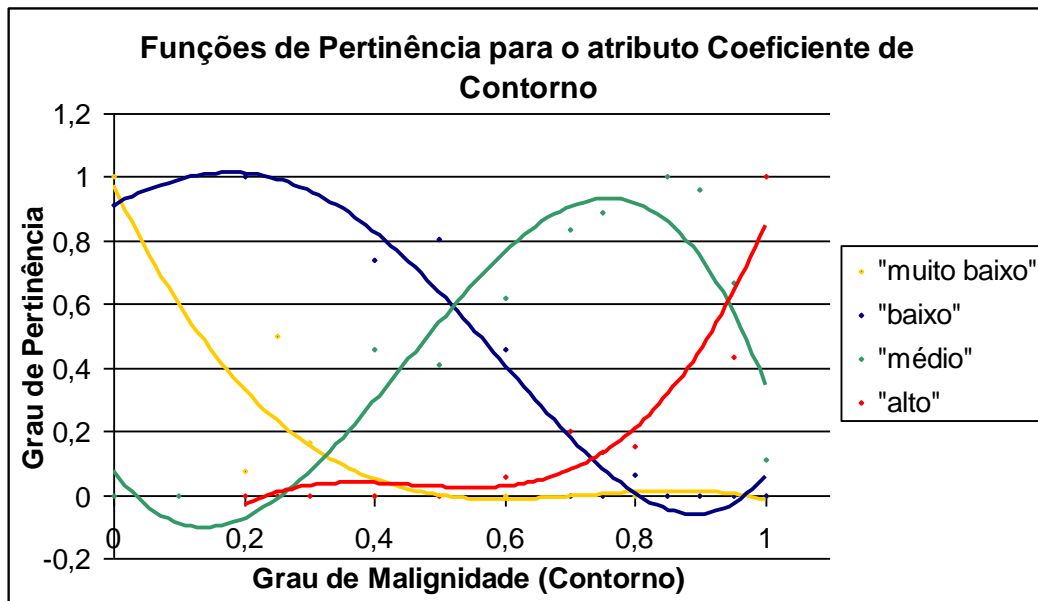


Figura 4.3. Funções de pertinência para o atributo Coeficiente de Contorno.

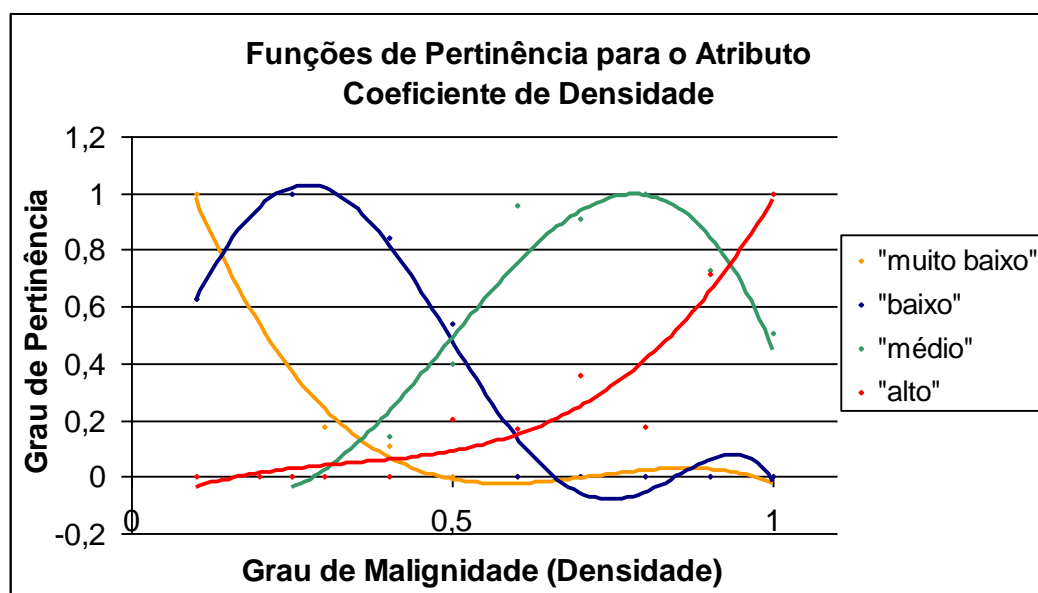


Figura 4.4. Funções de pertinência para o atributo Coeficiente de Densidade.

A base de regras trata todas as combinações possíveis dos atributos de entrada. Considerando 4 funções de pertinência para o atributo coeficiente de formato, 4 funções para o coeficiente de contorno e 4 funções para o coeficiente de densidade, teremos uma base com 64 regras, o que permite mapear combinações dos coeficientes para qualquer categoria BI-RADS.

A distribuição dos valores de entrada permite concluir que a maioria dos consequentes das regras terá a categoria 3 ou 4 associada. Durante os testes realizados, a quantidade de combinações utilizada mostrou-se redundante. Por isso decidiu-se utilizar 3 funções de pertinência para algum atributo entrada, diminuindo o esforço computacional no processo de inferência, já que durante o mesmo, todas as regras devem ser consultadas. O coeficiente de formato foi um candidato a utilizar apenas 3 funções de pertinência, já que suas funções correspondentes aos subconjuntos “muito baixo” e “baixo” têm grande área de intersecção e o máximo valor de pertinência para as funções é obtido com valores próximos de graus de malignidade. Os mesmos subconjuntos para os outros atributos de entrada geraram uma divisão ótima das categorias 2 e 3. Por isso, foram mantidas apenas as funções de pertinência características dos subconjuntos “baixo”, “médio” e “alto” para o coeficiente de formato (Figura 4.5). As equações polinomiais para cada uma das funções utilizadas neste trabalho podem ser encontradas no Apêndice C. Eventuais pontos de corte foram considerados para evitar a utilização de graus de pertinência negativos ou maiores que 1.

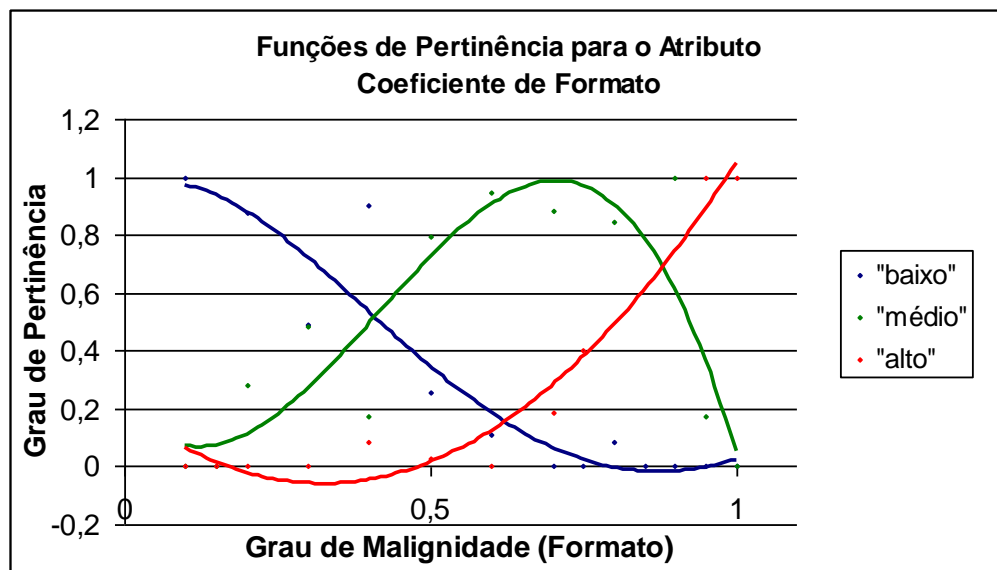


Figura 4.5. Funções finais de pertinência para o atributo Coeficiente de Formato.

Sejam os valores (0,5; 0,4; 0,6) exemplos de coeficientes de formato, contorno e densidade, com as funções definidas anteriormente, obtemos as seguintes possibilidades para este exemplo:

- Coeficiente de Formato = 0,5: 35% de pertinência ao subconjunto “baixo” e 72% de pertinência ao subconjunto “médio”;
- Coeficiente de Contorno = 0,4: 5% de pertinência ao subconjunto “muito baixo”, 83%

- de pertinência ao subconjunto “baixo” e 29% de pertinência ao subconjunto “médio”;
- Coeficiente de Densidade = 0,6: 14% de pertinência ao subconjunto “baixo”, 75% de pertinência ao subconjunto “médio” e 15% de pertinência ao subconjunto “alto”.

4.3. Base de Proposições Nebulosas

Combinando-se os três subconjuntos do coeficiente de formato, os quatro subconjuntos do coeficiente de contorno e os quatro subconjuntos do coeficiente de densidade e utilizando o procedimento descrito na seção 3.4, foram geradas 48 regras. A base de regras do sistema nebuloso proposto pode ser visualizada no Quadro 2.

Quadro 2. Base regras nebulosas.

Regra	Se Coeficiente de Formato	E Coeficiente de Contorno	E Coeficiente de Densidade	Então BI-RADS
1	alto	alto	alto	5
2	alto	alto	medio	5
3	alto	alto	baixo	4
4	alto	alto	muito baixo	4
5	alto	medio	alto	5
6	alto	medio	medio	4
7	alto	medio	baixo	4
8	alto	medio	muito baixo	4
9	alto	baixo	alto	4
10	alto	baixo	medio	4
11	alto	baixo	baixo	4
12	alto	baixo	muito baixo	3
13	alto	muito baixo	alto	4
14	alto	muito baixo	medio	4
15	alto	muito baixo	baixo	4
16	alto	muito baixo	muito baixo	4
17	medio	alto	alto	5
18	medio	alto	medio	4
19	medio	alto	baixo	4
20	medio	alto	muito baixo	4
21	medio	medio	alto	4
22	medio	medio	medio	4
23	medio	medio	baixo	4
24	medio	medio	muito baixo	3
25	medio	baixo	alto	4
26	medio	baixo	medio	4
27	medio	baixo	baixo	3
28	medio	baixo	muito baixo	3
29	medio	muito baixo	alto	4
30	medio	muito baixo	medio	3
31	medio	muito baixo	baixo	3
32	medio	muito baixo	muito baixo	3
33	baixo	alto	alto	4
34	baixo	alto	medio	4

35	baixo	alto	baixo	4
36	baixo	alto	muito baixo	4
37	baixo	medio	alto	4
38	baixo	medio	medio	4
39	baixo	medio	baixo	3
40	baixo	medio	muito baixo	3
41	baixo	baixo	alto	3
42	baixo	baixo	medio	3
43	baixo	baixo	baixo	3
44	baixo	baixo	muito baixo	3
45	baixo	muito baixo	alto	3
46	baixo	muito baixo	medio	3
47	baixo	muito baixo	baixo	2
48	baixo	muito baixo	muito baixo	2

4.4. Sistemas Nebulosos em Ambientes de Auxílio ao Diagnóstico

Considerando o contexto de auxílio ao diagnóstico de nódulos de mama, a utilização da lógica nebulosa deve fornecer uma análise quantitativa automatizada, funcionando como uma segunda opinião para auxiliar radiologistas (e também outros médicos) na tomada de decisão. Especificamente, no caso aqui estudado, foi desenvolvido um protótipo que fornece uma segunda opinião quantitativa da categoria BI-RADS associada a uma lesão de mama, utilizando o modelo nebuloso proposto neste trabalho.

4.4.1. Protótipo Implementado: Fuzzy BI-RADS

O sistema desenvolvido, denominado Fuzzy BI-RADS, foi implementado em linguagem Java⁴ [DEITEL, H. M. '06], utilizando o software *Netbeans*⁵ como ambiente integrado de desenvolvimento. A interface de usuário permite ao radiologista a escolha dos parâmetros referentes aos descritores de nódulos de mama, utilizados neste projeto, a fim de fornecer a entrada nebulosa para o sistema. Depois do processamento realizado pela máquina de inferência, o grau de pertinência às categorias de avaliação BI-RADS é retornado ao usuário através da mesma interface gráfica. Na Figura 4.6 é possível visualizar a interface implementada. A avaliação do formato e da densidade da lesão são similares. A escala representa o grau atribuído à malignidade, atuando como um peso. Para a avaliação do contorno, mais de uma característica poderá ser escolhida.

O sistema mapeia as características escolhidas na interface em subconjuntos nebulosos caracterizando a etapa de nebulização. As funções de pertinência utilizadas pelo

⁴ <http://java.sun.com/javase/downloads/?intcmp=1281>

⁵ <http://download.netbeans.org/netbeans/6.1/final/>

sistema foram descritas na seção 4.2. A partir da entrada fornecida, serão calculados valores conhecidos como coeficientes de ativação para cada regra presente na base de proposições nebulosas. O coeficiente de ativação é utilizado pelo método de *Mamdani* para a determinação dos graus de pertinência às categorias de saída do sistema. Um diagrama de classes do sistema Fuzzy BI-RADS implementado pode ser visualizado no Apêndice D.

Fuzzy BI-RADS

Consultar Diagnóstico

Selecione o formato para o nódulo visualizado na imagem

Redondo Oval Lobular Irregular

baixo Grau de Malignidade alto

Selecione a densidade do nódulo visualizado na imagem

Baixa Isodensa Alta

baixo Grau de Malignidade alto

Contém Gordura

Características associadas ao contorno do nódulo

totalmente circunscrito
 muito microlobulado
 parcialmente obscurecido
 pouco indistinto
 espiculado

Adicionar Remover

BI-RADS Automático

Figura 4.6. Interface gráfica do sistema desenvolvido.

A Figura 4.7 apresenta um exemplo de saída gerada pelo sistema. Foram adotados os seguintes valores como entrada: 0,6 como grau de malignidade associado ao formato lobular do nódulo, 0,8 como grau de malignidade associado à densidade alta e contorno muito espiculado e parcialmente indistinto. Essa combinação de características para o contorno levou a um peso final de 0,74, calculado segundo a fórmula e os pesos descritos na Seção 3.2.2, representativo do grau malignidade associado ao contorno.

Fuzzy BI-RADS

Consultar Diagnóstico

Selecione o formato para o nódulo visualizado na imagem

Redondo Oval Lobular Irregular

baixo Grau de Malignidade alto

Selecione a densidade do nódulo visualizado na imagem

Baixa Isodensa Alta

baixo Grau de Malignidade alto

Contém Gordura

Características associadas ao contorno do nódulo

totalmente circunscrito
 muito microlobulado
 parcialmente obscurecido
 pouco indistinto
 espiculado

muito espiculado
parcialmente indistinto

Adicionar Remover

BI-RADS Automático

BI-RADS 3: 10% de possibilidade
 BI-RADS 4: 91% de possibilidade
 BI-RADS 5: 12% de possibilidade

O sistema aconselha a adoção da categoria BI-RADS 4

Figura 4.7. Visualização da resposta gerada pelo sistema.

O valor 0,6 ativou os seguintes conjuntos nebulosos para o atributo coeficiente de formato, com os seguintes graus de pertinência: baixo (18%), médio (91%) e alto com (12%). Da mesma forma, o valor 0,8 ativou os seguintes subconjuntos para o atributo coeficiente de densidade: médio (99%) e alto (41%). Finalmente, quanto à avaliação do contorno foram obtidos os subconjuntos baixo (10%), médio (93%) e alto (12%) para o atributo coeficiente de contorno (Figura 4.8).

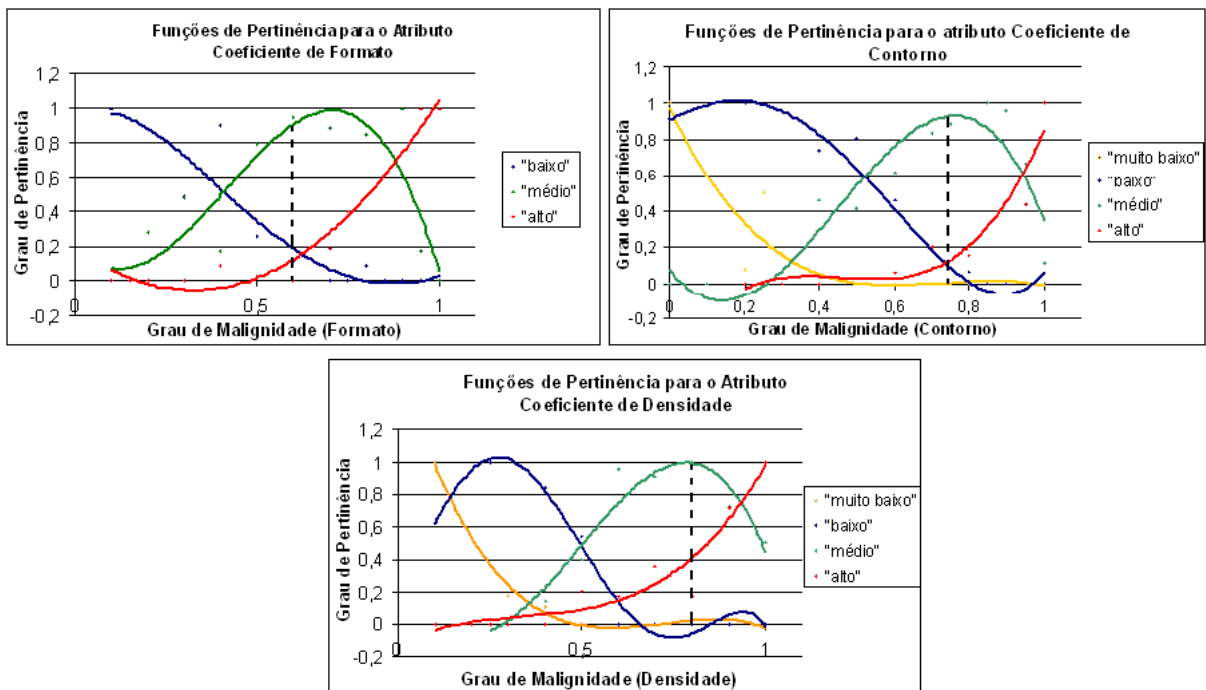


Figura 4.8. Subconjuntos nebulosos ativados.

As regras ativadas são formadas pela combinação dos subconjuntos nebulosos ativados para cada atributo. Neste exemplo foram ativadas 18 regras. Na Tabela 4 é possível visualizar as regras e os respectivos coeficientes de ativação. Por fim, foi obtido grau de pertinência de 10% à categoria BI-RADS 3, 91% à categoria BI-RADS 4 e 12% à categoria BI-RADS 5. Para a categoria 2 não foram obtidos valores de grau de pertinência.

Tabela 4. Regras ativadas para o exemplo em análise.

<i>Se Coeficiente de Formato</i>	<i>E Coeficiente de Contorno</i>	<i>E Coeficiente de Densidade</i>	<i>Então BI-RADS</i>	<i>Coeficiente de Ativação</i>
alto	alto	alto	5	0,12
alto	alto	medio	5	0,12
alto	medio	alto	5	0,12
alto	medio	medio	4	0,12
alto	baixo	alto	4	0,10
alto	baixo	medio	4	0,10
medio	alto	alto	5	0,12
medio	alto	medio	4	0,12
medio	medio	alto	4	0,41
medio	medio	medio	4	0,91
medio	baixo	alto	4	0,10
medio	baixo	medio	4	0,10
baixo	alto	alto	4	0,12
baixo	alto	medio	4	0,12
baixo	medio	alto	4	0,18
baixo	medio	medio	4	0,18
baixo	baixo	alto	3	0,10
baixo	baixo	medio	3	0,10

4.5. Validação dos resultados

Durante a etapa de testes, novas avaliações foram realizadas junto aos radiologistas. Seis novas imagens foram selecionadas da base de imagens [IARL, '08] (Figura 4.9) e apresentadas a cada um dos cinco radiologistas. A categoria BI-RADS atribuída durante a avaliação foi comparada com a saída gerada pelo sistema (Tabela 5), considerando-se os mesmos atributos de entrada obtidos das avaliações.

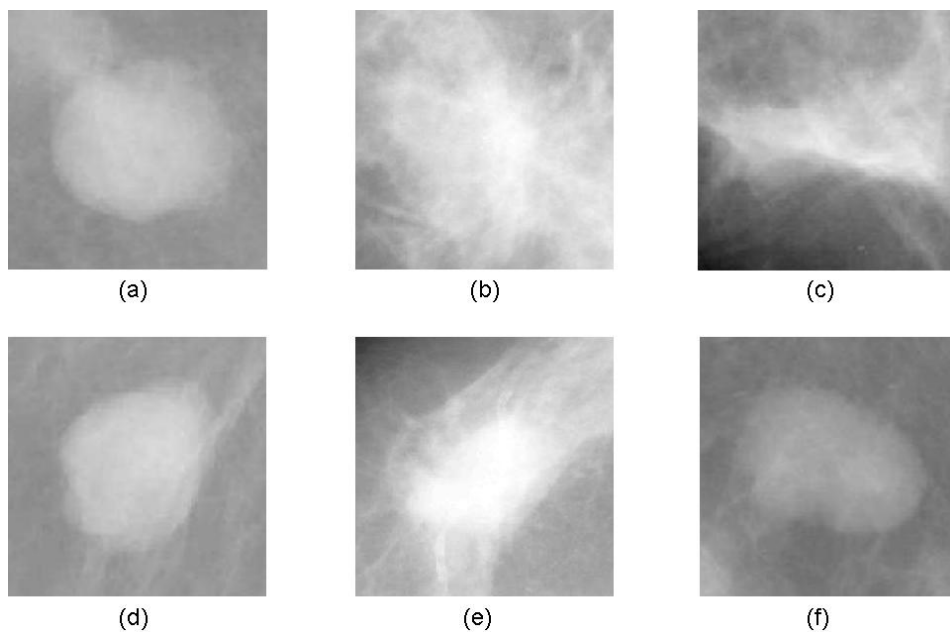


Figura 4.9. Imagem de referência (a) usf3424_LCC_L1_MS_B, (b) usf4077_LCC_L1_MS_M, (c) usf1187_RMLO_L1_MS_M, (d) usf0305_RMLO_L1_MS_B, (e) usf1820_LCC_L1_MS_M, (f) usf1394_RCC_L1_MS_B. [IARL, '08]

Para cada imagem, foi relacionada a avaliação de cada radiologista (R1, R2, R3, R4 e R5) à saída gerada pelo sistema. Nota-se que para uma mesma imagem foram obtidas avaliações diferentes quanto às características dos atributos e também quanto aos graus de malignidade atribuídos para o formato e a densidade. Porém é possível perceber um padrão de avaliação por imagem.

Tabela 5. Comparação dos resultados obtidos das avaliações com os resultados gerados pelo sistema.

Img	Avaliações			BI-RADS					Sistema
	formato	densidade	contorno	R1	R2	R3	R4	R5	
(a)	redondo(0,2)	alta(0,3)	parcialmente circunscrito	0					3(71%) 4(11%)
	redondo(0,5)	alta(0,7)	parcialmente circunscrito		0				3(35%) 4(71%)
	redondo(0,2)	alta(0,5)	parcialmente indistinto			4			3(48%) 4(49%) 5(4%)
	redondo(0,2)	alta(0,2)	muito circunscrito				3		2(9%) 3(88%) 4(11%)
	oval(0,4)	alta(0,8)	parcialmente circunscrito e parcialmente microlobulado					4	3(53%) 4(49%)
classe majoritária entre radiologistas				4					
(b)	irregular(0,9)	baixa(0,9)	totalmente indistinto	4					4(74%) 5(65%)
	irregular(0,8)	alta(0,7)	totalmente indistinto		4				4(91%) 5(25%)
	irregular(0,8)	alta(0,7)	muito espiculado			5			4(91%) 5(25%)
	irregular(0,8)	alta(0,5)	parcialmente indistinto e parcialmente espiculado				4		3(16%) 4(49%) 5(9%)
	irregular(0,8)	alta(0,8)	parcialmente indistinto e parcialmente espiculado					5	4(91%) 5(41%)
classe majoritária entre radiologistas				4					
(c)	irregular(1,0)	baixa(0,7)	totalmente indistinto	4					4(92%) 5(25%)
	irregular(0,6)	alta(0,6)	muito indistinto		4				3(18%) 4(75%) 5(12%)
	irregular(0,6)	alta(0,7)	muito indistinto			4			3(18%) 4(83%) 5(12%)
	irregular(0,2)	alta(0,5)	parcialmente indistinto				3		3(48%) 4(49%) 5(4%)
	irregular(0,8)	alta(0,7)	parcialmente indistinto e parcialmente espiculado					4	4(91%) 5(25%)
classe majoritária entre radiologistas				4					
(d)	redondo(0,2)	baixa(0,4)	totalmente circunscrito	0					2(25%) 3(82%) 4(11%)
	redondo(0,6)	alta(0,4)	parcialmente circunscrito		0				3(71%) 4(46%) 5(6%)
	oval(0,6)	alta(0,4)	muito microlobulado			0			3(18%) 4(82%) 5(12%)
	oval(0,2)	alta(0,2)	totalmente circunscrito				3		2(25%) 3(88%)
	oval(0,5)	alta(0,6)	parcialmente circunscrito e parcialmente microlobulado					4	3(35%) 4(72%)
classe majoritária entre radiologistas				0					
(e)	oval(0,3)	alta(0,6)	totalmente espiculado	4					3(14%) 4(69%) 5(15%)
	irregular(0,7)	alta(0,7)	muito obscurecido		4				3(6%) 4(83%) 5(25%)
	irregular(0,8)	alta(0,8)	muito espiculado			5			4(91%) 5(41%)
	irregular(0,8)	alta(0,7)	muito espiculado				4		4(91%) 5(25%)
	irregular(0,8)	alta(0,9)	muito espiculado					5	4(85%) 5(49%)
classe majoritária entre radiologistas				4					
(f)	oval(0,3)	alta(0,5)	parcialmente espiculado	0					3(48%) 4(49%) 5(9%)
	lobular(0,4)	alta(0,4)	parcialmente circunscrito		0				3(53%) 4(46%)
	oval(0,4)	isodensa(0,5)	pouco microlobulado			0			3(49%) 4(49%)
	oval(0,3)	alta(0,6)	muito microlobulado				4		3(14%) 4(72%) 5(13%)
	lobular(0,7)	alta(0,7)	parcialmente circunscrito e parcialmente microlobulado					4	3(6%) 4(73%) 5(25%)
classe majoritária entre radiologistas				0					

A imagem de referência (a), além da categoria BI-RADS 0 (não considerada como uma das saídas geradas pelo sistema), obteve classificação BI-RADS 3 e 4 entre os radiologistas. Essa intersecção entre categorias foi verificada nas saídas geradas pelo sistema para esta imagem.

Para a imagem (b), a classe majoritária entre os radiologistas foi a categoria 4. Os dois exemplos classificados como categoria 5, quando comparados às respectivas saídas geradas pelo sistema, não obtiveram pertinência máxima à mesma categoria. Porém foi identificado, em ambos os casos, um grau de pertinência menor à categoria 5, evidenciando que o exemplo possui características de categoria BI-RADS 5.

As imagens (c) e (e) também obtiveram categoria BI-RADS 4 como classe majoritária entre os radiologistas. Na imagem (c), para o exemplo destacado em vermelho, os graus de pertinência às categorias 3 e 4, gerados pelo sistema, foram bem próximos, indicando porcentagens praticamente idênticas. Na figura (e), a classe majoritária atribuída aos dois exemplos em vermelho foi diferente da categoria atribuída pelo radiologista, porém o grau de pertinência alcançado pela categoria 5 também foi expressivo para estes exemplos.

Por fim, as avaliações das imagens (d) e (e) obtiveram categoria majoritária 0. As saídas do sistema, apresentaram variações para cada exemplo. Observa-se graus de malignidade menores associados ao formato e densidade da lesão para os exemplos classificados como 0.

CAPÍTULO 5

Discussão e Conclusões

5.1. Discussão

Este trabalho apresenta um método para caracterização nebulosa dos principais descritores da terminologia lingüística utilizada na avaliação de nódulos de mama, assim como a categoria BI-RADS associada. O sistema desenvolvido contou 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, foi possível observar que dada uma mesma imagem de referência, diferentes avaliações entre os radiologistas foram obtidas. A Tabela 6 apresenta alguns exemplos de avaliações obtidas para diferentes imagens.

Tabela 6. Exemplos de avaliações realizadas.

Img	Radiolog	Forma	Contorno	Densidade	BI-RADS
Img 1	R1	irregular (0,4)	totalmente espiculado	alta (1,0)	5
	R2	irregular (0,9)	parcialmente obscuro e parcialmente indistinto	alta (0,6)	4
	R3	irregular (0,5)	parcialmente indistinto	alta (0,5)	4
	R4	irregular (0,7)	parcialmente indistinto	alta (0,6)	0
	R5	Irregular (0,7)	parcialmente obscuro e parcialmente indistinto	alta (0,8)	4
Img 2	R1	oval (0,5)	parcialmente obscuro	isodensa (0,5)	0
	R2	oval (0,4)	parcialmente circunscrito e parcialmente obscuro	isodensa (0,4)	3
	R3	oval (0,3)	totalmente circunscrito	baixa (0,1)	2
	R4	oval (0,2)	totalmente circunscrito	isodensa (0,2)	3
	R5	oval (0,1)	parcialmente circunscrito	isodensa (0,2)	3
Img 3	R1	redonda (0,25)	parcialmente espiculado	alta (0,5)	0
	R2	redonda (0,5)	totalmente microlobulado	alta (0,7)	4
	R3	oval (0,4)	parcialmente indistinto	isodensa (0,5)	3
	R4	oval (0,4)	totalmente circunscrito	alta (0,5)	3
	R5	oval (0,4)	parcialmente obscuro e parcialmente indistinto	alta (0,3)	4

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. Para a imagem de referência *Img 1*, em 60% das avaliações a categoria 4 foi atribuída ao exemplo, em 20% das avaliações a categoria 5 foi atribuída ao exemplo e com a mesma porcentagem a categoria 0 foi determinada. Para a imagem de referência *Img 2*, as categorias atribuídas variaram entre 0, 2 e 3, sendo a categoria 3 majoritária. Para a imagem de referência *Img 3* temos uma intersecção entre as categorias 3 e 4.

Nas avaliações realizadas com os radiologistas, preocupou-se 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. Essas escalas ficaram implícitas ao usuário na interface do sistema desenvolvido.

Quanto às funções de pertinência, verificou-se que a utilização dos algoritmos Omega e Fuzzy Omega adequou-se ao domínio de valores dos coeficientes. A proposta de um algoritmo que gere funções de pertinência de maneira automática, foi resultado da dificuldade de modelar as funções apenas com parâmetros lingüísticos ou baseados na subjetividade de um especialista. Considerando o caso de se trabalhar com atributos intrínsecos de imagens, quem representaria o especialista do domínio? Neste caso, o algoritmo Fuzzy-Omega seria bastante adequado.

As funções polinomiais obtidas refletiram a distribuição do conjunto de dados. Devido à pequena quantidade de exemplos pertencentes às categorias 2 e 5, as funções de pertinência geradas, utilizando-se essas classes como classes majoritárias, possuem intervalos menores de variação dos dados. Quando valores mais baixos eram adotados como grau de malignidade, geralmente a lesão era categorizada como BI-RADS 0 pelos radiologistas. Essa distribuição de valores também ficou evidente na base de proposições nebulosas, com uma grande quantidade de regras levando às categorias 3 e 4.

5.2. Possibilidade de Trabalhos Futuros

Neste trabalho foram utilizados apenas parâmetros de avaliação de nódulos de mama (massas). Como possibilidade de trabalho futuro, poderiam ser adicionados ao modelo proposto, parâmetros de avaliação de microcalcificações e achados adicionais, também padronizados pelo BI-RADS.

Outra possibilidade de continuidade deste trabalho seria a avaliação do grau de pertinência gerado pelo sistema em ambientes de recuperação de imagens. Dada uma saída nebulosa no seguinte formato $(X, \mu(x))$, no qual X representa uma das classes de saída do sistema e $\mu(x)$, o grau de pertinência do elemento à classe X , associar duas imagens segundo o cálculo de uma distância nebulosa entre elas. Essa distância levaria em consideração a diferença entre duas saídas nebulosas geradas pelo sistema para duas imagens distintas, assim como a diferença absoluta entre a pertinência máxima (100%) e o valor de $\mu(x)$ para cada uma das classes presentes na saída nebulosa.

Também há possibilidade de utilização de sistemas de inferência híbridos, como os sistemas *neuro-fuzzy* ou ANFIS, descritos em [SANDRI, S. '99] em trabalhos futuros. Esses sistemas utilizam tanto os paradigmas de sistemas nebulosos quanto os paradigmas de sistemas conexionistas, com o intuito de aprimorar os resultados obtidos. Nesta arquitetura, a fase de aprendizado é caracterizada pela utilização de redes neurais e a fase de processamento comporta-se como um sistema nebuloso. Os parâmetros e regras de um sistema nebuloso são aprendidos a partir do conjunto de pares (entrada, saída desejada).

Com o objetivo similar ao da utilização de sistemas *neuro-fuzzy*, uma arquitetura híbrida que utilize sistemas nebulosos e computação evolutiva (como os algoritmos genéticos), também poderia ser utilizada. Em especial, nestes sistemas, a computação evolutiva é utilizada na busca, aprendizado e otimização de componentes dos sistemas nebulosos enquanto que a experiência de especialistas em sistemas nebulosos pode ser utilizada no gerenciamento dos parâmetros de controle da computação evolutiva [FIALHO, A.R.S. '07].

5.3. Conclusões

A partir da ocorrência de divergências observadas nas avaliações dos achados mamográficos, durante as entrevistas realizadas com radiologistas, evidenciou-se a presença de subjetividade. A utilização da lógica nebulosa permitiu a representação dessa subjetividade.

Além disso, o método proposto possibilita a criação de um sistema CAD mais amigável, no qual os parâmetros de entrada são bastante próximos da forma natural como o radiologista avalia e qualifica os achados da imagem. Ao tratar os atributos formato, 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 e fornecer ao usuário uma análise quantitativa.

REFERÊNCIAS BIBLIOGRÁFICAS

[ALMEIDA, P. E. M. '03] ALMEIDA P. E. M.; EVSUKOFF A. G. Sistemas Fuzzy. In: Rezende, S. O. (org). *Sistemas Inteligentes: fundamentos e aplicações*. 1.ed. Barueri, SP: Manole, 2003. cap.7, p.169-200.

[AMENDOLA, M. '04] AMENDOLA, M.; SOUZA, A. *Manual do uso da teoria dos conjuntos Fuzzy no Matlab 6.5*. Campinas: CPG/FEAGRI/UNICAMP, 2004. 30p.

[ACR, '98] American College of Radiology. *Breast Imaging Reporting and Data System (BI-RADS®)*. 3rd ed. Reston, VA: American College of Radiology, 1998.

[ACR, '08] ACR. *Breast Imaging Reporting and Data System (BI-RADS®)*. <http://www.birads.at>. Acessado em 13/07/2008.

[ANDRÉ, T. C. S. S. '02] ANDRÉ, T. C. S. S.; SILVA-FILHO, A. C. R. Segmentation of Digitized Mammograms Using Self-Organizing Maps in a Breast Cancer Computer Aided Diagnosis System. *7th Brazilian Symposium on Neural Networks (SBRN 2002)*, Recife, Brazil, IEEE Computer Society. 11-14. 2002.

[BAKER, J. A. '96] BAKER J. A.; KORNGUTH P. J.; LO, J. Y.; FLOYD, C. E. Artificial neural network: improving the quality of breast biopsy recommendations. *Radiology*, 1996. 198(1):131-5.

[CASTLEMAN K. N. '96] CASTLEMAN, K. N. *Digital image processing*. New Jersey: Prentice-Hall, 1996.

[DEITEL, H. M. '06] DEITEL, H. M.; DEITEL, P. J. *Java: Como Programar*. 6.ed. São Paulo: Pearson Education do Brasil. 2006.

[DOI, K. '99] DOI, K. Computer-aided diagnosis and its potential impact on diagnostic radiology. In: DOI, K.; MACMAHON, H.; GIGER, M. L.; HOFFMANN, K. R.; eds. *Computer-aided diagnosis in medical imaging*. Amsterdam: Elsevier Science, 11–20. 1999.

[DOI, K. '05] DOI, K. Current status and future potential of computer-aided diagnosis in medical imaging. *British Journal of Radiology*, n. 78, pp. 3-19. 2005.

[DOI, K. '06] DOI, K. Diagnostic Imaging over the last 50 years: research and development in medical imaging science and technology. *Phys Med Biol*, vol. 51, n. 13, pp. 5-27. 2006.

[DUBOIS, D. '96] DUBOIS, D.; PRADE, H. What are fuzzy rules and how to use them. *Fuzzy Sets and Systems*, vol. 84, pp. 169-185. 1996.

[FIALHO, A. R. S. '07] FIALHO, A. R. S. *Exploração de Relações Entre as Técnicas Nebulosas e Evolutivas da Inteligência Computacional*. Tese de Mestrado. Escola Politécnica. Universidade de São Paulo, SP – Brasil. 2007.

[GIGER, M. L. '99] GIGER, M. L. Computer-aided diagnosis. In: *RSNA Categorical Course in Breast Imaging*, pp. 249–72. 1999.

[GIGER, M. L. '00] GIGER, M. L. Computer-aided diagnosis of breast lesions in medical images. *Computing in Science & Engineering*, v2, n.5, p. 39-45, 2000.

[IARL, '08] IARL. (2008) *Digital Database for Screening Mammography*. <http://marathon.csee.usf.edu/Mammography/>. Acessado em 13/07/2008

[JAFELICE, R. S. M. '05] JAFELICE, R. S. M.; LOPES, W. A.; BARROS, L. C. Diagnóstico Médico Fuzzy e Monitoramento da Pneumonia. *Biomatemática*. IMECC-UNICAMP, v. 15, p. 77-96, 2005.

[JOO, S. '04] JOO, S.; YANG, Y. S.; MOON, W. K. et al. Computer-aided diagnosis of solid breast nodules: use of an artificial neural network based on multiple sonographic features. *IEEE Trans. Med. Imaging* 23(10): 1292-1300. 2004.

[KLIR, G. J. '95] KLIR, G. J.; YUAN, B. *Fuzzy Sets And Fuzzy Logic: Theory and Applications*. Rio de Janeiro: Prentice-Hall. 1995.

[MARQUES, P. M. A. '01] MARQUES, P. M. A. Computer-aided diagnosis in radiology. *Radiol Bras*, vol.34, no.5, p.285-293. ISSN 0100-3984.Sept./Oct. 2001.

[MARSH, A. '97] MARSH, A. EUROMED - The Creation of a Telemedical Information Society. In *10' IEEE Symposium on Computer Based Medical Systems*. Maribor, Slovenia.1997.

[MENDEL, J. M. '95] MENDEL, J. M. *Fuzzy Logic Systems for Engeneering: a Tutorial*. Proc. IEEE, 3: 345-376. 1995.

[OBENAUER, S. '05] OBENAUER, S.; HERMANN, K. P.; GRABBE, E. Applications and literature review of the BI-RADS classification. *European Radiology*, vol.15, no.5, p.1027-1036. 2005.

[OLIVEIRA-JR, H. A. '99] OLIVEIRA-JR, H. A. "*Lógica Difusa: Aspectos Práticos e Aplicações*". Rio de Janeiro. Editora Interciência. 1999.

[RIBEIRO, M. X. '08] RIBEIRO, M. X.; FERREIRA, M. R. P.; TRAINA, A. J. M.; TRAINA, C. Data Pre-processing: A new algorithm for Feature Selection and Data Discretization. In: *International Conference on Soft Computing as Transdisciplinary Science and Technology*, 2008, Paris, France. Anais do CSTST2008, 2008. p. 1-6.

[SANDRI, S. '99] SANDRI, S.; CORREA, C. *Lógica Nebulosa*. V Escola de Redes Neurais, Promoção: Conselho Nacional de Redes Neurais pp. c073-c090, 19 de julho, 1999 – ITA, São José dos Campos - SP

[YIN, T.-K. '04] 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* 23(5): 639-654. 2004.

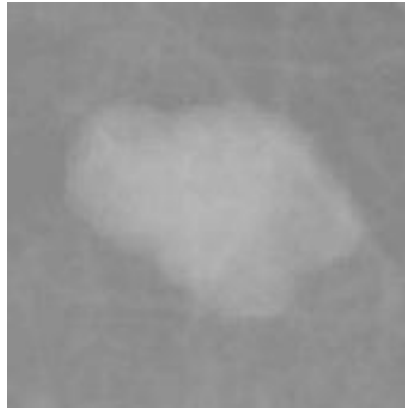
[ZADEH, L. A. '75] ZADEH, L. A. *Fuzzy logic and approximate reasoning (In memory of Grigore Moisil)*. Synthese, 30, pp. 407-428. 1975.

[ZADEH, L. A. '77] ZADEH, L. A.; BELLMAN, R. E. Local and Fuzzy Logics. In: Dunn, J. M.; Epstein, G. Eds, *Modern Uses of Multiple-Value Logics*, Dordrecht: Reidel 1977, pp. 106-107.

APÊNDICES

APÊNDICE A - Avaliação realizada com os radiologistas

Imagem de Referência: usf0305_RCC_L1_MS



- Marcar a opção escolhida para o atributo **forma** e o grau de malignidade associado à esta opção

- Redonda
 Oval
 Lobular
 Irregular

0	0.25	0.5	0.75	1
---	------	-----	------	---

- Marcar as opções escolhidas para o atributo **contorno**, relacioná-las com os qualificadores destacados abaixo (letras a, b, c ou d) e determinar o grau de malignidade associado às opções.

(a) totalmente (b) muito (c) parcialmente (d) pouco

___ Circunscrito

0	0.25	0.5	0.75	1
---	------	-----	------	---

___ Microlobulado

0	0.25	0.5	0.75	1
---	------	-----	------	---

___ Obscuro

0	0.25	0.5	0.75	1
---	------	-----	------	---

___ Indistinto

0	0.25	0.5	0.75	1
---	------	-----	------	---

___ Espiculado

0	0.25	0.5	0.75	1
---	------	-----	------	---

- Marcar a opção escolhida para o atributo **densidade** e o grau de malignidade associado à esta opção

- Alta
 Isodensa
 Baixa
 Contém gordura

0	0.25	0.5	0.75	1
---	------	-----	------	---

Categoria de Avaliação BI-RADS:

0 2 3 4 5

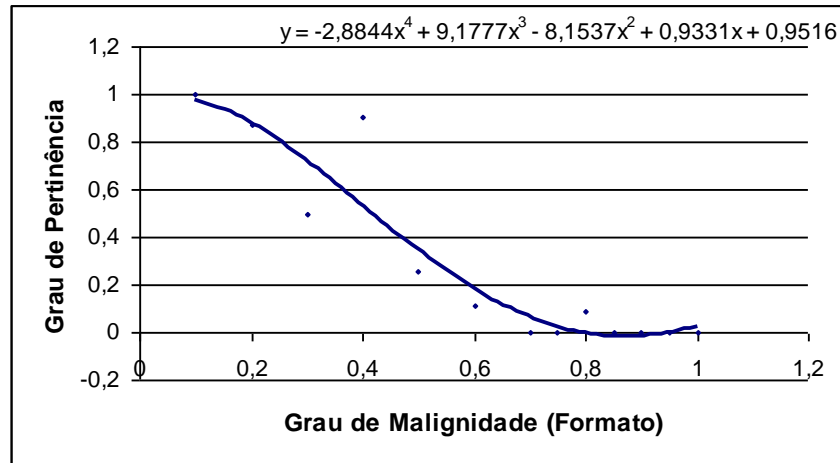
APÊNDICE B - Atributos de entrada do algoritmo Fuzzy-Omega

Coefficiente de Formato	BI-RADS	Coefficiente de Densidade	BI-RADS	Coefficiente de Contorno	BI-RADS
0,1	3	0,1	3	0	2
0,1	2	0,1	2	0	3
0,1	3	0,1	3	0,1	3
0,1	3	0,1	4	0,1	3
0,1	3	0,1	3	0,1	3
0,1	3	0,1	3	0,1	3
0,15	3	0,1	3	0,1	3
0,2	3	0,1	2	0,1	3
0,2	4	0,2	4	0,1	4
0,2	4	0,2	4	0,1	4
0,2	3	0,2	3	0,1	3
0,2	3	0,2	3	0,1	4
0,2	3	0,2	3	0,2	3
0,2	3	0,2	4	0,2	3
0,2	3	0,2	3	0,2	3
0,2	4	0,2	4	0,2	4
0,2	4	0,2	3	0,2	3
0,2	3	0,2	3	0,2	3
0,2	3	0,2	4	0,2	4
0,2	3	0,2	3	0,2	4
0,2	3	0,2	3	0,2	3
0,2	3	0,2	3	0,2	3
0,2	3	0,25	3	0,2	4
0,3	4	0,3	3	0,2	3
0,3	4	0,3	4	0,2	3
0,3	4	0,3	4	0,2	3
0,3	4	0,3	4	0,2	3
0,3	4	0,3	4	0,2	4
0,3	4	0,3	4	0,2	4
0,3	3	0,3	4	0,2	3
0,3	3	0,3	4	0,2	3
0,3	3	0,3	3	0,2	3
0,3	3	0,3	4	0,2	3
0,3	2	0,3	4	0,2	3
0,3	3	0,3	4	0,2	3
0,3	2	0,3	4	0,2	3
0,4	4	0,3	3	0,2	3
0,4	4	0,3	3	0,2	3
0,4	4	0,3	3	0,2	2
0,4	3	0,3	3	0,25	4
0,4	4	0,3	2	0,25	4
0,4	3	0,4	4	0,3	4
0,4	3	0,4	3	0,3	4
0,4	3	0,4	3	0,3	3
0,4	3	0,4	3	0,3	4
0,4	3	0,4	3	0,3	4
0,4	3	0,4	3	0,3	4
0,4	3	0,4	3	0,3	4
0,4	3	0,4	3	0,3	4
0,4	3	0,4	3	0,3	4
0,4	3	0,4	3	0,3	4
0,4	3	0,4	3	0,3	4
0,4	3	0,4	3	0,3	3
0,4	3	0,5	4	0,3	3
0,4	3	0,5	5	0,3	2
0,4	5	0,5	3	0,3	3
0,4	5	0,5	5	0,4	4
0,5	4	0,5	4	0,4	4
0,5	4	0,5	3	0,4	4
0,5	4	0,5	4	0,4	4
0,5	4	0,5	3	0,4	4
0,5	3	0,5	3	0,4	4
0,5	4	0,5	3	0,4	4

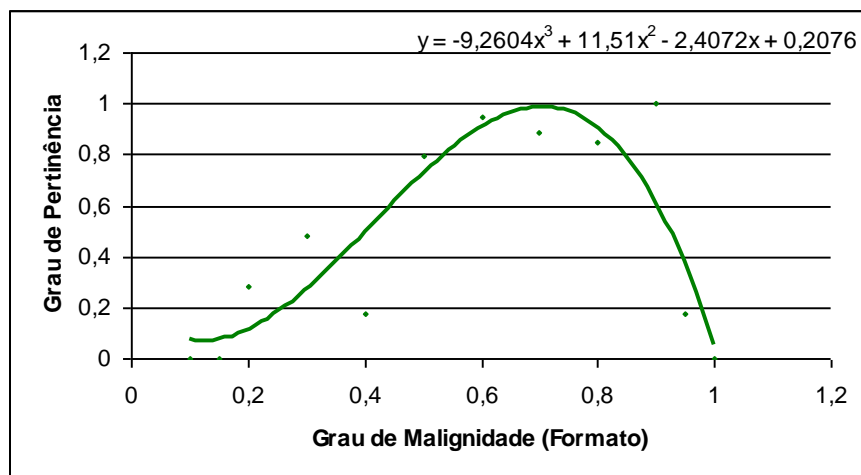
0,5	3	0,5	3	0,4	3
0,5	4	0,5	4	0,4	3
0,5	3	0,5	3	0,4	3
0,5	4	0,5	3	0,4	3
0,5	4	0,5	3	0,4	3
0,5	4	0,5	3	0,4	3
0,5	4	0,5	4	0,4	3
0,6	4	0,5	4	0,4	3
0,6	3	0,5	4	0,4	3
0,6	4	0,5	4	0,4	3
0,6	4	0,6	4	0,4	4
0,6	4	0,6	4	0,4	3
0,6	4	0,6	4	0,4	4
0,6	4	0,6	5	0,5	4
0,6	4	0,6	4	0,5	4
0,6	4	0,6	4	0,5	4
0,6	4	0,6	4	0,5	4
0,6	4	0,6	4	0,5	3
0,7	4	0,6	4	0,5	3
0,7	4	0,6	4	0,5	3
0,7	4	0,6	4	0,5	3
0,7	5	0,6	3	0,5	3
0,7	5	0,6	4	0,5	3
0,7	4	0,6	4	0,5	3
0,7	4	0,6	4	0,5	3
0,7	4	0,6	4	0,5	4
0,7	4	0,6	4	0,5	4
0,7	4	0,6	4	0,5	3
0,7	4	0,6	5	0,5	3
0,7	4	0,6	4	0,5	4
0,7	4	0,6	4	0,6	4
0,75	4	0,6	4	0,6	4
0,8	5	0,6	4	0,6	4
0,8	4	0,6	4	0,6	4
0,8	4	0,6	4	0,6	3
0,8	5	0,7	4	0,6	3
0,8	4	0,7	4	0,6	3
0,8	4	0,7	4	0,6	4
0,8	4	0,7	5	0,6	5
0,8	4	0,7	4	0,6	3
0,8	3	0,7	4	0,6	3
0,8	4	0,7	4	0,6	4
0,8	4	0,7	4	0,6	4
0,8	4	0,7	4	0,6	4
0,8	4	0,8	5	0,6	4
0,8	4	0,8	4	0,6	3
0,8	4	0,8	4	0,6	4
0,85	4	0,8	4	0,6	4
0,9	4	0,8	4	0,6	4
0,9	4	0,8	4	0,6	4
0,9	4	0,8	4	0,6	3
0,9	4	0,8	4	0,6	4
0,9	4	0,9	4	0,7	4
0,9	4	0,9	4	0,7	4
0,9	4	0,9	5	0,7	5
0,9	4	1	5	0,7	4
0,95	5	1	4	0,7	4
0,95	5	1	4	0,7	5
1	5	1	5	0,7	5
1	5	1	5	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,7	4
---	---	---	---	0,75	4
---	---	---	---	0,75	4
---	---	---	---	0,75	4
---	---	---	---	0,75	4
---	---	---	---	0,8	5
---	---	---	---	0,8	5

APÊNDICE C – Equações polinomiais das funções de pertinência

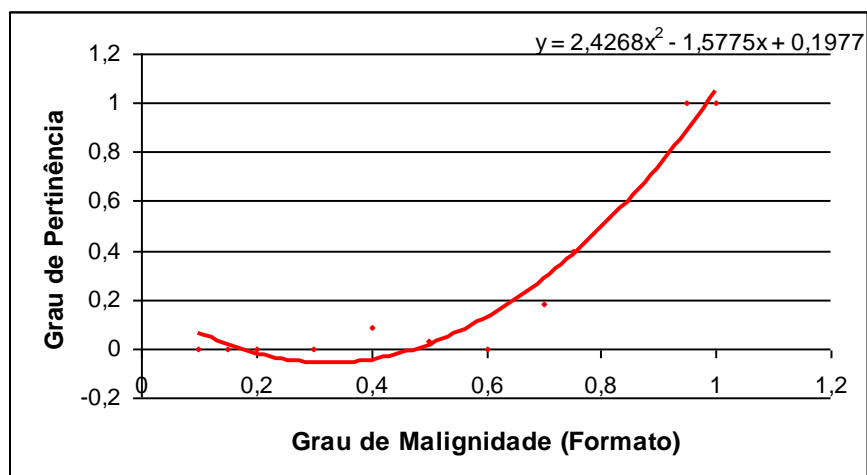
- **Coefficiente de Formato**



Função de pertinência do subconjunto "coeficiente de formato baixo"

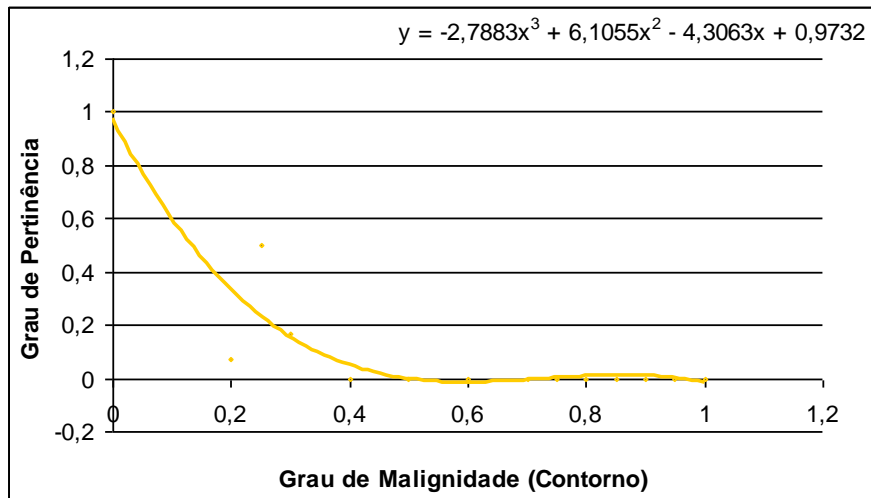


Função de pertinência do subconjunto "coeficiente de formato médio"

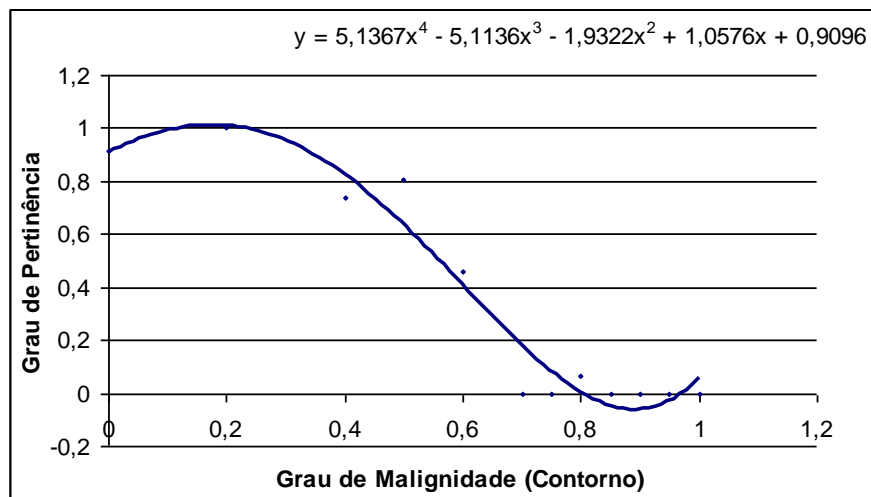


Função de pertinência do subconjunto "coeficiente de formato alto"

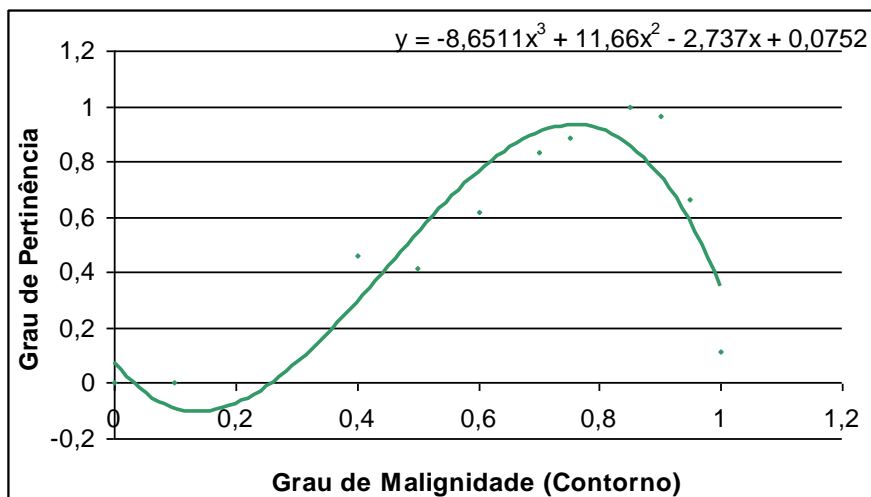
- **Coefficiente de Contorno**



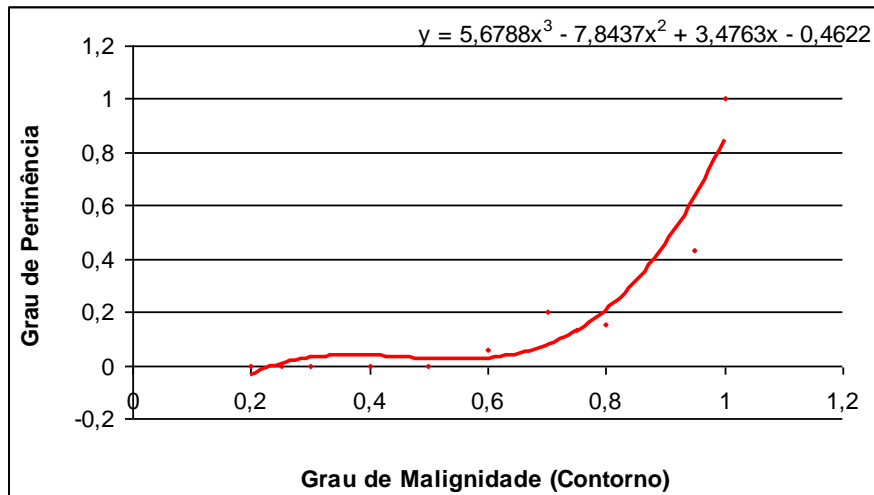
Função de pertinência do subconjunto “*coeficiente de contorno muito baixo*”



Função de pertinência do subconjunto “*coeficiente de contorno baixo*”

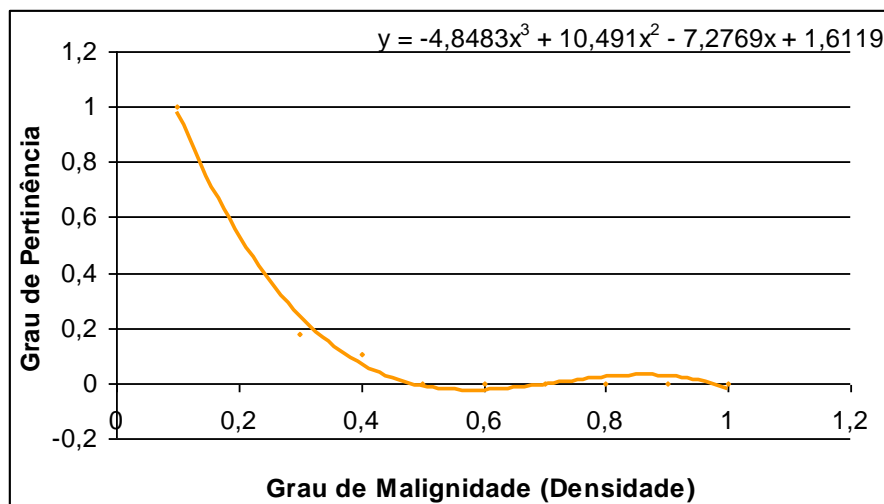


Função de pertinência do subconjunto “*coeficiente de contorno médio*”

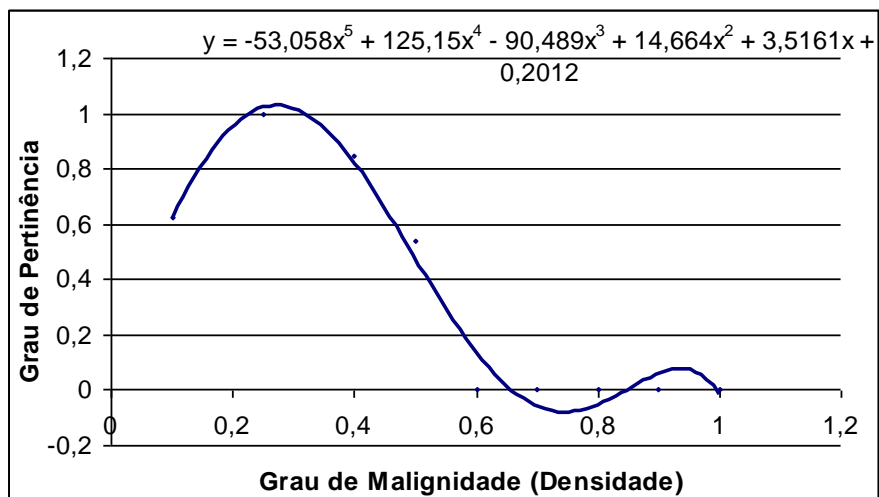


Função de pertinência do subconjunto "coeficiente de contorno alto"

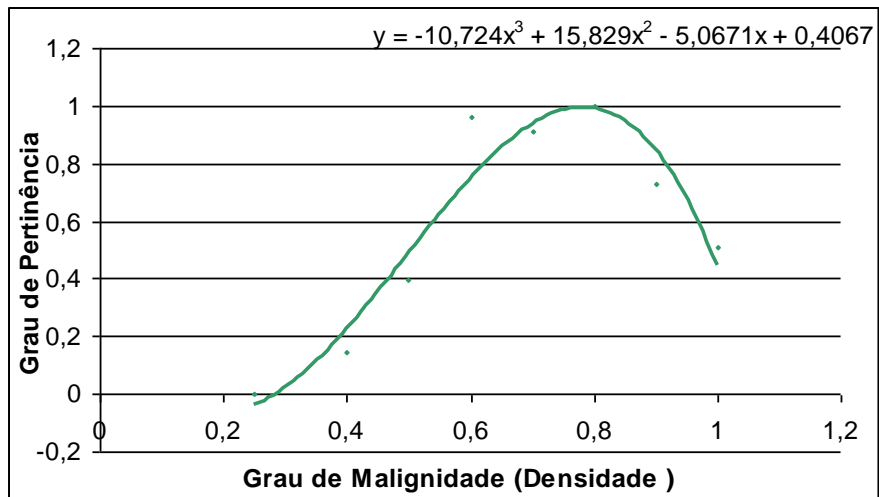
- **Coefficiente de Densidade**



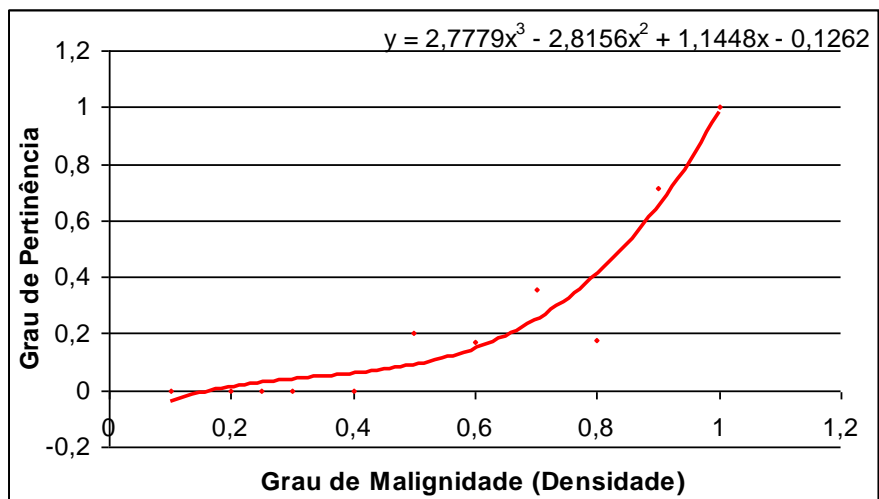
Função de pertinência do subconjunto "coeficiente de densidade muito baixo"



Função de pertinência do subconjunto "coeficiente de densidade baixo"



Função de pertinência do subconjunto "coeficiente de densidade médio"



Função de pertinência do subconjunto "coeficiente de densidade alto"

APÊNDICE D – Diagrama de Classes do Sistema Fuzzy-BI-RADS

