



**UNIVERSIDADE DA BEIRA INTERIOR**  
Faculdade de Ciências da Saúde

# Utilização da Lógica *Fuzzy* no Estudo de Nódulos Mamários

**Joana Martinha Reis Antunes**

Dissertação para a obtenção do grau Mestre em  
**Ciências Biomédicas**  
(2º ciclo de estudos)

Orientador: Prof. Doutor José António Menezes Felipe de Souza

**Covilhã, Outubro de 2016**



# Agradecimentos

Quero agradecer,

Especialmente a meus pais, Maria de Fátima Rodrigues dos Reis Antunes e Joaquim de Jesus Simões Antunes, que tudo fizeram para eu continuar a estudar, sem eles eu jamais poderia ter sequer pensado em começar um mestrado;

À minha irmã, Joselina Reis Antunes, que me apoiou nos momentos mais difíceis e por toda a ajuda na compreensão e desenvolvimento da dissertação;

Ao meu professor orientador, José António Menezes Felipe de Souza, que nunca me negou ajuda quando esta foi necessária;

Ao Ricardo Valente colega e amigo de longa data, que sempre esteve ao meu lado e me ajudou na correção da dissertação;

Aos meus amigos, que durante dois anos foram como uma família, visto que a minha se encontrava a grande distância. Em especial à Denise Seguro Carvalho, que esteve comigo nos momentos mais difíceis, apoiando-me e ajudando-me a ver as coisas de uma maneira mais simples.



# Resumo

Este trabalho aborda o funcionamento da Lógica *Fuzzy* no método de inferência de *Mamdani* e a sua capacidade de obter boas probabilidades na classificação, dependendo de características e regras implementadas.

Para o estudo foram usados dados de imagens médicas, mamografias, anteriormente classificadas por um profissional, onde foram feitas contagens e tabelados todos os valores encontrados. Depois foram calculadas para cada valor encontrado nas contagens, as suas devidas percentagens, para por fim, se proceder à passagem para o estudo em Lógica *Fuzzy*.

Para este trabalho foram feitos dois estudos onde foi usado o Matlab, na toolbox Lógica *Fuzzy* e usado o método de inferência de *Mamdani*, para a obtenção de resultados. No primeiro estudo foram usadas duas características das imagens como variáveis de entrada, o tipo de tecido e a anomalias encontradas no tecido mamário. No segundo estudo foram usadas três características como variáveis de entrada, foram elas as já usadas no estudo anterior, o tipo de tecido e as anomalias encontradas no tecido mamário e foi ainda aplicada uma terceira característica, o tamanho dos nódulos encontrados. Por fim foi usada a característica da severidade da doença, benigna ou maligna, como variável de saída, conseguindo assim perceber quais as probabilidades de um determinada tecido, cruzado com uma determinada anomalia e, no caso do segundo estudo, dependendo de um determinado tamanho, se obter uma massa maligna ou benigna.

No início deste trabalho é ainda explicado o que é o cancro da mama e o exame médico a ser feito para este tipo de doença. Esta explicação é necessária pois é importante conhecer a doença, entender todo o mecanismo da mesma e todo o seu tratamento, para que assim se consiga o melhor método para a obtenção de melhores resultados.

Após toda a análise feita, verifica-se a obtenção resultados bastante satisfatórios nos dois estudos, sendo que o primeiro obteve melhores valores probabilísticos, onde se obtiveram valores de 82.1% para malignidade em tecidos gordos para uma anomalia assimétrica (ASYM) do tecido mamário. Já o segundo estudo obteve maior pormenor de classificação, em que se provou que a malignidade está relacionada com o tamanho do nódulo, quanto maior for o tamanho do nódulo maior é a probabilidade do tumor ser maligno. Todo o processo é explicado no decorrer do trabalho.

## Palavras-chave

Lógica *Fuzzy*, método de inferência de *Mamdani*, cancro da mama, mamografia.



# Abstract

This paper discusses the operation of Fuzzy Logic in Mamdani inference method and its ability to get good odds in the standings, depending on features and implemented rules.

In this study, mammograms were used in medical imaging data, previously classified by a professional. Where a count was made and tabulated for the all values founds, and then calculated for each value found in scores their percentages, to then proceed to the passage to the study in Fuzzy Logic.

For the present work twostudies have been done where performed using Matlab in Fuzzy Logic toolbox and with Mamdani inference method. In the first study it was used two characteristics of images as input variables, the type of tissue and the anomalies found in breast tissue. In the second study were used three features as input variables, the ones used already used in the previous study, plus, the size of nodules found. Finally the characteristic of the severity of the disease , benign or malignant, has been used as output variable, thus obtaining, how likely a particular tissue, crossed with a specific anomaly and in the case of second study, depending on a size of nodule, obtaining a benign or malignant tumor.

At the beginning of this work is also explained what is breast cancer and the medical examination to be made for this type of disease. This explanation is necessary because it is important to know the disease, understand the whole mechanism of it and all its treatment, so that it gets the best method for achieving better results.

After all the analysis made, there has been obtained good results in both studies. The first study obtained the best probability values, where the malignancy has the probability of 82.1% for fatty tissue with an asymmetrical anomaly (ASYM) of breast tissue. In the second study, there was more detailed classification, which proved that the malignancy is related to the nodule size, the larger the nodule size is most likely to be malignant tumor. The whole process is explained in the course of the work.

# Keywords

*Fuzzy Logic, Mamdani inference method, breast cancer, mammography.*



# Índice

Agradecimentos .....	iii
Resumo .....	v
Abstract .....	viii
Índice .....	xi
Lista de Figuras.....	xiv
Lista de Tabelas .....	xvii
Lista de Acrónimos .....	xix
Capítulo I.....	1
Introdução .....	1
1.1 - Objetivos .....	4
1.2 - Estrutura e organização da dissertação .....	4
Capítulo II.....	6
Patologia médica abordada .....	6
2.1 - Anatomia da mama.....	7
2.2 - Fatores epidemiológicos .....	9
Capítulo III .....	11
Exame médico abordado .....	11
3.1 - Exame médico para o diagnóstico .....	12
3.2 - Estadiamento.....	17
3.3 - Tipos de anomalias dos tecidos mamários detetados na mamografia .....	18
Capítulo IV .....	21
Lógica Fuzzy .....	21
4.1 - Conjuntos <i>fuzzy</i> .....	23
4.2 - Teoria dos conjuntos <i>fuzzy</i> e a teoria de probabilidade .....	24
4.3 - Variáveis Linguísticas.....	25
4.4 - Regras linguísticas.....	26
4.5 - Funções pertinência .....	27

4.6 - Sistema de inferência .....	28
Capítulo V.....	31
Estudos feitos com Lógica fuzzy na área da saúde .....	31
Capítulo VI.....	36
Características usadas no estudo .....	36
Capítulo VII .....	43
Lógica <i>Fuzzy</i> utilizando o método de inferência <i>Mamdani</i> .....	43
7.1 - Comandos usados.....	45
7.1.1 - Primeiro estudo .....	48
7.1.2 - Segundo estudo.....	55
Capítulo VIII .....	62
Resultados e Discussão.....	62
8.1 - Resultados obtidos com o primeiro estudo.....	62
8.2 - Resultados obtidos com o segundo estudo .....	66
8.3 - Comparação entre os estudos realizados.....	73
Capítulo IX.....	75
Conclusão e trabalhos futuros.....	75
Bibliografia .....	78
Anexos.....	80



# Lista de Figuras

Figura 1 - Anatomia da mama, retirado de <a href="http://oncomastologia.com.br/noticias/a-mama">http://oncomastologia.com.br/noticias/a-mama</a> . (requerido em Setembro de 2016). .....	8
Figura 2 - Anatomia da mama aprofundada, retirado de <a href="http://oncomastologia.com.br/noticias/a-mama">http://oncomastologia.com.br/noticias/a-mama</a> (requerido em Setembro de 2016). .....	8
Figura 3 - Imagens demonstrativas das diferentes densidades mamárias, retirado de [16]. ...	12
Figura 4 - Compartimento e compressão, retirado de [15]. .....	14
Figura 5 - Procedimento durante o exame, vista lateral, e imagem obtida após o exame, retirado de [15]. .....	15
Figura 6 - Procedimento durante o exame, vista decima, e imagem obtida após o exame, retirado de [15]. .....	15
Figura 8 - Diferentes margens para contornas de um nódulo, retirado de [15]. .....	19
Figura 7 - Diversas formas de massas que um nódulo pode ter, retirado de [15]. .....	19
Figura 9 - a) copos com diferentes medidas, b) copos cheios, c) copos vazios, retirado de [7]. .....	22
Figura 10 - Varias linguística de velocidade de um determinado veículo, retirado de [21]. ...	26
Figura 11 - Formatos de funções de pertinência mais utilizadas, retirado de [7, 21]. .....	28
Figura 12 - Arquitetura padrão de um sistema fuzzy, retirado de [21]. .....	29
Figura 13 - Diagrama típico de um modelo de inferência de Mamdani, retirado de [36]. .....	44
Figura 14 - Exemplo de um sistema de inferência Mamdani, retirado de [34]. .....	45
Figura 15 - Aspeto da toolbox Lógica Fuzzy usando o método de inferência Mamdani, imagem próprias. ....	46
Figura 16 - Usando a toolbox Lógica Fuzzy usando o método de inferência Mamdani, para o primeiro estudo, imagem próprias. ....	48
Figura 17 - Visualização da primeira função de entrada na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias. ....	49
Figura 18 - Visualização da segunda função de entrada na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias. ....	50
Figura 19 - Visualização da função de saída na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias. ....	51
Figura 20 - Visualização do quadro das regras implementadas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias. ....	52
Figura 21 - Visualização do quadro das regras geradas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias. ....	53

Figura 22 - Visualização do quadro do gráfico gerado pela toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias. ....	54
Figura 23 - Usando a toolbox Lógica Fuzzy usando o método de inferência Mamdani, para o segundo estudo, imagem próprias. ....	55
Figura 24 - Visualização da terceira função de entrada na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no segundo estudo, imagens próprias. ....	56
Figura 25 - Visualização do quadro das regras implementadas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no segundo estudo, imagens próprias. ....	57
Figura 26 - Visualização do quadro das regras geradas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no segundo estudo, imagens próprias. ....	59
Figura 27 - Visualização do quadro do gráfico gerado pela toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias. ....	60
Figura 28 - Cruzamento entre o tipo de tecido e a anomalia apresentada, gráfico obtido com a toolbox Lógica Fuzzy. Imagem própria.....	65
Figura 29 - Cruzamento entre a anomalia apresentada e o tamanho, gráfico obtido com a toolbox Lógica Fuzzy. Imagem própria.....	72



## Lista de Tabelas

Tabela 1 - Divisão de grupos segundo o estadiamento TNM, retirado de [13]. .....	18
Tabela 2 - Contagens feitas para o primeiro estudo a partir da base de dados para a obtenção do cálculo percentual. Dados próprios. ....	37
Tabela 3 - Valores percentuais para o primeiro estudo. Dados próprios. ....	38
Tabela 4 - Divisão de grupos para o estudo. Dados próprios. ....	39
Tabela 5 - Contagens feitas para o primeiro estudo a partir da base de dados para a obtenção do cálculo percentual. Dados próprios. ....	40
Tabela 6 - Valores percentuais para o segundo estudo. Dados próprios. ....	41
Tabela 7 - Valores obtidos para o estudo 1, após todas as análises feitas na toolbox Lógica Fuzzy usando o método de inferência Mamdani. Valores próprios. ....	63
Tabela 8 - Valores obtidos para o estudo 2 após todas as análises feitas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, para o tecido gordo. Dados próprios. ....	66
Tabela 9 - Valores obtidos para o estudo 2 após todas as análises feitas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, para o tecido gordo glandular. Dados próprios. ....	68
Tabela 10 - Valores obtidos para o estudo 2 após todas as análises feitas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, para o tecido denso. Dados próprios. ....	70
Tabela 11 - Valores obtidos a partir da base de dados mini-MIAS database of mammograms. Retirado de: <a href="http://peipa.essex.ac.uk/info/mias.html">http://peipa.essex.ac.uk/info/mias.html</a> .....	80



# Lista de Acrónimos

kVp - Peak kilovoltage

LF - Fuzzy Logic (Lógica *fuzzy*)

MIAS - Mammographic Image Analysis Society

MLO - Incidência Oblicua

CC - Incidência Cârniocudal

LEM - linha de eixo mamilar

CR - Computed Radiography

DR - Digital Radiography

FDA - Food and Drogs Administration

TAC - Tomografia Axial Computadorizada

TNM - Tumor, Linfonodo, Metástase

NANDA-I - North American Nursing Diagnosis Association - International

ARCH - Distorção Arquitetural

ASYM - Assimetria

CALC - Calcificação

CIRC - Massa Circunscrita Bem Definida

MISC - Massa Mal Definida

SIPC - Massa Espiculada

ROI - Região de Interesse



# Capítulo I

## Introdução

Os computadores e os seus sistemas informáticos são nos dias de hoje cada vez mais importantes, poderá mesmo dizer-se que em determinados casos são mesmo imprescindíveis, na prática de medicina. A situação do sector de saúde seja ele público ao privado necessita da constante atualização dos seus sistemas informáticos e é de extrema importância que este esteja em constante crescimento e inovação, pois só assim é possível chegar de maneira mais rápida e mais eficiente aos constantes novos problemas das ciências médicas.

A área da visão computacional e imagem médica são áreas muito importantes na vida humana e para estas os computadores e os sistemas informáticos são imprescindíveis para a obtenção de resultados. As imagens médicas desempenham um papel muito importante no diagnóstico médico, pois a partir destas é possível avaliar patologias. Para a obtenção deste tipo de imagens em exames mamários é necessário proceder-se a exames médicos, sendo os mais usados a TAC [1], a ecografia [2] e as muito conhecidas Mamografias [3]. As mamografias são os exames mais conhecidos por serem de baixo custo e pela sua radiação consideravelmente inferior quando comparada com exames de TAC, por este espeto são estes os exames mamários mais fáceis de obter para realização de estudos, podendo assim conseguir uma grande e vasta coleção de imagem e dados sobre as mesmas. Por todos estes aspetos, serão as mamografias os exames estudados neste trabalho. Para o estudo que se pretende, as imagens não serão usadas diretamente, serão sim usados todos os dados dos exames feitos a partir das mesmas, fazendo um estudo estatístico da probabilidade da existência de massas malignas ou benignas para determinadas características apresentadas pelas imagens.

A área do processamento de dados tem inúmeras aplicações em diversas áreas distintas, como a científica, a médica, a comercial ou mesmo militares. O estudo feito neste trabalho foca-se em obter resultados na área científica e médica. Todas estas aplicações resultam da interação entre pesquisas científicas e o desenvolvimento de novas tecnologias, tecnologias essas consideradas de alto padrão. O desenvolvimento contínuo do estudo destas

metodologias tem vindo a dar bons resultados no que diz respeito ao processamento de dados médicos [4].

Existem diversas maneiras para conseguir estudar e trabalhar dados médicos, e construir a partir destes bases estatísticas, no entanto para este trabalho, é utilizada a Lógica *Fuzzy*.

A Lógica *Fuzzy* é vista de maneira diferente dependendo do autor, Segundo Lotfi A. Zadeh (1965) [5], a Lógica *Fuzzy* propõe-se a expressar matematicamente formulações do pensamento humano em linguagem natural sem que a potência expressiva destas seja diminuída, já Earl Cox (1994) [6], afirma que combinando a impressão associada aos eventos naturais e ao poder computacional das máquinas, é possível obter sistemas de respostas inteligentes, robustas e flexíveis. Enquanto que Boaventura (2010) [4], diz que o conceito de precisão ou imprecisão, dependendo do contexto aplicado, expressar-se-á numericamente, indicando assim a probabilidade de uma inferência estar correta ou incorreta.

Em conjuntos tradicionais o conceito de pertinência de um determinado elemento a um determinado conjunto é bem definido, enquanto que num conjunto *fuzzy*, está associada a ideia de um grau de pertinência que varia consoante o elemento em questão. A pertinência de um determinado elemento em relação a um determinado conjunto deverá ser entendida como a intensidade com que este mesmo elemento está relacionado ao conjunto em causa. Segundo Klir & Folger (1988), a variação do grau de pertinência poderá fazer um elemento, pertencer muito ou pouco ou até mesmo não pertencer ao conceito representado pelo conjunto *fuzzy*, podendo assim adquirir qualquer um dos valores intermediários apresentados. Esta característica de subjetividade permite à Lógica *Fuzzy* atuar fortemente em problemas que tem como objetivo modelar conceitos imprecisos [7].

Com todo este desenvolvimento e todas estas técnicas implementadas foi possível criar na medicina condições para deteção de doenças de maneira mais rápida e menos invasiva, facilitando o trabalho do técnico de saúde e ajudando ao bem-estar do doente. Atualmente existe um grande interesse na criação e desenvolvimento de metodologias que permitam a análise automática de estruturas anatómicas, pois estas possibilitam obter além de uma resposta mais rápida e simplificada, obter ainda uma menor intervenção por parte do técnico de saúde [4].

Para conseguir proceder ao estudo das técnicas a cima referidas, são necessárias imagens médicas e os dados referentes às mesmas, estas dependendo do formato pretendido não são simples de obter. É necessário saber que tipo de imagem se adequa ao estudo que se pretende realizar, que neste caso, para este trabalho serão utilizadas mamografias, pelos motivos já explicados.

Para obter imagens médicas é necessário recorrer a um hospital e pedir autorização ao responsável pelas mesmas. Existe também a possibilidade de aceder a bases de dados através da internet, sendo necessário ter em conta se esta é ou não fidedigna. Muitas das vezes é através destas bases de dados que muitos estudos são suportados, pois qualquer estudo para ser posto em prática necessita de ser testado e são na maioria das vezes estes dados que são usados por não serem necessárias autorizações e alguns deles podem ser obtidos sem qualquer valor monetário. No entanto, existem também as bases de dados e imagens taxadas.

Na área médica a tecnologia permite a geração e disponibilização de informações digitais rapidamente. Hospitais e clinicas têm por vista a melhora contínua da qualidade do seu serviço e o atendimento aos seus doentes. Para isto estes centros procuram integrar os seus sistemas de informação seguindo à risca os padrões adotados mundialmente como os sistemas de informação hospitalar, conhecido de HIS (*hospital information system*), os sistemas de informação em radiologia, conhecidos de RIS (*radiology information system*) e sistema de comunicação e armazenamento de imagens, conhecido de PACS (*picture archiving and communication system*) [4].

As bases de imagens que são referias a cima são de extrema importância e uma mais-valia para todos, tanto os que necessitam de imagens para conseguir fazer estudos com os programas que estão a implementar, como para os profissionais de saúde que mais tarde vão usufruir desses mesmo estudos e programas para conseguir, com mais facilidade e rapidez, chegar a um diagnóstico do doente em causa. A mamografia é uma radiografia feita à mama, onde o diagnóstico é feito por imagem e o prognóstico de saúde e doença são encontrados através do uso de tecnologias de produção e análise de imagens.

O cancro da mama é visto como uma doença que se for diagnosticada precocemente poderá ser tratada, no entanto é uma doença que afeta um grande número de mulheres não

só em Portugal como em todo o mundo, como é mostrado no capítulo II. Por isto embora sejam desenvolvidos constantes estudos para desenvolvimento de técnicas capazes de tratar esta doença, nos dias de hoje tenta-se conseguir maneiras de em vez de tratar, arranjar formas de deteção precoce da doença e assim evitar que esta seja evolua e se torne fatal.

## 1.1- Objetivos

O principal objetivo deste trabalho é usar a Lógica *Fuzzy* como ferramenta de estudo de uma determinada doença, neste caso o cancro da mama, e fazer um estudo probabilístico percebendo as probabilidades existentes de um nódulo apresentado numa mamografia, ser benigno ou maligno, consoante as características apresentadas pelo tipo de tecido, anomalias existentes e tamanho do nódulo existente. Para isso, foi necessário um estudo aprofundado da Lógica *Fuzzy*, também conhecida como Lógica Difusa ou Nebulosa, bem como um domínio na compreensão da mesma no seu todo, sendo necessário entender como e onde esta pode ser utilizada, quais as suas regras e como é todo o seu funcionamento.

Na escolha da doença teve-se em conta a sua incidência, isto porque para fazer um estudo probabilístico é imprescindível ter uma população alvo consistente. O cancro da mama é uma doença, infelizmente, bastante comum entre as mulheres, daí ser uma boa doença a estudar, conseguindo assim um estudo mais assertivo. Apesar de serem usadas mamografias, como imagem médicas, estas servirão apenas de referência para se obterem os dados dos exames feitos a partir de cada imagem. O importante a retirar destas imagens são os estudos feitos nelas, as suas características e ter em conta que para este estudo apenas serão usadas imagens com nódulos.

O estudo deverá ser realizado na Toolbox Lógica *Fuzzy*, usando o método de inferência *Mamdani*, tudo isto deverá ser implementado através do Matlab.

## 1.2- Estrutura e organização da dissertação

Este trabalho é constituído por nove capítulos, está assim distribuído de forma a que seja simples entender todo o processo aqui estudado, de forma sucinta.

O capítulo 1 refere-se a introdução e objetivos do estudo feito.

O capítulo 2 e 3 é referente à anatomia, patologia e ao exame médico abordado, nestes capítulos encontram-se todo o conhecimento necessário para o entendimento do estudo feito, no que diz respeito à doença, às causas para o seu aparecimento e as características apresentadas pelo órgão afetado pelo cancro da mama.

No capítulo 4 é apresentada a teoria usada para o estudo à doença em causa, onde é explicada toda a base utilizada pela teoria da Lógica *Fuzzy*.

No capítulo 5 são apresentados estudos feitos por outros autores para perceber segundo estes, o que de importante já foi estudado recorrendo à Lógica *Fuzzy*, podendo ou não estar relacionado com doença aqui estudada.

No capítulo 6 e 7 são explicadas as características utilizadas para os dois estudos feitos e todos os comandos usados durante o processo.

No capítulo 8 são explicados todos os resultados obtidos com os estudos e a comparação entre resultados do estudo 1 e do estudo 2.

Por fim no capítulo 9 é feita a conclusão de todo o trabalho são ainda referidos alguns aspetos a serem considerados para melhorar num trabalho futuro.

## Capítulo II

### Patologia médica abordada

Em Portugal o cancro da mama é o tipo de cancro mais comum entre as mulheres, correspondendo à segunda causa de morte por cancro em mulheres [8]. Estima-se que uma em casa 10 mil mulheres vá ter cancro da mama ao longo da sua vida. Calculando-se ainda que anualmente surjam cerca de 5000 novos casos e que esta incidência tenda a aumentar nas próximas décadas. Os dados estatísticos [9] apontam que cerca de 1500 mulheres em Portugal morrem todos os anos devido a esta doença sendo esta a principal causa de morte antes dos 70 anos em mulheres [9].

Esta é uma doença com grande impacto na nossa sociedade, isto porque para além de ser bastante frequente e associada a uma imagem de grande gravidade, esta é uma doença que agride um órgão cheio de simbolismo na maternidade e feminilidade.

O cancro da mama não afeta apenas mulheres, em Portugal esta doença já afeta 1% de todos os cancros de mama são nos homens. Todos os dias se trabalha para que seja possível oferecer às pessoas com esta doença novos meios de tratamento, novos métodos de prevenção e novas causas da doença para que assim seja possível conseguir melhores resultados aumentando assim a hipótese de sobrevivência dos afetados [8].

Existem 4 tipos de cancro de mama, o Ductal, que é originário das células dos ductos mamários, o lobular, este é o menos comum e tem origem nas células dos lóbulos mamários. O *in situ* este é limitado ao local, podendo evoluir para um tipo invasor, por fim existe ainda o invasor, este é o rompimento do nódulo e invasão de outros tecidos [10].

O cancro da mama é uma doença progressiva, ensaios clínicos demonstraram que um rastreamento mamográfico e posterior tratamento desta doença, quando presente, causa uma redução significativa na mortalidade por cancro de mama. Em casos da existência da doença em familiares diretos, como mãe, pai ou irmão, este rastreio é recomendado dez anos antes

da idade na qual a doença foi diagnosticada, no entanto não será necessária antes dos 30 anos, exceto em indivíduos com história de câncros da mama hereditários [11].

Não existe necessariamente uma idade recomendada para se deixar de fazer o rastreamento mamográfico, deve ter-se sempre em conta a expectativa de vida da mulher e a presença de outras doenças associadas.

## 2.1 - Anatomia da mama

É essencial perceber a constituição da mama, todas as suas estruturas. É ainda fundamental conhecer as patologias que possam incidir na mesma, só depois de entender isto é possível estudar a doença a fundo através de imagem médica.

A mama é um órgão existente no sexo masculino e no feminino, no entanto as glândulas mamárias são mais desenvolvidas nas mulheres exceto quando existe um problema hormonal, nesse caso poderão haver homens com mamas desenvolvidas tal como nas mulheres [12]. Esta contém aspetos peculiares em cada fase da vida de uma mulher, começando na infância, passando na adolescência, idade adulta, na gravidez, na lactação e por fim na senilidade. A mama é composta por tecido fibroglandular, esta é a parte mais importante da mama, onde é produzido o leite, por gordura e revestida de pele, tendo ainda na região central a aréola e o mamilo. A aréola possui aspeto circular e uma coloração diferente do resto da mama. A coloração da aréola apresenta-se de maneira diferente em momentos diferentes da vida de uma mulher, podendo apresentar-se mais escura durante a gravidez. Já o mamilo é composto por tecido mamário correspondendo à parte mais saliente da mama, é o mamilo que é usado para amamentar, onde se encontram os ductos lactíferos, responsáveis pela eliminação do leite. A mama feminina encontra-se na parte superior do tronco apoiada sobre o músculo peitoral maior, este por sua vez está apoiado nas costelas e na musculatura intercostal, como é visível na figura que se segue, figura1 [13].

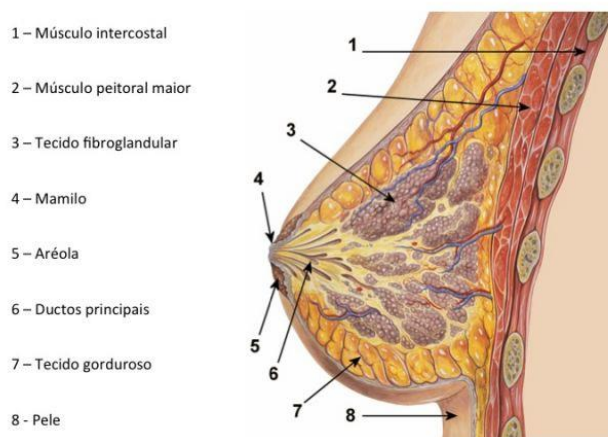


Figura 1 - Anatomia da mama, retirado de <http://oncomastologia.com.br/noticias/a-mama>. (requerido em Setembro de 2016).

Durante o ciclo menstrual e na gravidez ocorrem várias mudanças no tecido mamário devido às hormonas de estrogénio e progesterona. Apenas durante a gravidez os alvéolos aumentam, uma mulher não grávida tem os seus alvéolos pequenos e sólidos. Os alvéolos ao aumentar fazem com que as células sofram multiplicação rápida. O peito de uma mulher grávida é por norma maior e pendente, enquanto que numa mulher idosa, este é geralmente pequeno, devido à atrofia do tecido adiposo e glandular [12, 13].

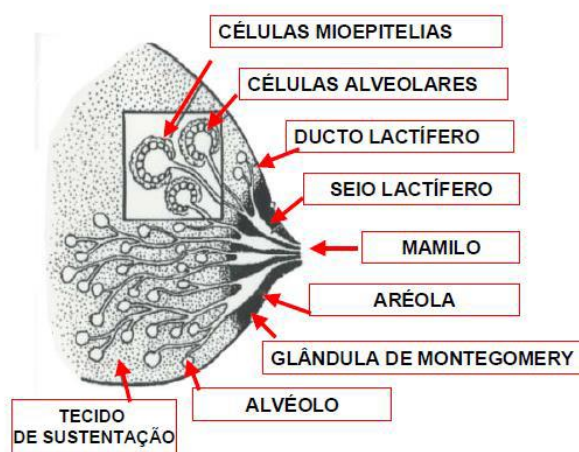


Figura 2 - Anatomia da mama aprofundada, retirado de <http://oncomastologia.com.br/noticias/a-mama> (requerido em Setembro de 2016).

Como é possível verificar na figura 2, a mama contém subdivisões, dividindo-se de 15 a 20 partes menores, estas são denominadas de lobos mamários [10]. Os lobos mamários por sua vez subdividem-se sendo constituídos por um conjunto de lóbulos, estes lóbulos são ainda constituídos por um conjunto de alvéolos. Este conjunto de subdivisões funciona como “um cacho de uvas”, um conjunto de pequenos ductos que se vão juntando formando ductos maiores até chegarem ao mamilo. Os ductos são revestidos por um epitélio especializado em proteger toda a sua superfície, é neste local onde se podem originar massas malignas que, são revestidas por células basais ou mioepiteliais. São estas que fazem a contração dos ductos auxiliando assim a ejeção do leite durante a lactação. Quando a célula maligna adquire a capacidade de ultrapassar a membrana basal além das células mioepiteliais, quer dizer que o tumor adquiriu características invasivas, tendo assim a capacidade de invadir os vasos linfáticos e os vasos sanguíneos e de se dispersar para os gânglios ou para outros órgãos [12, 13].

## 2.2 - Fatores epidemiológicos

Na maioria dos casos, os tumores que incidem na mulher são tumores esporádicos, é considerado que esta ocorrência se deva ao facto das decorrentes mutações somáticas predominantes associadas à exposição de fatores de risco ambientais. Fatores de risco como a idade, o sexo, a fatores hormonais e reprodutivos, ao estilo de vida, a hábitos alimentares e tabagismo, à etnia e claro à história pessoal e familiar, como a incidência direta desta doença no histórico familiar [14].

Fatores de risco como a idade e o sexo são dos mais importantes para o aparecimento do cancro da mama. As taxas de incidência do cancro de mama aumentam rapidamente até aos 40-50 anos, o que significa que com o envelhecer da população esta doença tem tendência a aumentar. Após os 50 anos este aumento ocorre de forma menos intensa, isto deve-se ao aparecimento da menopausa, que tem um grande impacto hormonal na mulher [13, 14].

Esta doença também se relaciona com o fator monetário, o cancro de mama afeta pessoas com menos dinheiro, pois tem menos acessos ao diagnóstico e tratamento da doença. É claro que fatores biológicos e genéticos contribuem também para o aparecimento da mesma. No caso da pessoa já ter apresentado cancro de mama anteriormente, esta tem um

risco acrescido de reaparecimento. Para cancro de mama *in situ* prévio, o risco em 10 anos é de 5% para o desenvolvimento de uma lesão invasora. Em mulheres com cancro de mama prévio, o risco é de 1% na pré-menopausa e de 0,5% na pós-menopausa, por ano [13].

No que diz respeito à história familiar de cancro da mama este é um fator de risco bastante importante, pois esta doença torna-se mais propícia quando a pessoa em causa tem familiares diretos, como pais e irmão, que tenham tido já contacto com esta doença. Nestes casos, quando existe um familiar direto que tenha cancro, pede-se aos familiares próximos que façam exames para se perceber se têm predisposição à doença. Existem vários casos de mães que apresentam com cancro de mama, e mais tarde as suas filhas desenvolvem a mesma doença [14].

A exposição prolongada de altas concentrações endógenas de estrogénio aumenta o risco de cancro de mama, ou seja na gestação e menopausa existe uma maior probabilidade de adquirir a doença. No que diz respeito aos fatores dietéticos, tem influência em todo o organismo, uma boa alimentação e uma vida o mais saudável que a pessoa consiga ter, diminui o risco de qualquer doença. No caso do álcool, este aparenta ser um risco acrescido, dependendo da quantidade que é ingerido, segundo *Fabiano Souza* [14], na sua tese de Doutoramento, ele afirma haver mais de quarenta estudos epidemiológicos a afirmar que existe uma associação entre o consumo excessivo de álcool e o aparecimento da doença [14].

## Capítulo III

### Exame médico abordado

No caso do cancro da mama, o principal exame feito é a mamografia, esta é considerada a melhor forma de deteção para este tipo de cancro. Com este exame estima-se que é possível conseguir reduzir a taxa de mortalidade entre 30 a 70%, estes são números que não foram ainda alcançados através de diagnóstico clínico nem mesmo através de medicação. Este é ainda um exame de fácil acesso e com custos reduzidos para o utente sendo ainda um método não invasivo com alta capacidade de reprodução, documentação e com tempo relativamente curto de execução, tornando-o assim o único método de rastreio apropriado [15].

A sensibilidade deste método está relacionado com a idade da mulher, este é muito mais sensível em mulheres jovens onde o tecido mamário apresentado é mais denso, ou então mulheres que fazem terapia de reposição hormonal [15].

Existem quatro padrões para identificar as diferentes densidades mamárias, estes podem ser vistos na figura 3. Poderão ser mamas com tecido predominantemente adiposo a), onde o tecido fibroglandular está presente em 25% da mama, mamas com tecido fibroglandular disperso b), que contém de 25 a 50% de tecido fibroglandular, mamas de densidade heterogénea c), que contém de 50 a 75% de tecido fibroglandular e por fim mamas extremamente densas d), que contém mais de 75% de tecido fibroglandular [16].

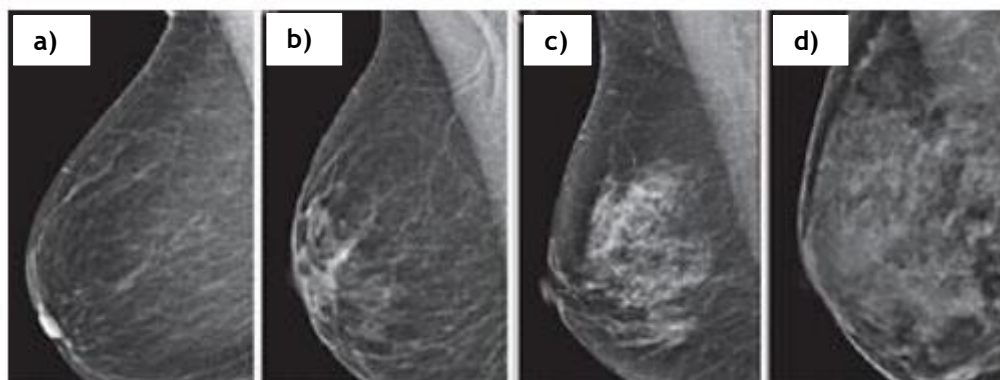


Figura 3 - Imagens demonstrativas das diferentes densidades mamárias, retirado de [16].

No momento da interpretação mamográfica, a densidade mamária é muito importante, isto porque a maioria dos nódulos tumorais apresentam densidades semelhantes à da água, confundindo-se assim com tecidos fibroglandulares, por este aspeto, as mamas gordurosas são mais fáceis de estudar [16].

Existe também dificuldade em analisar mamas densas devido a fatores físicos, como lesões mamárias que têm propriedades de atenuação de raio X similares às de tecidos glandulares e fibrosos, tornando a sua deteção mais difícil. A mama radiograficamente produz mais radiação espalhada, resultando um menor contraste nas imagens tornando assim muito complicado a exposição de forma ótima de todas as suas áreas [15, 16].

### 3.1 - Exame médico para o diagnóstico

A mamografia tem um papel muito importante na deteção de cancro da mama, isto porque se esta for detetada a tempo poderá salvar a vida da pessoa em causa, e mesmo evitar que esta seja submetida a determinados tratamentos, pois se for detetada no início da doença não serão necessários tratamentos tão agressivos [12, 14, 15].

A mamografia é um exame de imagem médica, este oferece informações sobre a morfologia, anatomia e patologia da mama. Este exame é normalmente usado para a deteção

e diagnóstico de cancro de mama e ainda para a avaliação de massas mamária e calcificações, este é um exame muito importante, pois esta doença se for detetada no seu estágio inicial poderá gerar o sucesso de todo um tratamento [14, 16].

A mamografia é uma forma particular de radiografia, neste exame trabalha-se com um tubo de gerador de raio X, que é basicamente formado por dois elétrodos que recebem uma tensão, na ordem de milhares de *volts*, de forma a gerar um campo elétrico. O eletrodo negativo, chamado de cátodo, quando aquecido pela passagem da corrente elétrica, liberta eletrões que são acelerados pelo campo elétrico até colidirem com o eletrodo positivo, chamado de ânodo, gerando assim um feixe de eletrões. A extremidade do ânodo, onde colidem os eletrões, é revestida de metal pesado, na colisão o que acontece é que 0,1% do feixe de eletrões é convertido em raios X e o restante é dissipado na forma de calor [15].

Os níveis de radiação são usados entre intervalos específicos, tendo como finalidade registar imagens do tecido mamário para diagnosticar uma eventual presença de estruturas que indiquem a existência da doença [15].

Para se conseguir um contraste adequado nas imagens de mamografia, os mamógrafos trabalham numa faixa de tensão mais baixa do que os equipamentos convencionais, com valores que variam os 20 a 35 kVp, enquanto que os equipamentos de raio X convencionais trabalham numa faixa de tensão geralmente de 40 a 120 kVp [15].

As mamas com tecidos densos são mais difíceis de avaliar do que mamas com tecidos gordos, isto porque a densidade numa mamografia, ou radiografia, aparece a branco. Assim ao existirem tecidos muito densos o que acontece é que ao ver a imagem obtida, se começa a confundir o tecido com as possíveis massas, parecendo tudo uma grande massa branca. Por este aspeto é extremamente difícil avaliar imagens mamárias de mulheres jovens, pois estas apresentam, por norma, tecido mamário mais denso [12, 14, 15].

O equipamento usado, o mamógrafo, contém um compartimento de compressão de acrílico, este compartimento tem como finalidade comprimir a mama. Esta compressão é importante, pois contribui para uma boa qualidade de imagem, dado que só assim se garante a imobilidade do doente durante todo o exame. Garantindo também desta forma a redução da espessura da mama, conseguindo diminuir a dose absorvida e o espalhamento dos fotões

de raio X no filme, durante o exame de mamografia. Todo este processo se for bem executado irá reduzir ao máximo o ruído da imagem, diminuindo ainda a sobreposição de imagens das estruturas da mama, permitindo com isto que lesões suspeitas sejam todas detetadas com mais facilidade e segurança [15].

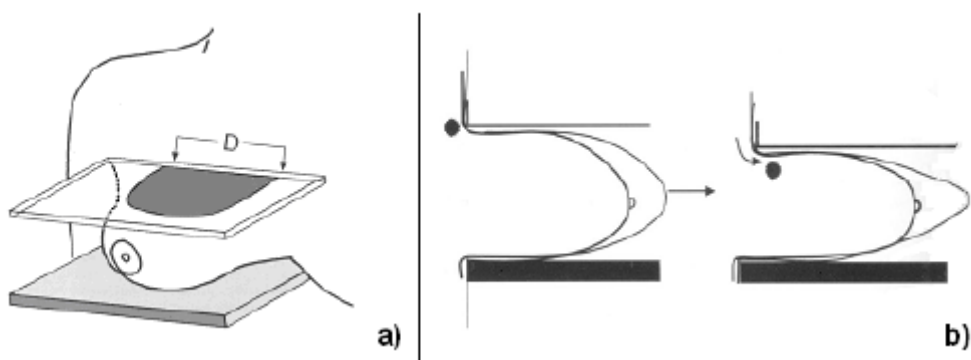


Figura 4 - Compartimento e compressão, retirado de [15].

Na figura 4 apresentada acima, pode ser visualizado na primeira imagem, a da esquerda a), a compressão da mama durante a mamografia e na segunda imagem, a da direita b), a compressão da mama facilitando a deteção de lesões com facilidade. Nesta figura consegue-se ver um nódulo, que é apontado por um ponto a preto. A imagem serve para nos mostrar, que apenas depois da compressão da mama, é possível ver este ponto, que antes disso ficava atrás da linha de imagem, não sendo detetado no diagnóstico. Por este aspeto é importante uma compressão bem feita, sendo só assim possível um diagnóstico completo e seguro [14, 15].

Com base na compressão, do conhecimento da anatomia e da mobilidade mamária, são usadas técnicas de posicionamento para assim ser mais fácil explorar a mama a fundo. As técnicas mais utilizadas são a incidência oblíqua (MLO). Esta técnica oferece uma melhor visualização de toda a mama numa única imagem. A incidência crâniocaudal, (CC), esta complementa a anterior conseguindo incluir todo o tecido posteromedial, o que a técnica anterior não consegue mostrar [15].

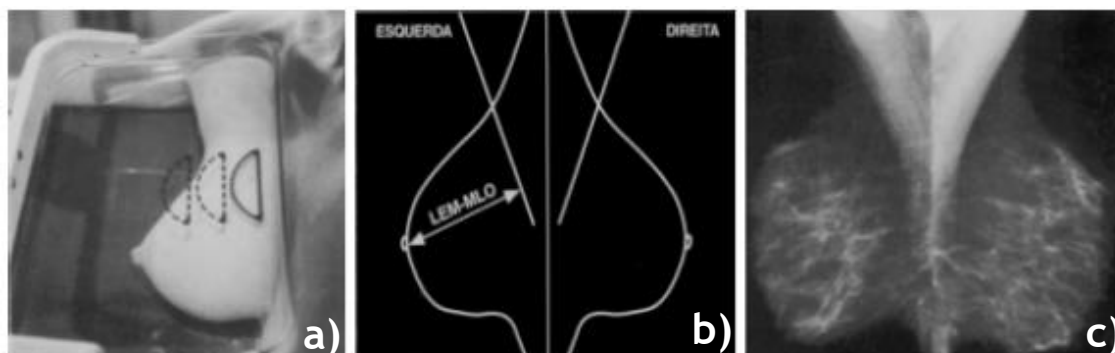


Figura 5 - Procedimento durante o exame, vista lateral, e imagem obtida após o exame, retirado de [15].

Na figura 5, são representadas três imagens, na primeira imagem, representada por a), está encontra-se a mamografia com projeção médio-lateral, oblíqua, na imagem do meio, representada por b), verifica-se o diagrama esquemático onde é indicada a visualização do peitoral maior na incidência MLO ao nível da linha do eixo da mama. Na última imagem, imagem de direita, representada por c), é mostrada a imagem resultante da mamografia com projeção MLO [14, 15].

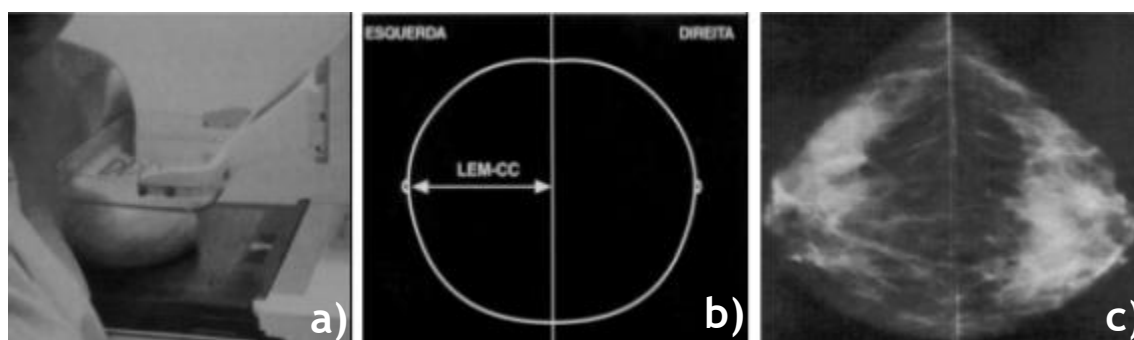


Figura 6 - Procedimento durante o exame, vista decima, e imagem obtida após o exame, retirado de [15].

Na figura 6, estão representadas três novas imagens, nestas a mama aparece a partir de cima, a primeira imagem, representada por a), mostra a mamografia com projeção crâniocaudal (CC), a imagem do meio, representada por b), mostra o diagrama esquemático

onde a linha feita por setas mostra a linha do mamilo até à parte posterior do filme na projeção CC. A linha deverá ser sempre 1 centímetro mais curta que a da LEM (linha do eixo mamilar) durante a projeção MLO, para poderem ser bem diferenciadas. Na última imagem, representada por c), mostra a imagem resultante da mamografia com projeção CC [12, 14, 15].

Há ainda outros dispositivos do mamógrafo, estes são o colimador, que é utilizado para evitar a radiação ionizante nas outras partes do corpo durante o exame, servindo ainda para a melhoria da imagem. Há ainda o chassi mamográfico, que se apresenta como um ecrã intensificador que é posicionado sobre o filme [14, 15].

A mamografia digital surgiu apenas no final da década de 80, em 1996, os *Computed Radiography* (CR) permitiram a troca da imagem de cassete anteriormente utilizada para este exame, por imagem digital. A partir de um sistema que efetua a digitalização dos chassis, no entanto, apenas em 2000 é que o primeiro equipamento mamográfico digital, *Digital Radiography* (DR) foi aprovado pelo departamento de saúde Norte Americano, *Food and Drugs Administration* (FDA) [14, 15].

O equipamento é semelhante ao anterior, com exceção do sistema de registo, onde o filme e o ecrã são trocados por sensores que absorvem a energia e a convertem em sinais elétricos, são então transformados em pulsos digitais e depois manipulados para a extração da informação. Uma imagem mamográfica é composta por pixéis, seja esta do mamógrafo digital ou da digitalização do filme mamográfico. As imagens deverão possuir resolução espacial geralmente de 300 a 600 DPI (pontos por polegadas), e uma resolução de contraste com uma quantidade de níveis de cinza suficientes para representar as diferenças entre os tecidos e estruturas que estejam incutidas na imagem, normalmente têm 12 ou mais bits [15].

A mamografia digital apresenta outros benefícios, como o de manipular a imagem, ampliando-a na região que se pretende estudar, pode alterar-se o brilho e o contraste facilitando a visualização das regiões que mais interessam, denominadas de regiões de interesse, (ROIs). Esta imagem é conseguida imediatamente após o exame, diminuindo assim o tempo de espera pelo processamento, sendo estas após conseguidas, enviadas

eletronicamente para o radiologista, estando assim dispensadas de operações manuais de transporte, colocações no negatoscópio e arquivamento analógico [14].

É importante salientar, que para o total uso dos recursos disponíveis deste tipo de diagnóstico é de extrema importância que os radiologistas sejam bem treinados, caso contrário não servirá de muito ter um bom sistema.

### 3.2 - Estadiamento

O estadiamento procura, com base em estudos, descrever aspetos do cancro, como a sua localização e se este está a afetar funções de outros órgãos do corpo para além da mama. Conhecer o estágio do tumor ajuda na definição do tipo de tratamento que deve ser feito para prever o prognóstico do doente.

O sistema de estadiamento utilizado para o cancro da mama é o sistema TNM da *America Joint Committee on Cancer* [13]. Este sistema utiliza três critérios para a sua avaliação, o próprio tumor, os linfonodos regionais ao redor do tumor e se o tumor se espalhou para outras áreas que não a mama.

A sigla TNM é a abreviatura para tumor, linfonodos e metástases, em que o “T” indica o tamanho do tumor primário e se este se disseminou para outras áreas, o “N” indica a descrição da existência da doença para os linfonodos regionais, ou se existem evidências de metáteses e o “M” indica a existência de presença de metástases noutras partes do corpo [13].

A tabela 1 que se segue mostra a divisão do estadiamento segundo o sistema TNM.

Tabela 1 - Divisão de grupos segundo o estadiamento TNM, retirado de [13].

Estadiamento TNM	Descrição
TX	Tumor principal não pode ser avaliado
T0	Nenhum sinal de tumor foi avaliado
Tis	Carcinoma <i>in situ</i>
T1	Tumor até 2cm de diâmetro
T2	Tumor entre 2 e 5 cm de diâmetro
T3	Tumor com mais de 5cm de diâmetro
T4	Tumor invadiu o tórax ou a pele

### 3.3 - Tipos de anomalias dos tecidos mamários detetados na mamografia

A mamografia é um exame útil para a avaliação de nódulos, calcificações e cistos mamários. Existe alguma similaridade inevitável entre forma, margens e densidade de muitas lesões tanto benignas como malignas. As lesões mais comuns nas mamografias são as microcalcificações e os nódulos. As microcalcificações são achados mamográficos encontrados entre 30 a 40% dos casos de cancro em lesões não palpáveis e podem apresentar sinal precoce de malignidade. [14, 15]

Uma análise de uma imagem de microcalcificações deverá incluir algumas observações importantes como o tamanho, o número, a forma, a densidade e a distribuição, estas são necessárias para que possa ser feito um bom e completo diagnóstico. Já os achados mamográficos são encontrados em 39% dos casos de cancro não palpáveis. Segundo *Boff*, *Ruaro* e *Schuh*, os nódulos definem-se como lesões que ocupam espaço visível mamograficamente em ao menos duas projeções diferentes [15]. A visibilidade destes nódulos vai ser melhor ou pior consoante a densidade mamária. Em mamas densas os nódulos podem ser pouco visíveis devido à semelhança da sua densidade com a do tecido mamário.

Na análise feita aos nódulos, deverão ter-se em conta alguns aspetos importantes, como o tamanho que no caso das lesões palpáveis é um parâmetro de importância relativa pois os nódulos diagnosticados apenas pela mamografia são normalmente de pequenas dimensões. A sua forma, estes podem apresentar-se de diversas formas, como regular, a lobulada e a espiculada. O seu grau de malignidade aumenta em função da ordem citada. Na figura seguinte, figura 7 estão representadas as várias formas que um determinado nódulo se pode apresentar [14, 15, 16].



Figura 7 - Diversas formas de massas que um nódulo pode ter, retirado de [15].

Deverá ainda ter-se em conta a margem, esta representa a relação do nódulo com as estruturas vizinhas, os limites têm de ser bem definidos, caso contrário poderão dar erro no diagnóstico. Na figura 8 que se segue, poderão ver-se as diferentes formas de margens que um nódulo pode apresentar [15].



Figura 8 - Diferentes margens para contornas de um nódulo, retirado de [15].

A densidade também é importante, os nódulos malignos são normalmente mais densos e às vezes de densidade média, mas raramente são de baixa densidade. Segundo o ACR (*American College of Radiology*), a margem entre uma lesão e o tecido circunvizinho, é um dos fatores mais importantes para se determinar se um nódulo é benigno ou maligno. As

margens da massa podem ser divididas em 5 tipos de contornos diferentes, o circunscrito, o obscurecido, o microlobulado, o mal definido e o espiculado [14, 15].

Numa massa circunscrita as margens formam uma transição nítida com o tecido circunvizinho, são quase sempre benignas, uma vez que algumas lesões com margens mal definidas parecem ser circunscritas a uma resolução baixa, pode ser valioso obter imagens com ampliação cujo ruído reduzido e nitidez melhorada possam aumentar a confiança da circunscrição, ou revelar uma margem bem menos definida, o que deve aumentar a preocupação [14, 15].

Numa massa obscurecida quando o tecido circunvizinho normal esconde a borda verdadeira da lesão, o profissional de radiologia, deverá decidir se a margem da lesão está obscurecida ou verdadeiramente mal definida devida a alguma infiltração, esta última aumenta o nível de preocupação [14, 15].

Uma massa microbulada reflete uma superfície irregular que pode ser produzida por um cancro de mama, embora os fibroadenomas e os quistos possam também ter margens idênticas. As saliências irregulares que nascem na borda de um tumor podem formar ondulações curtas na superfície da lesão quando esta é vista na mamografia. Estas são constituídas por uma massa lobulada com ondulações grandes [15].

Uma massa mal definida é a constituição da maior parte dos cancros que invadem os tecidos circunvizinhos, à medida que o cancro vai evoluindo, a massa vai tornando-se cada vez mais irregular. Isto produz uma margem verdadeiramente mal definida que deve gerar preocupação. A probabilidade de neoplasia maligna (massa maligna) é muito elevada em lesões com margens mal definidas devido ao confinamento de tecidos normais. Por fim nas massas espiculadas, o cancro de mama apresentado para estas lesões, contem por norma projecções fibrosas que se estendem desde a massa principal do tumor [15].

## Capítulo IV

### Lógica Fuzzy

A Lógica Fuzzy, também conhecida como Lógica difusa ou Nebulosa, veio auxiliar em diversos problemas de incerteza existentes na indústria, no comércio ou mesmo no dia-a-dia de toda uma sociedade. Esta lógica apesar de ter sido desenvolvida e utilizada com mais afinco a partir dos anos 60, já em 1920 era utilizada a sua teoria. Em 1920, o polaco de nome Jan Luasiewicz, apresentou pela primeira vez noções de lógica de conceitos vagos, onde é admissível um conjunto com valores não precisos. Já em 1927 era falado, por Heisenberg, no princípio da incerteza, este princípio serviu como alicerce principal para a teoria quântica. Mais tarde este princípio iria servir ainda de auxílio para o desenvolvimento da Lógica Fuzzy [17].

Foi em 1965 que o professor, Lofti Zadeh publicou um artigo de nome, *Fuzzy Sets* no *Journal Information and Control*, por esta altura a Lógica Fuzzy começaria a ser desenvolvida com mais profundidade. Earl Cox em 1994 vem afirmar que o que diferencia a Lógica Fuzzy da Lógica Booleana, também conhecida como Álgebra Booleana é a capacidade desta se conseguir aproximar do mundo real onde não existem somente respostas extremas. A Lógica Booleana é assim denominada em homenagem ao matemático George Boole. É composta por estruturas algébricas que conseguem captar as propriedades essenciais dos operadores lógicos e de conjuntos, conseguindo ainda oferecer uma estrutura para lidar com afirmações. Já a Lógica Fuzzy dá espaço para um meio-termo, não tem existir apenas extremos, pode existir o sim o não e o talvez. Consegue ainda a possibilidade de calcular o grau de aproximação da solução exata e assim deduzir algo que seja necessário [17, 18].

No livro de Gorge J.Klir e Bo Yuan, cujo título é *Fuzzy Sets And Fuzzy Logic*, é apresentada a principal diferença entre a Lógica Clássica e a Lógica Fuzzy, ou seja, enquanto a Lógica Clássica propõe que os valores encontrados sejam valores verdadeiros, a Lógica Fuzzy propõe que os mesmos sejam uma questão de grau [18].

A Lógica *Fuzzy* tentou adaptar sempre a sua teoria à realidade do mundo, apresentando sempre proximidade e adaptação aos problemas da atualidade. Devido a essa mesma proximidade, a Lógica *Fuzzy*, foi crescendo com o passar dos anos, tendo a sua grande expansão no Japão, principalmente, na década de 80. Esta é uma boa técnica para ser utilizada quando é apresentado um problema em que existe um grande grau de incerteza e a solução necessária passa por usar um modelo matemático capaz de considerar todos os aspectos e especificidade, que não possam ser considerados por lógicas tradicionais. No livro de *Earl Cox*, de nome, *The fuzzy systems handbook: a practitioner's guide to building, using, and maintaining fuzzy systems*, é referido que para estes casos em concreto é recomendada a Lógica *Fuzzy*, pois esta apresenta um modelo capaz de combinar a imprecisão associada aos eventos naturais e o poder computacional das máquinas produzindo assim sistemas de respostas inteligentes [18].

Segundo *Constantin Von Altrock*, no seu livro de nome, *"Fuzzy logic and neuro-fuzzy applications in busines and finance"*, um dos grandes objetivos da Lógica *Fuzzy* é aproximar-se o mais possível do raciocínio humano, conseguindo relacionar as informações obtidas, procurando assim respostas aproximadas aos problemas existentes [7].

Ainda no livro de *Earl Cox*, *"The fuzzy systems handbook: a practitioner's guide to building, using, and maintaining fuzzy systems"*, é explicado o funcionamento da Lógica *Fuzzy*, através de imagens. Estas imagens encontram-se na figura 9 que se segue.

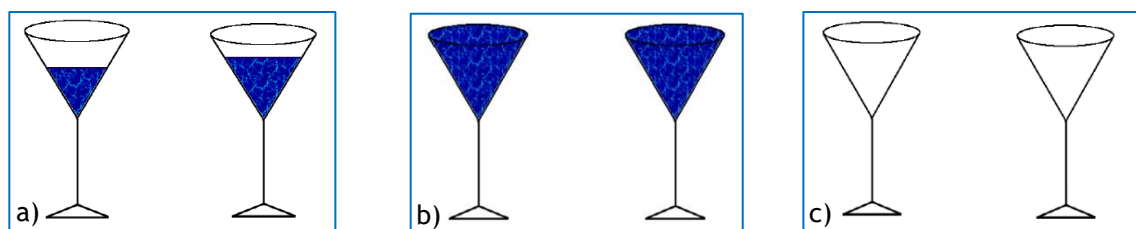


Figura 9 - a) copos com diferentes medidas, b) copos cheios, c) copos vazios, retirado de [7].

Na figura 9, encontram-se dois copos, a questão que se segue é se estes estão cheios ou vazios. Se esta pergunta for feita em relação a imagem c), tanto para o ser humano como para uma máquina devidamente programada, facilmente se chegará à conclusão que os copos se encontram vazios. A Resposta é bastante simples, tendo em conta que não tem qualquer líquido mostrado pela imagem. Já na imagem b), para a mesma questão a resposta seria o

oposto, pode verificar-se que os copos se encontram cheios. Para estas duas imagens não existe qualquer incerteza quanto à resposta a dar. No entanto não será tão evidente responder à mesma questão perante a imagem a). Para o ser humano é evidente que os copos se encontram com algum líquido, um contém mais que o outro, mas existe sempre um grau de incerteza. Perante esta dificuldade, existe a possibilidade de fazer várias afirmações sem que nenhuma delas esteja incorreta, como: o copo da direita encontra-se mas cheio que o da esquerda. O copo da direita encontra-se quase cheio. O copo da esquerda encontra-se praticamente vazio. É aqui que entra a Lógica Fuzzy, através de todas estas afirmações incertas embora corretas, foram criadas escalas de valores que procuram quantificar a quantidade de água de cada copo. Esta escala de valores é de extrema importância, pois irá permitir captar o grau de incerteza nas variáveis criadas permitindo gerar um modelo matemático [7, 17, 18].

Existem diversas áreas onde se aplica a Lógica Fuzzy, seja na área da engenharia, como na área da medicina, na economia e gestão, na investigação, entre outras. Para este estudo interessa focar a área de investigação na saúde. Pretende-se verificar alguns estudos feitos na área da saúde, tentando perceber o que já foi conseguido, o que está no momento a ser estudado e o que deverá ser, ainda, explorado para ajudar a ciência na saúde a evoluir, conseguindo prevenir e diagnosticando doenças antes de estas se tornarem graves necessitando de tratamentos mais “agressivos”. Muitos dos estudos feitos nesta área, são estudos de classificação, onde se tenta perceber se com a Lógica Fuzzy é realmente possível chegar a uma resposta mais rápida e precisa de um determinado problema.

## 4.1 - Conjuntos fuzzy

Os conjuntos clássicos contêm elementos que satisfazem propriedades precisas, estes são classificados em categorias, podendo ser enumeráveis ou não. Nestes conjuntos não deverão existir dúvidas quanto à categoria a que estes pertencem, o conceito de pertinência de um determinado elemento a um determinado conjunto tem de ser bem definido. Imaginando uma tarefa hipotética, classificar doentes de uma enfermaria como estando vivos ou mortos. O raciocínio subentendido no processo é simples e espera-se que a precisão dos dados recolhidos seja grande. Na lógica clássica, caberá a um profissional de saúde responsável pelo serviço, atribuir aos doentes vivos o valor “1”, e aos doentes mortos o valor “0”. Ao ser usada a linguagem matemática, o facto de um ser humano, ( $x$ ), ser membro do

conjunto “vivo”/”morto”, ( $A$ ), pode ser indicado através de uma função característica, ( $\mu_A(x)$ ), onde o valor de  $x$  pertence ou não a  $A$  [19]. Ou seja:

$$\mu_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \notin A \end{cases} \quad (1)$$

Ao contrário dos conjuntos tradicionais/clássicos, os conjuntos *Fuzzy*, são capazes de incorporar tanto o conhecimento objetivo a partir dos dados numéricos, como o conhecimento subjetivo a partir de informações imprecisas, conseguindo assim suportar os modos de raciocínio aproximado e conseguindo respeitar os critérios subjetivos. OS conjuntos *Fuzzy* conseguem implementar funções com a finalidade de representar os seus elementos, que assumem valores reais entre  $[0, 1]$  [20]. Ou seja, a ideia de um conjunto *fuzzy* está associada a um grau de pertinência  $\mu$ , em que este varia de acordo com o elemento em questão. Sendo assim, de acordo com esta variação, de grau de pertinência, um determinado elemento pode pertencer muito ou pouco e até mesmo não pertencer a um determinado conjunto *fuzzy*, podendo assim assumir qualquer um dos valores intermédios [19-21].

## 4.2 - Teoria dos conjuntos *fuzzy* e a teoria de probabilidade

Deve ser bem entendida a distinção existente entre teoria dos conjuntos *fuzzy* e a da probabilidade. Na teoria da probabilidade o acontecimento é claramente definido e a dúvida reside na sua ocorrência ou falta dela, após o acontecimento essa dúvida desaparecerá. No caso da teoria dos conjuntos *fuzzy* é feita uma manipulação de limites imprecisos, ao invés de restritos, usando ainda um grau de pertinência como a possibilidade de um determinado elemento ser membro do conjunto *fuzzy*, podendo também esta ser descrita como teoria das probabilidades. Ou seja, no caso de se estar a falar numa caixa de bolas brancas e pretas, e se pergunta qual a probabilidade da primeira bola retirada ser branca. Na teoria das probabilidades, a probabilidade poderá ser calculada, quando se deseja sortear na caixa composta por bolas brancas e pretas, uma bola de cor específica, ao terminar o sorteio, a incerteza desaparecerá. No entanto se em vez de bolas pretas e brancas, a caixa contivesse bolas em vários tons de branco e vários tons de preto, a pergunta teria de ser reformulada,

pois quando se tirasse a bola da caixa não se iria conseguir perceber que tipo de cor seria a bola retirada, para esta questão será usada a probabilidade com conjuntos *fuzzy* [20].

Para sistemas que utilizam Lógica Fuzzy, o processamento de informação consiste normalmente em operações que são realizadas sobre os seus conjuntos. As operações principais são as de união, de intersecção, de completação e de diferença entre conjunto *fuzzy*. Já em operações básicas de união, intersecção e complemento são normalmente definidas em função dos operadores *máximo* (*max*) e *mínimo* (*min*) [18, 22].

### 4.3 - Variáveis Linguísticas

Define-se como variável linguística aquela cujos valores dão nomes de elementos de conjuntos Fuzzy. Uma variável linguística admite como termos apenas expressão linguísticas, como por exemplo, “muito grande”, “muito baixo”, “pequeno”, “aproximadamente”. Imaginado a velocidade de um determinado veículo, a variável linguística pode assumir nomes como, “lenta” (*slow*), “média” (*medium*), ou “rápida” (*fast*). Na figura 10 apresentada em abaixo é demonstrado o referido no texto, em que é dado os menores valores à velocidade lenta, os valores médios à velocidade média e os elevados à maior velocidade atingida, são ainda atribuídos valores de pertinência que variam entre 0 e 1. A intersecção das variáveis é importante isto porque como é mostrado na figura 10, uma velocidade de 40km/h, é considerada sem dúvidas nenhuma baixa, uma velocidade de 55 km/h considerada média e uma velocidade de 70 km/h é considerada rápida. No entanto existem velocidades entre estas, como a de 50 e a de 65 km/h que está entre as velocidades baixa/média e média/alta. Estas velocidades dependendo do seu grau de pertinência poderão ser consideradas baixa, média ou alta, estando mais perto do 0 ou do 1 respetivamente [21].

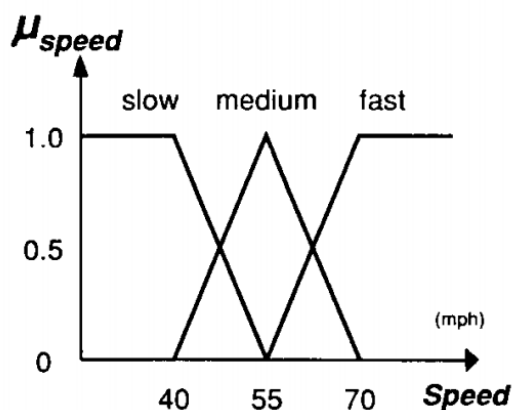


Figura 10 - Varias linguística de velocidade de um determinado veículo, retirado de [21].

A principal função de uma variável linguística é fornecer, dentro de uma função pertinência, uma maneira sistemática para uma caracterização aproximada de fenômenos complexos ou mal definidos onde os valores poderão ser sentenças de uma linguagem especificada [21].

#### 4.4 - Regras linguísticas

Na sua forma mais comum, as regras linguísticas são do tipo condição consequência, como por exemplo, “*se < antecedente >, então < consequente >*”. Nestas regras, um conjunto de condições descrevendo uma parcela observável das saídas do processo é associado a consequência que irá manter ou levar o processo às condições de operações desejadas [20, 21, 23].

Para que a regra seja considerada linguística, os termos antecedente e consequente, devem conter uma ou mais cláusulas linguísticas. Quando satisfeito o antecedente, mesmo que parcialmente, é determinado através de um mecanismo de inferência durante o seu processamento. Assim o consequente representa um conjunto de ações ou consequências [7, 17, 18, 21].

A intenção destas regras é representar o conhecimento a partir de regras nas quais as condições das implicações são dadas a partir de um conjunto de termos linguísticos associados às variáveis de entrada do processo e as consequências ou saídas expressam-se de um modo similar para cada variável. As regras de um sistema *fuzzy* são expressas em forma de proposição, como:

*SE Banda 1 é alta < conitivo 1 > Banda 2 é alta, ENTÃO x é Classe A < conectivo 2 >*

*SE Banda 1 é média < conitivo 1 > Banda 2 é média, ENTÃO x é Classe B*

Onde Banda 1 e Banda 2 são variáveis de entrada,  $x$  é uma variável de saída, Alta é um valor linguístico com a respetiva função de pertinência definida nas bandas 1 e 2. Médio é um valor linguístico com a sua respetiva função pertinência definida nas bandas 1 e 2. A Classe A e B, são valores linguísticos com as suas respetivas funções pertinência definidas em  $x$ . Por fim o conectivo 1 e 2 são os operadores lógicos escolhidos para se expressar a inferência desejada. O conectivo 1 está a ser utilizado no processo de agregação da regra, já o conectivo 2 está a ser usado no processo de combinação das regras. A primeira regra usada após o “Se” é também chamada de premissa e a segunda parte, depois do “Então”, é a conclusão ou consequente da implicação. [7, 21]

## 4.5 - Funções pertinência

A função pertinência associa os elementos numéricos reais que estão contidos no intervalo de  $0 \leq x \leq 1$  aos elementos  $x \in A$ , deduzindo desta forma o grau de pertinência do elemento  $x$  no conjunto  $A$ . A pertinência de um determinado elemento a um determinado conjunto deverá ser entendida como a intensidade com que o elemento em causa está relacionado ao conjunto [21]. A definição de pertinência é muito subjetiva, considerando que um conjunto poderá apresentar diversos resultados em diferentes usuários e em diferentes contextos. Para além de definir o conceito, esta função representa ainda os limites do conjunto. Por isto é importante mapear corretamente a forma de variação e os limites de um conjunto fuzzy através da definição adequada da função pertinência [23].

Uma definição pertinência pode adquirir vários formatos, no entanto as funções triangulares, trapezoidais e gaussianas são as mais comuns para modelar as distribuições dos dados *fuzzy*.

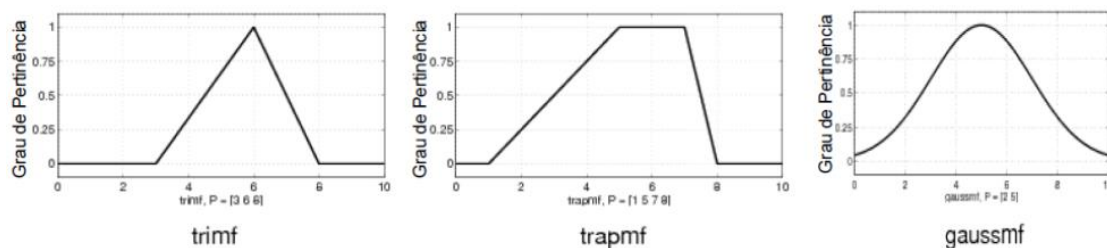


Figura 11 - Formatos de funções de pertinência mais utilizadas, retirado de [7, 21].

Na figura 11 estão representadas as funções mais utilizadas, a *trimf*, é a função triangular, a *trapmf* é a função trapezoidal e a *gaussmf* é a função gaussiana. No caso de conjunto de dados concentrados numa determinada região, as melhores funções de pertinência serão as triangulares e as trapezoidais. No caso de conjuntos mais dispersos então são usadas as gaussianas. No entanto estas regras não são exatas, devido à dependência da distribuição dos dados e a natureza do problema, a forma de obtenção dos parâmetros que definem uma função pertinência, varia de acordo com a aplicação, não havendo, uma metodologia formal para as suas elaborações [23].

Uma função trapezoidal é expressa através de quatro parâmetros, estes descrevem os quatro vértices de um trapézio. Entre os vértices, dois possuem pertinência igual a 1 e os outros dois possuem uma pertinência igual a 0. Já uma função triangular, existe apenas 1 parâmetro e este possui um grau de pertinência igual a 1 [7].

## 4.6 - Sistema de inferência

Um sistema de inferência *fuzzy* é um sistema baseado em regras *fuzzy*. Os sistemas *fuzzy* estimam funções com descrição parcial do comportamento do sistema, em que especialistas podem fornecer o conhecimento para encontrar a resolução do problema ou então este poderá ainda ser deduzido a partir de dados de entrada - saída do sistema. Pode

então dizer-se que os sistemas *fuzzy* são sistemas baseados em regras que utilizam variáveis linguísticas para executar processos de tomada de decisão [7, 21].

Na arquitetura padrão de um sistema *fuzzy* é possível verificar os componentes que fazem parte desse sistema. Este possui cinco componentes básicos, a interface de fuzzificação, a base do conhecimento, a base de dados/base de regras, o procedimento de inferência e a interface de defuzzificação. Todos os passos estão representados na figura 12.

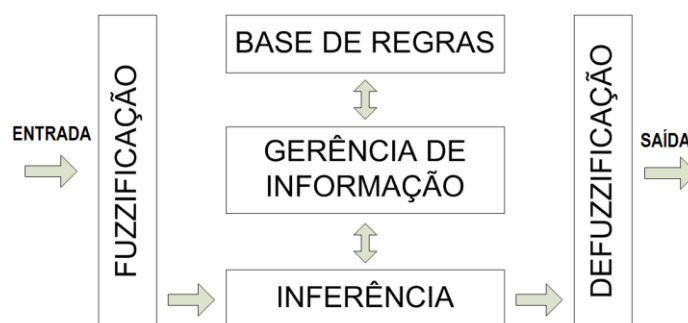


Figura 12 - Arquitetura padrão de um sistema fuzzy, retirado de [21].

A interface de fuzzificação recebe valores de entrada, condiciona os valores recebidos ao universo de discurso normalizado e fuzzifica-os, transformando-os em elementos de conjunto que possam tornar-se instâncias de variáveis linguísticas. A interface de fuzzificação está diretamente ligada às variáveis, aos seus respectivos universos de discurso e ainda ao que cada variável pode assumir.

Nesta fase, é recebido um valor de discurso sendo depois retornados os graus de pertinência aos respectivos conjuntos *fuzzy*. No entanto o tipo e a quantidade de funções de pertinência usados num sistema dependem de alguns fatores, sendo estes, a precisão, a estabilidade, a facilidade de desenvolvimento, entre outros. Assim a fim de colaborar na construção das funções pertinência para a descrição de uma determinada entrada, é importante a atuação de um especialista na área do objeto que poderá ser então modelado.

O procedimento de inferência é o componente do sistema em que as proposições são definidas e depois examinadas paralelamente. É nesta fase que são realizadas as operações com conjuntos *fuzzy* propriamente ditas [7, 21]. O procedimento de inferência processa os dados fuzzificados de acordo com as regras *fuzzy* definidas com o objetivo de deduzir as

ações de controlo. Neste componente as regras são utilizadas para se obter a relação *fuzzy* previamente modeladas, tendo bastante importância, por fornecer a saída a partir de cada entrada, e é deste que depende o sucesso do sistema inteiro. Os conjuntos de entrada, relativos aos antecedentes das regras e o de saída, referentes aos consequentes, podem ser definidos previamente ou gerados automaticamente a partir dos dados [18, 21]. Para realizar estas operações, existem métodos de inferência *fuzzy*, como o método de *Mamdani* e o método de *Tkagi-Sugeno-Kang* [21].

O gerenciador de informação obtém da base de regras as normas aplicáveis para entradas específicas. Cada antecedente  $\langle Se \rangle$  tem um determinado grau de pertinência e cada ação da regra  $\langle Então \rangle$  representa a saída da regra. Durante uma avaliação a intensidade referida pela regra é calculada com base nos valores dos antecedentes sendo então são indicadas pelas referidas regras [23].

A base de regras ou base de dados é composta por um conjunto de condições  $\langle SE \rangle$ , onde se usa o E, o OU e o NÃO, uma conclusão  $\langle ENTÃO \rangle$  e uma conclusão opcional  $\langle SENÃO \rangle$ . Estas regras são aplicadas com base num processo intermédio chamado de propagação [22]. Estas poderão ser elaboradas por especialistas do domínio, em forma de sentenças linguísticas, definindo assim a estratégia e o controlo do sistema que irá constituir um aspeto fundamental no seu desempenho. Alternativamente ao uso dos especialistas podem ser usados métodos automáticos e extração de regras a partir de dados numéricos que são particularmente uteis em problemas de classificação [7, 18, 23].

A interface de defuzzificação, processa as saídas *fuzzy*, do processamento anterior em saídas não-*fuzzy* através de um escalamento. Isto é feito como objetivo de compatibilizar os valores normalizados com os valores dos universos de discurso reais das variáveis. Neste processo ocorre a etapa de relação funcional entre regiões *fuzzy* e o valor esperado. Por outras palavras, como o algoritmo de controlo faz com que o processo das variáveis linguísticas de entrada resultem num valor da variável de saída. O processo de defuzzificação consiste em selecionar um valor numérico específico que represente o resultado *fuzzy* da variável de saída que é produzido pelo conjunto de regras *fuzzy*. Existem vários métodos de defuzzificação, uns mais conhecidos que outros, os mais usados são, o centróide, a média de máximos, o critério máximo, o método da altura e a barra vertical [18, 21, 23].

## Capítulo V

### Estudos feitos com Lógica fuzzy na área da saúde

Foram já realizados vários estudos utilizando a lógica *fuzzy*. Na área da saúde, os estudos desenvolvidos baseados em dados estatísticos são os mais encontrados, embora estejam a aumentar os estudos feitos em imagem médica, são com os de base estatística os mais usados por profissionais de saúde. Para o desenvolvimento deste trabalho foi necessário perceber bem todos os pontos da doença, mas não só, foi ainda imprescindível perceber como funciona a Lógica *Fuzzy* e onde pode esta ser aplicada da melhor forma, para se obterem os melhores resultados. Para isto foi necessário estudar trabalhos já desenvolvidos, apesar de alguns deles não estarem relacionados diretamente com a doença, o que é pretendido é perceber o princípio de funcionamento da Lógica *Fuzzy* para então aplicar o necessário obtendo assim os melhores resultados no estudo que se pretende realizar.

Um dos artigos estudados tem como título “*Aplicação do Sistema de Interferência Nubloso na Determinação de Medicina Preventiva no Sistema de Saúde Complementar*” e foi desenvolvido na Faculdade IBMEC no Brasil, por Igor Medeiros e Maria Augusta Machado. Este trabalho tem como objetivo direcionar os custos médicos de operadoras de planos de saúde mediante o estudo de variáveis, previamente, selecionadas por grupos médicos de atuação no mercado. Esta medida irá influenciar as decisões pós entrevista qualificada através do mapeamento do risco dos novos clientes. Mostra ainda a atual situação do mercado de saúde privado tendo como premissa os escassos recursos da saúde pública. Este estudo ficou bastante limitado, pois não foi possível analisar todas as variáveis necessárias para uma melhor definição dos problemas de saúde. Foi então concluído que existem várias vias para a geração de novos canais de distribuição dos custos médicos através de técnicas administrativas que podem diminuir o custo das operadoras em pesquisa, gerando assim mais eficiência. Quanto à capacidade de se medir o risco dos novos clientes demonstrou-se muito eficiente e a sua automação permitiu uma maior velocidade no tempo de ação [22]. Este artigo prova a versatilidade da lógica *fuzzy*, pois apesar do objetivo do trabalho estar ligado à medicina, a aplicação da Lógica *Fuzzy* foi em economia.

O artigo seguinte tem como autores, Graciele Paraguaia Silveira, Laércio Luis Vendite, Laércio Carvalho de Barros e como título: “*Nomograma com Lógica Fuzzy para o Câncer de Próstata - NFCP*”. Teve como objetivo inicial desenvolver um modelo matemático *Fuzzy* capaz de prever o estado patológico do cancro na próstata. O modelo usado também chamado de “*nomograma*” consiste num sistema baseado em regras *Fuzzy*, que combina dados pré-cirúrgicos do doente, levando-se em conta um conjunto de regras linguísticas, elaboradas a partir dos *nomogramas* de *Kattan et al*, já existentes. A saída do sistema irá fornecer a resposta à pergunta, em que estado da doença se encontra, localizado, localmente avançado, ou então metastizado. Foram feitas simulações com os dados reais, (doentes do Hospital UNICAMP), comparando-os com as probabilidades de *Kattan*. Por fim foi feita a validação dos resultados obtidos da análise por meio e análise da Curva de ROC (*Receiver Operating Characteristic*), usando a base de dados do hospital referido [24].

O estudo que se segue é sobre tumores superficiais na bexiga, tem como título: “*Modelagem Fuzzy para Predizer os Riscos de Recidiva e Progressão de Tumores Superficiais de Bexiga*” e autores, Kenia D. Savergnini, Laércio L. Vendite e Wagner E. Matheus, sendo este um estudo feito na UNICAM no Brasil. Para este estudo foram elaborados dois modelos matemáticos para auxiliar o especialista na tomada de decisão. A ferramenta usada foi a teoria dos conjuntos *fuzzy*, devido à sua capacidade de lidar com incertezas inerentes aos conceitos médicos. No primeiro modelo, estado, tamanho e grau do tumor são combinados para obter o risco de recidiva de um tumor superficial de bexiga. No segundo modelo são também combinados o estado, tamanho, grau do tumor e ainda o *carcinoma in situ* para conseguir obter os riscos de progressão de tumores superficiais. Foram feitas simulações para os dois modelos e a partir do banco de dados e das possibilidades/probabilidades, foi possível gerar a probabilidade do caso real de cada conjunto *fuzzy* de saída. Foi concluído que como era de esperar os doentes com *carcinoma in situ*, apresentam maiores chances do tumor progredir, desta forma os resultados do novo sistema baseado em regras *fuzzy*, foram mais otimistas e mais coerente com a realidade dos doentes [25].

O próximo artigo foi um estudo feito em São Paulo, no Brasil e teve como autores Maria Lopes, Heimar Marin, Neli Ortega e Eduardo Massad, e tem o título: “*Modelo Baseado em Lógica Fuzzy para o Diagnóstico Diferencial de Enfermagem em Alterações na Eliminação Urinária*”. O objetivo deste trabalho foi apresentar um modelo baseado em Lógica *Fuzzy* para o diagnóstico diferencial de alterações na eliminação urinária, para isto foram considerados os diagnósticos de enfermagem aprovados pela *North American Nursing Diagnosis Association*

- *International* (NANDA-I). Para o desenvolvimento do modelo foi usada a composição máximo-mínimo. Este modelo foi testado em 195 casos, este foi capaz de determinar o diagnóstico em tal concordância comum painel e três especialistas em 79,5% dos casos. O modelo conseguiu diagnosticar 19% dos casos concordância parcial com o painel de especialistas, somente em 1,5% dos casos é que o modelo apresentou discordância com o painel de especialistas. Este estudo concluiu que o modelo apresentado é simples e apresenta um bom desempenho. No entanto os autores recomendam que sejam feitos mais testes antes que este modelo seja usado amplamente como suporte na decisão clínica [26].

“*Um Sistema Fuzzy para Auxilio ao Diagnostico da Densidade Mineral Óssea*”, foi um estudo feito por Frederico Silva, André Moraes, Graçaliz Dimuro, Marilton Aguiar e foi realizado na Universidade Católica de Pelotas no Brasil, que veio propor um modelo computacional para o cálculo e visualização da densidade mineral óssea a partir de imagens convencionais de raios-x. A ideia foi obter um intervalo de valores resultante do processamento e segmentação das imagens médicas e aplica-las como valores de entrada num sistema *fuzzy* intervalar modelado especificamente para essa mesma necessidade. A proposta feita faz parte do projeto M-PIE (*Medical Information Processing Enviroment*) e tem como objetivo principal fornecer o processamento e simulação de aspetos dinâmicos baseados em informações médicas. As pesquisas em desenvolvimento nesta área indicam ser bastante promissoras quanto ao uso do sistema de diagnóstico auxiliado por computador, como a segunda opinião no processo de diagnóstico por imagem. Os autores chegaram então à conclusão que o modelo atingiu resultados significativos e promissores. Este projeto foi ainda elogiado pelo ministério da saúde, pois apresenta a premissa de software livre e objetiva-se a disponibilizar ferramentas resultantes para os hospitais e postos de saúde sem quaisquer custos, o que facilita o diagnóstico a quaisquer pessoas independentemente da sua classe social [27].

Os autores Sami Dhahbi, Waalid Barhoumi e Ezzeddine Zagrouba, no seu artigo “*Breastcancer diagnosis in digitized mammograms using curvelet moments*”, propõem um método de extração de característica, baseado em *curvelet transform* e *moment theory* para assim conseguirem a descrição completa da mamografia. Para este estudo foram usadas mamografias com tumores malignos e benignos, através da base de dados *Mammographic Image Analysis Society* (mini-MIAS). Os resultados obtidos provaram que o método usado é eficiente e superior aos restantes métodos usados para a mesmo efeito [28].

O estudo que se segue foi realizado com o objetivo de provar que os sistemas informatizados são imprescindíveis na medicina, que sem eles seria impossível ter as condições que existem no sistema de saúde e mesmo no estudo e desenvolvimento da ciência. Com base neste pensamento, M. Mahfouf, M. F. Abbod e D. A. Linkens, desenvolveram um estudo de nome: “*A survey of fuzzy logic monitoring and control utilisation in medicine*”, onde reviram vários trabalhos, de modo a explicar como a utilização do controlo da Lógica Fuzzy, pode melhorar todos os sistemas, no desenvolvimento das ciências médicas e o desenvolvimento de estudos ligados à investigação. [29]

O.W. Samuel, M.O. Omisore e B.A. Ojokoh, usaram a febre tifóide como estudo, como esta envolve muitas variáveis para poder ser estudada, o uso a Lógica Fuzzy seria ótima neste problema. O artigo tem de nome: “*A web based decision support system driven by fuzzy logic for the diagnosis of typhoid fever*”, onde foi proposto um sistema baseado em apoio à decisão web, (WBDSS) impulsionado por Lógica Fuzzy, para o diagnóstico. Foi usado o método de sistema de inferência mais usual, com fuzzificação, motor de inferência e desfuzzificação. O objetivo seria determinar o melhor diagnóstico possível. Para o estudo foram usados os dados médicos de doentes obtidos a partir do Centro Federal de Medicina, Owo, Ondo Estado-Nigéria, durante um período de 6 meses. Os resultados foram adequados às exigências pré-definidas por um médico da área. Foram obtidos bons resultados, o sistema proposto teve uma eficácia de 94% na obtenção de um diagnóstico preciso [30].

O estudo seguinte defende que o cancro como doença do século deverá ser diagnosticado para que não sejam necessários tratamentos agressivos, e para isso será necessário criar uma base de dados capaz de criar um prognóstico e verificar as probabilidades de um aparecimento desta doença o mais precocemente possível. O estudo tem como nome: “*Cancer risk analysis by fuzzy logic approach and performance status of the model*” e foi usada a Lógica Fuzzy, onde desenvolveram um modelo de inferência com o método da *Mamdani* com os riscos dos tipos de cancro e diagnósticos preliminares, por fim foram cruzados valores e assim calculadas as maiores probabilidades da incidência desta doença consoante as variáveis usadas. Os seus autores são, A. Yilmaz e K. Ayan [31].

Este é um estudo que cruza a medicina tradicional e a complementar, onde explica que para o bom funcionamento a medicina tradicional tem de ser ajustada para assim cumprir todos os critérios necessários e para isto terá de se usar técnicas inteligentes de modo a conseguir avaliar de forma mais abrangente todos os problemas da medicina. Tem como

autores K.J. Kumar e S.P. Bhatnagar e como título, “*Fuzzy Rule Based Systems in Classification of Terminalia Chebula Fruits of Complementary and Alternative Medicine*”. Neste estudo procurou-se com base na Lógica Fuzzy estudar um medicamento tradicional usado na Índia e assim verificar se este seria ou não de qualidade. Os resultados indicam que a Lógica Fuzzy como ferramenta moderna pôde, sem problemas alguns executar com precisão o estudo pretendido e apoiar assim a decisão do profissional de saúde no controlo de qualidade automatizado [32].

Para finalizar é falado de um estudo que pretendeu utilizar conceitos médicos usados por profissionais de saúde com anos de experiência e modela-los para serem usados numa base de dados. O artigo foi feito por M. Kwiatkowska e K. Kielan e tem como título “*Fuzzy logic and semiotic methods in modeling of medical concepts*”. As variáveis usadas são várias doenças, sintomas, anomalias entre outros problemas que já foram identificados antes por médicos em consultas de rotina. Com estes dados é então criado através da Lógica Fuzzy, um sistema de inferência capaz de avaliar todos os dados obtidos verificando as várias probabilidades para diversas doenças ou anomalias, podendo assim ajudar a obter um diagnóstico mais rápido e por vezes até mais eficaz. O método obteve bons resultados [33].

Após esta esta revisão bibliográfica dos estudos aqui referidos é de prever que este método, Lógica Fuzzy, será sem dúvida uma boa escolha para o estudo que é proposto. Estes artigos não se referem todos a uma única doença, nem têm todos o mesmo foco de estudo, no entanto todos eles seguem o mesmo caminho na resolução de um problema, obtendo bons resultados nos objetivos a que se propõem. Este trabalho segue o mesmo princípio de funcionamento no que diz respeito à Lógica Fuzzy, onde é usado o método de inferência *Mamdani* e assim estudada a probabilidade de ocorrência de uma determinada doença, consoante um conjunto de características, procurando assim tal como os estudos referidos neste capítulo encontrar os melhores resultados possíveis para o estudo proposto.

## Capítulo VI

### Características usadas no estudo

A primeira fase do estudo foi perceber quais as melhores características a serem estudadas, e quais dessas deveriam ser cruzadas para se obterem os melhores resultados. Para isto foi necessário perceber o que poderia influenciar uma massa maligna ou benigna a ser gerada. Após esta reflexão decidi então começar por estudar o tipo de tecido e a anomalia encontrada no tecido mamário e então avaliados os resultados.

Para proceder ao estudo foi usada a base de dados miniMIAS, (*The Mammographic Image Analysis Society Digital Mammogram Database*), esta é uma organização de grupos de pesquisa no Reino Unido (UK) interessados na compreensão de mamografias. Esta base de dados tem como objetivo a disponibilização de imagens para facilitar e auxiliar pesquisas e desenvolvimentos ligados ao estudo de mamografias de rotina. As mamografias disponibilizadas são de diversos doentes e contém severidades malignas e benignas, sendo todas elas previamente diagnosticadas por um profissional de saúde no departamento de Radiologia da UFF, Hospital Universitário Antónia Pedro e pelo IRSA.

A base de dados usada contém 322 imagens e todas as características das mesmas. Das 322 imagens foram usadas apenas 121, pois as outras 201 imagens não têm qualquer anomalia, logo não poderiam ser estudadas segundo os mesmos parâmetros. Estes dados podem ser verificados nos anexos, na tabela 11, páginas 67, 68, 69.

A tabela 2 que se segue diz respeito às contagens da tabela 11 em anexo, onde se encontra na primeira coluna o tipo de tecido, na segunda o tipo de anomalia do nódulo, que está descrita em siglas da mesma forma que foi retirada da base de dados e que significa:

- ARCH - Distorção Arquitetural
- ASYM - Assimetria
- CALC - Calcificação

- CIRC - Massa Circunscrita Bem Definida
- MISC - Massa Mal Definida
- SIPC - Massa Espiculada

Na terceira e quarta coluna encontram-se os valores obtidos nas contagens das imagens com nódulos benignos e malignos encontradas para cada tipo de anomalia e para cada tipo de tecido, e por fim na quinta e última coluna encontram-se a junção das imagens malignas e benignas para cada tipo de anomalia do nódulo e para cada tipo de tecido.

Tabela 2 - Contagens feitas para o primeiro estudo a partir da base de dados para a obtenção do cálculo percentual. Dados próprios.

Tipo de tecido	Anomalia do nódulo existente	Número de imagens benignas	Número de imagens malignas	Total de imagens contadas
<b>Gordo</b>	ARCH	4	2	6
	ASYM	1	3	4
	CALC	2	4	6
	CIRC	12	2	14
	MISC	3	5	8
	SPIC	2	3	5
<b>Gordo Glandular</b>	ARCH	2	4	6
	ASYM	2	2	4
	CALC	3	4	7
	CIRC	6	2	8
	MISC	3	2	5
	SPIC	4	3	7
<b>Denso</b>	ARCH	3	4	7
	ASYM	3	4	7
	CALC	8	7	15
	CIRC	3	0	3
	MISC	1	1	2
	SPIC	5	2	7

A tabela 2 é importante pois é a partir dela que se conseguem obter os resultados das percentagens para cada tipo de tecido e mais tarde usar estes valores no estudo que se pretende realizar.

Na tabela 3 que se segue, verificam-se os valores percentuais dentro de cada tipo de tecido para cada tipo de anomalia do nódulo existente. Para a obtenção destas percentagens foram usados os valores da tabela 2 e a seguinte fórmula simples de cálculo de percentagem:

$$\frac{\text{número de imagens} * 100}{\text{total de imagens contadas}} = \% \quad (2)$$

Após o cálculo, foram obtidos os valores tabelados que se seguem na tabela 3.

Tabela 3 - Valores percentuais para o primeiro estudo. Dados próprios.

Tipo de tecido	Anomalia do nódulo existente	Percentagem de imagens benignas (em %)	Percentagem de imagens malignas (em %)
<b>Gordo</b>	ARCH	<u>66.66</u>	33.33
	ASYM	25	<u>75</u>
	CALC	33.33	<u>66.66</u>
	CIRC	<u>85.71</u>	14.28
	MISC	37.5	<u>62.50</u>
	SPIC	40	60
<b>Gordo Glandular</b>	ARCH	33.33	<u>66.66</u>
	ASYM	<u>50</u>	50
	CALC	42.85	<u>57.14</u>
	CIRC	<u>75</u>	25
	MISC	<u>60</u>	40
	SPIC	<u>57.14</u>	42.85
<b>Denso</b>	ARCH	42.85	<u>57.14</u>
	ASYM	42.85	<u>57.14</u>
	CALC	<u>53.33</u>	46.66
	CIRC	<u>100</u>	0
	MISC	50	<u>50</u>
	SPIC	<u>71.42</u>	28.57

Os valores percentuais descritos acima, foram usados no primeiro estudo, onde se verificou a probabilidade de que um determinado tecido, juntamente com um determinado tipo de anomalia do nódulo, teria para ser maligno ou benigno.

Os valores obtidos com estas tabelas serão usados em cruzamento de dados durante a utilização da toolbox Lógica Fuzzy.

O estudo seguinte irá cruzar mais uma característica, o tamanho do nódulo. Pretende-se com isto perceber se caso se juntar mais uma característica ao cruzamento de dados a probabilidade irá aumentar ou diminuir para cada severidade, maligna e benigna. O estudo desta característica teve de ser formulado de maneira diferente, não são como as características anteriores, em que são usados nomes e grupos para as identificar, estes são valores numéricos, dispersos e em grande quantidade. É primeiro necessário agrupa-los antes de os usar, para isso foi usado o sistema de estadiamento TMN, já referido no capítulo III na tabela 1. Ao verificar a mesma é possível ver que apenas interessa dividir os tamanhos da anomalia em 3 grupos, estes encontram-se no T1, no T2 e no T3, pois sabe-se que o estudo foi feito apenas para imagens com nódulos, logo terão de existir tumores, e que eles variam entre 1,3cm e 19,7cm e que estas massas/nódulos estão ainda concentradas no tecido mamário não invadindo outros tecidos. Para descrever estes grupos foram apenas usados números, descritos na tabela 4 que se segue, sendo assim mais simples identifica-los no estudo.

Tabela 4 - Divisão de grupos para o estudo. Dados próprios.

Divisão de grupos segundo o estadiamento TNM	Divisão de grupos segundo o estudo a realizar
T1	1
T2	2
T3	3

É ainda importante referir que existem valores desconhecidos, se a tabela 11 que se encontra nos anexos nas páginas 67, 68, 69, for analisada, é possível verificar valores de tamanhos de alguns nódulos em falta, estes poderão não ter sido medidos corretamente ou serem pequenos demais para serem medidos, pois se a massa/nódulo foi identificada, significa que esta existe para quem estudou a imagem, logo terá de existir uma medida, no entanto como não é dada, para este estudo eu considereei que para valores desconhecidos estes estariam associados a massas/nódulos inferiores a 2cm. A falta destes valores poderá ser verificada na imagem 59, imagem 216, imagem 245 e imagem 233.

Após perceber a melhor maneira de agrupar estes valores procede-se de novo à contagem da tabela 11 em anexos, obtendo de a tabela 5 que se segue. Nesta é já possível encontrar as contagens referentes aos tamanhos dos nódulos existentes, aqui referidos como tamanho 1, 2 e 3, como já explicado na tabela 4.

Para esta contagem foi usado o mesmo tipo de pensamento incluindo agora o tamanho dos nódulos, ou seja, para cada tipo de tecido existente (denso, gordo ou glandular), foram divididas as seis anomalias do nódulo encontradas, sendo então contadas as imagens para cada grupo de tamanhos (1, 2 e 3), dentro de cada anomalia, para imagens benignas e malignas.

Tabela 5 - Contagens feitas para o primeiro estudo a partir da base de dados para a obtenção do cálculo percentual. Dados próprios.

Tipo de tecido	Anomalia do nódulo existente	Número de imagens benignas			Número de imagens malignas			Total de imagens contadas
		1	2	3	1	2	3	
Gordo	ARCH	0	2	2	0	0	2	6
	ASYM	0	1	0	0	3	0	4
	CALC	1	1	0	2	2	0	6
	CIRC	4	6	2	0	1	1	14
	MISC	1	2	0	0	2	3	8
	SPIC	0	2	0	1	0	2	5

<b>Gordo Glandular</b>	ARCH	0	1	1	0	1	3	6
	ASYM	0	1	1	0	2	0	4
	CALC	2	1	0	2	1	1	7
	CIRC	0	3	3	0	1	1	8
	MISC	0	2	1	0	1	1	5
	SPIC	0	3	1	0	1	2	7
<b>Denso</b>	ARCH	0	3	0	0	1	3	7
	ASYM	0	2	1	0	1	3	7
	CALC	5	3	0	1	5	1	15
	CIRC	0	1	2	0	0	0	3
	MISC	0	1	0	0	1	0	2
	SPIC	1	2	2	0	1	1	7

Após a contagem procedeu-se de novo ao cálculo das percentagens, onde foi usada de novo a fórmula (1) já descrita a cima, mas para este caso não é usada para imagem benigna e maligna mas sim usada para cada grupo de tamanhos dentro das imagens malignas e benignas, depois multiplicada por 100 e por fim dividida pelo total de imagens do grupo em causa, por fim é então calculada a percentagem para cada um dos valores de medida do nódulo, descritos na tabela 5, onde se obteve a tabela 6 que se segue.

Tabela 6 - Valores percentuais para o segundo estudo. Dados próprios.

Tipo de tecido	Anomalia do nódulo existente	Percentagem de imagens benignas (em %)			Percentagem de imagens malignas (em %)		
		1	2	3	1	2	3
<b>Gordo</b>	ARCH	0	<u>33.33</u>	33.33	0	0	<u>33.33</u>
	ASYM	0	25	0	0	<u>75</u>	0
	CALC	16.66	16.66	0	<u>33.33</u>	<u>33.33</u>	0
	CIRC	<u>28.57</u>	<u>42.85</u>	<u>14.28</u>	0	7.14	7.14
	MISC	<u>12.5</u>	25	0	0	<u>25</u>	<u>37.5</u>
	SPIC	0	40	0	<u>20</u>	0	<u>40</u>
<b>Gordo</b>	ARCH	0	<u>16.66</u>	16.66	0	16.66	<u>50</u>

Glandular	ASYM	0	25	<u>25</u>	0	<u>50</u>	0
	CALC	28.57	<u>14.28</u>	0	<u>28.57</u>	14.28	<u>14.28</u>
	CIRC	0	<u>37.5</u>	<u>37.5</u>	0	12.5	12.5
	MISC	0	<u>40</u>	20	0	20	<u>20</u>
	SPIC	0	<u>42.85</u>	14.28	0	14.28	<u>28.57</u>
Denso	ARCH	0	<u>42.85</u>	0	0	14.28	<u>42.85</u>
	ASYM	0	<u>28.57</u>	14.28	0	14.28	<u>42.85</u>
	CALC	<u>33.33</u>	20	0	6.66	<u>33.33</u>	<u>6.66</u>
	CIRC	0	<u>33.33</u>	<u>66.66</u>	0	0	0
	MISC	0	50	0	0	<u>50</u>	0
	SPIC	<u>14.28</u>	<u>28.57</u>	<u>28.57</u>	0	14.28	14.28

A tabela 6 é referente ao segundo estudo, esta será também ela usada, tal como a tabela 3 descrita neste capítulo para o primeiro estudo, para o cruzamento de dados durante a utilização da toolbox Lógica Fuzzy. Pretende-se perceber com estes dois estudos, em que num se usam apenas duas características e no seguinte se usam três, qual destes conseguirá obter melhores valores. Ou seja qual dos dois estudos conseguirá chegar mais perto do valor ótimo para cada severidade, maligna ou benigna. Tudo isto, usando a toolbox Lógica Fuzzy.

Nas tabelas 3 e 6, de valores percentuais encontra-se algumas percentagens a negrito e sublinhado, esta diferenciação serve para fazer as regras de inferência na toolbox. Os valores usados para as regras serão os que estão a negrito e sublinhado, pois não existe necessidade de se usar todos os valores obtidos. Para a toolbox é apenas necessário atribuir a maior probabilidade de ocorrência, neste caso a maior quantidade de imagens dos parâmetros estudados, sendo por isso usadas apenas as percentagens de maior valor. O programa irá concluir que se a percentagem atribuída é referente a uma determinada severidade, maligna ou benigna, a percentagem em falta será o contrário que não tem significado estatístico, logo não será retido para avaliação, sendo assim mais simples de retirar conclusões dos resultados finais, pois não são criados tanto dados desnecessários, é obtido um resultados mais “limpo”.

## Capítulo VII

### Lógica *Fuzzy* utilizando o método de inferência *Mamdani*

O método de inferência de *Mamdani* é o método mais usado na Lógica *Fuzzy* por ser considerado um método simples de ser entendido e simples de ser aplicado, é ainda intuitivo e bastante adequado para linguagem humana, o que significa que este método irá aceitar variáveis como, lento, rápido, mediano, ou frio, quente, ameno. Este foi dos primeiros sistemas de controlo a ser construído usando a teoria dos conjuntos *fuzzy*, foi proposto por *Ebrahim Mamdani* em 1975 tendo como objetivo controlar uma máquina a vapor combinada com uma caldeira, segundo um conjunto de regras de controlo linguístico. O método de *Mamdani* é baseado no trabalho “*Fuzzy algorithms for complex systems and decision processes*” elaborado por *Zadeh* em 1973 [23].

O método de inferência *Mamdani* usado na *toolbox* Lógica *Fuzzy* tem como base a teoria dos conjuntos nublados, também conhecida como a teoria dos conjuntos *fuzzy* ou ainda *fuzzy set*, esta é a teoria que mais se adequa para o estudo de “imperfeições de informação”, ou seja esta é capaz de capturar informações vagas, como a linguagem humana e converter essa linguagem para um formato numérico [34, 35].

A imagem 13 que se segue diz respeito ao diagrama típico de um sistema de processamento de conhecimento *fuzzy* que utiliza o modelo de inferência de *Mamdani*. Este inclui módulos de interface que transformam as variáveis de entrada em conjuntos *fuzzy* equivalentes e depois as variáveis *fuzzy* geradas em variáveis numéricas proporcionais, adequadas ao sistema de atuação existente [34, 36].

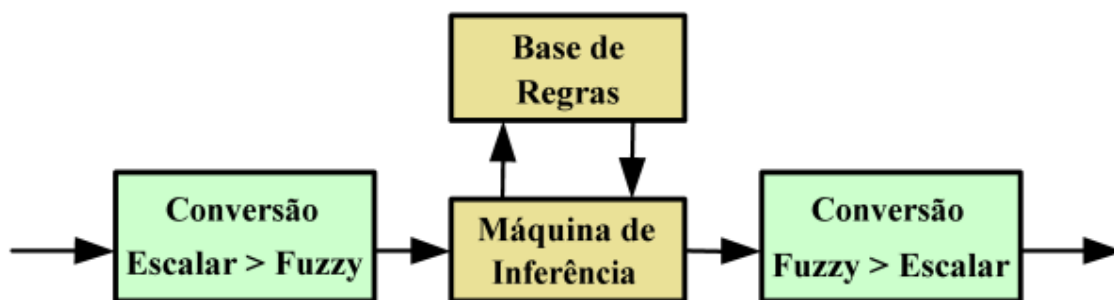


Figura 13 - Diagrama típico de um modelo de inferência de Mamdani, retirado de [36].

Na figura 13, no diagrama apresentado, a máquina de inferência recebe os valores *fuzzy* que vêm da interface de entrada, (o primeiro quadro, à esquerda), e irá processar as regras existentes na base de conhecimento gerando assim um conjunto *fuzzy* de saída, (o primeiro quadro, à direita), para o módulo de interface correspondente, segundo todas as regras que lhe foram atribuídas [36].

A regra tradicionalmente utilizada para o processamento de inferências com o modelo *Mamdani* é chamada de inferência máx - min, esta é utilizada para operações de união e de interseção entre conjuntos da mesma maneira que foi formulada por *Zadeh*, através dos operadores máximos e mínimo [34].

Durante o processo de conversão escalar - *fuzzy*, os antecedentes de cada regra são processados através de interseção *fuzzy* entre o grau de pertinência das entradas atuais nos termos primários definidos em cada uma e com isto gerar um grau de pertinência para cada regra de produção, este processo é conhecido como fuzzificação [34, 36].

A conversão *fuzzy* - escalar irá transformar as informações qualitativas, em informações quantitativas. Este é um processo de especificação o chamado processo defuzzificação [34, 36].

Na figura 14, é possível verificar um exemplo de um sistema de inferência *Mamdani*.

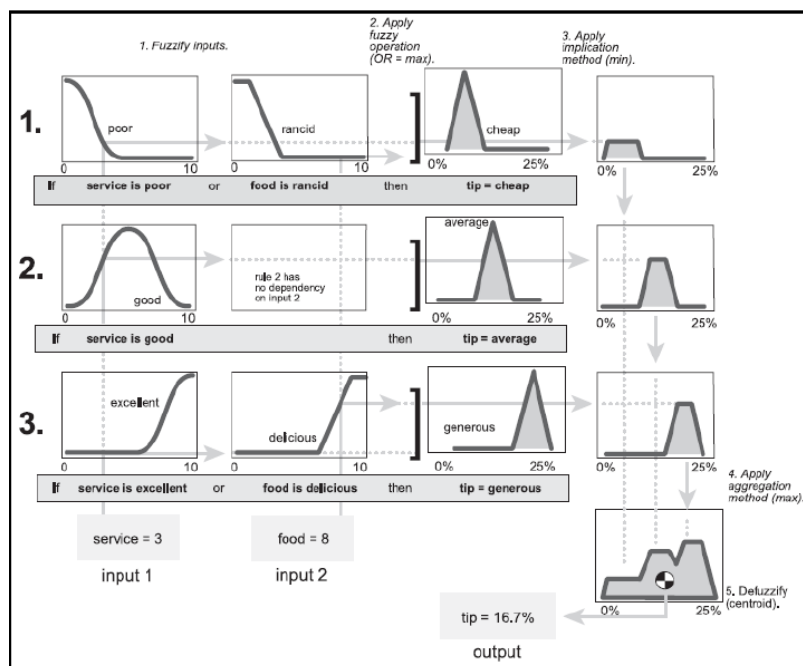


Figura 14 - Exemplo de um sistema de inferência Mamdani, retirado de [34].

A figura 14 mostra o tipo de gorjeta dada num determinado restaurante, consoante o serviço prestado e a qualidade da comida servida. Na primeira coluna é apresentado o serviço prestado, onde pode ser verificado que este foi dividido em três parâmetros, fraco, bom ou excelente, em que o fraco se encontra mais próximo do 0 e o excelente mais próximo do 10. Na segunda coluna é apresentada a qualidade da comida servida, esta foi dividida em dois parâmetros, fraca e deliciosa, em que a fraca se encontra mais próximo do 0 e a deliciosa mais próximo do 10. A terceira coluna mostra a gorjeta oferecida dependendo do serviço e comida, se fraco a gorjeta é mais próxima do zero se mediana a gorjeta é mediana e se bom a gorjeta fica mais próxima do 25% que é o máximo para este estudo. O exemplo dado é, para um serviço avaliado em 3 e uma comida avaliada em 8 verifica-se que a gorjeta é avaliada em 16,7% o que significa que se encontra mais perto do 25% logo é uma gorjeta generosa [34].

## 7.1 - Comandos usados

Para começar o estudo usando a Lógica Fuzzy com o método de inferência Mamdani, foi necessário criar as tabelas com as características, já descritas já no capítulo VI, os valores

descritos nas mesmas são imprescindíveis para ser possível a criação de regras de inferência e assim obter valores. Estes valores são apresentados em percentagem, isto porque os valores que são implementados na toolbox necessitam de estar compreendidos entre 0 e 1, caso contrário a toolbox não os aceitará. Por este aspeto foi necessário arranjar uma maneira de usar os dados atribuindo-lhes valores que depois no cruzamento de dados seguindo as regras de inferência, que estes tivessem algum significado para depois conseguir fazer a interpretação dos dados.

Após a conclusão das tabelas e todos os valores percentuais calculados é necessário ir ao Matlab e abrir a toolbox, para isso é usado o comando que se encontra no quadro seguinte:

```
fuzzy
```

Este comando é o primeiro passo a ser executado no Matlab, existem outras maneiras de chamar a toolbox no entanto esta é a maneira mais simples e rápida de o fazer.

Após colocar o comando fuzzy, este irá abrir a toolbox, apresentando assim um quadro com o seguinte aspeto, (figura 15):

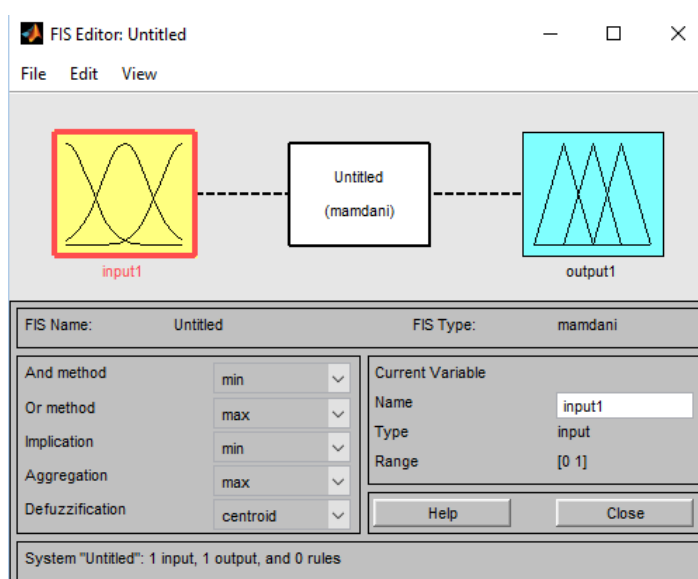


Figura 15 - Aspeto da toolbox Lógica Fuzzy usando o método de inferência Mamdani, imagem próprias.

Este quadro é a base da toolbox, é possível identificar três “caixas” no sistema apresentado, são estas caixas que vão conter todos os valores de entrada, de saída e as regras para o cruzamento de dados. A caixa amarela onde se encontra escrito *input* diz respeito às variáveis de entrada, estas servirão como ponto de partida para o cruzamento entre dados depois da avaliação segundo as regras impostas. O nome desta caixa pode ser modificado para que seja mais fácil identificar as variáveis que são introduzidas, para mudar o nome basta escrever o que se pretende no retângulo branco do lado direito, onde está escrito *Name*.

A caixa branca que se encontra ao centro, é referente às regras que serão criadas para o estudo ser orientado da maneira que se pretende. O nome desta caixa é o nome de todo projeto, estando ainda identificado o método usado, neste caso é o método de inferência *Mamdani*. Por fim a caixa azul diz respeito às variáveis de saída, representada como *output*. É nesta caixa que se encontram os valores que queremos cruzar no final do estudo, o que queremos saber, imaginando que o estudo estava relacionado com alturas e o que queremos saber é se as alturas se encontram mais concentradas nas pessoas altas ou nas pessoas baixas, esta caixa vai ter o valor das alturas. Nesta caixa é também possível modificar o nome para a identificar com mais facilidade, este processo é feito da mesma forma que na caixa de *input*, basta carregar com o rato do computador, em cima da caixa azul e mudar o nome no retângulo branco à direita.

Os pontos que se seguem, 7.1.1 e 7.1.2., dizem respeito aos comandos usados em dois estudos separados. Estes foram realizados com o objetivo de perceber o que é necessário para obter melhores resultados no final do estudo.

Foi então considerado que seria interessante perceber, se usando uma maior quantidade de variáveis de entrada (*inputs*), se obteriam melhores ou piores resultados (mais ou menos assertivos), ou até mesmo se os resultados seriam idênticos independentemente da quantidade de *inputs*. Ou ainda apesar de se obterem diferentes valores, que estes sejam tão parecidos que não tenham significância estatística para se conseguir dizer que um estudo é mais assertivo que o outro.

Normalmente quanto mais *inputs* oferecemos ao estudo mais específico ele se torna, logo deveria dar melhores valores, ou seja, pegando no exemplo dado anteriormente do restaurante, na figura 14, é dado um *input* para o serviço, outro para a comida, caso fosse

dados mais um *input*, por exemplo para o vinho, outro para os cuidados de higiene do restaurante, o resultado final teria de ser um cruzamento por todos estes valores de entrada, tornaria o estudo muito mais específico, obtendo valores mais próximos do que queremos. No entanto pode não se verificar isto, porque se os valores não tiverem grandes diferenças entre eles, podem não chegar para alterar algo no resultado final. É isto que pretendo perceber cruzando os valores no final de cada estudo.

### 7.1.1 - Primeiro estudo

No primeiro estudo são usados dois *inputs*, como pode ser vista na figura 16, um deles é o tipo de tecido e o outro é a anomalia do tecido.

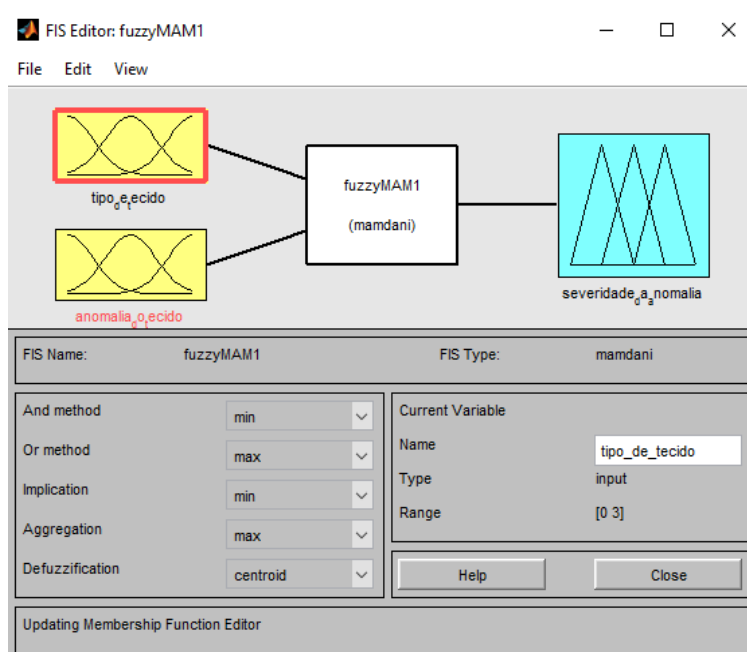


Figura 16 - Usando a toolbox Lógica Fuzzy usando o método de inferência Mamdani, para o primeiro estudo, imagem próprias.

Para serem colocados os *inputs* é necessário carregar duas vezes em cima de cada caixa amarela, depois disso irá aparecer um novo quadro com o nome, *membership function*

editor, para se poder começar a colocar o que se pretende fazer no estudo que pode ser visto nas próximas figuras.

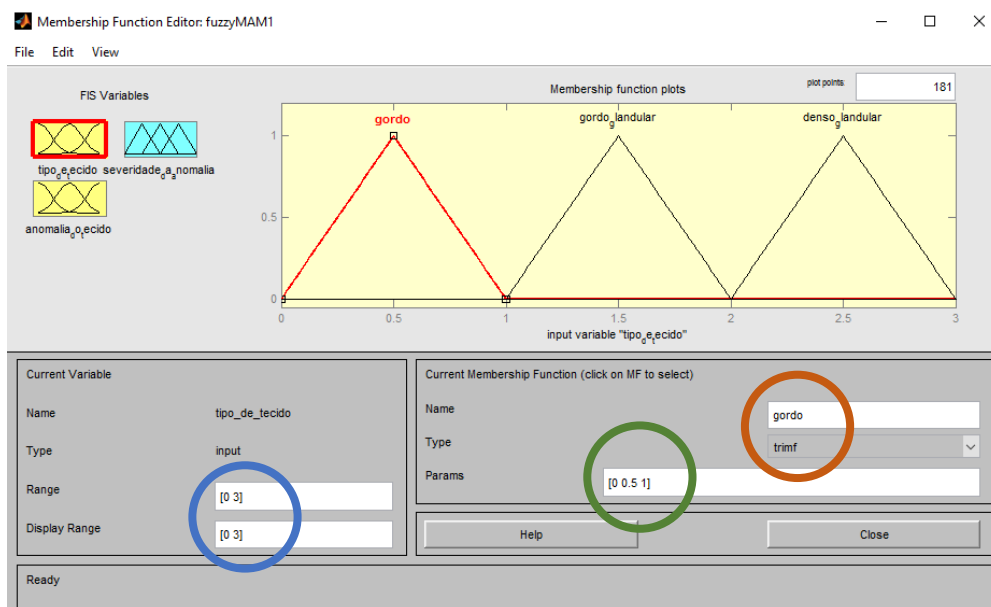


Figura 17 - Visualização da primeira função de entrada na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias.

Na figura 17, encontram-se 3 circunferências, uma azul, uma verde e uma castanha, estas dizem respeito aos vários valores a ter-se em conta. A circunferência azul diz respeito ao eixo dos  $xx$ , que é escolhido pela pessoa consoante o que lhe interessa estudar, neste caso eu escolhi que fosse entre  $[0\ 3]$ , porque tenho 3 parâmetros com o mesmo peso, encontrando-se o tecido gordo do 0 ao 1, o glandular do 1 ao 2 e o denso do 2 ao 3. Poderia estar ao contrário, eu coloquei nesta ordem porque me dava jeito assim. A circunferência castanha diz respeito ao nome, a operação é feita exatamente como já explicada para a figura 15, e a função que se pretende usar, as funções mais usadas são a triangular, a trapezoidal e a gaussiana. As que foram utilizadas foram a trapezoidal e a triangular, pois são as que mais se adequam aos estudos pretendidos e são as que normalmente são usadas neste tipo de análises. Sendo ainda simples de aplicar, objetivas e precisas. Nestes casos é necessários limitar as curvas dizer onde começam e onde acabam e obter picos de valores para poder retirar o valor ótimo do cruzamento.

A circunferência verde diz respeito aos parâmetros usados na curva apresentada com a linha vermelha. Cada função tem parâmetros específicos, no caso da usada, de nome triangular, tem um mínimo, um pico, (visto no eixo dos  $yy$ ), e um máximo. Ou seja para o tecido gordo, (linha apresentada a vermelho do gráfico apresentada na figura 17), encontra-se um valor mínimo de 0, um pico no 0.5 onde atinge o seu máximo no eixo dos  $yy$  e um máximo no 1. Para as outras funções foi realizado o mesmo processo, modificando-se o nome para um mais adequado e acertando-se os parâmetros consoante o necessário. Para dar nomes e parâmetros diferentes a cada uma das funções, basta carregar em cima da linha da função que se pretende mudar e acertar para o que se deseja.

A figura 18 que se segue é referente ao segundo *input*.

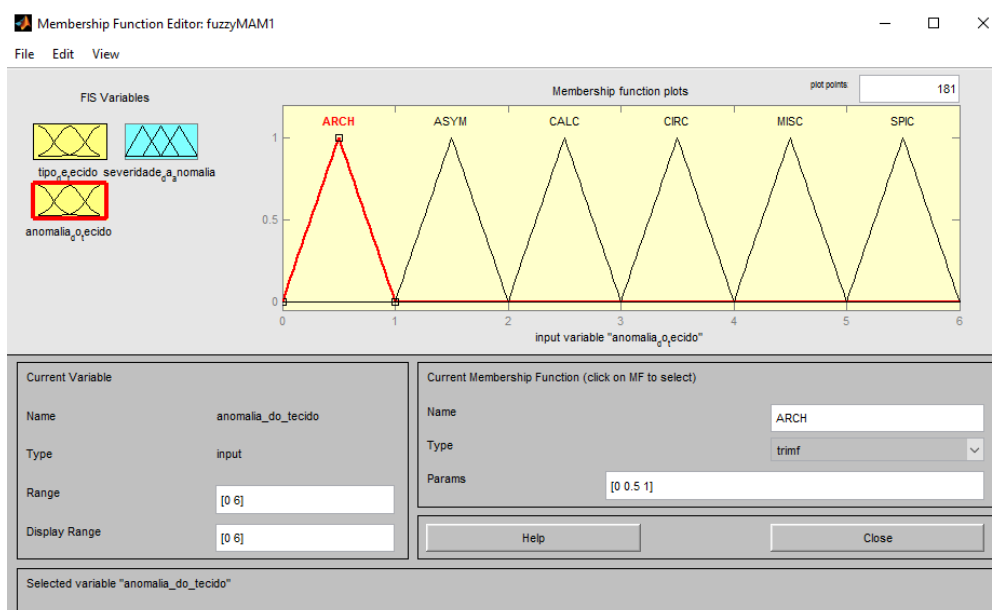


Figura 18 - Visualização da segunda função de entrada na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias.

Para a introdução deste segundo *input* são seguidos os mesmos passos descritos no anterior, a grande diferença é que neste são usados mais parâmetros. No anterior o eixo do  $xx$  ia de 0 a 3, para este os valores vão de 0 a 6, todos eles têm o mesmo peso tal como no anterior. Estes parâmetros são referentes às anomalias dos nódulos, as siglas inglesas foram mantidas para que fosse mais simples de atribuir nomes aos parâmetros e que estes fossem vistos no gráfico de forma legível, todas elas foram já descritas no capítulo VI.

Na figura 19 que se segue é apresentado o *output* que é referente à severidade da anomalia encontrada.

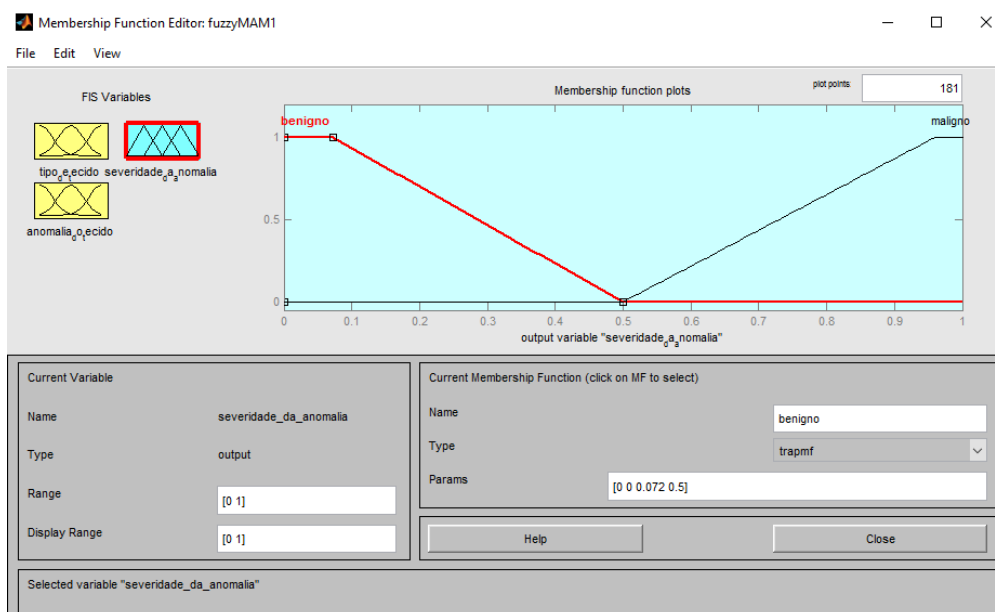


Figura 19 - Visualização da função de saída na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias.

Este gráfico é um trapezoidal, funciona da mesma forma que o anterior, mas apresenta maior quantidade de valores nos parâmetro. Foi usada esta função em vez da triangular pelo facto de esta não apresentar um pico, pois não teria sentido usar um valor máximo quando se pretende saber qual a possibilidade de uma determinada doença se apresentar em determinadas condições, mas sabendo à partida que nunca existem certezas absolutas que a doença se encontra presente. Existe a certeza nas imagens da base de dados utilizada, pois essas são confirmadas por um profissional de saúde, no entanto este estudo baseia-se nessas imagens e num conjunto de características que após serem cruzadas irão facilitar a construção de uma base estatística que poderá ajudar a perceber se uma mama com um nódulo e com determinadas características tem mais probabilidade de ser maligna ou benigna, mas nunca tendo certezas absolutas, essa certeza apenas poderá ser dada por um profissional de saúde, após exames específicos.

Para este gráfico são apenas usados valores de 0 a 1, sendo o 0 o valor mais elevado para benignidade e o 1 o mais elevado para malignidade. Isto serve para facilitar na leitura

dos dados, no final do cruzamento, se os valores estiverem de 0 a 1 será mais simples avaliar o valor percentual que este representa evitando assim mais cálculos.

Após finalizar a introdução das funções de *input* e *output*, é necessário construir um grupo de regras que vai ajudar no fim a decisão. Este conjunto de regras é feito na caixa branca que se encontra ao centro da imagem na figura 16, que tem como nome fuzzyMAM1, que foi o nome que dei ao meu primeiro estudo.

Para construir um conjunto de regras é necessário abrir o comando das regras, para isso basta carregar duas vezes em cima da caixa branca, da mesma forma que foi feito para as anteriores. O comando que irá abrir tem o nome de *rule editor*, é aqui que irão ser usados os valores percentuais calculados anteriormente e apresentados nas tabelas 3 e 6, do capítulo VI. Este comando pode ser visto na figura 20.

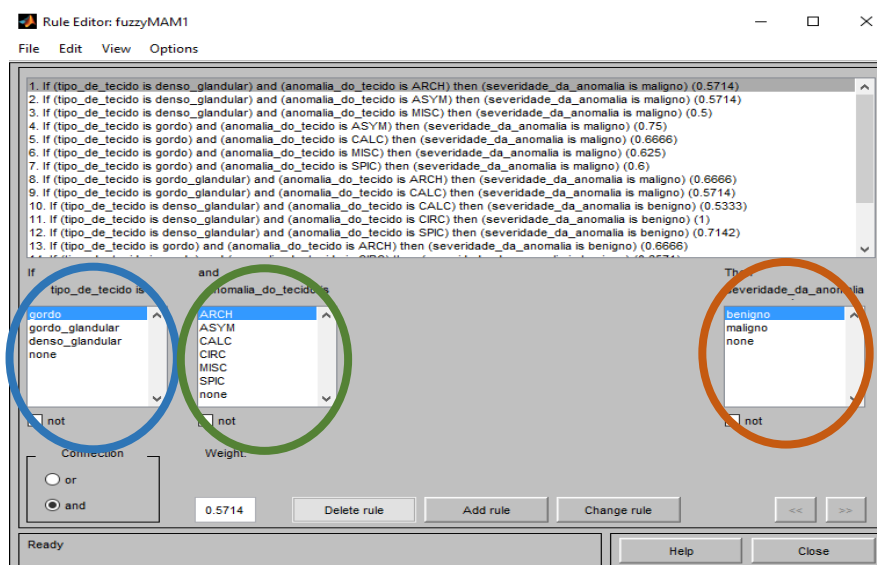


Figura 20 - Visualização do quadro das regras implementadas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias.

Na figura 20 estão apresentadas 18 regras, apesar de estas não se verem todas, todas elas podem ser consultadas com pormenor nos anexos página 70 e 71. Este quadro está apresentado de uma maneira muito simples de ser usado, a circunferência azul e a verde são referentes aos valores de entrada, enquanto que a castanha é referente as valores de saída. O que aqui é feito é cruzar estes valores, para isso é usado o *if* (circunferência azul/tipo de

tecido), o *and* (circunferência verde/anomalia do tecido) e o *then* (circunferência castanha/severidade da anomalia). Tudo isto se processa da seguinte maneira:

Se (*if*) tipo de tecido e (*and*) anomalia do tecido então (*then*) severidade da anomalia tem uma determinada percentagem.

Este processo é feito para cada tipo de tecido, para cada anomalia e por fim atribuído um determinado valor para cada severidade, sendo assim possível avaliar a probabilidade do nódulo a estudar ser maligno ou benigno, consoante um determinado tecido, em conjunto com uma determinada anomalia. Todas estas 18 regras serão avaliadas pela toolbox que irá construir um novo quadro onde são apresentados gráficos e valores estatísticos para cada regra construída. Esses gráficos são apresentados na figura 21.

Para ir buscar estes gráficos é necessário ir a *view* no quadro apresentado na figura 20 (*rule editor*), e carregar em *Rules*, após estes passos deverá aparecer o quadro apresentado na figura 21.

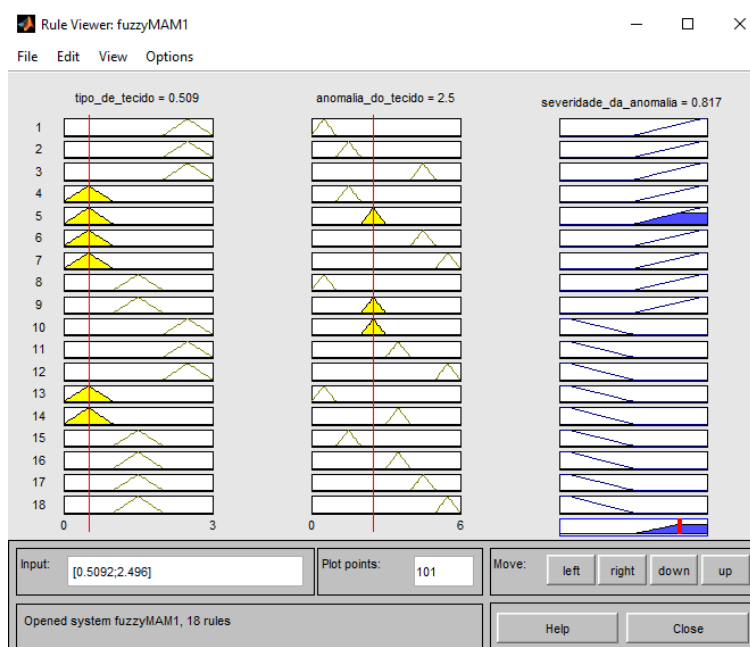


Figura 21 - Visualização do quadro das regras geradas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias.

O quadro apresentado tem todas as informações dos cruzamentos feitos com as regras implementadas, onde podem ser vistas 18 regras. É então agora necessário analisar o quadro com os gráficos e procurar a informação desejada para o estudo, que neste caso será perceber todos os cruzamentos feitos e verificar quais as regras que apontam para nódulos malignos e quais apontam para nódulos benignos. Para obter esta informação é necessário andar de um lado para o outro, com as linhas que se encontram a vermelho, estas apenas se encontram na primeira e segunda coluna que dizem respeito aos *inputs*, o que faz sentido pois queremos perceber o resultado tendo em conta as características dadas como variáveis de entrada. A terceira coluna diz respeito ao *output* este irá modificar-se e apresentar diferentes valores, consoantes as variáveis de entrada que são apontadas pela linha vermelha.

Na primeira coluna verificam-se três linhas, sendo estas as funções triangulares, e dizem respeito aos três tipos de tecidos, na segunda encontram-se as funções triangulares das anomalias, e por fim a avaliação da sua malignidade na terceira coluna. No exemplo que está mostrado, verifica-se que para tecidos gordos em conjunto com a anomalia calcificação, a severidade da anomalia encontra-se nos 0.817, significa que está mais perto do 1, ou seja a probabilidade de ser maligno é muito elevada.

A análise pode ainda ser feita através do gráfico *surface*, este encontra-se da mesma forma que o *rule view*, para isso basta ir ao botão *view* e de seguida carregar em *Surface*. Após este processo irá aparecer um quadro com um gráfico igual ao da figura 22.

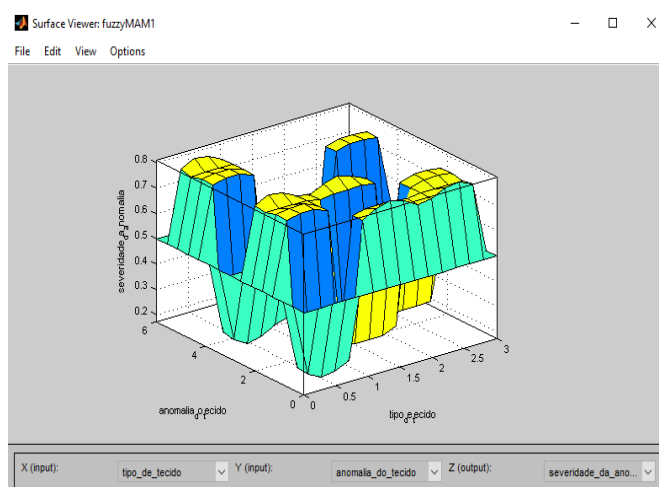


Figura 22 - Visualização do quadro do gráfico gerado pela toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias.

O gráfico apresentado, representa todos os valores obtidos com os cruzamentos, onde o eixo do  $xx$  representa o tipo de tecido, o eixo dos  $yy$  representa os valores das anomalias e o eixo dos  $zz$  representa os cruzamentos feitos para a obtenção da severidade.

Foram retirados os valores obtidos para cada uma das regras geradas, verificando assim a probabilidade que cada uma tem para benignidade e malignidade, todos estes valores foram tabelados e serão discutidos no capítulo VIII, assim como este gráfico.

### 7.1.2 - Segundo estudo

O segundo estudo usa tudo o que já foi feito, mostrado e explicado no primeiro estudo, com a diferença das variáveis de entrada, que em vez de duas (tipo de tecido e anomalia do nódulo), utiliza três variáveis, as duas já usadas anteriormente e uma nova variável que é o tamanho do nódulo apresentado na imagem. Esta mudança pode ser vista na figura 23.

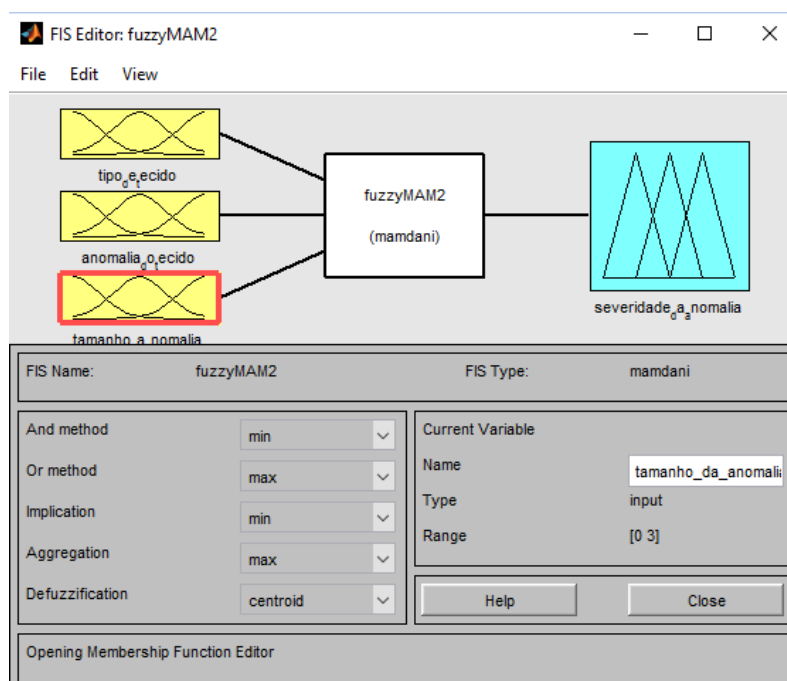


Figura 23 - Usando a toolbox Lógica Fuzzy usando o método de inferência Mamdani, para o segundo estudo, imagem próprias.

A *toolbox* ao ser iniciada para um novo estudo, tem como regra automática apresentar duas caixas de *input*, o que significa que se forem precisas mais do que as apresentadas, como é o caso, é necessário fazer o procedimento para que esta apareça no estudo, para isso basta ir ao botão *Edit*, e carregar em *Add Variable - input*. Após estes passos irá aparecer uma nova caixa como é mostrado na figura 23, onde é então atribuído um nome, este processo é exatamente igual ao já explicado para as outras duas variáveis anteriores. Depois de tudo concluído é então passado para dentro das caixas das variáveis, onde lhes são alterados todos os valores identificando assim as funções desejadas.

Nesta nova variável as funções são as mesmas que as usadas nas variáveis anteriores, aliás as funções serão idênticas às usadas anteriormente nos tipos de tecidos. Pode ser visto o gráfico gerado na figura 24.

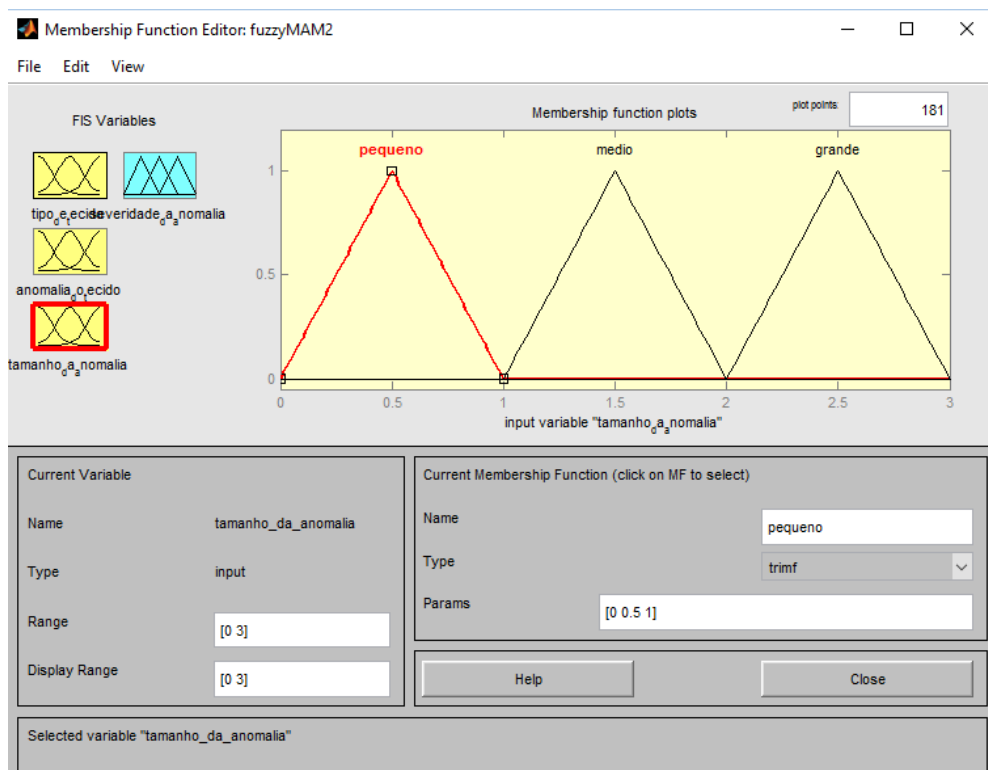


Figura 24 - Visualização da terceira função de entrada na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no segundo estudo, imagens próprias.

Esta variável contém três funções, são elas o pequeno que diz respeito aos valores mais baixos dos tamanhos de nódulos encontrados nas imagens, o médio, que diz respeito ao intervalo de valores médios e o grande que diz respeito aos maiores valores encontrados. Como foi explicado no capítulo VI, estes valores foram divididos em três grupos, (1, 2 e 3), para os gráficos como já se tem de trabalhar com vários números e como é possível dar nomes às funções foi então atribuído o nome pequeno, médio e grande, sendo assim mais fácil fazer a análise no final. O tipo de gráfico e os valores usados nas funções, são idênticos ao feito na primeira variável usada, o gráfico triangular e os valores de 0 a 3, cada função tem o mesmo peso.

A variável de saída, *output*, é feita exatamente da mesma forma que a já feita no estudo 1, nada é alterado. Após concluídas todas as funções, procede-se então a implementação das regras, no estudo 1 foram implementadas apenas 18 regras e com isto foi possível estudar tudo o necessário para fazer a avaliação no final da análise, já para o estudo 2 foram necessárias muitas mais, isto porque a informação era maior e por isso houve necessidade de gerar mais regras, sendo usadas 40 regras neste estudo.

O quadro com a implementação de todas as 40 regras pode ser visto na figura 25.

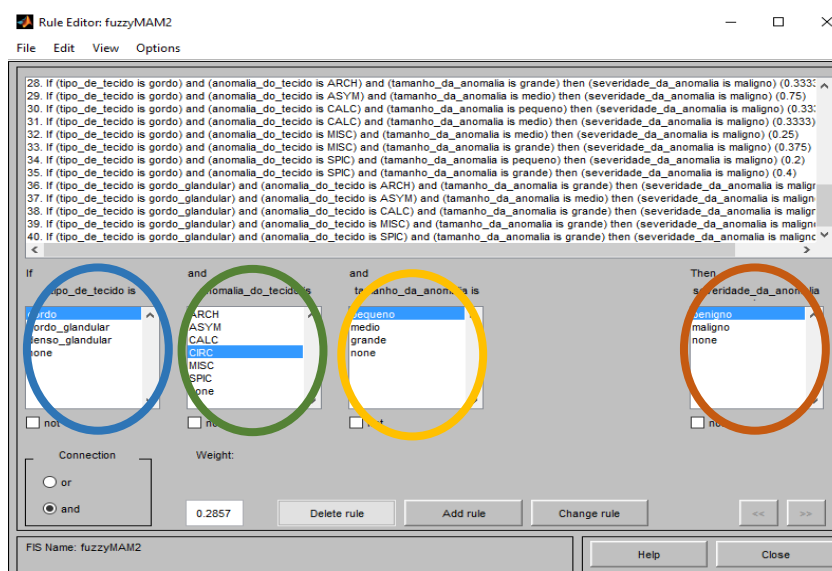


Figura 25 - Visualização do quadro das regras implementadas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no segundo estudo, imagens próprias.

Nesta figura encontra-se de novo as circunferência já usadas anteriormente, onde a azul, a verde e a castanha significam o mesmo que as já usadas na figura 20, já a amarela é referente às novas funções usadas para o tamanho dos nódulos existentes. Todo o processo é feito exatamente da mesma maneira já explicada no estudo 1, com a diferença que agora em vez de duas variáveis, usam-se três, pois tem de se juntar o tamanho do nódulo às regras. Imaginando que se está a fazer a regra para um tecido denso, uma anomalia CALC e um tamanho de nódulo médio, para se verificar a probabilidade de ter tecido maligno iria ficar a seguinte regra:

*If tecido denso and anomalia CALC and tamanho de nódulo médio then severidade maligna tem uma determinada percentagem.*

Volto a referir que todas as percentagens usadas se encontram tabeladas e devidamente identificadas no capítulo VI, já as regras poderão ser vistas e analisadas uma por uma nos anexos, na página 70, 71, 72, 73.

Depois de todas as regras implementadas passa-se então ao quadro da análise das regras já implementadas (*Rule View*), este quadro pode ser visto na figura 26, no entanto este é constituído por tantas regras que é impossível colocar todos os gráficos visíveis.

O mais importante a ser retirado deste quadro para além de perceber o seu funcionamento e todos os seus gráficos, são os seus valores percentuais. São estes valores que servirão para fazer a análise de todo o estudo. Todos os valores percentuais encontrados serão retirados e tabelados para depois serem analisados e cruzados com os que foram encontrados no estudo 1 e então discutidos no capítulo VIII.

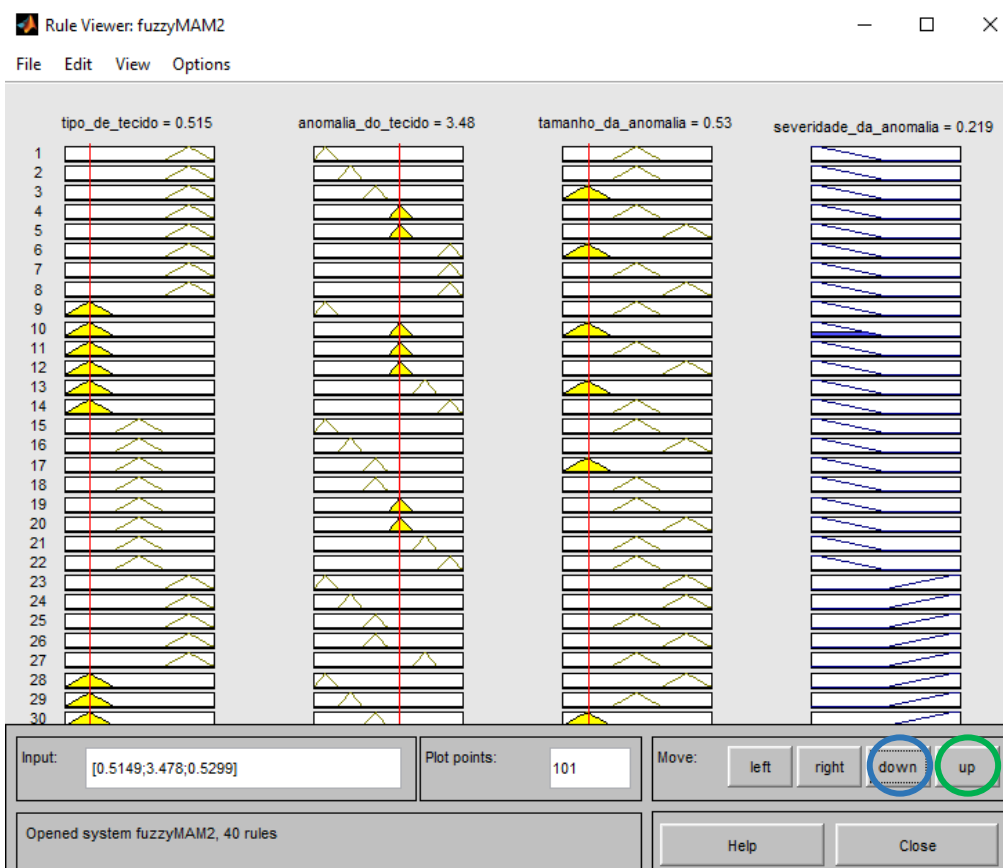


Figura 26 - Visualização do quadro das regras geradas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, no segundo estudo, imagens próprias.

Nesta figura 26, é ainda possível ver duas circunferências, uma azul e uma verde, estas são referentes aos botões subir (*up* - verde) e descer (*down* - azul). Estes botões servem para visualizar todas as regras feitas. Se olharmos para a figura apenas se veem 30 elementos, no entanto eles são 40, ao carregar no botão *down*, o quadro irá mover-se mostrando os restantes 10 elementos em falta na imagem.

A análise deste quadro é feita da mesma forma que o do estudo 1, colocam-se as linhas vermelhas em cima dos pontos que se pretendes estudar e observa-se os gráficos da coluna de *output*, quarta coluna, percebendo assim se o nódulo é benigno ou maligno.

Tal como na figura 21 a figura 26 foi deixada com as linhas vermelhas em sítios específicos propositadamente, para que assim pudesse ser explicado como foi feita esta análise. Na primeira coluna encontram-se os tipos de tecidos onde a linha vermelha se situa

em cima da primeira função que é referente aos tecidos gordos, na segunda coluna encontram-se as anomalias, onde a linha vermelha se situa em cima da quarta função, a anomalia CIRC (massa circunscrita bem definida), e a terceira coluna encontram-se os tamanhos dos nódulos, onde a linha vermelha se situa em cima da primeira função, referente aos tamanhos pequenos. O cruzamento destas funções é verificado na quarta tabela, referente às severidades da anomalia. Deste cruzamento obtém-se o valor 0.219, o que significa que a probabilidade da severidade para este cruzamento será benigna pois este valor encontra-se mais perto do 0.

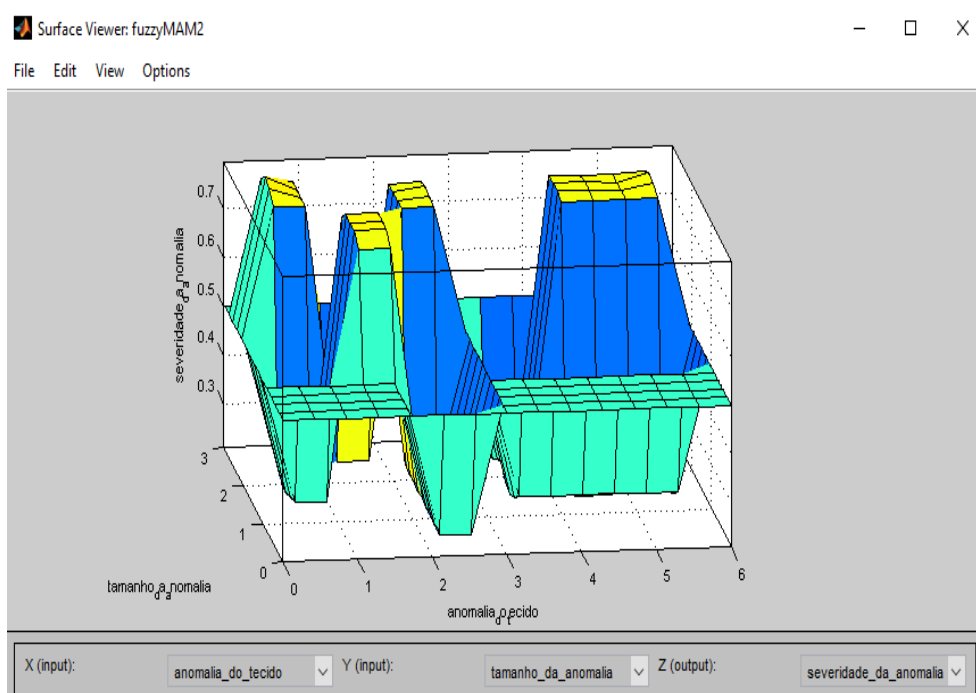


Figura 27 - Visualização do quadro do gráfico gerado pela toolbox Lógica Fuzzy usando o método de inferência Mamdani, no primeiro estudo, imagens próprias.

A obtenção do gráfico representado na figura 27, é feita exatamente da mesma forma que o já referido e mostrado na figura 22. Este gráfico é ligeiramente diferente do anteriormente obtido, isto porque as regras para além de serem diferentes são usadas com uma nova variável. Aqui é avaliada a probabilidade de se obter um determinado tamanho em função de uma determinada anomalia e por isto os valores usados nos eixos serão diferentes. O eixo do  $xx$  representa os valores obtidos para a anomalia, o eixo dos  $yy$  representa os valores obtidos para o tamanho e por fim o eixo dos  $zz$  representa os cruzamentos obtidos para a severidade em função destas duas variáveis.

No gráfico não está representado o tipo de tecido, isto porque este já foi estudado anteriormente, para além do gráfico não poder apresentar este valor pois os seus eixos já estão ocupados, não faria sentido voltar a referir esta variável, pois ela está incluída na variável das anomalias. Nas tabelas mostradas no capítulo VI, para o segundo estudo, é facilmente verificado o agrupamento feito para o cruzamento, pode ser visto que todas as variáveis são divididas consoante o tipo de tecido, ou seja, todas as anomalias são divididas consoante o tipo de tecido a estudar, logo não faria sentido mostrar a variável, tipo de tecido, neste gráfico pois o pretendido é que sejam estudados os tamanhos e para isto basta o cruzamento com a variável das anomalias.

## Capítulo VIII

### Resultados e Discussão

A Lógica Fuzzy foi usada neste trabalho com o objetivo de estudar a probabilidade existente, de uma determinada massa encontrada num exame de mamografia, ser maligna ou benigna, dependendo do tecido da mama, da anomalia existente no tecido e do tamanho do nódulo existente. Após a análise feita aos resultados obtidos através do sistema de inferência fuzzy com o método de inferência *Mamdani*, utilizando a toolbox Lógica Fuzzy, que se encontra inserida no Matlab e depois da análise feita às tabelas do capítulo V e dos gráficos obtidos com as mesmas representados no capítulo VI, foi possível obter os valores das tabelas descritas neste capítulo.

#### 8.1 - Resultados obtidos com o primeiro estudo

A tabela 7, diz respeito ao primeiro estudo, onde foram usados apenas valores do tipo de tecido e do tipo de anomalia existente. Esta tabela foi contruída seguindo o mesmo pensamento do estudo mostrado no capítulo anterior, apontando qual a probabilidade do tipo de tecido e da anomalia do nódulo para a obtenção de uma massa maligna ou benigna. Todos os valores obtidos com a *rule view*, são agora desritos nesta tabela, verificando quais apontam para benignidade e quais apontam para malignidade.

Foram ainda colocados nas tabelas deste capítulo mais uma coluna, esta é referente aos valores da percentagem referente ao valor obtido com a toolbox para que as diferenças sejam vistas com mais facilidade, para a obtenção destes valores foi feita uma conversão para percentagem, os valores das massas malignas é simples pois basta multiplicar por 100, já para as benignas é necessário ter-se em conta que aqui o valor máximo é o 0 e o mínimo o 1, e obter assim o valor percentual. Encontram-se ainda valores de cores diferentes, estes foram colocados desta maneira para que seja facilmente distinguidos os valores para benignidade (valores sublinhados) e malignidade (valores não sublinhados).

Tabela 7 - Valores obtidos para o estudo 1, após todas as análises feitas na toolbox Lógica Fuzzy usando o método de inferência Mamdani. Valores próprios.

Tipo de tecido	Anomalia do nódulo	Severidade da anomalia	Valor obtido com o estudo na toolbox	Porcentagem sobre o valor obtido (valor em %)
<b>Gordo</b>	ARCH	Benigna	0.186	<u>81.4</u>
	ASYM	Maligna	0.821	82.1
	CALC	Maligna	0.815	81.5
	CIRC	Benigna	0.176	<u>82.4</u>
	MISC	Maligna	0.812	81.2
	SPIC	Maligna	0.809	80.9
<b>Gordo Glandular</b>	ARCH	Maligna	0.817	81.7
	ASYM	Benigna	0.201	<u>79.9</u>
	CALC	Maligna	0.81	81
	CIRC	Benigna	0.184	<u>81.6</u>
	MISC	Benigna	0.193	<u>80.7</u>
	SPIC	Benigna	0.194	<u>80.6</u>
<b>Denso Glandular</b>	ARCH	Maligna	0.81	81
	ASYM	Maligna	0.808	80.8
	CALC	Benigna	0.197	<u>80.3</u>
	CIRC	Benigna	0.167	<u>83.3</u>
	MISC	Maligna	0.802	80.2
	SPIC	Benigna	0.183	<u>81.3</u>

Após a análise da tabela verifica-se que os valores se encontram muito próximos, isto quer dizer duas coisas, a primeira é que a toolbox conseguiu valores distintos para malignidade e benignidade, ou seja é sem sombra de dúvida, com base nos valores percentuais, fácil de verificar que quando se diz que a probabilidade é ser maligna refere-se a valores a cima de 80% o que não deixa dúvidas na sua classificação. O mesmo se verifica para a benignidade, todos os valores obtidos nesta tabela não deixam dúvidas disso. A segunda é que quando se procura saber a probabilidade entre anomalias esta é mais complexa, isto porque os valores obtidos são muito próximos. Se forem observados todos os valores obtidos para benignidade, verifica-se que estes variam em décimas entre eles e algumas das vezes nem isso, verificando-se o mesmo para os valores obtidos para malignidade, o que dificulta o estudo caso se quisesse verificar dentro das anomalias qual ou

quais seriam as mais propícias a gerar nódulos malignos e quais as mais propícias a gerar nódulos benignos.

No entanto pode ser afirmado, consoante os valores obtidos, que para casos de tecidos gordos, a maior probabilidade de se obter um nódulo maligno encontra-se na anomalia ASYM, (assimétrica) com o valor percentual de 82.1%. Para casos de benignidade dentro do mesmo tipo de tecido, a maior probabilidade encontra-se na anomalia CIRC, (massa circunscrita bem definida), com o valor percentual de 82.4%.

No caso do tipo de tecido ser glandular, a probabilidade de se obter um nódulo maligno é maior na anomalia ARCH (distorção arquitetural), com o valor percentual de 81.7%. Para nódulos benignos a probabilidade verifica-se maior na anomalia CIRC, (massa circunscrita bem definida) com o valor percentual 81.6%.

Para o caso do tecido ser denso, a maior probabilidade para um nódulo maligno encontra-se na anomalia ARCH (distorção arquitetural), com o valor percentual de 81%. Já a maior probabilidade de obter um nódulo benigno é na anomalia SIPC (massa espiculada), onde se encontra um valor percentual de 81.3%.

Vendo a tabela de ponto de vista geral, não há uma grande significância probabilística, ou seja não existe grande diferença entre probabilidades benignas e malignas, logo não se consegue afirmar que existe uma grande probabilidade do nódulo ser maligno ou benigno. No entanto os valores embora com muito pouca diferença entre eles, verifica-se que para tecidos gordos a probabilidade maior é de ser benigno com 82.4% contra a probabilidade de ser maligno com 82.1%. Para tecidos glandulares a probabilidade é maior para malignidade com 81.7% contra a benignidade com valor de 81.6%. Por fim para tecidos densos, a probabilidade é maior para nódulos benignos com valor de 81.3%, contra a probabilidade de 81% para nódulos malignos.

A figura 28 que se segue diz respeito ao gráfico mostrado no capítulo VII, este gráfico mostra o cruzamento entre os valores do tipo de tecido com os valores das anomalias. O gráfico mostra o mesmo que a tabela já descrita, mas de uma forma mais rápida e simples de visualizar e mais sucinta. O gráfico é uma réplica do apresentado na figura 22, no entanto este é visto de cima, por isso não parece 3D. Os valores para as severidade são apresentados por

cores, sendo a cor amarela a que apresenta a maior probabilidade para malignidade e a cor azul escura a maior probabilidade para benignidade. A cor azul clara apresentada no gráfico refere-se aos valores de percentagem 0.

Analisando o gráfico da figura 28, no que diz respeito à probabilidade de severidade maligna esta é encontrada em tecidos gordos com anomalias ASYM, CALC, MISC e SPIC. Em tecidos glandulares com anomalias ARCH e CALC. Já no tecido denso esta verifica-se em anomalias como a ARCH, a ASYM e a MISC.

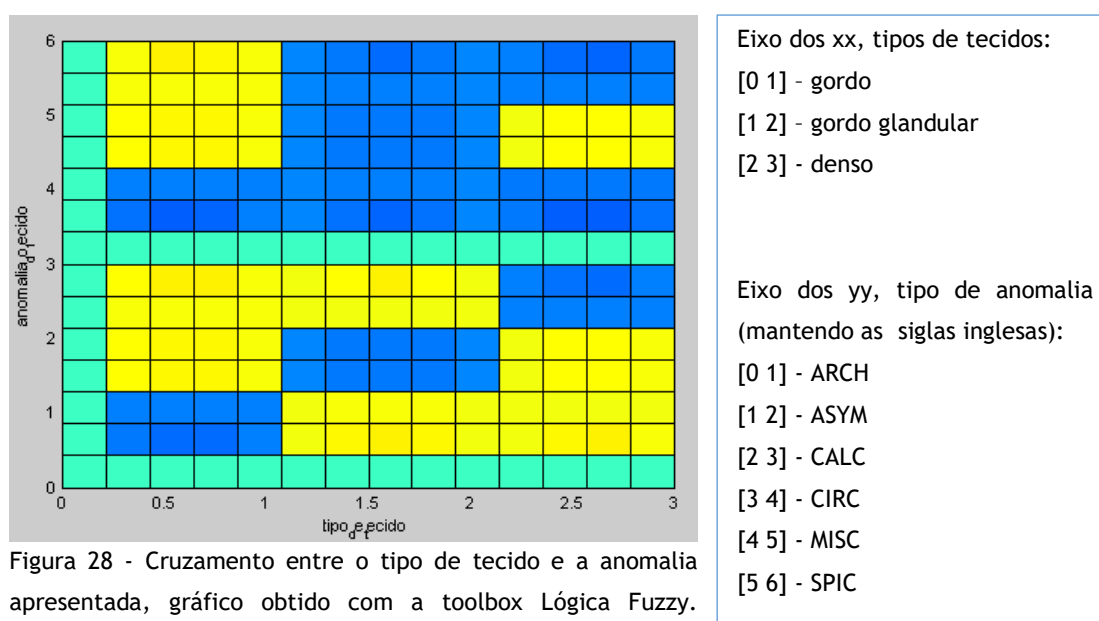


Figura 28 - Cruzamento entre o tipo de tecido e a anomalia apresentada, gráfico obtido com a toolbox Lógica Fuzzy. Imagem própria.

No que diz respeito à probabilidade de severidade benigna, esta é encontrada no tecido gordo em anomalias como a ARCH e a CIRC. Em tecidos glandulares esta encontra-se em anomalias como ASYM, CIRC, MISC e SPIC. No tecido denso a probabilidade será maior em anomalias como a CALC, a CIRC e a SPIC.

Ao analisar o gráfico olhando apenas para os tipos de tecidos, verifica-se que para a população em estudo, a maior probabilidade de se obter um nódulo maligno será em pessoas com tecidos mamários gordos e a maior probabilidade da obtenção de um nódulo com severidade benigna será em tecidos glandulares. No que diz respeito às pessoas com tecidos

mamários densos estes também apresentam uma maior probabilidade para a severidade maligna tal como o tecido gordo, no entanto a probabilidade não é tão alta como neste.

## 8.2 - Resultados obtidos com o segundo estudo

As 3 tabelas seguintes fazem todas parte da mesma, no entanto para ser mais simples de analisar, optei então por fazer três e explicar os valores de cada parte. Estas dizem respeito ao segundo estudo feito neste trabalho, onde foram usados valores dos tipos de tecidos, das anomalias existentes e dos tamanhos dos nódulos existentes. Tal como a tabela anterior, também estas são contruídas seguindo o mesmo pensamento do estudo mostrado no capítulo anterior, mostrando qual a probabilidade do tipo de tecido, do tipo de anomalia e do tamanho do nódulo existente, para a abtenção de uma massa maligna ou uma massa benigna. Para além dos valores sublinhados, benignos e os malignos, esta tabela é constituída ainda por valores de “0”, estes são os valores que não servem para tomar qualquer decisão, são valores de 0%. Estes zeros dizem respeito aos valores que não existem e também eles foram calculados na toolbox, pois esta consegue identificar através do que foi colocado as percentagens restantes. Se voltarmos às tabelas apresentadas no capítulo IV, verificamos que existem zeros, estes são a inexistência de nódulos nas imagens estudadas para aquele determinado tamanho.

A tabela 8 que se segue diz respeito aos valores obtidos em relação ao tecido gordo.

Tabela 8 - Valores obtidos para o estudo 2 após todas as análises feitas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, para o tecido gordo. Dados próprios.

Tipo de tecido	Anomalia do nódulo	Tamanho da anomalia	Severidade da anomalia	Valor obtido com o estudo na toolbox	Percentagem sobre o valor obtido (valor em %)
Gordo	ARCH	1	--	0.5	0
		2	Benigno	0.214	<u>78.6</u>
		3	Maligno	0.786	<u>78.6</u>
	ASYM	1	--	0.5	0

		2	Maligno	0.821	82.1
		3	--	0.5	0
		1	Maligno	0.786	78.6
	CALC	2	Maligno	0.784	78.4
		3	--	0.5	0
		1	Benigno	0.218	<u>78.2</u>
	CIRC	2	Benigno	0.205	<u>79.5</u>
		3	Benigno	0.233	<u>76.7</u>
		1	Benigno	0.236	<u>76.4</u>
	MISC	2	Maligno	0.776	77.6
		3	Maligno	0.787	78.7
		1	Maligno	0.773	77.3
	SPIC	2	Benigno	0.207	<u>79.3</u>
		3	Maligno	0.793	79.3

Ao analisar a tabela voltam-se a verificar valores muito próximos, tal como no estudo 1 e quando comparadas as percentagens, estas verificam-se mais baixas. No primeiro grupo, para a anomalia ARCH, encontra-se igual probabilidade para a severidade, ou seja, tanto poderá ser maligno, como benigno e pode ainda ter um tamanho considerado médio ou grande. No caso da anomalia ASYM, apenas se verifica percentagens para malignidade, esta anomalia apresenta maior probabilidade para nódulos malignos e ainda apresenta maior probabilidade para ser um nódulo de tamanho médio com um valor de 82.1%.

Na anomalia CALC a probabilidade de malignidade é maior, no entanto não há grande diferença de valores nos tamanhos dos nódulos, obtendo-se valores para para tamanho pequeno de 78.6% e para valores de tamanho médio de 78.4%. Já a anomalia seguinte refere-se à CIRC e tem maior probabilidade de ser um nódulo benigno, com valores mais elevados para nódulos de tamanho médio, com probabilidade de 79.5%, contra a probabilidade de 78.2% e 76.7%, para tamanhos pequenos e grandes respetivamente.

A anomalia que se segue é a MISC, esta apresenta maior probabilidade de ser um nódulo maligno e de tamanho grande, com valores de 78.7%, contra o valor de 77.6% para tamanhos médios. Apresentando ainda uma probabilidade de 76.4% para nódulo benigno de tamanho pequenos. Por último é apresentada a anomalia SIPC, esta apresenta maior probabilidade de malignidade, com valores de 77.3% para tamanhos pequenos e 79.3% para

tamanhos grandes, apresentando ainda valores de 79.3% de para nódulos benignos de tamanho médio.

Ao analisar a tabela de um modo geral, é verificado que para tecidos gordos, existe uma maior probabilidade de ter uma anomalia do tipo ASYM, de tamanho médio e que esta seja maligna, verificando-se um valor de 82.1%. Já de um ponto de vista invertido, a menor probabilidade seria obter uma anomalia tipo ASYM também esta maligna mas para tamanho pequeno, verificando-se um valor de 51.7%.

A tabela 9 que se segue diz respeito aos valores conseguidos como o estudo do tecido gordo glandular. Analisando a tabela verifica-se o mesmo problema da anterior, os valores encontram-se muito próximos.

Tabela 9 - Valores obtidos para o estudo 2 após todas as análises feitas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, para o tecido gordo glandular. Dados próprios.

Tipo de tecido	Anomalia do nódulo	Tamanho da anomalia	Severidade da anomalia	Valor obtido com o estudo na toolbox	Percentagem sobre o valor obtido (valor em %)
Gordo Glandular	ARCH	1	--	0.5	0
		2	Benigno	0.23	<u>77</u>
		3	Maligno	0.803	80.3
	ASYM	1	--	0.5	0
		2	Maligno	0.803	80.3
		3	Benigno	0.221	<u>77.9</u>
	CALC	1	Benigno	0.217	<u>78.3</u>
		2	Benigno	0.232	<u>76.8</u>
		3	Maligno	0.768	76.8
	CIRC	1	--	0.5	0
		2	Benigno	0.209	<u>79.1</u>
		3	Benigno	0.209	<u>79.1</u>
	MISC	1	--	0.5	0

		2	Benigno	0.206	<u>79.4</u>
		3	Maligno	0.774	77.4
	SPIC	1	--	0.5	0
		2	Benigno	0.204	<u>79.6</u>
		3	Maligno	0.783	78.3

Ao verificar os valores obtidos para a primeira anomalia, a ARCH, encontra-se uma maior probabilidade para malignidade em tamanhos grandes, com um valor de 80.3% e uma probabilidade de 77%, mais reduzida, para benignidade em tamanhos médios. Na anomalia ASYM encontram-se valores muito parecidos mas para tamanhos diferentes. Para esta anomalia obtiveram-se valores mais elevados para a probabilidade de malignidade em tamanhos médios, com um valor de 80.3% e uma probabilidade menor para benignidade em tamanhos grandes, com um valor de 77.9%.

No caso da anomalia CALC, esta apresenta uma maior probabilidade para tumores benignos de tamanhos pequenos, apresentando um valor de 78.3% contra um valor de 76.8 para tamanhos médios, apresentando ainda uma probabilidade de 76.8% para nódulos malignos de tamanhos grandes. A anomalia CIRC apresenta valores idênticos para benignidade em tamanhos médios e grandes, com um valor 79.1%, não apresentando qualquer valor para tamanhos pequenos.

A probabilidade para a anomalia MISC é maior na severidade benigna, apresentando um valor de 79.4% para tamanhos médios. Tendo ainda uma probabilidade de 77.4%, para apresentar malignidade em tamanhos grandes. Por fim a anomalia SPIC, esta apresenta também uma maior probabilidade para benignidade em tamanhos médios, podendo ainda apresentar com 78.3% de probabilidade, malignidade em tamanhos grandes.

Analisando esta tabela de ponto de vista geral, encontra-se uma maior probabilidade de obtenção da severidade maligna em anomalias ARCH e ASYM, para tamanhos grande e médio respetivamente com valores idênticos de 80.3%. A menor probabilidade encontrada, é para a anomalia CALC onde se encontram valores de 76.8%, tanto para a severidade maligna em tamanho médio como benigna em tamanho grande.

A tabela 10 que se segue é referente à ultima parte deste estudo e diz respeito aos valores obtidos para o tipo de tecido denso.

Tabela 10 - Valores obtidos para o estudo 2 após todas as análises feitas na toolbox Lógica Fuzzy usando o método de inferência Mamdani, para o tecido denso. Dados próprios.

Tipo de tecido	Anomalia do nódulo	Tamanho da anomalia	Severidade da anomalia	Valor obtido com o estudo na toolbox	Percentagem sobre o valor obtido (valor em %)
Denso	ARCH	1	--	0.5	0
		2	Benigno	0.204	<u>79.6</u>
		3	Maligno	0.796	79.6
	ASYM	1	--	0.5	0
		2	Benigno	0.217	<u>78.3</u>
		3	Maligno	0.796	79.6
	CALC	1	Benigno	0.213	<u>78.7</u>
		2	Maligno	0.787	78.7
		3	Maligno	0.817	81.7
	CIRC	1	--	0.5	0
		2	Benigno	0.213	<u>78.7</u>
		3	Benigno	0.183	<u>81.7</u>
	MISC	1	--	0.5	0
		2	Maligno	0.803	80.3
		3	--	0.5	0
	SPIC	1	Benigno	0.232	<u>76.8</u>
		2	Benigno	0.217	<u>78.3</u>
		3	Benigno	0.217	<u>78.3</u>

Analisando esta tabela, para a primeira anomalia, encontra-se e mesma probabilidade para benignidade e malignidade com valores iguais em tamanhos médio e grande, em que o médio é referente à severidade benigna e o grande à maligna, os seus valores são de 79.6%. Para a anomalia ASYM existe maior probabilidade de que o nódulo seja maligno e para

tamanhos grandes, este tem um valor de 79.6%. Obtendo ainda uma probabilidade de benignidade de 78.3% para tamanhos médios.

A anomalia CALC apresenta uma grande probabilidade de malignidade para tamanhos grandes com 81.7%, contra uma probabilidade de 78.7% para malignidade em tamanhos médios. Apresenta ainda um probabilidade idêntica (78.7%) aos tamanhos médios, mas para tamanhos pequenos, no entanto esta probabilidade refere-se a benignidade. Para a anomalia CIRC a maior probabilidade é em obter um nódulo benigno, apresentando um valor de 81.7% para tamanhos grandes, contra um valor de 78.7%, para tamanhos médios.

A anomalia MISC apresenta apenas probabilidade para a obtenção de severidade maligna em 80.3% para tamanhos médios. Por fim a anomalia SPIC apresenta uma maior probabilidade do nódulo ser benigno, tendo valores iguais para tamanhos grandes e médios com valor de 78.3%, contra um valor mais baixo para tamanhos pequenos de 76.8%.

Avaliando a tabela por inteiro, de uma maneira geral, esta é a tabela que obteve mais valores iguais o que dificulta a análise. No entanto tem também valores que sobressaem perante os outros, estes são os valores de 81.7%, um deles para benignidade e o outro para malignidade. O valor para a severidade maligna, encontra-se na anomalia CALC, para nódulos de tamanhos grandes. Já o valor para a severidade benigna, encontra-se na anomalia CIRC, também este para nódulos de tamanhos grandes.

Se as três partes (tabelas) forem avaliadas num todo, verifica-se que o maior e o menor valor se encontra na tabela do tecido gordo para uma anomalia ASYM, com um valor de 82.1% para nódulos de tamanho médio e o menor valor se encontra na anomalia MISC para uma probabilidade de 76.4% em nódulos de tamanho pequeno. Ou seja no geral será maior a probabilidade de se obter uma severidade maligna num tecido gordo, com um anomalia ASYM, para nódulos de tamanhos médios (valores que vão dos 2 centímetros até aos 5 centímetros de diâmetro).

Tal como para o estudo 1, para este estudo foi feito um gráfico através da toolbox e tal como para o estudo 1 este gráfico foi mostrado no capítulo VI, a figura 27 é a réplica desse mesmo gráfico mas visto de cima.

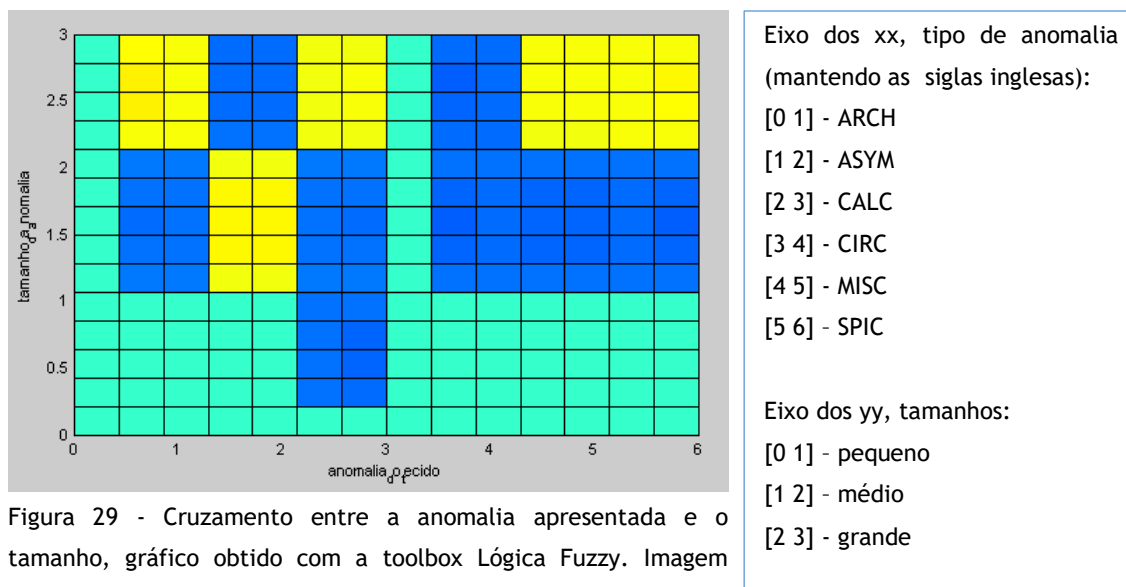


Figura 29 - Cruzamento entre a anomalia apresentada e o tamanho, gráfico obtido com a toolbox Lógica Fuzzy. Imagem própria.

Na figura 29 estão representados todos os valores usados nas tabelas 8, 9 e 10, este segue os mesmos parâmetros que o já descrito no estudo anterior, mas em vez de cruzar valores para tipo de tecidos e anomalias, o cruzamento é feito com as anomalias e os tamanhos dos nódulos presentes.

Olhando para o gráfico verifica-se logo à partida que a severidade maligna está concentrada nos tamanhos para nódulos grandes. Existindo ainda uma grande probabilidade de severidade maligna em nódulos de tamanhos médios para a anomalia ASYM. Verifica-se ainda que a probabilidade em anomalias do tipo MISC, é maior para benignidade. Em tamanhos pequenos apenas são apresentados valores de significancia estatística para a anomalia CALC em que a severidade é benigna. Para todas as outras anomalias é verificada que a probabilidade de se obter um nódulo, seja ele benigno ou maligno, é muito parecida ou mesmo igual (valores a azul claro). É importante voltar a referir que todo o estudo recai sobre uma população, logo estes valores são apenas referentes para o grupo de estudo usado.

## 8.3 - Comparação entre os estudos realizados

A técnica usada, Lógica Fuzzy é por norma uma técnica que consegue obter bons resultados neste tipo de estudos, o que também aconteceu neste trabalho.

Os estudos obtidos são apresentados nas tabelas 7, 8, 9 e 10 deste capítulo, estes demonstraram bons resultados em todos os cruzamentos feitos, apresentando sempre valores a cima dos 76%. No primeiro estudo onde são cruzadas apenas duas características os resultados apresentados verificam-se superiores aos encontrados no segundo estudo, encontrando-se sempre valores entre os 80% a 90%. No entanto tanto um estudo como o outro apresentam valores a cima dos 76%. Isto significa que a classificação foi bem conseguida, os valores obtidos não deixam dúvidas, pois independentemente do cruzamento feito, o resultado obtido, a probabilidade gerada, é suficientemente elevada para se dizer com certeza quando é que a severidade encontrada é maligna e benigna dependendo do tipo de características que a mama apresenta.

O facto do segundo estudo ter apresentado valores percentuais mais baixos pode estar relacionado com o facto de terem sido cruzadas mais características e estas não apresentarem diferença substancial entre elas. Ou seja apesar de lhe estar a ser dado mais uma característica para o cruzamento, se esta não tiver valores suficientemente elevados para existir uma barreira entre o que se quer estudar, neste caso, diferentes tamanhos de nódulos existentes, os resultados obtidos serão mais baixos. No entanto, os resultados que se obtiveram apresentaram bons resultados para o pretendido, quando comparado com os estudos já realizados.

Comparando os dois estudos, apesar de se ter obtido melhores valores no primeiro estudo, o segundo verifica-se mais pormenorizado. O primeiro apresenta como média de percentagens os 81.22%, e o segundo apresenta apenas uma média de 76.82%. No entanto no primeiro estudo é apenas possível verificar, se um tipo de tecido e uma determinada anomalia apresentada poderá ou não apresentar maior probabilidade de malignidade ou benignidade. Enquanto que no segundo estudo é possível verificar que um tipo de tecido, uma determinada anomalia apresentada e um tamanho do nódulo poderá ter maiores probabilidades de ser maligno ou benigno, apresentando uma tendência de malignidade

quanto maior for o seu tamanho. O segundo estudo apresenta maior profundidade, ou seja consegue uma avaliação mais pormenorizada, especificando assim mais o estudo.

É ainda importante perceber que os resultados obtidos dependem muito da população a estudar, pois esta poderá apresentar características diferentes consoante o estilo de vida, o nível de vida, a alimentação, a constituição, o peso a idade e mesmo o local onde esta vive. Tudo isto são factores para que uma determinada doença se possa manifestar de diferentes maneiras numa determinada população. Todos estes aspetos terão de se ter em conta, a quando a realização deste tipo de estudos.

## Capítulo IX

### Conclusão e trabalhos futuros

Este trabalho teve como objetivo estudar a Lógica *Fuzzy* no seu todo e as suas aplicações na medicina, após perceber o que esta poderia fazer pela área da saúde era pretendido usa-la para fazer um estudo de interesse.

O objetivo do trabalho foi conseguido, a Lógica *Fuzzy* é um método usado em diversas áreas, aqui importa referir as áreas da saúde em que esta pode ser aplicada, pois é esta a área de interesse. Mesmo na área da saúde a Lógica *Fuzzy* pode ser implementada de maneiras diferentes dependendo do que é pretendido com o estudo. Esta pode ser utilizada em estudos de imagem médica, podendo mesmo dentro desta área ser usada de diversas maneiras. E pode ser usada para estudos idênticos ao feito neste trabalho, como a avaliação de uma doença, ou o funcionamento de um medicamento, seja na sua constituição ou o seu funcionamento no organismo humano, ou o funcionamento de internamentos de um hospital, entre muitas outras coisas, onde são procuradas as melhores probabilidades de obter bons resultados numa determinada classificação.

Como já foi dito neste trabalho, é importante escolher o método a usar consoante o estudo que se pretende, pois a Lógica *Fuzzy*, como qualquer outro método de trabalho, obtém bons resultados se lhe forem fornecidas boas características. Ou seja, não se pode pedir bons resultados se não forem oferecidas boas formas de classificação. Por este aspecto antes de ser utilizado um método que utilize a Lógica *Fuzzy*, é necessário entender bem o que se pretende estudar. Para este tipo de estudos é necessário uma boa população, se não se tiverem dados em quantidades suficientes para fazer o estudo não se obterão resultados sequer satisfatórios. Estes dados poderão ser obtidos de várias formas, sendo uma delas um hospital, onde se faz um estudo de uma determinada população alvo e se estudam essas pessoas durante um determinado período de tempo, ou então como foi feito para este trabalho. Onde se usou uma base de dados criada por especialistas, onde se encontram dados de uma população alvo, neste caso são doentes com cancro de mama. Assim todos os dados terão de ser fidedignos para que o estudo seja também fidedigno.

Para se proceder a este estudo foi imprescindível criar tarefas a realizar. Primeiro foi necessário criar tabelas de contagens para se poderem criar conjuntos a serem estudados, depois das contagens feitas foi necessário usa-las para criar tabelas de percentagens desses grupos de valores, pois só assim seria possível usar os valores na toolbox Lógica Fuzzy. Tudo isto foi feito para os dois estudos, em que o primeiro utiliza duas características e o segundo utilizada três características para o cruzamento de dados. É através deste cruzamento que se conseguem obter os resultados, estes têm de avaliar a probabilidade da existência de um nódulo maligno ou de um nódulo benigno consoante os dados usados no cruzamento.

Após todos os dados tabelados, é então começado o primeiro estudo onde são usadas as características do tipo de tecido e a anomalia detetada no tecido mamário. Este estudo tem como objetivo verificar a probabilidade da existencia de um nódulo de severidade maligna ou benigna consoante o que lhe é implementado nas características e segundo regras impostas onde são usados os valores das tabelas feitas anteriormente. Já no segundo estudo são realizadas as mesmas operações mas neste, em vez de serem usadas apenas o tipo de tecido e as anomalias, é lhe aplicada uma nova característica que oferece mais pormenor ao estudo, esta é o tamanho do nódulo apresentado pela imagem retida da mamografia.

Os estudos implementados forneceram resultados bastante satisfatórios. O primeiro ofereceu melhores resultados percentuais e o segundo estudo consegue obter maior pormenor na classificação, sendo assim possível verificar através dos resultados obtidos, que dependendo do tamanho do nódulo apresentado pode significar benignidade ou malignidade.

Apesar do trabalho realizado ter obtido resultados satisfatórios, o trabalho poderia ser melhorado, pois quanto maior for a população melhores serão os resultados obtidos. No futuro deverá ter-se em conta este pormenor, o trabalho realizado com um maior número de doentes seria mais vantajoso na obtenção de melhores resultados. Seria ainda interessante fazer este estudo com mais características de maneira a pormenorizar o estudo ao máximo. Poderiam ser feitos dois novos estudos para serem cruzados entre eles, estudando primeiro as características apresentadas pela imagem e exames feitos, e outro estudo que avaliasse a pessoa consoante o seu estilo de vida, como por exemplo, se fuma, se bebe, que tipo de alimentação está sujeita, a que tipo de stress está sujeita no dia a dia, entre outras. Por fim fazer a classificação e verificar as maiores probabilidades de obtenção de nódulos malignos ou benignos. Seria ainda interessante fazer este tipo de estudos mas para pessoas com e sem doença, fazendo assim uma avaliação da probabilidade de obtenção de doença em

determinadas características, para isto deveriam ser coletados dados de pessoas em exames com e sem doença e verificar quais as características apresentadas por cada tipo.

Este tipo de estudos poderão ser uma mais valia podendo auxiliar um profissional de saúde em determinadas ocasiões, ajudando como auxílio do diagnóstico, ajudando assim a determinar uma maior ou menor probabilidade de obtenção da doença dependendo das características apresentadas pelo doente.

## Bibliografia

1. Nóbrega, A.I.d., *Técnicas de imagem por Tomografia Computadorizada*, 2014, Universidade São Camilo: Hospital Santa Catarina / Hospital Alemão Osvaldo Cruz. p. 98.
2. Segura, E., et al., *ECOGRAFIA ABDOMINAL E PLEURO-PULMONAR NA URGÊNCIA - PROTOCOLO e-FAST (FOCUSED ASSESSMENT WITH SONOGRAPHY FOR TRAUMA)*, in *Rev Soc Port Anestesiologia* 2014: Rev Soc Port Anestesiologia. p. 6.
3. Grilo, A.M. and M.C. Santos, *A mamografia: experiência psicológica, atitudes e comportamentos do técnico de radiologia facilitadores da colaboração da mulher*. 2013. 9: p. 9.
4. Boaventura, I.A.G., *Números Fuzzy em Processamento de Imagens Digitais e Suas Aplicações na Detecção de Bordas*, in *Departamento de engenharia* 2010, Universidade de São Paulo. p. 218.
5. Zadeh, L.A., *Fuzzy Sets*. Vol. 8. 1965. 6.
6. Cox, E., *The Fuzzy Systems Handbook Second Edition*. 1999: A.P.
7. Yen, J. and R. Langari, *Fuzzy Logic: intelligence, control and information* Tom Robbins. 548.
8. Cancro. L.P.C.o. *Cancro da Mama*. 2015; Available from: [www.ligacontracancro.pt/cancro-da-mama/](http://www.ligacontracancro.pt/cancro-da-mama/). (requerido em Setembro de 2016).
9. TAGGED. *Mama Help*. 2015; Available from: <http://www.mamahelp.pt/index.php/temas/item/49-b-estatisticas>. (requerido em Setembro de 2016).
10. Ficagna and P.E.N., *Mamografia: importância e seus métodos*. 2016.
11. REGATTIERI and N.A.T., *O exame de mamografia: abordagem da usuária, posicionamento e incidências*, 2009, Universidade Tecnológica Federal de Panamá: Colégio Brasileiro de Radiologia.
12. Teixeira and R.F.d.S., *Análise Automática de Imagens de Mamografia: distinção de incidências*, 2012, Universidade do Porto.
13. Oncologia, I.P.d. *Cancro da mama*. 2014 05/10/2014; Available from: <http://www.oncoguia.org.br/conteudo/exames-de-imagem-para-o-diagnostico-do-cancer-de-mama/1389/264/>. (requerido em Setembro de 2016).
14. Souza and F.H., *Mamografia Digital em Comparação com Mamografia Convencional no Rastreamento da Câncer de Mama no Brasil: Revisão Sistemática Custo da Doença e Análise de Custos-Efetividade do Sistema Único de Saúde*, in *Faculdade de medicina* 2012, Universidade Federal do Rio Grande do Sul. p. 212.
15. Ribeiro and P.B., *Esquema CAD para classificação de nódulos em imagens mamográficas digitais baseado na segmentação pelo modelo EICAMM*, in *Escla de Engenharia* 2013, Universidade de São Paulo.
16. Ribeiro, G.H.d.P.T., *Classificação Automática da Densidade Mamária em Tomossíntese*, in *Faculdade de ciências departamento de física* 2016, Universidade de Lisboa.
17. Aguado, A.G. and M.A.C., *Lógica Fuzzy*, in *Faculdade de Tecnologia*, Universidade Estadual de Campinas. p. 12.
18. Simões, M.G. and I.S.S., *Controle e modelagem fuzzy*. Vol. 2. 1999: FAPESP - Fundação da Amparo à Pesquisa do Estado de São Paulo.
19. Oliveira, D.C.d., *Elaboração de modelos linguísticos baseados na teoria dos conjuntos fuzzy para mensuração de qualidade de vida relacionada à saúde*, in *Departamento de medicina preventiva* 2004, Universidade Federal do Rio de Janeiro. p. 135.

20. Bellucci, D.P., *Sistemas Baseados em Regras Fuzzy e Aplicações*, in *Matemática aplicada*2009, Universidade Federal da ABC. p. 57.
21. A.M.P.G., V., *Técnica para Segmentação Automática de Imagens Microscópicas de Componentes Sanguíneos e Classificação Diferencial de Leucócitos Baseada em Lógica Fuzzy*, in *Engenharia Elétrica e de Computadores*2014, Universidade Federal do Rio Grande do Norte. p. 143.
22. Silva, F.C.d., *Processamento de imagens baseado em algoritmos genéticos fuzzy intervalares*, in *Departamento de Informática*2009, Universidade Católica de Pelotas. p. 146.
23. S. N. Silvanandam, S.S., S. N. Deepa, *Introduction to Fuzzy Logic using Matlab*. Springer. 442.
24. Silveira, G.P., L.L.V, and L.C.d. Barros, *Nomograma Com Lógica Fuzzy Para o Câncer De Próstata - NFCP*, in *REVISTA ELETRÔNICA DE MATEMÁTICA*2010: REMat.
25. Savergnini, K.D., L.L.V., and W.E. Matheus, *Modelagem Fuzzy para Predizer os Riscos de Recidiva e Progressão de Tumores Superficiais de Bexiga*. Biomatemática IMECC, 2010: p. 10.
26. Lopes, M.H.B.d.M., et al., *Modelo Baseado em Lógica Fuzzy para o Diagnóstico Diferencial de Enfermagem em Alterações na Eliminação Urinária*, in *Faculdade de Ciências Médicas*, Universidade Estadual de Campinas. p. 5.
27. SILVA, F.C.D., et al., *UM SISTEMA FUZZY PARA AUXÍLIO AO DIAGNÓSTICO DA DENSIDADE MINERAL ÓSSEA.*, in *Departamento de Ciências da Computação*2007. p. 18.
28. Sami Dhahbi, W.B. and E. Zagrouba, *Breast cancer diagnosis in digitized mammograms using curvelet moments*. *Computers in Biology and Medicine*. 2015: p. 12.
29. Mahfouf, M., M.F. Abbod, and D.A. Linkens, *A survey of fuzzy logic monitoring and control utilisation in medicine*, in *Department of Automatic Control and Systems Engineering*2000, The University of Sheffield: *Artificial intelligence in Medicine*. p. 16.
30. Samuel, O.W., M.O. Omisore, and B.A. Ojokoh, *A web based decision support system driven by fuzzy logic for the diagnosis of typhoid fever*. ELSEVIER: *Expert Systems with Applications*, 2013: p. 8.
31. YILMAZ, A. and K. AYAN, *Cancer risk analysis by fuzzy logic approach and performance status of the model*. *Turkish Journal of Electrical Engineering & Computer Sciences*, 2013: p. 17.
32. Kumar, K.J. and S.P. Bhatnagar, *Fuzzy Rule Based Systems in Classification of Terminalia Chebula Fruits of Complementary and Alternative Medicine*. *Journal of Complementary and Integrative Medicine*, 2007. 4(1).
33. Kwiatkowska, M. and K. Kielan, *Fuzzy logic and semiotic methods in modeling of medical concepts*. *Sci Verse Science Direct*, 2012: p. 16.
34. *Matlab, Fuzzy Logic Toolbox™ User's Guide*. 2015: MathWorks. 368.
35. Igor Barone de Medeiros, M.A.S.M., *Aplicação do Sistema de Inferência Nebuloso na Determinação de Medicina Preventiva no Sistema de Saúde Complementar*, in *RESI-Revista Eletrônica de Sistemas de Informação*2006. p. 10.
36. Almeida, P.E.M.d. and A.G. Evsukoff, *Sistemas Fuzzy*. 2000. 22.

## Anexos

A tabela11 que se segue apresenta todos os valores usados nos estudos, para este trabalho, para imagens com nódulos. Foram retiradas as imagens de doente que não apresentavam patologia, pois estes não foram usados neste trabalho, logo não seriam significantes.

Tabela 11 - Valores obtidos a partir da base de dados mini-MIAS database of mammograms. Retirado de: <http://peipa.essex.ac.uk/info/mias.html>

IMAGENS	TIPO DE TECIDO	ANOMALIA DO NÓDULO EXISTENTE	TAMANHO DA ANOMALIA EM MILÍMETROS	SEVERIDADE DA ANOMALIA
MDB126	D	ARCH	23	B
MDB165	D	ARCH	42	B
MDB163	D	ARCH	50	B
MDB099	D	ASYM	23	B
MDB104	D	ASYM	50	B
MDB107	D	ASYM	111	B
MDB223	D	CALC	6	B
MDB226	D	CALC	7	B
MDB226	D	CALC	8	B
MDB236	D	CALC	14	B
MDB222	D	CALC	17	B
MDB240	D	CALC	23	B
MDB226	D	CALC	25	B
MDB223	D	CALC	29	B
MDB290	D	CIRC	45	B
MDB244	D	CIRC	52	B
MDB315	D	CIRC	93	B
MDB063	D	MISC	33	B
MDB207	D	SPIC	19	B
MDB199	D	SPIC	31	B
MDB145	D	SPIC	49	B
MDB198	D	SPIC	93	B
MDB193	D	SPIC	132	B
MDB167	F	ARCH	35	B

MDB152	F	ARCH	48	B
MDB160	F	ARCH	61	B
MDB150	F	ARCH	62	B
MDB097	F	ASYM	34	B
MDB248	F	CALC	10	B
MDB252	F	CALC	23	B
MDB059	F	CIRC		B
MDB132	F	CIRC	18	B
MDB080	F	CIRC	20	B
MDB091	F	CIRC	20	B
MDB005	F	CIRC	26	B
MDB142	F	CIRC	26	B
MDB005	F	CIRC	30	B
MDB010	F	CIRC	33	B
MDB012	F	CIRC	40	B
MDB069	F	CIRC	44	B
MDB132	F	CIRC	52	B
MDB025	F	CIRC	79	B
MDB312	F	MISC	20	B
MDB144	F	MISC	29	B
MDB314	F	MISC	39	B
MDB204	F	SPIC	21	B
MDB195	F	SPIC	26	B
MDB127	G	ARCH	48	B
MDB121	G	ARCH	87	B
MDB083	G	ASYM	38	B
MDB081	G	ASYM	131	B
MDB218	G	CALC	8	B
MDB227	G	CALC	9	B
MDB219	G	CALC	29	B
MDB017	G	CIRC	48	B
MDB019	G	CIRC	49	B
MDB021	G	CIRC	49	B
MDB015	G	CIRC	68	B
MDB002	G	CIRC	69	B
MDB001	G	CIRC	197	B
MDB013	G	MISC	31	B
MDB030	G	MISC	43	B
MDB032	G	MISC	66	B
MDB190	G	SPIC	31	B
MDB175	G	SPIC	33	B

MDB191	G	SPIC	41	B
MDB188	G	SPIC	61	B
MDB130	D	ARCH	28	M
MDB125	D	ARCH	60	M
MDB171	D	ARCH	62	M
MDB170	D	ARCH	82	M
MDB102	D	ASYM	38	M
MDB110	D	ASYM	51	M
MDB105	D	ASYM	98	M
MDB111	D	ASYM	107	M
MDB239	D	CALC	25	M
MDB253	D	CALC	28	M
MDB241	D	CALC	38	M
MDB239	D	CALC	40	M
MDB249	D	CALC	48	M
MDB249	D	CALC	64	M
MDB216	D	CALC		M
MDB058	D	MISC	27	M
MDB202	D	SPIC	37	M
MDB179	D	SPIC	67	M
MDB158	F	ARCH	88	M
MDB155	F	ARCH	95	M
MDB075	F	ASYM	23	M
MDB095	F	ASYM	29	M
MDB092	F	ASYM	43	M
MDB238	F	CALC	17	M
MDB256	F	CALC	37	M
MDB231	F	CALC	44	M
MDB245	F	CALC		M
MDB141	F	CIRC	29	M
MDB028	F	CIRC	56	M
MDB144	F	MISC	27	M
MDB134	F	MISC	49	M
MDB267	F	MISC	56	M
MDB271	F	MISC	68	M
MDB274	F	MISC	123	M
MDB206	F	SPIC	17	M
MDB184	F	SPIC	114	M
MDB148	F	SPIC	174	M
MDB124	G	ARCH	33	M
MDB120	G	ARCH	79	M

MDB117	G	ARCH	84	M
MDB115	G	ARCH	117	M
MDB072	G	ASYM	28	M
MDB090	G	ASYM	49	M
MDB211	G	CALC	13	M
MDB213	G	CALC	45	M
MDB209	G	CALC	87	M
MDB233	G	CALC		M
MDB023	G	CIRC	29	M
MDB270	G	CIRC	72	M
MDB264	G	MISC	36	M
MDB265	G	MISC	60	M
MDB186	G	SPIC	47	M
MDB181	G	SPIC	54	M
MDB178	G	SPIC	70	M

Na tabela apresentada encontram-se todas as características usadas no estudo. A primeira coluna diz respeito às imagens, não se encontram valores seguidos, isto é porque a tabela já foi trabalhada no Excel para que se agrupem os valores de forma a simplificar o estudo dos mesmos. Esta foi agrupada consoante o tipo de tecido, de seguida o tipo de anomalia da massa/nódulo mamário, depois o valor do tamanho da massa/nódulo apresentado no exame e por fim a separação consoante a severidade da anomalia.

A tabela está exatamente como foi retirada da base de dados e por isto o seu conteúdo está descrito em inglês, na segunda coluna, onde se encontram os tipos de tecido, a separação é feita por tecido D, que em português é o tecido denso glandular, de seguida o tecido F, que em português é o tecido gordo e por fim o tecido G, que em português é o tecido gordo glandular. Na terceira coluna encontram-se as anomalias das massas, que neste estudo são separadas em seis tipos já descritos no capítulo VI. Estas seis anomalias encontram-se separadas por ordem alfabética e separadas por grupos dentro de cada tipo de tecido, ou seja no caso do tecido denso, existem seis anomalias presentes e cada uma delas se encontra separada por ordem alfabética. Na quarta coluna encontram-se os tamanhos das anomalias, estes tamanhos foram separados de acordo com a separação anterior e por ordem crescente. Para cada tipo de tecido e cada tipo de anomalia existe um determinado tamanho. Estes tamanhos foram estudados e agrupados segundo o sistema de estadiamento TMN, já

referido no capítulo III na tabela 1. Por fim a quinta coluna da encontra-se a severidade da anomalia, esta pode ser benigna ou maligna, dividindo-se assim em dois grupos.

A seguir encontram-se todas as regras usadas na toolbox. As 18 regras referidas dizem respeito ao primeiro estudo realizado, onde foram usadas como características o tipo de tecido e as anomalias existentes nos tecidos mamários para determinar a probabilidade de uma severidade maligna ou benigna.

- 1. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is ARCH) then (severidade\_da\_anomalia is maligno) (0.5714)
- 2. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is ASYM) then (severidade\_da\_anomalia is maligno) (0.5714)
- 3. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is MISC) then (severidade\_da\_anomalia is maligno) (0.5)
- 4. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is ASYM) then (severidade\_da\_anomalia is maligno) (0.75)
- 5. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is CALC) then (severidade\_da\_anomalia is maligno) (0.6666)
- 6. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is MISC) then (severidade\_da\_anomalia is maligno) (0.625)
- 7. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is SPIC) then (severidade\_da\_anomalia is maligno) (0.6)
- 8. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is ARCH) then (severidade\_da\_anomalia is maligno) (0.6666)
- 9. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is CALC) then (severidade\_da\_anomalia is maligno) (0.5714)
- 10. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is CALC) then (severidade\_da\_anomalia is benigno) (0.5333)
- 11. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is CIRC) then (severidade\_da\_anomalia is benigno) (1)
- 12. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is SPIC) then (severidade\_da\_anomalia is benigno) (0.7142)
- 13. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is ARCH) then (severidade\_da\_anomalia is benigno) (0.6666)
- 14. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is CIRC) then (severidade\_da\_anomalia is benigno) (0.8571)
- 15. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is ASYM) then (severidade\_da\_anomalia is benigno) (0.5)
- 16. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is CIRC) then (severidade\_da\_anomalia is benigno) (0.75)
- 17. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is MISC) then (severidade\_da\_anomalia is benigno) (0.6)

- 18. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is SPIC) then (severidade\_da\_anomalia is benigno) (0.5714)

As 40 regras apresentadas a referem-se ao segundo estudo realizado, onde foram usadas como características o tipo de tecido, as anomalias apresentadas pela tecido mamário e o tamanho dos nódulos existentes, para assim determinar a probabilidade de uma severidade maligna ou benigna.

- 1. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is ARCH) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.4285)
- 2. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is ASYM) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.2857)
- 3. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is CALC) and (tamanho\_da\_anomalia is pequeno) then (severidade\_da\_anomalia is benigno) (0.3333)
- 4. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is CIRC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.3333)
- 5. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is CIRC) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is benigno) (0.6666)
- 6. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is SPIC) and (tamanho\_da\_anomalia is pequeno) then (severidade\_da\_anomalia is benigno) (0.1428)
- 7. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is SPIC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.2857)
- 8. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is SPIC) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is benigno) (0.2857)
- 9. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is ARCH) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.3333)
- 10. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is CIRC) and (tamanho\_da\_anomalia is pequeno) then (severidade\_da\_anomalia is benigno) (0.2857)

- 11. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is CIRC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.4285)
- 12. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is CIRC) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is benigno) (0.1428)
- 13. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is MISC) and (tamanho\_da\_anomalia is pequeno) then (severidade\_da\_anomalia is benigno) (0.125)
- 14. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is SPIC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.4)
- 15. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is ARCH) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.1666)
- 16. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is ASYM) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is benigno) (0.25)
- 17. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is CALC) and (tamanho\_da\_anomalia is pequeno) then (severidade\_da\_anomalia is benigno) (0.2857)
- 18. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is CALC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.1428)
- 19. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is CIRC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.375)
- 20. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is CIRC) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is benigno) (0.375)
- 21. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is MISC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.4)
- 22. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is SPIC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is benigno) (0.4285)
- 23. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is ARCH) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is maligno) (0.4285)
- 24. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is ASYM) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is maligno) (0.4285)
- 25. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is CALC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is maligno) (0.3333)

- 26. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is CALC) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is maligno) (0.6666)
- 27. If (tipo\_de\_tecido is denso\_glandular) and (anomalia\_do\_tecido is MISC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is maligno) (0.5)
- 28. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is ARCH) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is maligno) (0.3333)
- 29. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is ASYM) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is maligno) (0.75)
- 30. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is CALC) and (tamanho\_da\_anomalia is pequeno) then (severidade\_da\_anomalia is maligno) (0.3333)
- 31. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is CALC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is maligno) (0.3333)
- 32. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is MISC) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is maligno) (0.25)
- 33. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is MISC) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is maligno) (0.375)
- 34. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is SPIC) and (tamanho\_da\_anomalia is pequeno) then (severidade\_da\_anomalia is maligno) (0.2)
- 35. If (tipo\_de\_tecido is gordo) and (anomalia\_do\_tecido is SPIC) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is maligno) (0.4)
- 36. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is ARCH) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is maligno) (0.5)
- 37. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is ASYM) and (tamanho\_da\_anomalia is medio) then (severidade\_da\_anomalia is maligno) (0.5)
- 38. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is CALC) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is maligno) (0.1428)
- 39. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is MISC) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is maligno) (0.2)
- 40. If (tipo\_de\_tecido is gordo\_glandular) and (anomalia\_do\_tecido is SPIC) and (tamanho\_da\_anomalia is grande) then (severidade\_da\_anomalia is maligno) (0.2857)