



UNIVERSIDADE DA BEIRA INTERIOR
Engenharia

Determinação de parâmetros em operadores para segmentação de nódulos mamários com o auxílio de Algoritmos Genéticos

Emanuel José Soares Bordalo Teixeira

Dissertação para obtenção do Grau de Mestre em
Engenharia Eletrotécnica e de Computadores: Sistemas Biónicos
(2º ciclo de estudos)

Orientador: Prof. Doutor José António Menezes Felippe de Souza

Covilhã, outubro de 2014

Agradecimentos

Agradeço a Deus por todo o meu percurso, toda a minha energia e auxílio.

Aos meus pais, que estiveram presentes em todos os momentos, foram sempre a minha referência e os meus alicerces para nunca desistir, foram o meu apoio incondicional, o meu incentivo, motivação, sem os quais nada seria possível, pois proporcionaram-me as melhores condições para alcançar mais esta etapa...

Ao meu Orientador o Professor Doutor José António Menezes Felipe de Souza, agradeço a sabedoria transmitida, a enorme disponibilidade que sempre demonstrou, a valiosa ajuda em apoio bibliográfico, pela forma como orientou o meu trabalho que contribuiu para a realização desta dissertação e pela amizade então demonstradas.

A todos os meus colegas e amigos, pelo companheirismo e entreaajuda e apoio moral.

Ao Professor Doutor António Manuel Gonçalves Pinheiro pelas instalações cedidas e pela ajuda em apoio bibliográfico. Aos Professores: Doutor José Alberto Fonseca Moutinho e o Professor Doutor Mário Marques Freire pelos conteúdos cedidos.

A toda a minha família...

Resumo

O cancro é uma das causas de morte mais frequente em todo o mundo e o cancro da mama é o mais comum em mulheres. Logo, é de extrema importância a detecção atempada, pois o diagnóstico precoce e o tratamento do cancro mamário aumentam significativamente as probabilidades de tratamento bem-sucedido.

O exame da mama através de imagens por ultrassom (US) é o principal coadjuvante da mamografia, tornando-se deste modo um exame ainda mais importante, complementado por sistemas CAD, detetando e classificando nódulos mamários.

Na presente dissertação foi desenvolvida uma metodologia, baseada em Operadores Morfológicos, Detecção de contornos, Métodos de Thresholding entre outros, para segmentar automaticamente imagens de US mamários, que determina diversos parâmetros com auxílio de Algoritmos Genéticos (AG).

Palavras-chave

Algoritmos Genéticos, Cancro da mama, CAD, Segmentação, Ultrassom mamário.

Abstract

Cancer is one of the most frequent causes of death all over the world and breast cancer is the most common in women. Therefore, it's extremely important beforehand detection, since early diagnosis and treatment of breast cancer increases significantly the chances of successful treatment.

The breast examination through ultrasound images (US) is the main assistant method to mammography, therefore becoming an even more important exam, complemented by CAD systems, detecting and classifying breast masses.

In present dissertation it was developed a methodology, based on Morphological Operators, Contour Detection, Thresholding Methods among others, for automatically segment US breasts pictures, that determines several parameters with the support of Genetic Algorithms (GA).

Keywords

Genetic Algorithms, Breast cancer, CAD, Segmentation, Breast ultrasound.

Índice

Capítulo 1	1
Introdução.....	1
1.1.Enquadramento	1
1.2. Estrutura do Trabalho.....	2
Capítulo 2	3
Fundamentos teóricos.....	3
2.1. Cancro.....	3
2.1.1. O cancro da mama.....	3
2.2. Detecção do cancro da mama	4
2.2.1. Exame clínico da mama	5
2.2.2. Mamografia de diagnóstico	6
2.2.3. Ecografia	6
2.2.4. Ressonância magnética.....	6
2.2.5. Biópsia.....	6
Capítulo 3	7
Diagnóstico por Ultrassons e Sistemas CAD	7
3.1. Ultrassons	7
3.1.1. Transdutores de ultrassons -materiais piezoelétricos.....	7
3.1.2. Modos de visualização da informação obtida através dos ecos de ultrassons	8
3.2. Ecografia mamária.....	8
3.2.1. Alcance e limitação dos Ultrassons	9
3.3. Diagnóstico assistido por computador – CAD (Computer-Aided Diagnosis)	9
3.3.1. Os benefícios do CAD.....	10
3.3.2. CADs comerciais	10
3.3.3. Exemplo de um sistema de CAD.....	10
Capítulo 4	11
Processamento de imagens	11
4.1. Processamento digital de imagens.....	11
4.2. Técnicas para minimização de ruído speckle.....	11
4.3. Técnicas de segmentação	12
4.3.1. Detecção de contornos.....	12
4.3.1.1. Filtro Sobel	14
4.3.2. Métodos de Clustering e de Thresholding.....	14

4.4. Morfologia Matemática.....	15
4.4.1. Tipos de Operadores Morfológicos	16
4.4.2. Elemento Estruturante	16
4.4.3. Definições Fundamentais.....	17
4.4.4. Morfologia binária	18
4.4.5. Operadores básicos	18
4.4.5.1. Erosão	18
4.4.5.2. Dilatação	18
4.4.5.3. Abertura.....	20
4.4.5.4. Fecho.....	20
Capítulo 5	21
Algoritmos Genéticos	21
5.1. Algoritmos Genéticos	21
5.1.1. Biologia e Evolução em problemas computacionais	22
5.1.2. Termos técnicos da Biologia aplicados nos Algoritmos Genéticos.....	24
5.1.3. Componentes de um Algoritmo Genético	25
5.1.4. Estrutura de um Algoritmo Genético Simples.....	27
5.2. Optimization Toolbox™ - Toolbox Algoritmo Genético [31].....	28
5.2.1. Escrever uma <i>M-File</i> para a função que se pretende otimizar	28
5.2.2. Usar a Toolbox.....	28
5.2.3. Invocar a função “ga” na linha de comandos.....	28
5.2.4. Uso da Toolbox, com a interface gráfica.....	29
5.2.5. Exibir Gráficos.....	31
5.2.6. Definir Opções para o AG.....	32
5.2.7. Diversidade	33
5.2.8. Opções da População	33
5.2.9. Opções de dimensionamento da aptidão.....	33
5.2.10. Opções de Seleção.....	34
5.2.11. Opções de reprodução.....	36
5.2.12. Mutação e Crossover	36
5.2.13. Opções de mutação	37
5.2.14. A fração Crossover / Recombinação.....	37
Capítulo 6	39
Determinação de parâmetros dos operadores no algoritmo desenvolvido com auxílio dos	
Algoritmos Genéticos	39
6.1. Determinação de parâmetros dos operadores com auxílio dos Algoritmos Genéticos...39	

6.2. Algoritmo	40
6.2.1. Funcionamento do Algoritmo Genético	46
Capítulo 7	49
Parte prática - Resultados.....	49
7.1. Resultados.....	49
Capítulo 8	61
Conclusão	61
8.1. Trabalho futuro	62
Referências	63
Anexos	67
O cancro da mama	67
Constituição do órgão: A mama	67
Nódulos mamários	68
Deteção do cancro da mama	69
Exame clínico da mama	70
Biópsia	70
Ultrassons.....	72
Transdutores de ultrassons -materiais piezoelétricos	72
Modos de visualização da informação obtida através dos ecos de ultrassons	76
Alcance e limitação dos Ultrassons	78
Bibliografia adicional.....	79

Lista de Figuras

Figura 1 - Gráfico - Os cinco cancros mais frequentes nas mulheres [4].....	3
Figura 2 - Visão de uma mamografia [6].....	6
Figura 3 - Arquitetura típica de um sistema CAD. O quadrado cinza simboliza o bloco <i>off-line</i> , enquanto, o resto da figura pertence ao bloco on-line. [12]	10
Figura 4 Limiar em histograma bimodal [19].....	15
Figura 5 - Exemplo de elementos estruturantes: (a) em coluna, (b) em linha, (c) em cruz, (d) em quadrado, (e) em octógono, (f) em disco [24].	17
Figura 6 - Operações de Erosão e dilatação [23], [25].....	19
Figura 7 - Operações de abertura e fecho [23], [25].	20
Figura 8 - Semelhanças entre o ciclo da evolução e o ciclo de um Algoritmo Genético.....	25
Figura 9 - Representação de um Algoritmo Genético Simples.....	26
Figura 10 - Interface gráfica da Toolbox Algoritmo Genético	29
Figura 11 - Campo do número de variáveis da Toolbox Algoritmo Genético.....	30
Figura 12 - Campo Run Solver da Toolbox Algoritmo Genético	30
Figura 13 - Campo Status and Results da Toolbox Algoritmo Genético	30
Figura 14 - Campo Plots da Toolbox Algoritmo Genético.....	31
Figura 15 - Gráficos do melhores valores médios da função de aptidão em cada geração da Toolbox Algoritmo Genético	31
Figura 16 - Campo Stopping Criteria da Toolbox Algoritmo Genético	32
Figura 17 - Representação de Seleção por Torneio binário	35
Figura 18 - Representação da Seleção por roleta do Rácio de Fitness	35
Figura 19 - Um exemplo de crossover de um ponto: (a) dois indivíduos são escolhidos. (b) um ponto (4) de crossover é escolhido. (c) são recombinadas as características, gerando dois novos indivíduos. (d) Recombinação por crossover de um ponto com ponto de corte no gene 2	38
Figura 20 - Recombinação por crossover de dois pontos.....	38
Figura 21 - Recombinação por crossover uniforme.....	38
Figura 22 - Mutação no gene 3 do cromossoma.....	38
Figura 23 - Nódulos manualmente detetadas.....	40
Figura 24 - Imagem original	41
Figura 25 - Máscara binária do gradiente	41
Figura 26 - Máscara de uma dilatação.....	42
Figura 27 - Máscara da inversão.....	42
Figura 28 - Máscara do preenchimento de lacunas de Interiores	43
Figura 29 - Máscara da remoção de objetos conectados ao contorno	43

Figura 30 - Máscara da erosão	44
Figura 31 - Máscara das Operações intermédias.....	44
Figura 32 - Máscara da apresentação da maior área	45
Figura 33 - Máscara obtida a partir da segmentação manual e máscara obtida a partir do algoritmo	45
Figura 34 - Imagem original com o contorno à volta do nódulo	46
Figura 35 - Fluxograma ilustrando o funcionamento do sistema AG a respetiva função fitness.	47
Figura 36 - 1ª Imagem.....	49
Figura 37 - Exemplo de duas imagens de ultrassom mamário com marca	49
Figura 38 - Exemplo de quatro imagens de ultrassom mamário com marca, segmentadas	50
Figura 39 - Segmentação inicial.....	51
Figura 40- Interface gráfica	51
Figura 41 - Diversos gráficos que podem ser produzidos pela Toolbox	52
Figura 42 - “Average Distance Between Individuals”	52
Figura 43 - Imagem ilustrativa mostrando que a memória foi excedida	52
Figura 44 - Gráficos obtidos a partir da opção de seleção Stochastic uniform.....	55
Figura 45 - Gráficos obtidos a partir da opção de seleção Reminder	55
Figura 46 - Gráficos obtidos a partir da opção de seleção Uniform.....	55
Figura 47 - Gráficos obtidos a partir da opção de seleção Roleta	56
Figura 48 - Gráficos obtidos a partir da opção de seleção Torneio	56
Figura 49 - Gráficos de Best Fitness e Current Best Individual.....	57
Figura 50 - Imagens segmentadas com o contorno à volta do nódulo	58
Figura 51 - Imagens segmentadas sem nódulo	59
Figura 52 - Constituição de uma mama/secção transversal [36]	67
Figura 53 - Irrigação sanguínea e drenagem linfática da mama [37].....	67
Figura 54 - Comportamento de um material piezoelétrico [39].....	73
Figura 55 - Diferentes tipos de transdutores: Endoscópico, convexo, lineares, laparoscópicos...[40].....	73
Figura 56 - Medição de distâncias por intermédio do princípio do tempo de trânsito [41].....	74
Figura 57 - Os princípios físicos da imagem e os artefactos na ecografia em modo B. [42]	75
Figura 58 a) e b) - Sistema de ultrassons e leitura para operação no Modo A. [42] [43].....	76
Figura 59 a) e b) - diagrama simplificado de um sistema de ultrassonografia típico [42] e ultrassom fetal Modo-B. [43]	77
Figura 603 - Ultrassom Modo-M. [44]	77

Lista de Tabelas

Tabela 1 - Termos técnicos da Biologia aplicados nos Algoritmos Genéticos	24
Tabela 2 - Testes usando Stall Time 50	53
Tabela 3 - Testes usando Stall Time 100	54
Tabela 4 - Diversidade dos Tempos de execução.....	54
Tabela 5 - Testes de Tempo para o número de imagens 5, 10, 20 (Stall Time 50, População 25)	54

Lista de Acrónimos

UBI	Universidade da Beira Interior
AG	Algoritmo Genético
CAD	Computer Aided Diagnosis
OM	Operadores morfológicos
US	Ultrassom

Capítulo 1

Introdução

1.1. Enquadramento

Como aluno do ramo do mestrado Sistemas Biónicos de Engenharia Eletrotécnica e de Computadores, não podia deixar de me interessar por um tema tão pertinente como este. E o meio para solucionar o mesmo é baseado em determinados processos biológicos dos seres vivos, isto é, baseado em sistemas bioinspirados, os Algoritmos Genéticos (AG).

Em todo o mundo, milhões de pessoas são diagnosticadas com cancro e o cancro da mama é o mais comum. A investigação nesta área é muito importante e é inquestionavelmente necessária. O progresso aliado da tecnologia tem decifrado algumas das suas causas, sobre a forma como se desenvolve e propaga. Também são estudadas novas formas de o prevenir, detetar e tratar, tendo sempre em atenção a melhoria da qualidade de vida das pessoas com cancro, durante e após o tratamento.

Sendo de extrema importância a deteção atempada, pois o diagnóstico precoce e o tratamento do carcinoma mamário aumentam significativamente as possibilidades de tratamento, tendo deste modo um maior benefício social e económico do que o custo do tratamento da doença na fase mais avançada.

A mamografia é a técnica de diagnóstico que mais tem contribuído, através de um programa de acompanhamento periódico, na redução da mortalidade por cancro da mama. Porém, estudos recentes determinam que a probabilidade de mulheres morrerem de cancro da mama era igual à das mulheres que não fizeram o exame. [1]

Isso ilustra a importância das ecografias como exame para auxiliar as mamografias. O exame da mama através de imagens por ultrassom (US) é o mais importante coadjuvante da mamografia, tornando-se deste modo um exame ainda mais importante, complementado por sistemas CAD, detetando e classificando nódulos mamários (entre outros), reduzindo a maioria de falsos Positivos e Negativos, evitando tratamentos e outras ações desnecessárias.

Na presente dissertação foi desenvolvida uma metodologia, baseada em Operadores Morfológicos (OM), Deteção de contornos, Métodos de Thresholding entre outros, para segmentar automaticamente imagens de US B-mode mamários. Sem processamento prévio da imagem, não perdendo a maioria das informações que podem ser dissipadas com algumas técnicas [2], [3], determinando diversos parâmetros com auxílio de Algoritmos Genéticos.

1.2. Estrutura do Trabalho

Este trabalho é constituído por 8 capítulos. No capítulo 2 é feita a descrição do problema e os meios para deteção.

No capítulo 3 é apresentado com maior detalhe o método de deteção por ultrassons e os Sistemas CAD.

No capítulo 4 é abordado o processamento de imagens bem como algumas técnicas de segmentação incluindo as necessárias aos procedimentos do trabalho.

Tendo em vista o uso que será feito dos AG, o capítulo 5 é dedicado a uma abordagem mais extensiva aos AG tal como da *Optimization Toolbox*[™] onde está inserida a *Toolbox do Algoritmo Genético*.

No capítulo 6 fazemos a determinação de parâmetros dos operadores no algoritmo com auxílio dos AG. Desta forma o algoritmo que desenvolvemos torna-se a função fitness do AG.

Já no capítulo 7 consiste o procedimento prático.

Por fim no último cap. Ocorre a reflexão de todo o trabalho, bem como as dificuldades sentidas e trabalho futuro que possa ser desenvolvido.

Apresentando ainda as Referências e os Anexos que auxiliaram a compreensão de algumas questões.

Capítulo 2

Fundamentos teóricos

2.1. Cancro

O cancro é a proliferação anormal de células. Naturalmente, as células crescem e dividem-se para formar novas. Durante ciclo celular, estas envelhecem, morrem (apoptose) e são substituídas por outras novas.

Em certas circunstâncias, este processo controlado não corre da melhor forma: formam-se células novas, sem que o organismo necessite e, ao mesmo tempo, as células velhas não morrem. Este conjunto de células extra formam um tumor. Nem todos os tumores correspondem a cancro. Os tumores podem ser benignos ou malignos.

Os tumores benignos não são cancro: Raramente põem a vida em risco, na sua maioria, podem ser removidos e, muitas vezes, regredem. As células destes não disseminam para os tecidos em volta ou para outras partes do organismo - a metastização.

Os tumores malignos são cancro: Normalmente são mais graves que os tumores benignos, podendo colocar a vida em risco; podem ser removidos, embora possam voltar a surgir. Estas podem invadir e danificar os tecidos e órgãos circundantes e ainda, libertar-se do tumor primitivo e entrar na corrente sanguínea ou no sistema linfático - processo de metastização das células cancerosas, a partir do cancro original (tumor primário), formando novos tumores noutros órgãos.

2.1.1. O cancro da mama

O cancro da mama é o tipo de cancro mais comum entre as mulheres (não considerando o cancro da pele), e corresponde à segunda causa de morte por cancro, na mulher.

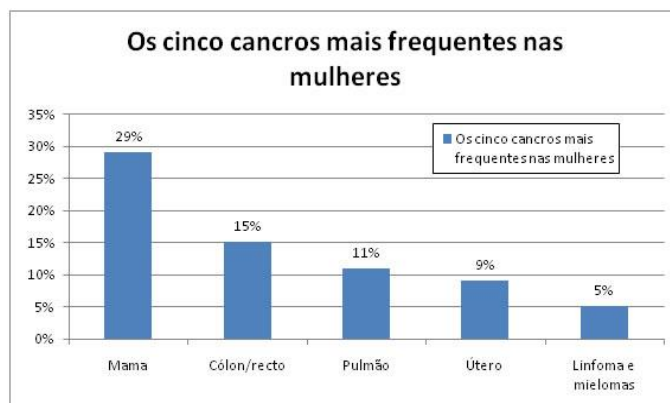


Figura 1 - Gráfico - Os cinco cancros mais frequentes nas mulheres [4]

Anualmente surgem mais de 1,7 milhões de novos casos, com 500mil mortes, números que subiram 11% e 8,4%, respetivamente, desde 2008. A incidência do cancro da mama está a aumentar de ano para ano, segundo estimativas da Organização Mundial da Saúde (OMS). Só em Portugal, anualmente são detetados cerca de 4500 novos casos de cancro da mama, e 1500 mulheres morrem com esta doença.

O cancro da mama é uma das doenças com maior impacto na nossa sociedade, não só por ser muito frequente, e associado a uma imagem de grande gravidade, mas também porque agride um órgão cheio de simbolismo, na maternidade e na feminilidade. Em Portugal, cerca de 1% a 2% de todos os cancros da mama são no homem. [5]

O nódulo, ao aparecer na mama de uma mulher adulta, não deve ser considerado normal e sempre deve ser verificado por um médico. Embora a maioria não apresente perigo (menos de um quarto de todos os nódulos mamários são considerados cancerosos), o diagnóstico logo no início e o tratamento do carcinoma mamário aumentam significativamente as possibilidades de um bom resultado.

2.2. Detecção do cancro da mama

É muito importante fazer exames de rastreio, antes de surgirem quaisquer sinais ou sintomas; para a medicina ajudar a detetar e tratar precocemente o cancro. Se o cancro for detetado precocemente, a probabilidade do tratamento ser eficaz e bem-sucedido é muito mais elevada.

O médico pode sugerir a realização de exames de rastreio do cancro da mama, antes de se desenvolverem quaisquer sintomas: Mamografias de rastreio, Exames clínicos da mama, Mamografia de Diagnóstico.

A mamografia mostra, muitas vezes, um nódulo (ou caroço) na mama, antes que este possa ser sentido ou palpado. Pode, também, mostrar uma agregação de pequenas partículas de cálcio. As partículas chamam-se microcalcificações. Tanto os caroços como estas agregações podem ser sinais de cancro.

Na mamografia, se o médico identificar uma área anormal pode pedir que seja repetida a mamografia. Pode, ainda, ser necessário fazer uma biópsia. A biópsia é o único processo através do qual se podem ter certezas quanto à existência de cancro.

A mamografia tem sido a melhor "ferramenta" de que os médicos dispõem para descobrir o cancro em fase precoce. No entanto:

- A mamografia pode não detetar alguns cancros que estejam já presentes; são os chamados "falsos negativos".

- A mamografia pode detetar algo que, mais tarde, se verifique não ser um cancro: são os chamados "falsos positivos".
- Alguns tumores, de crescimento rápido, podem já ter-se metastizado para outras partes do corpo, antes que a mamografia os tenha detetado.
- A mamografia (bem como os raios-X aos dentes, e outros raios-X de rotina) usa doses muito pequenas de radiação. Em geral, os benefícios superam os riscos, a exposição repetida aos raios-X pode ser nociva.

Porém, estudos recentes (publicados em fevereiro de 2014) determinam que a probabilidade de mulheres morrerem de cancro da mama era igual à das mulheres que não fizeram o exame. [1]

Realizado por investigadores canadianos da Universidade de Toronto - "Toronto Centre for Medical Imaging", que acompanhou 90 mil mulheres, durante 25 anos. Concluindo que no grupo de mulheres que se submeteram a uma mamografia todos os anos, durante cinco anos, a probabilidade de morrerem de cancro da mama era igual à das mulheres que não fizeram o exame. Mais, 22% das anomalias registadas não eram malignas. Estas pessoas foram tratadas, tendo-se percebido, posteriormente, que a lesão descoberta não era cancerosa.

"Estima-se que, só nos EUA, em 30 anos, 1,3 milhões de pacientes tenham sido sobrediagnosticadas (50% dos cancros detetados) e falsamente classificadas como pacientes oncológicas. "Há situações que se resolvem sozinhas. Se a mulher repetir o exame dali a seis meses, já não se vê nada na ecografia", nota o médico António Vaz Carneiro, diretor do Centro de Estudos de Medicina Baseada na Evidência, da Faculdade de Medicina de Lisboa."

"É preciso rastrear 2 mil mulheres, durante dez anos, para salvar uma vida. Neste lote, dez delas vão ser tratadas sem ser preciso e 200 terão um falso alarme (um falso positivo, descartado com outros exames).[1]

Isso ilustra a importância das ecografias como exame para auxiliar as mamografias. O exame da mama através de imagens por ultrassom (US) é o mais importante coadjuvante da mamografia, tornando-se deste modo um exame ainda mais importante, complementado por sistemas CAD, detetando e classificando nódulos mamários (entre outros), limitando a maioria de falsos Positivos e Negativos, evitando tratamentos e outras ações desnecessárias, que se podem tornar tão mortais como um tumor maligno!

2.2.1. Exame clínico da mama

O médico procura quaisquer diferenças entre as mamas, incluindo diferenças invulgares de tamanho ou forma. Na pele, é verificada a presença de vermelhidão, depressões cutâneas ou outros sinais anormais. Os mamilos devem ser pressionadas para verificar se existe alguma secreção ou perda de líquido.

Deverá fazer um exame físico. O médico irá fazer perguntas relacionadas com a história clínica e familiar. Pode precisar de fazer uma biópsia, para verificar se na zona suspeita existem células cancerosas.

2.2.2. Mamografia de diagnóstico

A mamografia de diagnóstico é constituída por imagens de raio-X da mama, para que sejam obtidas imagens mais claras e detalhadas de qualquer área que pareça suspeita ou anormal.

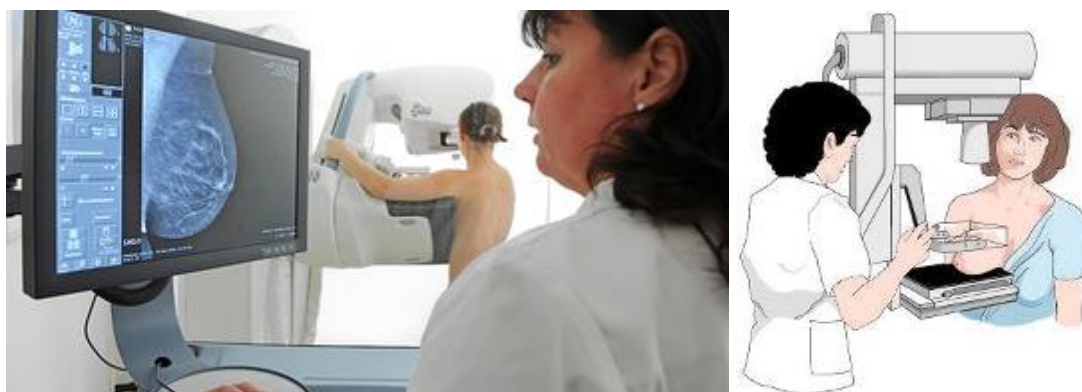


Figura 2 - Visão de uma mamografia [6]

2.2.3. Ecografia

Através de ondas de som de alta-frequência, a ecografia (ultrassons) pode, frequentemente, mostrar se um nódulo é um quisto, cheio de líquido, ou uma massa sólida que pode, ou não, ser cancerosa. O médico vê estas imagens num monitor. Após o exame, as imagens podem ser impressas ou gravadas em vídeo. Este exame pode ser usado em conjunto com a mamografia, como complemento imagiológico.

2.2.4. Ressonância magnética

Na ressonância magnética (RM), liga-se um potente íman a um computador, para que sejam produzidas imagens mais detalhadas dos tecidos internos da mama. O médico vê estas imagens num monitor ou pode gravá-las. A RM pode ser usada em conjunto com a mamografia.

2.2.5. Biópsia

Muitas vezes é necessário retirar tecido ou líquido, para ajudar o médico a perceber se é um cancro. Este procedimento é denominado por biópsia.

Capítulo 3

Diagnóstico por Ultrassons e Sistemas CAD

3.1. Ultrassons

Ultrassons são vibrações mecânicas com frequência acima da audição humana que varia entre 20 e 20.000Hz. Em aplicações médicas as frequências empregadas situam-se na gama dos 500 kHz e 100 MHz.

Sinais acústicos com estas frequências são direcionados para os tecidos humanos onde se propagam à velocidade do som. À medida que viajam pelos vários tecidos, as ondas sonoras sofrem processos de absorção, refração, reflexão e dispersão.

Processo de propagação das ondas sonoras depende das propriedades acústicas dos tecidos (densidade, velocidade do som, coeficiente de absorção e homogeneidade) e das alterações dessas propriedades nas zonas de interface. A velocidade de propagação do som (c) na água, no sangue e na maior parte dos tecidos do corpo humano é aproximadamente 1500 ± 100 m/s.

O ultrassom é uma onda mecânica logo depende de um meio para se transmitir, ao invés da energia iónica (i.e raio X) que também pode ser transmitido no vácuo. Pelo facto de os ultrassons não serem uma energia iónica, estes não apresentam efeitos biológicos nocivos com as intensidades utilizadas na prática clínica (ao invés das intensidades utilizadas no tratamento de lesões desportivas e até como contraceptivos masculinos) [7]. Tendo a possibilidade de obter imagens em tempo real, prover informações das propriedades elásticas dos tecidos, o ser um método não invasivo e não ionizante, tornam o uso cada vez mais frequente em aplicações médicas.

3.1.1. Transdutores de ultrassons -materiais piezoelétricos

Um equipamento de ultrassons é constituído por uma unidade de geração e transmissão de pulsos eléctricos para excitação dos transdutores, uma unidade de receção e amplificação dos sinais captados, uma unidade de controlo e processamento que é utilizada para configurar os parâmetros das unidades de transmissão e receção, e uma unidade para visualização dos resultados do processamento.

O transdutor converte energia eléctrica em vibração mecânica e vice-versa, e define a direcção, frequência e geometria do feixe de som.

Os materiais piezoelétricos podem ser cristais de quartzo, que possuem uma elevada sensibilidade e reduzida largura de banda.

Existem diversos tipos de transdutores, cada um adequado para um diferente tipo de exame. O transdutor convexo é o mais adequado para o exame de abdómen e pélvis, e o linear, para o exame da tiroide e mamas. Os transdutores de frequência elevada (maiores que 7,5MHz) são utilizados para exame de tecidos superficiais, como a mama... Quanto maior a frequência do transdutor, menor o comprimento da onda sonora e melhor a resolução espacial.

3.1.2. Modos de visualização da informação obtida através dos ecos de ultrassons

A técnica de ultrassons para diagnóstico compreende basicamente três modos: o Modo A (de Amplitude), Modo B (de Brilho) e o Modo M (de Movimento).

O Modo B é amplamente utilizado em obstetrícia, e anormalidades podem ser mapeadas sem o risco da exposição do feto à radiação ionizante. Outras aplicações importantes: exames abdominais para localização de tumores e análise da imagem de órgãos internos; em ginecologia, para localização de quistos e tumores no ovário e mamas, no músculo-esquelético para observação dos membros inferiores e superiores; e em cardiologia, na visualização das paredes e válvulas cardíacas.

3.2. Ecografia mamária

A ecografia mamária é um exame de ecografia destinado ao estudo da mama. É um ótimo exame para estudo das glândulas mamárias quer, em algumas situações, como exame diagnóstico de primeira linha, quer como complemento de outros exames, em particular da mamografia.

Em algumas situações a ecografia é o exame de primeira linha no estudo da mama. É o caso das mulheres jovens e das grávidas. É de grande importância como complemento da mamografia, especialmente em mamas de elevada densidade radiológica ou para ajudar a esclarecer algumas alterações observadas na mamografia. Muitas vezes deteta pequenas lesões (até mesmo cancros), antes de serem detetados clinicamente ou por vezes até ocultos na mamografia.

Para realizar ecografia mamária não é necessária qualquer preparação. O aparelho de ecografia mamário é idêntico ao utilizado para as restantes ecografias. O exame é realizado com a paciente deitada, é-lhe colocado um gel próprio para aumentar a condutividade entre o transdutor e o corpo. O exame não é doloroso. Não há habitualmente inconveniente em ser executada em qualquer altura do ciclo menstrual. O exame é executado por médicos radiologistas com experiência em patologias mamárias. Como dito anteriormente por não ser uma radiação ionizante entre outros fatores não se conhecem riscos para realização da ecografia mamária.

3.2.1. Alcance e limitação dos Ultrassons

A ultrassonografia mamária diagnóstica e intervencionista tem grande participação como exame complementar à mamografia e à clínica, tornando-se um método valioso e bem estabelecido no diagnóstico das doenças mamárias

É um exame que permite obter informações da textura e conteúdo de nódulos mamários. Outra limitação importante é a identificação de tumores menores que 1 cm, quando estes estão localizados profundamente em mamas volumosas e com grande quantidade de tecido adiposo, a identificação torna-se precária, porque pequenos tumores e porções de gordura geram imagens semelhantes. Portanto, na maioria dos casos é usado como método complementar da mamografia.

A ultrassonografia tem grande aplicabilidade na diferenciação entre tumores císticos e sólidos e também é capaz de identificar lesões no interior de um quisto, indicando a remoção através de cirurgia. Este método apresenta as vantagens de ser um procedimento não-invasivo, bem tolerado pelos pacientes e que pode trazer informações importantes, complementares ao exame físico e à mamografia. Amplamente usado como auxiliar no diagnóstico médico dentro da obstetrícia e ginecologia, oftalmologia, neurologia e cardiologia além da sua utilização como ferramenta comum em procedimentos terapêuticos.

3.3. Diagnóstico assistido por computador - CAD (Computer-Aided Diagnosis)

A ideia de sistemas auxiliarem radiologistas por parte de um computador para detetarem cancro de mama não é recente [8]. No entanto, hoje em dia o desenvolvimento rápido de sistemas digitais mamográficos completos tem sido acompanhado pelo aumento natural de tais sistemas. Um sistema de (CAD) é um conjunto de ferramentas automáticas ou semidesenvolvidas para ajudar os médicos radiologistas na deteção e/ou avaliação de imagens médicas [8],[9].

Como Taylor [10] sugeriu, o desenvolvimento de métodos automáticos de classificação do tecido mamário são justificados, pelo menos, por dois fatores:

- Para permitir uma melhor rentabilização do tempo e as habilitações dos observadores especializados, permitindo que as mamografias deixem de ser dificilmente examinados pelos leitores mais experientes.
- Para aumentar as possibilidades de deteção de anormalidades com o CAD será mais fácil.

3.3.1. Os benefícios do CAD

Em 2001 Freer e Ulissey usaram um banco de dados contendo 12,860 pacientes concluíram que o uso do CAD na interpretação de imagens médicas pode aumentar a detecção de neoplasias malignas em estágio inicial [10]. Os benefícios de sistemas CAD serão vistos no futuro, quando os algoritmos internos de detecção de anormalidades, diagnóstico e preparação de especialistas sejam melhorados.

3.3.2. CADs comerciais

O pioneiro foi o Image Checker da R2 Technologies Inc. [11]. Kodak Mamografia Motor CAD em 2004.

3.3.3. Exemplo de um sistema de CAD

Mais detalhadamente, um sistema CAD é composto por:

- **Interface simples** - A interação entre o computador e o radiologista tem de ser tão simples quanto possível. Além disso, a interface fornece um grande número de ferramentas para ajudar, por exemplo, a detetar anormalidades, a medir as áreas de interesse, a comparar as imagens...
- **Base de dados médicas** - (mamográfica e/ou ecográfica). Um sistema de CAD contém uma base de dados de imagens médicas previamente diagnosticadas, a fim de ser capaz de retornar os casos mais semelhantes ao observador. Esta base de dados constitui o conhecimento do sistema. Como casos mais conhecidos (e corretamente diagnosticados) estão contidos na base de dados.
- **Caso Desconhecido** - O caso desconhecido é o novo caso que o observador deseja diagnosticar.
- **Segmentação** - Segmentação refere-se à detecção de anormalidades nas imagens.
- **Caracterização** - A caracterização das imagens é feita de acordo com as características extraídas das anormalidades
- **Feedback Relevância**. O objetivo deste passo final é a refinar o resultado obtido.

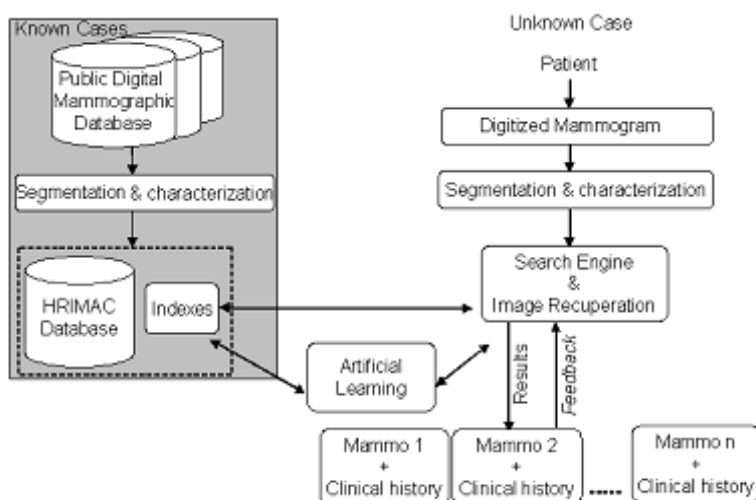


Figura 3 - Arquitetura típica de um sistema CAD. O quadrado cinza simboliza o bloco *off-line*, enquanto, o resto da figura pertence ao bloco *on-line*. [12]

Capítulo 4

Processamento de imagens

4.1. Processamento digital de imagens

O objetivo do processamento digital de imagens consiste em melhorar o aspeto visual de certas configurações estruturais, a fim de fornecer informações para que o analista humano possa realizar a interpretação da imagem de maneira confiável, além do facto de gerar imagens que possam ser posteriormente submetidas a processamentos realizados de forma automática, como a segmentação e a classificação.

Técnicas responsáveis pelo processamento de imagens têm tido um crescimento abundante nos últimos anos, porque hoje dispomos de um conjunto cada vez maior de imagens, provenientes de diversos tipos de fontes. A medicina é uma das áreas que mais se tem beneficiado da interação com a computação, através da análise de imagens principalmente para diagnóstico.

O interesse em métodos de processamento digital de imagens surgiu, principalmente, da necessidade de melhorar a qualidade da informação dos *pixéis* para interpretação humana. Como por exemplo a Equalização do histograma, Mapeamento das variações de cinza, Correção Gama, Aumento de Contraste, Técnicas de realce...

4.2. Técnicas para minimização de ruído speckle

Todas as técnicas de obtenção de imagens estão sujeitas a algum tipo de degradação, seja causada por defeitos do equipamento ou por distorções no processamento dos sinais.

Os ruídos introduzidos pelos artefactos da obtenção de sinais biomédicos podem causar degradação do sinal em imagem de diferentes modalidades. Entre os ruídos predominantes nas imagens formadas através de energia coerente, como os ultrassons, são destacados o ruído branco e o ruído do tipo speckle [13]. Este é um ruído multiplicativo proporcional à intensidade do sinal recebido, está sempre associado a sistemas de imagiologia coerente, tais como os obtidos por micro-ondas, laser, ou ultrassonografia.

O efeito visual deste ruído proporciona uma textura granulosa que pode dificultar a interpretação das imagens, reduzindo a separabilidade entre os objetos, impossibilitando muitas vezes a análise automática dessas imagens.

Para reduzir o ruído vários filtros têm sido desenvolvidos, com vistas a melhorar a separabilidade da imagem, com o mínimo de perda de dados. Os filtros devem preservar os contornos presentes

na imagem e as informações de textura. Como por exemplo o Filtro de Wiener, Filtro de Lee, Filtro de Frost e Wavelets.

Contudo essa perda de dados ainda é significativa, perdendo certas informações que podem ser dissipadas com algumas técnicas acima mencionadas. Isto é o speckle pode ser considerado uma característica [2], [3].

4.3. Técnicas de segmentação

A segmentação ou detecção de anormalidades é o passo inicial no diagnóstico assistido por CAD. Um algoritmo de segmentação, em contexto de imagens médicas, é um algoritmo usado para detectar algo, geralmente em toda a mama, num tipo específico de anormalidades.

Além de que existe uma grande variabilidade de lesões, o que se reflete na variação morfológica (forma e tamanho das lesões), e também no grande número de características que têm sido utilizados para as detectar e classificar [14].

4.3.1. Detecção de contornos

A separação dos pixels relativos a cada objeto, ou região, é uma etapa fundamental para o sucesso do processo de análise da imagem.

Entre as diversas formas e categorias que podem ser alvo da aplicação de um processo de segmentação, a detecção de contornos é uma técnica para determinar pontos de uma imagem digital em que a intensidade luminosa muda repentinamente. A detecção de contornos é um dos processos mais utilizados em processamento de imagem. A definição correta dos contornos das regiões e os seus limites são de extrema importância para a avaliação e classificação das mesmas.

A utilização reduz significativamente a quantidade de dados a serem processados, e descarta informação que é considerada menos relevante, ainda que preservando importantes propriedades estruturais de uma imagem.

O processo de dividir a imagem permite a simplificação na representação, facilitando assim a análise. Como resultado, temos um conjunto de contornos ou regiões extraídos da imagem. Cada um dos pixels de uma região é similar com referência a alguma característica ou propriedade computacional, cor, intensidade, textura ou continuidade. Regiões adjacentes devem possuir diferenças significativas com respeito à mesma característica. [15]

Identificar regiões com base na informação das arestas está longe de ser trivial, uma vez que os algoritmos para detecção de contornos geralmente não possuem a capacidade do sistema de visão humana para completar arestas interrompidas com experiência e informação de contexto. Portanto, por vezes, os contornos não são detetados pelo que a transição a partir de uma região a outra os contornos não são corretamente detetados e têm, frequentemente

lacunas. Assim, os contornos detetados podem não necessariamente formar um conjunto de curvas fechadas ligadas que rodeiam regiões distintas. [16]

Algoritmos típicos para encontrar arestas são baseados em filtrar a imagem, a fim de melhorar os contornos relevantes antes da fase de detecção. [17], [18].

Este processo consiste na detecção de variações de intensidade com recurso a operações de primeira e segunda derivadas. Estes operadores irão atuar na imagem, calculando estas variações que serão traduzidas no gradiente. Sabendo que um contorno é normalmente definido por uma passagem de uma região para outra, será de esperar que nesse local, existirá uma variação mais acentuada da intensidade do que no restante da imagem. Isto faz com que ao longo de um gradiente num dado local, o máximo estará colocado na posição do contorno [19].

O cálculo do gradiente ao longo da imagem será realizado em duas direções distintas. Por um lado seguindo a variação da intensidade na direção do eixo horizontal e por outro lado seguindo-a na direção do eixo vertical.

A magnitude do gradiente é a grandeza da variação da intensidade é dada pela seguinte equação:

$$G = \sqrt{I_x^2 + I_y^2} \tag{1}$$

A orientação do gradiente, que é sentido onde está a ocorrer a passagem de uma intensidade maior para uma menor, é dada pela seguinte expressão:

$$\theta = \tan^{-1} \frac{I_y}{I_x} \tag{2}$$

São utilizados operadores de gradiente para calcular uma aproximação do gradiente ao longo da imagem. Como por exemplo o operador Sobel que é um dos mais utilizados e conhecidos.

$$I_x \sim \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad I_y \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{3}$$

Os detetores de contornos normalmente obedecem a um conjunto de passos, até chegarem à detecção final. Devido aos operadores de gradiente serem muito sensíveis ao ruído, inicia-se por se aplicar uma filtragem à imagem original de modo a diminuir a razão entre o sinal e o ruído. Porém poderá levar a um enfraquecimento dos pontos dos contornos, que levará a uma fraca definição destes. De seguida será aplicado um operador de gradiente, de modo a fazer um realce nos locais com uma variação de intensidade mais significativa. No final deste ponto, para a detecção dos diversos pontos de contorno ser feita, é necessário fazer uma seleção e etiquetagem destes, onde o método de thresholding é uma escolha geralmente usada. [19]

4.3.1.1. Filtro Sobel

Em termos técnicos o filtro Sobel, consiste num operador que calcula diferenças finitas, dando uma aproximação do gradiente da intensidade dos pixels da imagem. Em cada ponto da imagem, o resultado da aplicação do filtro devolve o gradiente ou a norma deste vetor.

Este calcula o gradiente da intensidade da imagem em cada ponto, dando a direcção da maior variação de claro para escuro e a quantidade de variação nessa direcção. Assim, obtém-se uma noção de como varia a luminosidade em cada ponto, de forma mais suave ou abrupta.

Sendo possível estimar a presença de uma transição claro-escuro e qual a orientação desta. Como as variações claro-escuro intensas correspondem a fronteiras bem definidas entre objetos, será possível fazer a deteção de contornos.

Matematicamente este operador utiliza duas matrizes 3x3 que são convoluídas com a imagem original para calcular aproximações das derivadas - uma para as variações horizontais e uma para as verticais. Sendo A a imagem inicial então, G_x e G_y serão duas imagens que em cada ponto contem uma aproximação às derivadas horizontal e vertical de A.

$$\mathbf{G}_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * \mathbf{A} \quad \text{e} \quad \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A} \quad (4)$$

Portanto a magnitude, G, e a direcção, Θ , do gradiente são dados seguintes expressões:

$$G = \sqrt{G_x^2 + G_y^2} \quad \Theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (5)$$

4.3.2. Métodos de Clustering e de Thresholding

Métodos de agrupamento são uma das técnicas mais utilizadas na segmentação de imagens. Técnicas de agrupamento podem ser divididas em algoritmos hierárquicos e divisivos.

Os métodos de limiar (threshold) como um caso especial de métodos de agrupamento particionais, onde apenas dois *clusters* são considerados. Por exemplo utilizados diferentes valores de limiar de nível de cinzento, dependendo do tipo de tecido da mama com base na análise de histograma. [20]

Consiste na definição de um limiar para o qual a intensidade de cada pixel da imagem é comparada. Caso a intensidade seja inferior, o pixel será considerado como parte do fundo da imagem, caso seja maior, o pixel será considerado parte de alguma região ou objecto. Na prática o thresholding resulta numa imagem binária, onde os pixels brancos correspondem aos pixels da imagem original de intensidade superior ao limiar (os objectos), e os pixels pretos aos inferiores (o fundo). [21]

Muitas vezes as imagens de escala de cinzentos onde se aplicam estes limiares, são imagens bimodais, ou seja, no seu histograma notam-se duas distribuições. Quando existe a sobreposição destas distribuições a solução convencionalmente usada para contornar este facto, será a aplicação de um limiar algures no fundo do vale criado pelas duas distribuições

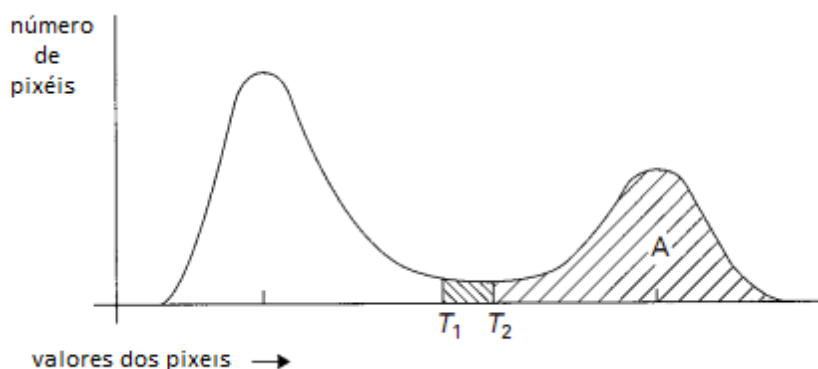


Figura 4 Limiar em histograma bimodal [19]

Várias situações poderão tornar a aplicação de um limiar global à imagem difícil. O fraco contraste pode levar a uma grande sobreposição de distribuições, tornando difícil a distinção. Uma grande variação na intensidade dos objetos alvos ou do fundo pode tornar impossível, a aplicação de um limiar global que não contenha erros no final. Levando a que para cada situação se tenham de tomar diferentes abordagens e até um pré-processamento da imagem, para melhor aplicação dos limiares. [19]

É também usada limiarização multinível para detetar contornos fechados. Nesta abordagem um grupo de contornos concêntricos representará a propagação de informações de densidade a partir da porção central de núcleo de um objeto ou região de tecido na imagem para os tecidos circundantes. Este algoritmo pode ser considerado como uma região de crescimento, onde, em cada iteração com vizinhos semelhantes de nível cinza de valores estão agrupados.

Em alguns casos, o limiar não será aplicado diretamente na imagem médica, mas numa versão melhorada da imagem original. Aplicando um filtro iris concebido para melhorar as opacidades e sendo insensível às estruturas finas anatómicas. Em vez de melhoria da imagem, uma abordagem diferente é primeiro extrair algumas características (textura) a partir da imagem e do limiar. [22]

4.4. Morfologia Matemática

A análise morfológica centraliza-se no estudo da estrutura geométrica das imagens. Morfologia é a forma e estrutura de um objeto ou os arranjos e inter-relacionamentos entre as partes de um objeto. Pode ser aplicada em enumeras áreas de processamento de imagens, como realce, filtragem, segmentação, esqueletização entre outras... Morfologia digital é um meio para descrever ou analisar a forma de um objeto digital.

A ideia de morfologia digital vem de uma imagem que consiste num conjunto de elementos, os *pixéis* que são reunidos em grupos tendo uma estrutura bidimensional (forma). Certas operações matemáticas em conjuntos de *pixéis* podem ser usadas para sobressair aspetos específicos das formas permitindo que sejam contabilizados ou reconhecidos.

O princípio da morfologia consiste em extrair duma imagem desconhecida a sua geometria através da utilização da transformação de uma outra imagem completamente definida, ou seja, consiste em extrair as informações reativas à geometria e à topologia de um conjunto desconhecido pela transformação através de outro conjunto bem-definido, chamado elemento estruturante. Sendo importante a utilização de teoria dos conjuntos, pois esta é a base utilizada na morfologia, assim é com esta teoria que será descrita e apresentada uma imagem. As operações básicas da morfologia digital são a erosão, em que os *pixéis* que não atendem a um dado padrão são apagados da imagem, e na dilatação, em que uma pequena área relacionada a um *pixel* é alterada para um dado padrão. Contudo, dependendo do tipo de imagem sendo processada (preto e branco, tons de cinza ou RGB) a definição destas operações é alterada, assim cada tipo deve ser considerado separadamente.[23]

4.4.1. Tipos de Operadores Morfológicos

Operações morfológicas estão divididas em operações morfológicas binárias e operações morfológicas sobre imagens coloridas ou de tons de cinza. Havendo diferenças em cada grupo.

Operações morfológicas binárias são definidas em imagens com *pixéis* somente pretos e brancos isto é *são utilizadas como operações de “hit or miss” numa imagem, ou seja, um elemento estruturante é aplicado bit a bit a uma imagem em que o bit resultante é '1' se o elemento for completamente encontrado na imagem ou '0' se não.* Um objeto é considerado um conjunto matemático de *pixéis* pretos, cada *pixel* é identificado pelos seus índices de linha e coluna sendo chamado de um ponto no espaço bidimensional. Nestas, um *pixel*, quando for afetado por uma operação, será substituído pelo seu valor oposto.

Nas operações morfológicas sobre imagens com tons de cinza ou cores, as operações de conjunto realizadas não implicam a subtração ou inclusão de um *pixel* (substituição do valor por preto ou branco), mas sim na modificação parcial do valor.[23]

4.4.2. Elemento Estruturante

O elemento estruturante é um conjunto definido e conhecido (forma e tamanho), que é comparado, a partir de uma transformação, ao conjunto desconhecido da imagem. Pode assumir várias formas e a sua origem pode ser definida em qualquer ponto. Não necessariamente o elemento estruturante é uma matriz retangular ou quadrada, podem ser por exemplo:

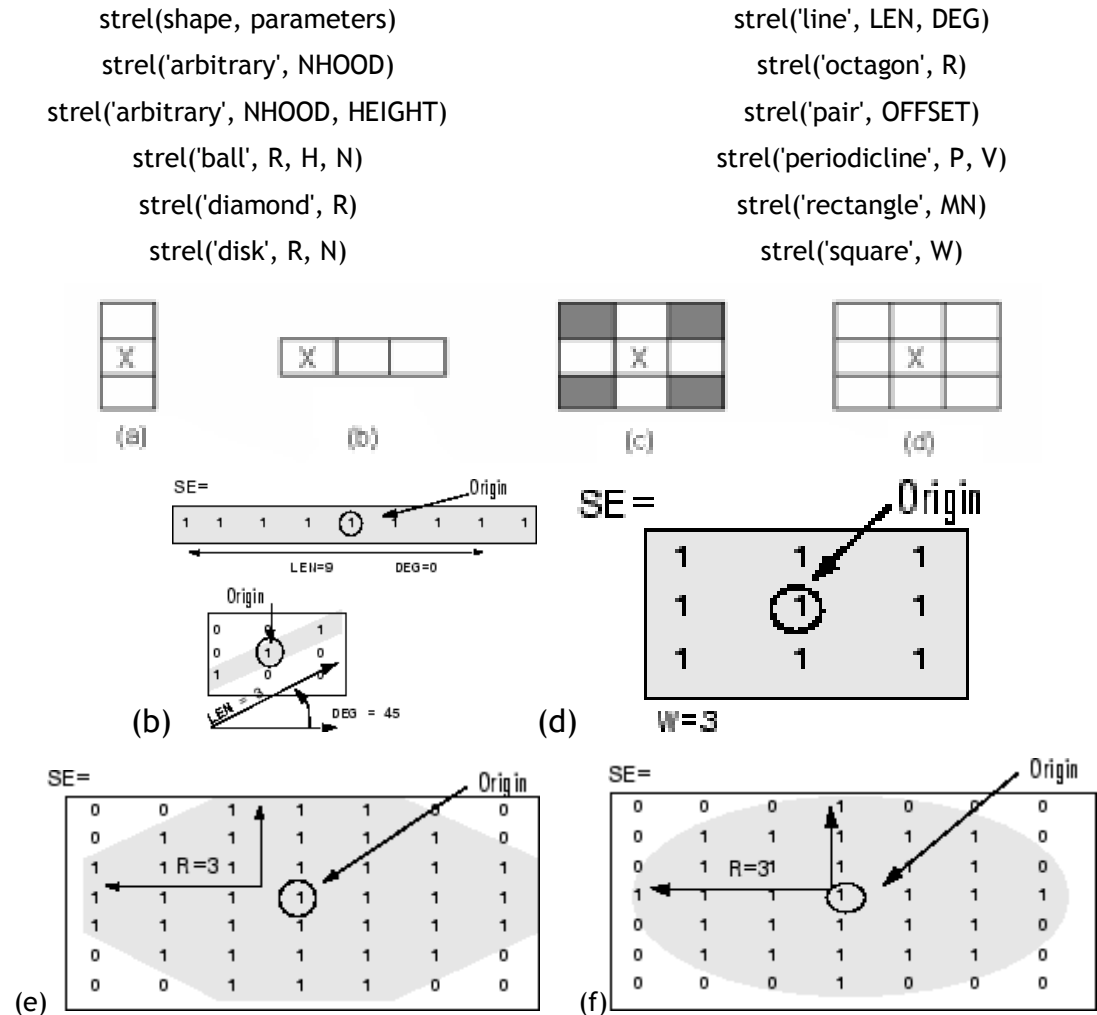


Figura 5 - Exemplo de elementos estruturantes: (a) em coluna, (b) em linha, (c) em cruz, (d) em quadrado, (e) em octógono, (f) em disco [24].

4.4.3. Definições Fundamentais

As operações fundamentais associadas com um objeto são o conjunto padrão de operações: união $\{\cup\}$, interseção $\{\cap\}$, e complemento $\{\complement\}$ com transação:

Transação: Dado um vetor x e um conjunto A , a transação, $A + x$, é definida como:

$$A + x = \{x + a / a \in A\} \tag{6}$$

Em que em imagens digitais compostas de pontos com coordenadas inteiras (Z^2), implicam restrições no vetor de transação x .

O conjunto básico de operações de Minkowski a adição e a subtração são definidos, baseado nas considerações anteriores. Dados dois conjuntos A e B :

Adição de Minkowski:
$$A \oplus B = \bigcup_{b \in B} (A + b) \tag{7}$$

Subtração de Minkowski:

$$A \ominus B = \bigcap_{B \in B} (A + \tilde{B}) \quad (8)$$

4.4.4. Morfologia binária

Em morfologia binária uma imagem é vista como um subconjunto de um espaço Euclidiano \mathbb{R}^d ou uma quadrícula inteira \mathbb{Z}^d , para alguma dimensão d. [23]

4.4.5. Operadores básicos

As operações básicas são operadores de permutação-invariante estreitamente relacionados com a soma de Minkowski. Seja E um espaço euclidiano ou uma quadrícula inteira e A uma imagem binária em E.[23]

4.4.5.1. Erosão

A erosão de uma imagem binária A pelo elemento estruturante B é definida por:

$$A \ominus B = \{z \in E | B_z \subseteq A\} \quad (9)$$

onde B_z é a translação de B pelo vetor z, isto é,

$$B_z = \{b + z | b \in B\}, \forall z \in E \quad (10)$$

ou então:

$$A \ominus B = \bigcap_{b \in B} A_{-b} \quad (11)$$

Quando o elemento estruturante B tem um centro (por exemplo, B é um disco ou um quadrado) e este centro se encontra na origem de E, então a erosão de A por B pode-se entender como o lugar geométrico dos pontos alcançados pelo centro de B quando B se move dentro de A. [23], [24].

4.4.5.2. Dilatação

A dilatação pode ser descrita como um crescimento de pixéis, isto é, marca-se com 1 a parte do fundo da imagem que contacte um pixel que faz parte da região. Permitindo que aumente um pixel ao redor da circunferência de cada região e assim poder incrementar dimensões, no qual ajuda a encher lacunas dentro da região.

A dilatação de A pelo elemento estruturante B define-se por:

$$A \oplus B = \bigcup_{b \in B} A_b \quad (12)$$

Esta é comutativa, sendo dada também por:

$$A \oplus B = B \oplus A = \bigcup_{a \in A} B_a \quad (13)$$

Se B tem um centro na origem, como antes, então a dilatação de A por B pode-se entender como o lugar geométrico dos pontos cobertos por B quando o centro de B se move dentro de A.

Ainda se pode obter por:

$$A \oplus B = \{z \in E | (B^s)_z \cap A \neq \emptyset\} \quad (14)$$

onde B^s denota a simetria de B , que é,

$$B^s = \{x \in E | -x \in B\} \quad (15)$$

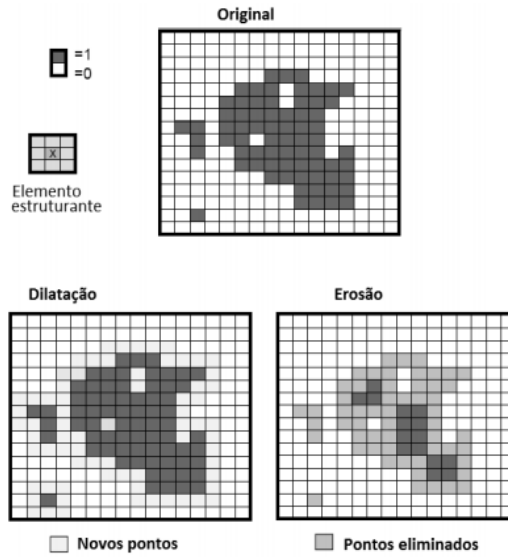


Figura 6 - Operações de Erosão e dilatação [23], [25].

A dilatação e a erosão possuem as seguintes propriedades:

Comutatividade:

$$D(A, B) = A \oplus B = B \oplus A = D(B, A) \quad (16)$$

Não-Comutatividade:

$$E(A, B) \neq E(B, A) \quad (17)$$

Associatividade:

$$A \oplus (B \oplus C) = (A \oplus B) \oplus C \quad (18)$$

Invariante de Transação:

$$A \oplus (B + x) = (A \oplus B) + x \quad (19)$$

Dualidade:

$$D^C(A, B) = E(A^C, -B) \quad (20)$$

$$E^C(A, B) = D(A^C, -B)$$

Combinando a dilatação e erosão, podem-se criar variados métodos de operação morfológica. Dois dos mais relevantes são a abertura e o fecho, muito utilizados para limpar artefactos contidos na imagem. [23], [24].

4.4.5.3. Abertura

A abertura de A por B obtém-se pela erosão de A por B, seguida por a dilatação da imagem resultante por B:

$$A \circ B = (A \ominus B) \oplus B \quad (21)$$

A abertura também é dada por:

$$A \circ B = \bigcup_{B_x \subseteq A} B_x \quad (22)$$

O que significa que é o lugar geométrico das translações do elemento estruturante B dentro da imagem A. [23], [24].

4.4.5.4. Fecho

O fecho de A por B obtém-se pela dilatação de A por B, seguida pela erosão da estrutura resultante por B:

$$A \bullet B = (A \oplus B) \ominus B \quad (23)$$

Também se pode obter por:

$$A \bullet B = (A^c \circ B^s)^c \quad (24)$$

Onde X^c denota o complemento de X referente a E,

$$X^c = \{x \in E | x \notin X\} \quad (25)$$

Significa que o fecho é complemento do lugar geométrico das translações da simetria do elemento estruturante fora da imagem A. [23], [24].

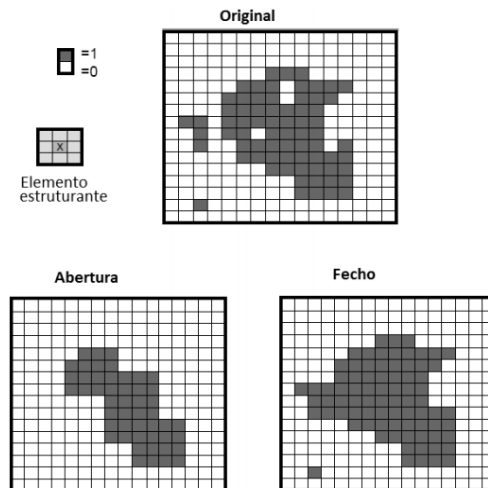


Figura 7 - Operações de abertura e fecho [23], [25].

Capítulo 5

Algoritmos Genéticos

5.1. Algoritmos Genéticos

Os AG são processos de procura baseados na mecânica de seleção natural e na genética, utilizados na ciência da computação para achar soluções aproximadas em problemas de otimização. Combinam a sobrevivência a do mais apto em estruturas de strings com uma troca estruturada, mas no entanto aleatória, de informação para criar um algoritmo de pesquisa [26].

Estes são implementados como uma simulação de computador em que uma população de representações abstratas de solução é selecionada em busca de soluções melhores. A evolução geralmente inicia-se a partir de um conjunto de soluções criadas aleatoriamente e é executada ao “longo de gerações”. A cada geração, a adaptação de cada solução na população é avaliada, alguns indivíduos são selecionados para a próxima geração, e recombinados ou mutados para formar uma nova. A nova população então é utilizada como entrada para a próxima iteração do algoritmo.

Foram criados por *John Holland* nos anos 60 e desenvolvidos por ele, pelos seus colegas e alunos nos anos 70 e, em contraste das estratégias evolucionistas e da programação evolucionaria, o objetivo de *Holland* não era desenhar algoritmos para resolver problemas específicos mas sim, abstrair e explicar o fenómeno da adaptação do modo que ela ocorre na natureza e desenvolver maneiras para que os mecanismos de adaptação pudessem ser trazidos para sistemas artificiais de computação [27].

A função-objetivo é o objeto da otimização. Pode ser um problema de otimização, um conjunto de teste para identificar os indivíduos mais aptos, onde nos retorna um valor que queremos otimizar. A grande vantagem dos AG está no facto de não precisarmos saber como funciona esta função objetivo, apenas tê-la disponível para ser aplicada aos indivíduos e comparar os resultados.

O indivíduo é meramente um portador do código genético. O código genético é uma representação do espaço de busca do problema a ser resolvido, em geral na forma de sequências de bits.

Os AG diferem dos métodos tradicionais de pesquisa e otimização em vários pontos-chave, o que os torna numa ferramenta superior, com uma robustez mais elevada. Possuem um

conjunto de propriedades que os tornam especialmente robustos para efetuar pesquisas em espaços de grande dimensão e complexidade. Algumas das principais características que contribuem para esta situação são:

- O facto de efetuarem o processamento simultâneo de uma população de pontos de potenciais soluções para o problema e não de um ponto apenas o que permite que diferentes áreas do espaço de procura sejam analisadas em paralelo o que vai levar a uma redução da probabilidade de convergência prematura para um local ótimo,
- Usam regras de transição probabilísticas e não determinísticas e não recorrem a informação específica sobre o problema que estão a resolver necessitando apenas do resultado da função de aptidão e não de derivadas ou de outros conhecimentos auxiliares
- Existe um equilíbrio entre exploração de novas regiões do espaço e conservação da informação já adquirida onde a seleção atua como principal força de conservação, permitindo manter e consolidar a informação descoberta anteriormente e,
- Através de transformações efetuadas nos elementos selecionados em conjunto com os mecanismos de busca paralela e adaptativa promovem a exploração de novas regiões do espaço o que vai fazer com que maus pressupostos não afetem negativamente o conjunto de soluções finais e que até mesmo soluções não ideais poderão conduzir à descoberta de novas regiões, onde se encontrem soluções de qualidade superior.

Mas o modo de funcionamento geral de um AG leva a que estes sejam algoritmos lentos que apesar de encontrarem sempre o conjunto das melhores soluções, não conseguem encontrar sempre a solução exata tendo assim um elevado risco de apresentarem uma solução não ótima.[28]

5.1.1. Biologia e Evolução em problemas computacionais

Inúmeros problemas computacionais requerem procurar soluções entre grandes quantidades de possibilidades como por exemplo um problema da engenharia de proteínas onde um algoritmo pesquisa entre vastas sequências de aminoácidos por uma proteína com propriedades específicas.

Portanto, em vez de se pesquisar uma sequência de aminoácidos individualmente, é muito mais rápido e vantajoso avaliar várias sequências de aminoácidos simultaneamente recorrendo ao uso de paralelismo no qual são testadas simultaneamente várias possibilidades de uma forma mais eficiente [29].

Para além do uso de paralelismo computacional que permite que sejam avaliadas várias sequências ao mesmo tempo, também é necessária uma estratégia inteligente que indique qual o próximo conjunto de sequências a ser analisado. Estas são as peças chave para a resolução deste tipo de problemas. A evolução biológica constitui uma fonte de inspiração para a resolução destes problemas. A evolução é um método de busca entre um número

enorme de possibilidades de soluções. Em biologia, o enorme conjunto de possibilidades é o conjunto de possíveis sequências genéticas, e as soluções desejadas são organismos bem aptos a sobreviver e se reproduzir nos seus ambientes.

A procura de soluções em face de mudanças nas condições é precisamente o que é necessário para programas de computador adaptativos. Além de que, a evolução é um método de pesquisa paralela em grande escala, que em vez de trabalhar numa espécie de cada vez, testa e muda milhões de espécies em paralelo. A evolução também pode ser vista como um método para a conceção de soluções inovadores para os problemas complexos. Os mecanismos de evolução podem inspirar deste modo métodos de pesquisa computacional. [28],[29].

De acordo com os pressupostos apresentados por Charles Darwin em 1859, o processo de evolução ocorre do seguinte modo:

- Os seres vivos da mesma espécie apresentam variações entre si.
- As populações têm tendência para crescer em progressão aritmética.
- O número de indivíduos de uma espécie geralmente não altera muito a cada geração.
- Em cada geração uma boa parte dos indivíduos é eliminada para que se estabeleça entre eles uma luta pela sobrevivência devido à competição por alimento, pelo refúgio ou espaço e a capacidade de fuga aos predadores.
- Sobrevivem os indivíduos que estiverem mais bem adaptados, isto é, os que possuírem as características que lhes conferem qualquer vantagem em relação aos restantes.
- Os indivíduos menos aptos ao longo do tempo são eliminados progressivamente.
- Os indivíduos mais bem adaptados vivem durante mais tempo e reproduzem-se mais transmitindo as suas características à sua descendência.
- A acumulação de pequenas variações determina, a longo prazo, a transformação e o aparecimento de novas espécies [30].

Podemos então observar que existem principalmente dois mecanismos associados a esta teoria que são:

- Seleção que garante que os indivíduos mais aptos têm maior probabilidade de sobreviver e procriar;

- Reprodução com variação que vai levar a que os descendentes não sejam cópias exatas dos progenitores.

A combinação destes dois mecanismos faz com que, com o tempo, as populações evoluam num sentido de maior adaptabilidade face ao meio onde se encontram.

Com o contributo de *Gregor Mendel* na formulação da teoria genética da hereditariedade em 1866 foi possível integrar o conceito de gene no processo evolutivo ajudando a compreender como são introduzidas alterações a uma população e, a variabilidade das espécies passou a ser explicada através da aplicação de operadores genéticos como o *crossingover* e a mutação.

5.1.2. Termos técnicos da Biologia aplicados nos Algoritmos Genéticos

Os AG recorrendo à inspiração na natureza processam conjuntos de elementos do espaço de procura que são potenciais soluções para o problema. Estes conjuntos, usualmente denominados populações, vão evoluindo ao longo de sucessivas iterações com o objetivo de encontrar uma solução de qualidade elevada para o problema que está a ser resolvido. Estes usam diversa terminologia que é aplicada em Biologia de modo análogo. Todos os organismos vivos são compostos por células e cada célula contém um ou mais cromossomas, que são cadeias de DNA que por sua vez são as instruções como o organismo e/ou um composto funciona.

Nos AG o termo cromossoma tipicamente refere-se a uma solução candidata para um problema e é escrita sobre a forma de uma *string* de *bits*. O total de todas as *strings* é chamado de estrutura. Estas são compostas por características que podem tomar valores como 0 ou 1. [28],[29]. A tabela seguinte contém uma versão condensada das características a cima descritas.

Tabela 1 - Termos técnicos da Biologia aplicados nos Algoritmos Genéticos

<u>Natureza</u>	<u>Algoritmos Genéticos</u>
<i>Cromossoma</i>	<i>String</i>
<i>Gene</i>	Elemento da <i>String</i>
<i>Alelos</i>	Valor do Elemento da <i>String</i>
<i>Genótipo</i>	Estrutura
<i>Fenótipo</i>	Estrutura descodificada

A par dos termos em comum também existe uma evidente semelhança no funcionamento básico do ciclo evolutivo e no ciclo de funcionamento de um AG.

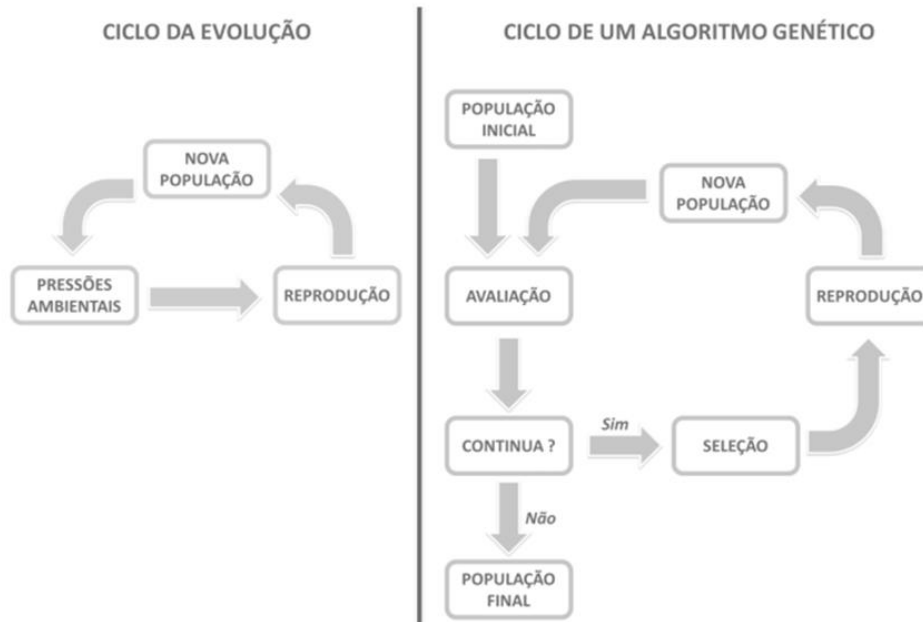


Figura 8 - Semelhanças entre o ciclo da evolução e o ciclo de um Algoritmo Genético

Do lado esquerdo, temos o ciclo da evolução “natural” onde os elementos da nova população durante o seu tempo de vida estão sujeitos a pressões ambientais servindo como mecanismo de seleção em que os indivíduos menos adaptados de uma população acabem por morrer. Isto vai levar a que só os elementos mais adaptados tenham possibilidade de se reproduzirem dando origem a uma nova população com as características dos seus progenitores ou seja, mais adaptada ao ambiente.

Num AG, representado do lado direito, é criada uma população inicial que através do uso de regras e equações matemáticas vai ser avaliada na qualidade individual de cada elemento de uma população. Se esta população não necessitar de ser melhorada então o algoritmo continua, se necessitar nesse caso o elemento Seleção vai mimetizar o comportamento das pressões ambientais no ciclo da evolução, dando aos elementos mais adaptados a probabilidade de se reproduzir e vai levar ao desaparecimento dos elementos menos adaptados. Os melhores elementos reproduzem-se, seguindo os métodos naturais. Do mesmo modo, também estes novos elementos herdam as características dos progenitores criando assim uma nova população mais bem adaptada que a anterior. Esta nova população é agora avaliada e o ciclo recomeça. [28], [29].

5.1.3. Componentes de um Algoritmo Genético

Um AG regue-se pela seguinte ordem de funcionamento. No início são escolhidas aleatoriamente várias soluções. A partir deste conjunto inicial, e de acordo com os dois mecanismos básicos utilizados pela evolução natural reprodução, com recombinação e mutação, as transformações na população de soluções são efetuadas do seguinte modo:

- Os elementos mais aptos de uma determinada geração são selecionados de modo probabilístico para servirem de progenitores das soluções da geração seguinte.
- Operadores de transformação, designados operadores genéticos, atuam sobre os elementos selecionadas originando as novas soluções.
- São aplicados sobre os elementos selecionados operadores genéticos originando novas soluções.

Existem duas decisões cruciais que é necessário tomar quando se pretende aplicar um AG na resolução de um problema. A primeira decisão é a escolha da representação para as soluções que fazem parte do espaço de procura. *John Holland* propôs a utilização de código binário para efetuar a representação das possíveis soluções de um problema pois nesta forma, os indivíduos processados por um AG clássico são simplesmente sequências de 0s e 1s codificando a informação necessária para representar um ponto do espaço de procura. A segunda decisão é a definição da função de aptidão que associe a cada solução uma medida de qualidade capaz de representar a sua capacidade para resolver o problema em causa. [28],[31].

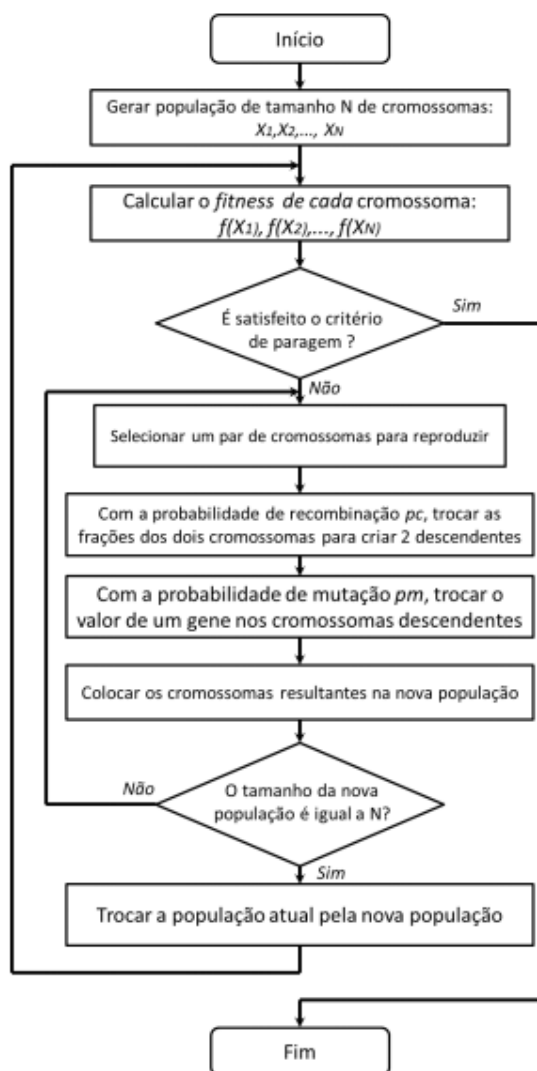


Figura 9 - Representação de um Algoritmo Genético Simples

5.1.4. Estrutura de um Algoritmo Genético Simples

Devido ao mecanismo de seleção, a qualidade média dos elementos que constituem a população tem tendência para aumentar ao longo do tempo. Os operadores genéticos são os responsáveis pela obtenção de novas soluções, ao mesmo tempo que tentam garantir que o processo mantenha um nível adequado de diversidade. Um AG possui três tipos de operadores genéticos fundamentais: seleção, recombinação e mutação.

▫ **Seleção:** Este operador seleciona cromossomas da população para reprodução dando aos cromossomas mais adaptados, uma maior probabilidade de serem selecionados para se reproduzir do que aos menos aptos [28].

▫ **Recombinação:** Operador estocástico que efetua a troca de material genético entre dois progenitores, conduzindo à criação de dois novos indivíduos. Estes serão formados por sequências genéticas parciais de cada um dos elementos originais. O operador de recombinação é considerado o principal responsável pela pesquisa efetuada por um AG. A sua função é atuar sobre as soluções com algum potencial pois a combinação de características de indivíduos promissores poderá originar novas soluções que integrem as vantagens de ambos e que possuam uma aptidão acrescida para o problema que está a ser resolvido [27].

▫ **Mutação:** Operador que atua sobre as soluções resultantes da recombinação e que altera ligeiramente algumas das suas características. Num AG, a mutação é um processo completamente aleatório e tem como objetivo manter um nível de diversidade adequado na população para que esta não caia num ótimo local o que a impediria de avançar. À medida que o número de gerações aumenta, um AG converge gradualmente para regiões do espaço de procura onde se encontram soluções promissoras. A otimização termina quando um determinado critério de terminação é atingido sendo os mais comuns:

- Número limite de gerações atingido;
- Descoberta de uma solução com qualidade pretendida;
- Inexistência de melhoria durante um determinado período de tempo.

Ao atingir um critério de paragem, o AG devolve o resultado final da otimização. Existem algoritmos que devolvem a melhor solução encontrada ao da longo otimização e há aqueles que devolvem um conjunto de indivíduos de qualidade elevada como todos os indivíduos que fazem parte da última geração.

5.2. Optimization Toolbox™ - Toolbox Algoritmo Genético [31]

Para uma mais fácil análise neste trabalho foi usada a *Toolbox* de AG.

A *Toolbox* de AG agora faz parte da *Optimization Toolbox™* é uma coleção de funções que ampliam os recursos da *Toolbox* de Otimização e o ambiente de computação do *MATLAB®*. Esta *Toolbox* inclui rotinas para resolver problemas de otimização utilizando AG.

Esses algoritmos permitem resolver uma variedade de problemas de otimização que se encontram fora do âmbito da caixa de ferramentas de otimização padrão.

É possível estender as capacidades da *Toolbox* escrevendo os devidos *M-files*, ou usando a caixa em combinação com outras, ou com *Simulink®*.

5.2.1. Escrever uma *M-File* para a função que se pretende otimizar

Para usar a *Toolbox*, é necessário escrever um *M-file* que descreve a função que se deseja otimizar. O *M-file* deve aceitar um vetor linha, cujo comprimento é o número de variáveis independentes para a função objetivo, e devolver um escalar.

$$f(x_1, x_2) = x_1^2 - 2x_1x_2 + 6x_1 + x_2^2 - 6x_2 \quad (26)$$

Maximizando Versus Minimizando

As funções de otimização da *Toolbox* minimizam a função-objetivo ou “*fitness function*”. Ou seja, são resolvidos da seguinte forma:

$$\underset{x}{\text{minimize}} \quad f(x) \quad (27)$$

Para maximizar $f(x)$, pode-se fazê-lo através da minimização $-f(x)$, porque o ponto em que o mínimo de $-f(x)$ ocorre é o mesmo que o ponto em que o máximo de $f(x)$ ocorre. [26]

5.2.2. Usar a *Toolbox*

Há duas maneiras de usar a *Toolbox* com o AG:

- Invocar a função “*ga*” do AG na linha de comandos,
- ou
- Usando a *Toolbox*, como interface gráfica.

5.2.3. Invocar a função “*ga*” na linha de comandos

Para usar a *Toolbox* na linha de comandos, é necessário invocar a função *ga* com a sintaxe:

$$[x \text{ fval}] = \text{ga}(@\text{fitnessfun}, \text{nvars}, \text{options})$$

Onde:

- *@fitnessfun* - é a fitness function.
- *nvars* - é o número de variáveis independentes da *fitness function*.
- *options* - é a estrutura que contem as opções do Algoritmo. (passando este passo as definições permanecem por defeito)

Os resultados são apresentados por:

- *fval* – Valor final da *fitness function*
- *x* – ponto em que o valor final é atingido onde *fitness function* foi minimizada.

5.2.4. Uso da Toolbox, com a interface gráfica

Usando a Toolbox, como interface gráfica permite usar a ferramenta sem trabalhar na linha de comandos. Bastando escrever “*gatool*” que abre a ferramenta, como mostrado na figura.

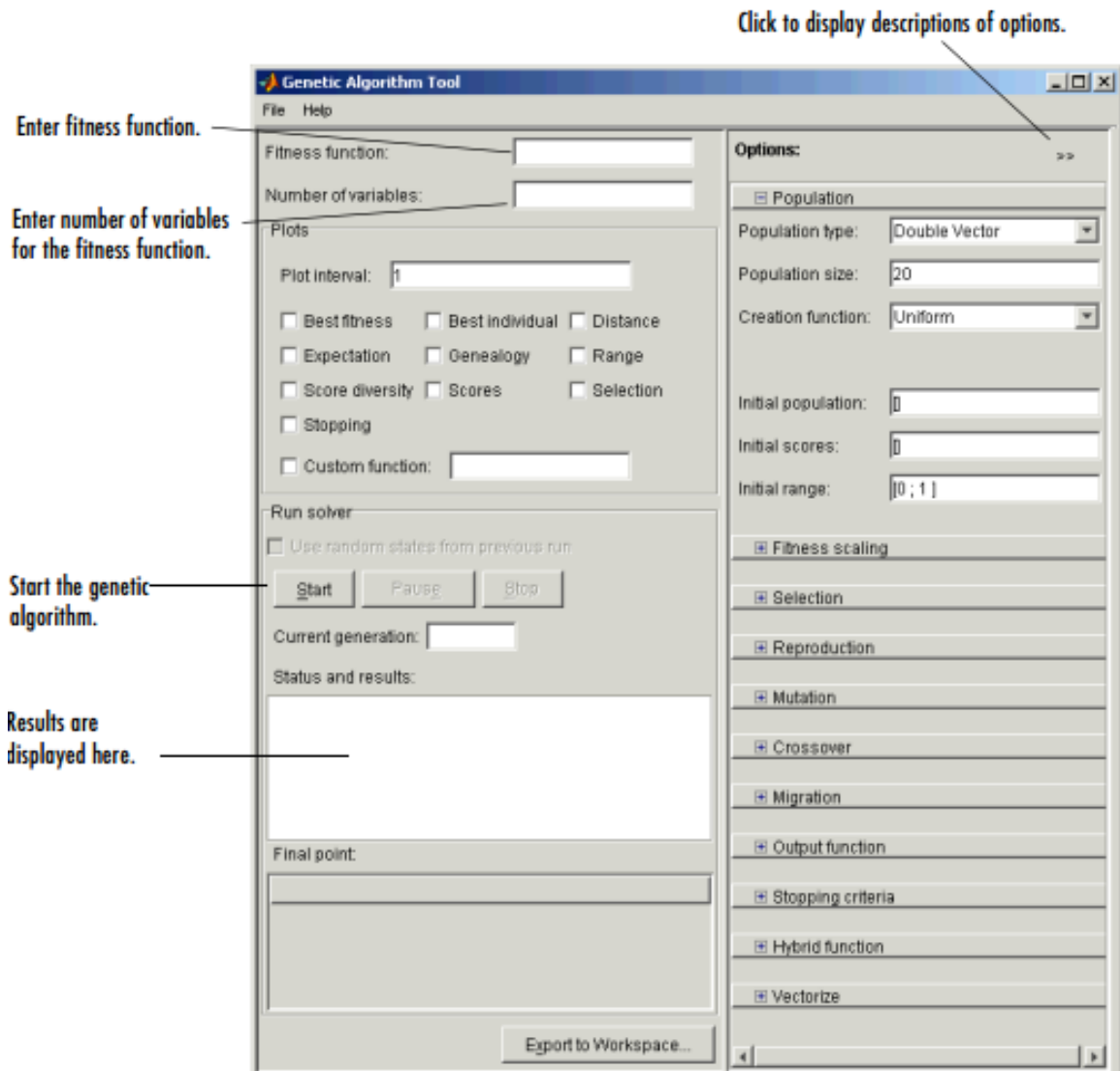


Figura 10 - Interface gráfica da Toolbox Algoritmo Genético

A função de *fitness* e o campo do número de variáveis deve aparecer como mostrado na figura 10.

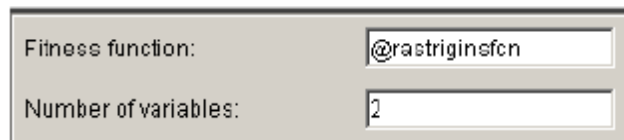


Figura 11 - Campo do número de variáveis da Toolbox Algoritmo Genético

Clicando no botão **Start (Iniciar)** no painel **Run solver**, como mostrado na figura 12, todas as definições permanecem por defeito.

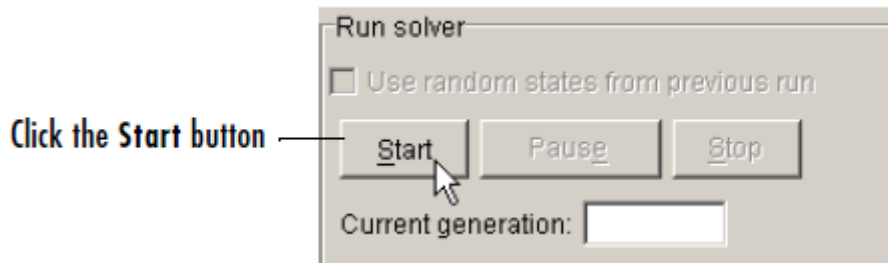


Figura 12 - Campo Run Solver da Toolbox Algoritmo Genético

Enquanto o algoritmo é executado, o campo “**Current generation**” exibe o número da geração atual. É possível pausar temporariamente o algoritmo, clicando no botão de **pausa**. Clicando em **Resume** o algoritmo continua a partir do ponto em que se parou.

Quando o algoritmo é concluído, o painel de “**Status and results**” é apresentado da seguinte forma:

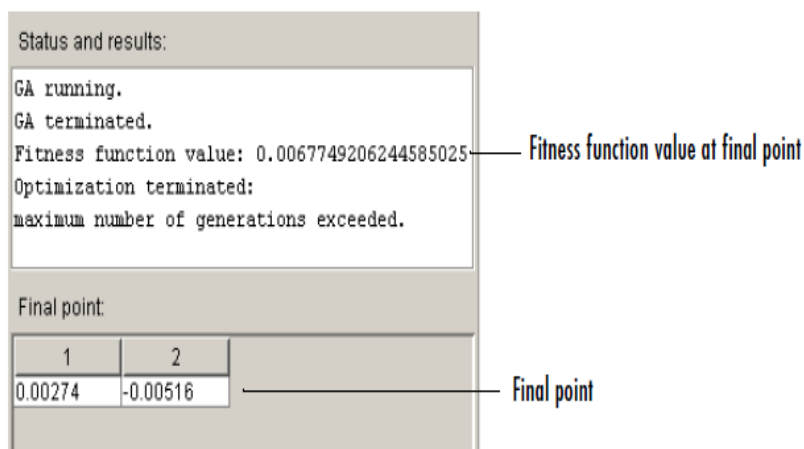


Figura 13 - Campo Status and Results da Toolbox Algoritmo Genético

Este painel apresenta as seguintes informações:

- O valor final da *fitness* function quando o algoritmo termina.
- Razão porque o algoritmo terminou.
- Ponto final

5.2.5. Exibir Gráficos

O painel “*Plots*” permite exibir vários gráficos que fornecem informações sobre o AG, enquanto está a ser executado. Estas informações podem ajudar a alterar as opções para melhorar o desempenho do algoritmo. Por exemplo, para traçar as melhores médias da função *fitness* a cada geração, basta marcar a opção “*Best fitness*”:

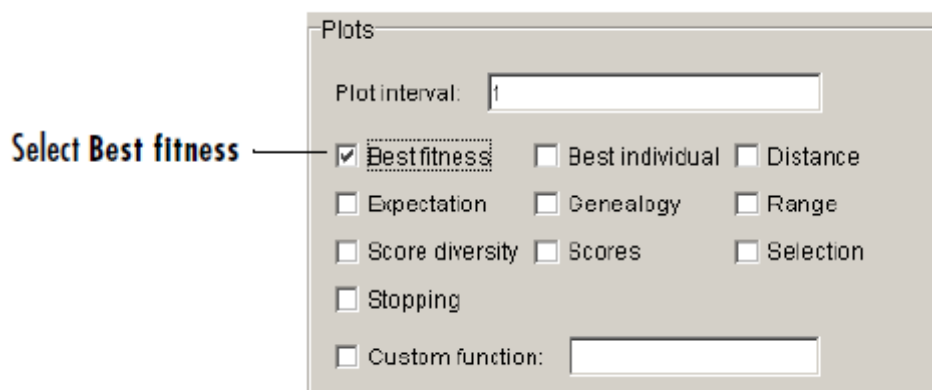


Figura 14 - Campo Plots da Toolbox Algoritmo Genético

Ao clicar em Iniciar, a ferramenta exibe uma série dos melhores valores médios da função de aptidão em cada geração.

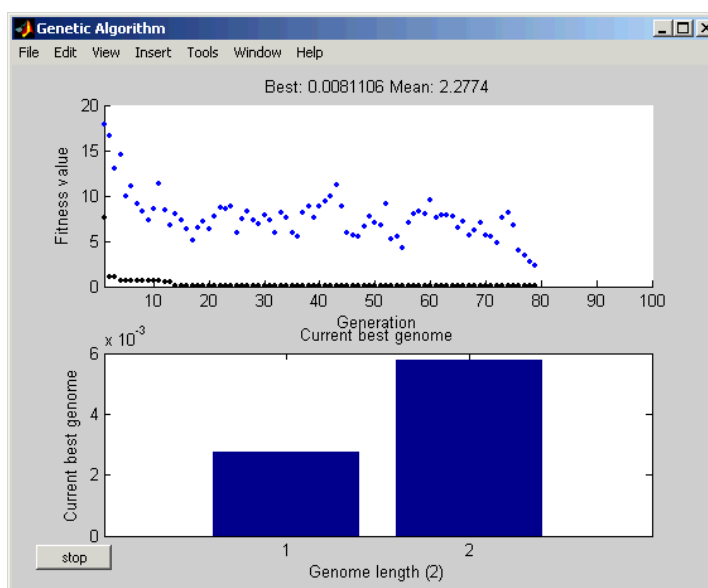


Figura 15 - Gráficos do melhores valores médios da função de aptidão em cada geração da Toolbox Algoritmo Genético

5.2.6. Condições de paragem do Algoritmo

O AG utiliza cinco condições a seguir para determinar quando deve parar:

- **Generations** - O algoritmo termina quando o número de gerações atinge o valor definido.
- **Time limit** - O algoritmo termina depois de correr uma porção de tempo em segundos, igual ao limite de tempo.
- **Fitness limit** - O algoritmo termina quando o valor da função de aptidão para o melhor ponto da população atual é inferior ou igual a limite de aptidão.
- **Stall generations** - O algoritmo termina se não houver melhoria na função objetivo para uma sequência de gerações consecutivas.
- **Stall time limit** - O algoritmo termina se não houver melhoria na função objetivo durante um intervalo de tempo em segundos.

O algoritmo termina assim que qualquer uma destas cinco condições seja verificada. Pode-se especificar os valores destes critérios. Os valores padrão são mostrados na figura abaixo.

Stopping criteria	
Generations:	100
Time limit:	Inf
Fitness limit:	-Inf
Stall generations:	50
Stall time limit:	20

Figura 16 - Campo Stopping Criteria da Toolbox Algoritmo Genético

O painel de *status* exibe o critério que causou a paragem do algoritmo.

Se o algoritmo parar devido a uma dessas condições, é possível melhorar os resultados, aumentando os critérios de paragem.

5.2.7. Definir Opções para o AG

Para obter os melhores resultados do AG, pode ser necessário experimentar diferentes opções. Selecionando as melhores opções para um problema envolve tentativa e erro.

Opções:

- “Diversity”
- “Population Options”
- “Fitness Scaling Options”

- “*Selection Options*”
- “*Reproduction Options*”
- “*Mutation and Crossover*”
- “*Mutation Options*”
- “*The Crossover Fraction*” ...

5.2.8. Diversidade

Um dos fatores mais importantes para determinar o quão bom é o AG, é a diversidade da população. Se a distância média entre os indivíduos é grande, a diversidade é alta; se a distância média é pequena, a diversidade é baixa. A diversidade afeta o desempenho do AG. Se a diversidade for muito grande ou muito pequena, o algoritmo pode não ter um bom desempenho.

É possível controlar a diversidade na população por várias configurações de opções, incluindo a “*Initial range*” e a porção de mutação em “*Mutation*”.

5.2.9. Opções da População

- **Opções de controlo populacional** das características dos indivíduos da população. A especificação da gama inicial de modo que contenha uma estimativa para a solução, pode aumentar a velocidade do algoritmo.

No entanto, este pode encontrar soluções, mesmo que não se encontrem no intervalo inicial, desde que os grupos possuem diversidade suficiente.

- **Definir o tamanho da população**

O campo “*Size*” em “*Population*” determina o tamanho da população em cada geração. Aumentar o tamanho da população permite que o AG procure mais pontos e assim obter um melhor resultado. No entanto, quanto maior for o tamanho da população, mais tempo o AG irá demorar para calcular cada geração.

5.2.10. Opções de dimensionamento da aptidão

O dimensionamento converte as pontuações de aptidão que são devolvidos pela função de aptidão para valores num intervalo que é adequado para a função de seleção. A função de seleção usa os valores de aptidão calibrados para selecionar os pais da próxima geração. A função de seleção atribui uma maior probabilidade de seleção de indivíduos com valores em escala alta.

A gama de valores escalados afeta o desempenho do AG.

Se os valores em escala variarem muito amplamente, os indivíduos com maiores valores escalados reproduzem-se muito rapidamente, tomando conta da pool genética da população muito rapidamente, impedindo o algoritmo de procurar outras áreas do espaço de solução. Por outro lado, se os valores em escala variarem pouco, todos os indivíduos têm, aproximadamente, a mesma possibilidade de reprodução e a pesquisa vai progredir muito lentamente.

A função aptidão padrão, “*Rank*”, dimensiona os resultados brutos com base na classificação de cada indivíduo. O ranking de um indivíduo é a posição na pontuação ordenadamente: a classificação do indivíduo mais apto é 1, o próximo mais apto é 2, e assim por diante. A função de classificação de escala atribui valores escalonados de modo que

- O valor de escala de um indivíduo com grau n é proporcional a $1/(f^n)$.
- A soma dos valores escalonados ao longo de toda a população é igual ao número de pais necessárias para criar a próxima geração.

5.2.11. Opções de Seleção

Esta função escolhe os pais para a próxima geração com base nos valores em função da aptidão. Um indivíduo pode ser selecionado mais de uma vez como um pai, caso contribua com os seus genes para mais de um filho. A função de seleção padrão, “*Stochastic uniform*”, estabelece uma linha em que cada um dos pais corresponde a uma secção da linha de comprimento proporcional ao seu valor escalado. O algoritmo move-se ao longo da linha em passos de igual tamanho. A cada passo, o algoritmo aloca um pai na seleção.

A função de seleção mais determinista é “*Reminder*”, que realiza dois passos:

- No primeiro passo, a função seleciona os pais deterministicamente de acordo com a parte inteira do valor escalonado para cada indivíduo. Por exemplo, se o valor em escala de um indivíduo é de 2.3, a função seleciona o indivíduo duas vezes como um pai.
- No segundo, a função de seleção seleciona pais adicionais usando as partes fracionárias dos valores em escala, como na seleção uniforme estocástica. A função define uma linha em secções, cujos comprimentos são proporcionais à parte fracionária do valor escalonado dos indivíduos, e move-se ao longo da linha em etapas iguais, para selecionar os pais. [26], [28]

5.2.11.1. Seleção por torneio

Na Seleção por torneio a seleção do cromossoma a reproduzir consiste em escolher k elementos de forma aleatória da população e ver qual deles é a melhor. Este método é um dos métodos mais usados na atualidade, sendo mais comum o torneio binário, ou seja com $k = 2$.

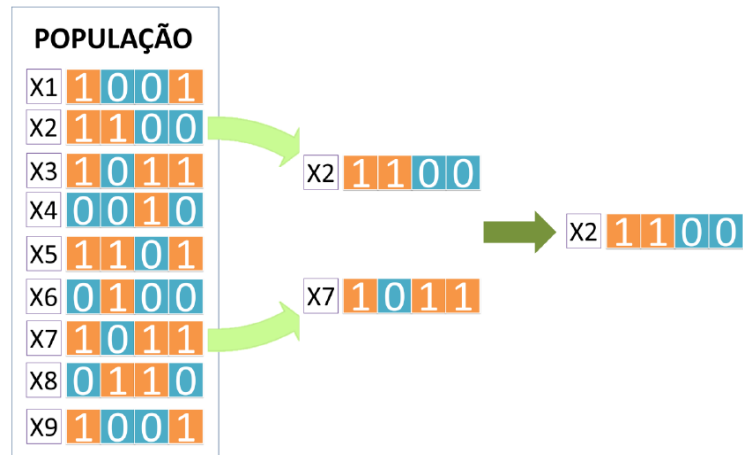


Figura 17 - Representação de Seleção por Torneio binário

5.2.11.2. Seleção por truncatura

Neste modo de seleção, é definido o valor τ que representa o limite dos indivíduos que vão ser selecionados. São então selecionados os $\tau \cdot N$ melhores elementos, em função do seu valor de aptidão. A vantagem deste método é que escolhe sempre os melhores elementos da população. Em norma o valor de τ varia entre 0,3 e 0,5, ou seja, os 30% melhores elementos da população ou 50% respetivamente.

5.2.11.3. Seleção por rácio da aptidão

Esta técnica depende do rácio da aptidão para determinar a probabilidade de um cromossoma ser selecionado para se reproduzir. Nesta roleta cada cromossoma tem uma fatia cuja área é igual ao rácio da aptidão do mesmo. Como tal os cromossomas mais adaptados tem uma área muito maior que os menos adaptados. Para selecionar um cromossoma para cruzamento é gerado um número aleatório entre 0 e 100 e o cromossoma que contiver esse número na sua fatia é selecionado.

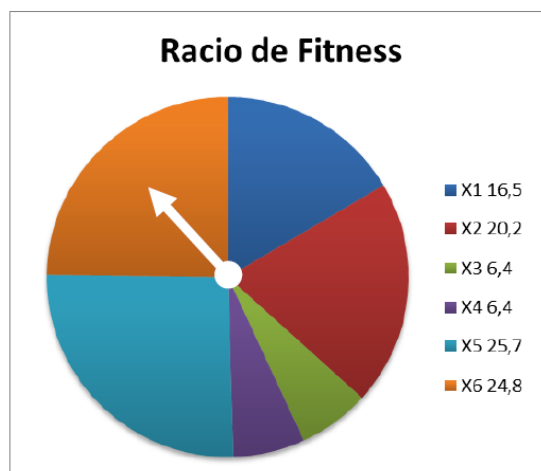


Figura 18 - Representação da Seleção por roleta do Rácio de Fitness

5.2.12. Opções de reprodução

Opções de reprodução controlam a forma como o AG cria a próxima geração. As opções são:

- “*Elite count*” - O número de indivíduos com os melhores valores de aptidão da geração atual que são os mais aptos a sobreviver à próxima geração.

Estes indivíduos são chamados “*elite children*”. O valor padrão é 2.

Quando a “*Elite count*” é pelo menos 1, o melhor valor de *fitness* só pode diminuir de uma geração para a seguinte. Isto é o que se quer que aconteça, uma vez que o AG minimiza a função de aptidão. Definir esta para um valor alto faz com que os indivíduos mais aptos dominem a população, tornando a busca menos eficaz.

- “*Crossover fraction*” - A fração de indivíduos na próxima geração serão criados por crossover.

5.2.13. Mutação e Crossover

O AG utiliza os indivíduos da geração atual para criar os filhos que compõem a próxima geração. Além das “*elite children*”, que correspondem aos indivíduos na atual geração com os melhores valores de aptidão, o algoritmo cria

- “*Crossover children*”, selecionando as entradas do vetor, ou genes, a partir de um par de indivíduos da geração atual e combina-os para formar uma “*child*”.
- “*Mutation children*”, aplicando mudanças aleatórias num único indivíduo na geração atual para criar uma “*child*”.

Ambos os processos são essenciais para o algoritmo. O Crossover permite que o algoritmo extraia os melhores genes de diferentes indivíduos e os recombine em “*child*” potencialmente superiores. A mutação aumenta a diversidade de uma população e, assim, aumenta a probabilidade de que o algoritmo irá gerar indivíduos com melhores valores de aptidão. Sem mutação, o algoritmo apenas pode produzir indivíduos cujos genes eram um subconjunto dos genes combinados na população inicial.

É possível especificar o número de cada tipo de “*children*”, em:

- “*Elite count*”, em “*Reproduction*”, especifica o número de “*elite children*”.
- “*Crossover fraction*”, em “*Reproduction*”, especifica a fração da população, que não sejam filhos da elite, e que são filhos de crossover. [32].

5.2.14. Opções de mutação

O AG aplica mutações utilizando a função especificada no campo “*Mutation function*”. A função de mutação padrão é “*Gaussian*”, que adiciona um número aleatório, ou uma mutação, escolhido a partir de uma distribuição de Gaussiana, a cada entrada do vetor pai. Tipicamente, a “quantidade” de mutação, que é proporcional ao desvio padrão da distribuição, diminui em cada nova geração. Pode-se controlar a quantidade média de mutação que o algoritmo aplica a um progenitor em cada geração por meio das opções “*Scale and Shrink*”:

- “*Scale*” controla o desvio padrão da mutação na primeira geração, que é multiplicado pela escala da gama da população inicial, que é especificado pela opção “*Initial range*”.
- “*Shrink*” controla a taxa em que a quantidade média de mutação diminui. O desvio padrão diminui linearmente de modo a que o seu valor final seja igual a $1 - Shrink$.

5.2.14. A fração Crossover / Recombinação

O campo de “*Crossover fraction*”, nas opções de reprodução, especifica a fração de cada população, exceto “*elite children*”, que são feitas por “*elite children*”. Uma fração de crossover de 1 significa que para além de todos os outros “*children*”, são “*crossover children*”, enquanto uma fração de crossover de 0 significa que todas as “*children*” são “*mutation children*”.

O operador genético crossover começa por efetuar um teste de probabilidade para ver se o crossover irá ocorrer. Se o crossover não ocorrer então os descendentes vão ser cópias exatas nos cromossomas progenitores. Regra geral o valor de 0,7 para a probabilidade de ocorrer crossover costuma produzir bons resultados, mas esta probabilidade depende do problema em questão. Os tipos de recombinação são o crossover de um ponto, de dois pontos e o uniforme, existindo no entanto mais tipos.

5.2.14.1. Crossover de um ponto

Escolhe-se aleatoriamente um ponto para em seguida efetuar a separação e troca de informação dos dois cromossomas criando assim os cromossomas descendentes.

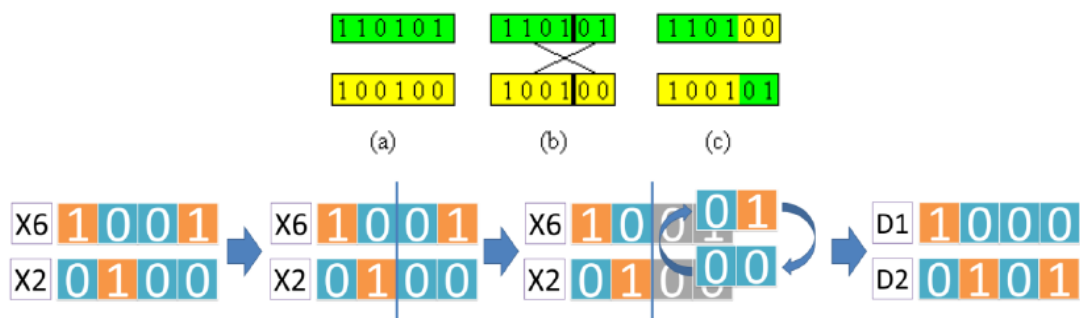


Figura 19 - Um exemplo de crossover de um ponto: (a) dois indivíduos são escolhidos. (b) um ponto (4) de crossover é escolhido. (c) são recombinadas as características, gerando dois novos indivíduos. (d) Recombinação por crossover de um ponto com ponto de corte no gene 2

5.2.14.2. Crossover de dois pontos

A diferença desta técnica para a anterior é que nesta são escolhidos aleatoriamente dois pontos e troca de informação dos dois cromossomas é efetuada no intervalo entre os dois pontos apenas.

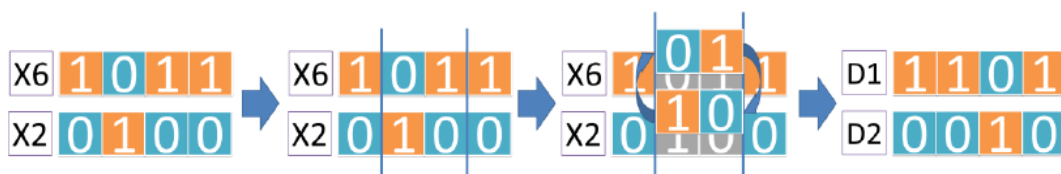


Figura 20 - Recombinação por crossover de dois pontos

5.2.14.3. Crossover uniforme

Nesta técnica, cada bit da sequência tem 50% de hipótese de ser trocado com o outro cromossoma e cada bit é testado individualmente para a sua eventual troca.



Figura 21 - Recombinação por crossover uniforme

5.2.14.4. Mutação

O operador genético mutação altera o valor num gene escolhido aleatoriamente dentro de um cromossoma. A mutação ocorre nos genes com alguma probabilidade que costuma ter valores bastante pequenos entre 0.001 e 0.01 pois se a probabilidade de mutação ocorrer for muito elevada isto vai criar uma instabilidade na população. Alguns livros indicam que a probabilidade de a mutação ocorrer num dado bit é de onde representa o tamanho do cromossoma. [26], [28], [32].



Figura 22 - Mutação no gene 3 do cromossoma

Capítulo 6

Determinação de parâmetros dos operadores no algoritmo desenvolvido com auxílio dos Algoritmos Genéticos

6.1. Determinação de parâmetros dos operadores com auxílio dos Algoritmos Genéticos

Na prática um técnico ou médico que analisa imagens de ecografias mamárias verifica a existência de nódulos manualmente. Ou seja, ele faz uma segmentação manual da imagem. Esta forma humana de fazer a detecção do nódulo pode ser feita com o auxílio de técnicas computacionais.

Métodos de threshold, de detecção de contornos e operadores morfológicos, ambos já mencionados no capítulo 4, são técnicas de processamento de imagens que podem auxiliar nesta detecção. Os elementos estruturantes dos operadores morfológicos possuem diversas formas e tamanhos tendo sido ilustrado no capítulo 4.

Entretanto diferentes formas e tamanhos dos elementos estruturantes dos operadores morfológicos aplicados à imagem podem-nos fornecer resultados diferentes para detecção na imagem. Uns melhores, outros piores. A variedade de formas e tamanhos possíveis é enorme (∞) e a seleção dos mesmos aplicados para melhor detetar um nódulo torna-se uma tarefa árdua de ser realizada manualmente. Além de que estas formas e tamanhos irão variar de imagem para imagem ou seja de uma ecografia para outra. Em outras palavras, a melhor seleção para uma imagem pode não ser tão favorável para outra

Neste presente trabalho procuramos determinar os melhores parâmetros dos métodos de threshold e dos elementos estruturantes dos operadores morfológicos (formas e tamanhos) para a segmentação de nódulos em imagens dentre os inúmeros possíveis, fazendo uso de AG. Esta ideia é certamente diferente de outras formas de detetar nódulos mamários com técnicas computacionais... Foi dado um novo enfoque para lidar com este problema da detecção dos nódulos mamários! Talvez o trabalho que mais se aproxime desta forma de abordar o problema seja a investigação de Alvarenga e seu grupo [33], [34]. Mesmo este, pouco se assemelha ao que foi desenvolvido no presente trabalho.

Note que os AG não são aqui usados para tratar a imagem, como normalmente são usados outros métodos. Os AG neste trabalho apenas são o meio para determinar os valores necessários à forma ótima para detetar possíveis nódulos de uma ecografia.

6.2. Algoritmo

O algoritmo para a segmentação de nódulos utiliza métodos de threshold, de deteção de contornos e operadores morfológicos, supra citados no capítulo 4.

Aqui os AG determinam os melhores parâmetros dos métodos de threshold e dos elementos estruturantes dos operadores morfológicos. Neste caso este algoritmo torna-se a função fitness a ser utilizada pelo AG.

Uma dissemelhança desta abordagem em relação a outras é a segmentação da imagem sem ser pré-processada, como mencionado nos anteriores capítulos, retirar ruído por exemplo, pode subtrair alguma informação que possa ser fundamental para a classificação [2], [3]. O algoritmo promove a não eliminação de qualquer informação, portanto os critérios de diagnóstico são preservados.

Alguns passos do algoritmo seguem o esquema apresentado em baixo: A comparação dos resultados deste é feita a partir de nódulos manualmente segmentadas.



Figura 23 - Nódulos manualmente detetadas.

Passo 1: Leitura da imagem

Leitura da ecografia mamária.

imagem original

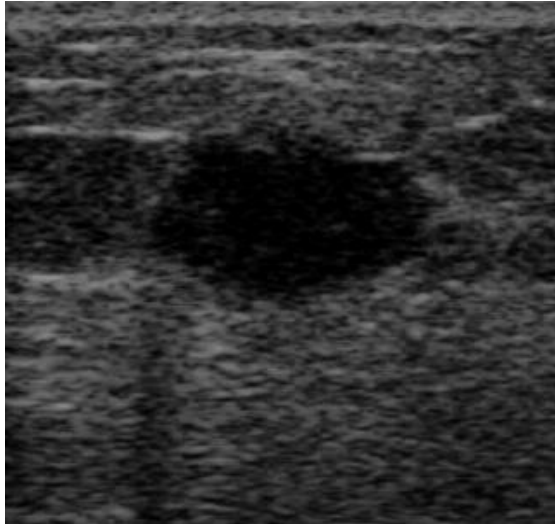


Figura 24 - Imagem original

Passo 2: Detecção do nódulo

O objeto a ser segmentado é muito diferente em contraste com a imagem de fundo. Alterações do contraste podem ser detetadas pelos operadores que calculam o gradiente de uma imagem. O gradiente da imagem pode ser calculado com um Treshold que é aplicado para criar uma máscara binária que contém o nódulo a ser segmentado. Pode-se usar a função *edge* e o operador *Sobel* para calcular o valor do limiar. Em seguida, ajusta-se o valor do limiar e usa-se novamente a função *edge* para se obter uma máscara binária que contém o nódulo parcialmente segmentado.

binary gradient mask

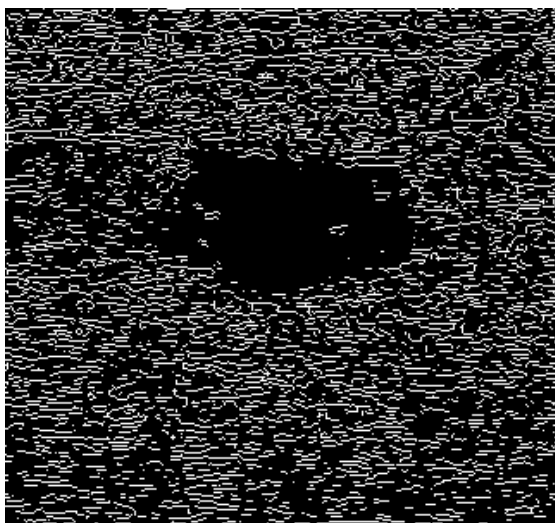


Figura 25 - Máscara binária do gradiente

Passo 3: Dilação da Imagem

A máscara binária do gradiente mostra linhas de elevado contraste na imagem. As linhas não chegam a delinear totalmente o nódulo. É possível observar lacunas nas linhas que rodeiam o nódulo na máscara de gradiente. As lacunas podem desaparecer se a imagem for dilatada com elementos estruturantes lineares por exemplo, que podem ser criados com a função Strel.

A máscara binária do gradiente é dilatada usando por exemplo elemento estruturante vertical seguido pelo elemento estruturante. A função imdilate dilata a imagem.

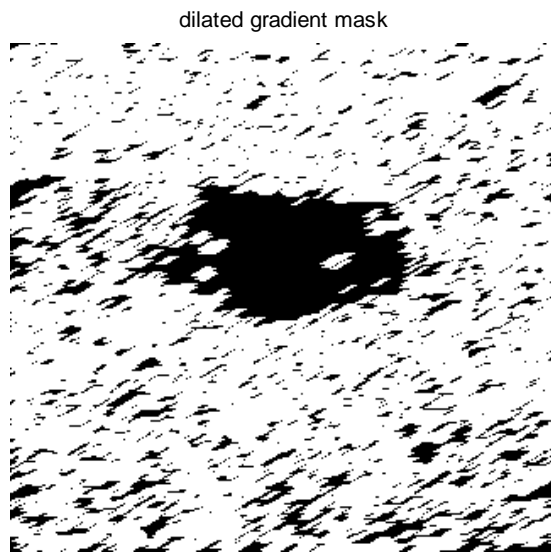


Figura 26 - Máscara de uma dilatação

Passo 4: Inversão da Imagem

É aplicada à imagem ou à máscara uma inversão da mesma, já que esta agora é uma imagem binária, todos os 0s são trocados por 1s e vice-versa.

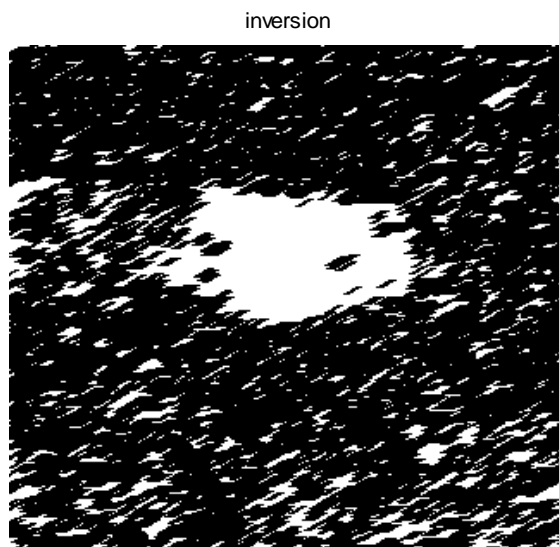


Figura 27 - Máscara da inversão

Passo 5: Preenchimento de lacunas de Interiores

A máscara de gradiente dilatada mostra o contorno do nódulo bastante bem, mas ainda existem lacunas no interior do mesmo. Para as preencher usa-se a função *imfill*.

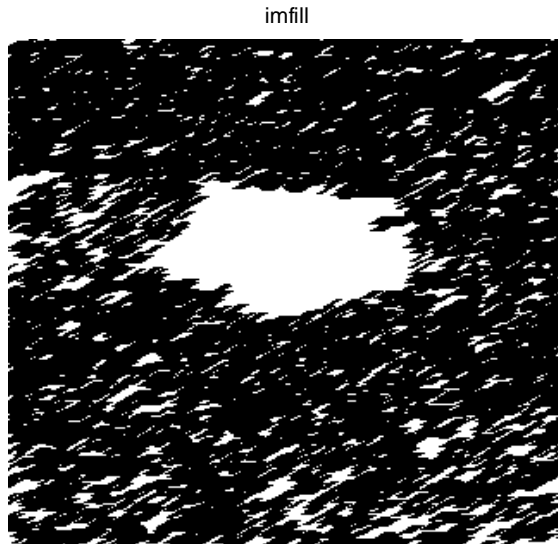


Figura 28 - Máscara do preenchimento de lacunas de Interiores

Passo 6: Remover objetos conectados ao contorno

O nódulo não é o único objeto que é observado. Todos os objetos que estão ligados ao contorno da imagem devem ser removidos usando a função *imclearborder*. A conectividade na função *imclearborder* também é ajustada para remover ligações diagonais.

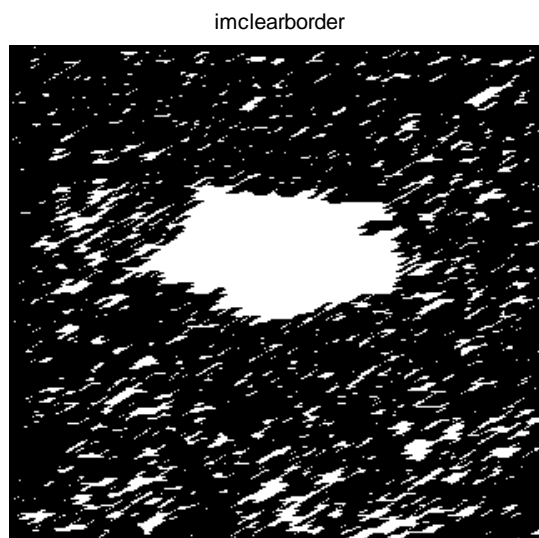


Figura 29 - Máscara da remoção de objetos conectados ao contorno

Passo 7: suavizar o objeto

Para tornar o objeto segmentado mais “limpo”, suaviza-se o objeto, erodindo a imagem com um elemento estruturante. Criado a partir da função *Strel*.

erosion



Figura 30 - Máscara da erosão

Passo 8: Operações intermédias

Outras operações são calculadas

outras



Figura 31 - Máscara das Operações intermédias

Passo 9: Mostrar apenas maior área

É calculada a área dos objetos da imagem e apenas são apresentados os de mais relevante.

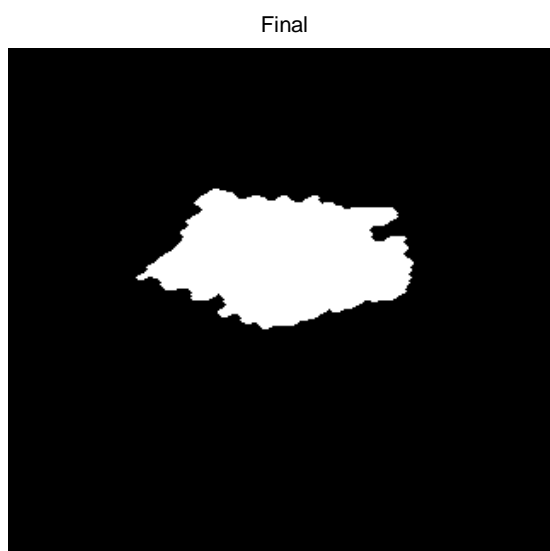


Figura 32 - Máscara da apresentação da maior área

Passo 10: Comparação com segmentação manual

Para o AG atuar e “sortir efeito” é necessário comparar a máscara obtida a partir da segmentação manual com a máscara obtida a partir do algoritmo até aqui procedido.

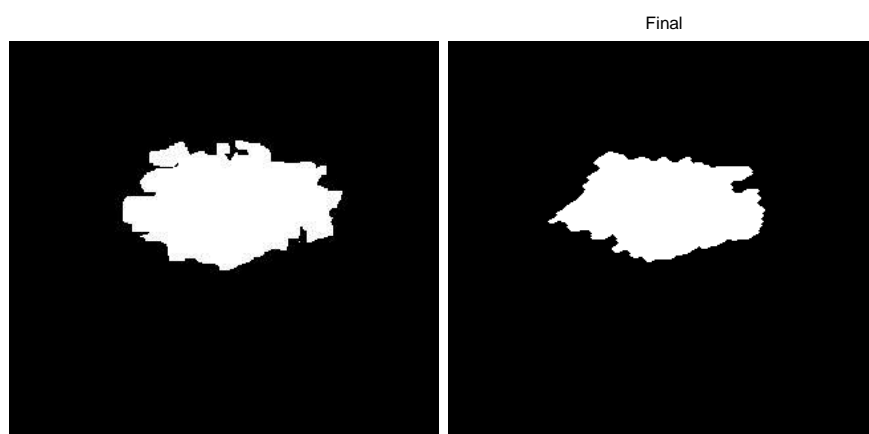


Figura 33 - Máscara obtida a partir da segmentação manual e máscara obtida a partir do algoritmo

Passo 11: Guardar as imagens (Máscara) e/ou Imagem segmentada

Por fim as imagens/máscaras são/podem ser guardadas para posterior análise/classificação.

Pode ser guardada também a imagem original já com o contorno à volta do nódulo. Isto é com o contorno à volta do nódulo segmentado.

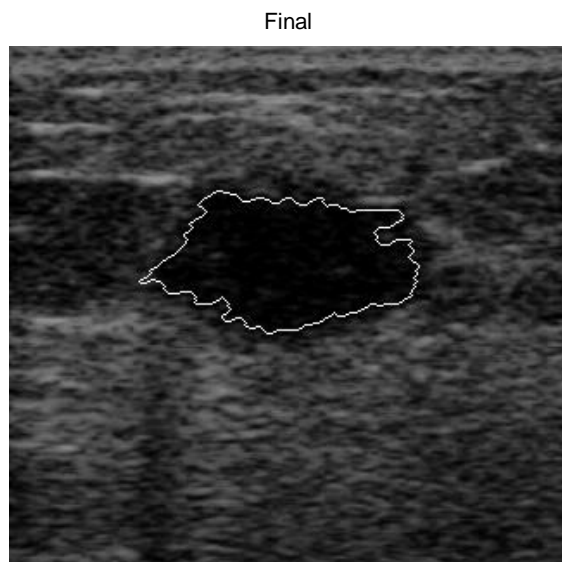


Figura 34 - Imagem original com o contorno à volta do nódulo

6.2.1. Funcionamento do Algoritmo Genético

O algoritmo começa com a criação de uma população inicial aleatória. Ocorrendo a primeira execução da função fitness de acordo com a estratégia acima indicada.

O algoritmo cria uma sequência de novas populações, ou gerações.

A cada passo, o algoritmo usa os indivíduos da geração atual para criar a próxima geração, procedendo outra execução da função fitness. Para criar a nova geração, o algoritmo executa as seguintes tarefas:

- Pontua cada membro da população atual calculando o seu valor de fitness.
- Escala as pontuações de aptidão e converte-as numa gama mais utilizável de valores.
- Seleciona os pais com base na aptidão.
- Produz filhos dos pais. Crianças são produzidos fazendo mudanças aleatórias num único pai - as mutações - ou pela combinação das entradas do vetor de um par de pais - o crossover.
- Substitui a população atual com as crianças para formar a próxima geração.
- O algoritmo termina quando um dos critérios de paragem for satisfeito.

O diagrama da figura X, mostra um fluxograma adaptado de [35] que ilustra o funcionamento do AG com o algoritmo acima como função fitness, conforme aqui foi procedido.

O bloco GA - refere-se ao funcionamento do AG em si, já o bloco maior refere-se ao funcionamento geral já com a função fitness incluída, bem como a entrada das imagens do conjunto de treino e saída dos resultados a serem aplicados no conjunto de teste.

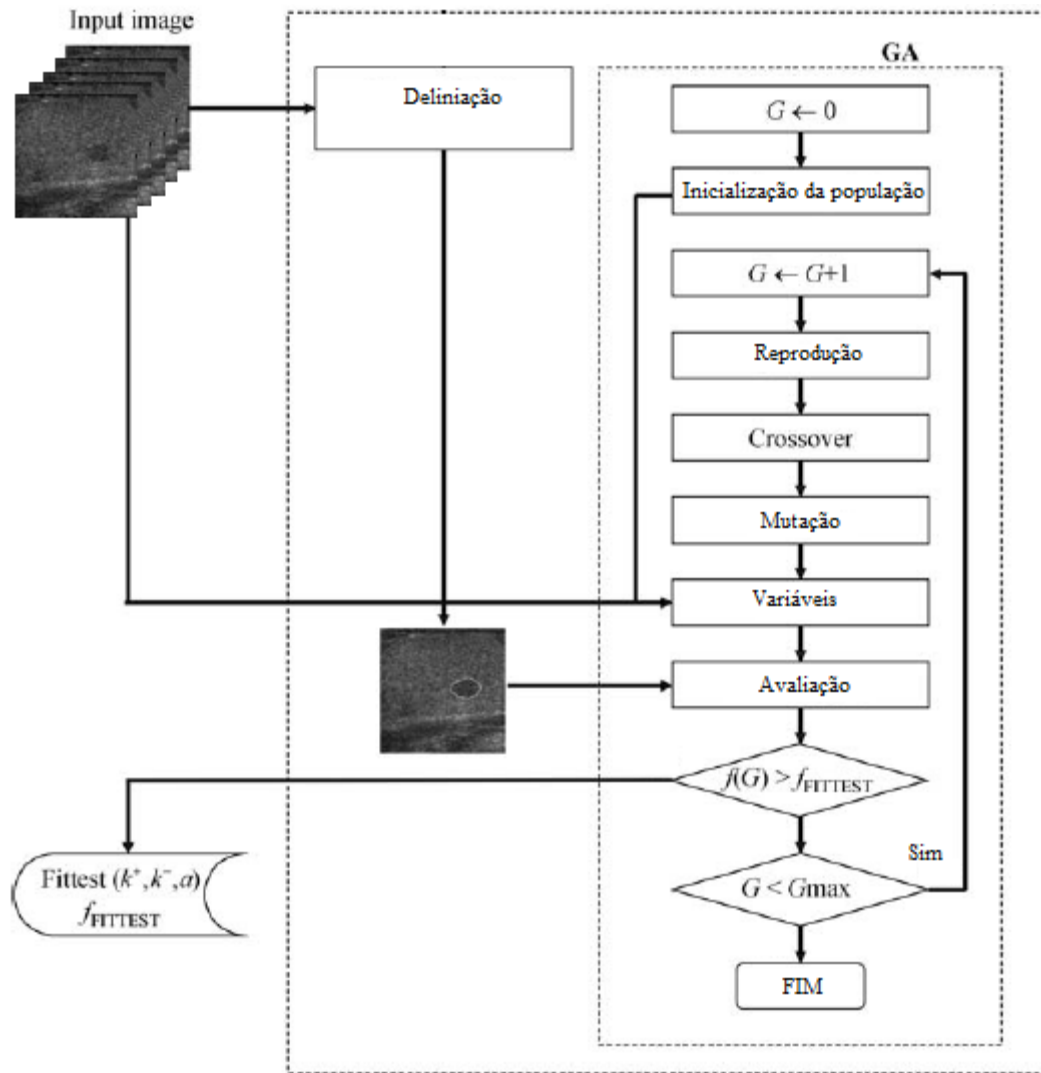


Figura 35 - Fluxograma ilustrando o funcionamento do sistema AG a respectiva função fitness.

Capítulo 7

Parte prática - Resultados

7.1. Resultados

Inicialmente foi desenvolvido um algoritmo para a detecção automática de nódulos mamários, apenas para a única imagem que então possuíamos, apresentada na Figura 36. Esse algoritmo era relativamente bom para aquela imagem.



Figura 36 - 1ª Imagem

Mais tarde, já na posse de um conjunto novas imagens de ultrassom B-mode da mama, obtidas no Hospital da Covilhã, gentilmente cedidas e obtidas pelo Professor Doutor José Alberto Fonseca Moutinho e através do Professor Doutor Mário Marques Freire, passámos a desenvolver novos procedimentos. Em que 20 imagens foram consideradas como conjunto de treino e 40 como conjunto de teste (20 destas com nódulos as outras 20 sem). Estas continham marcas aplicadas pelo médico. A Figura 37 mostra um exemplo de duas imagens deste conjunto.

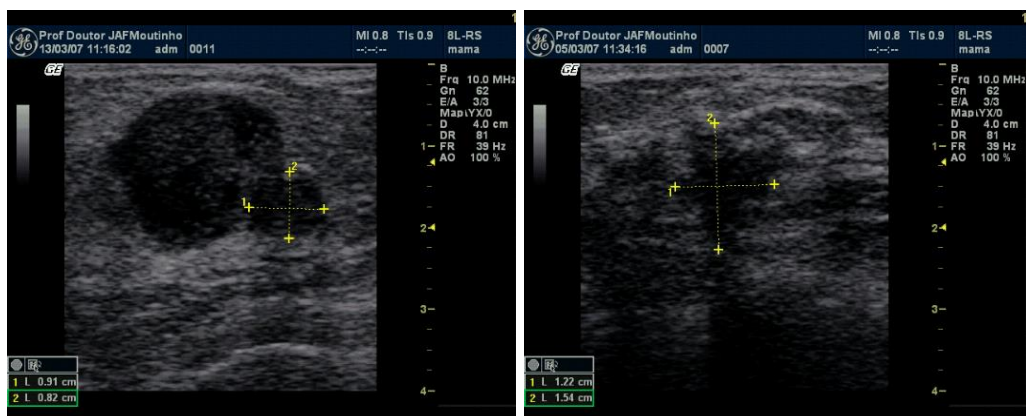


Figura 37 - Exemplo de duas imagens de ultrassom mamário com marca

O algoritmo foi aplicado ao conjunto de novas imagens. Constatou-se que as marcas impediam o bom funcionamento do algoritmo. Como é possível ver na Figura 38.

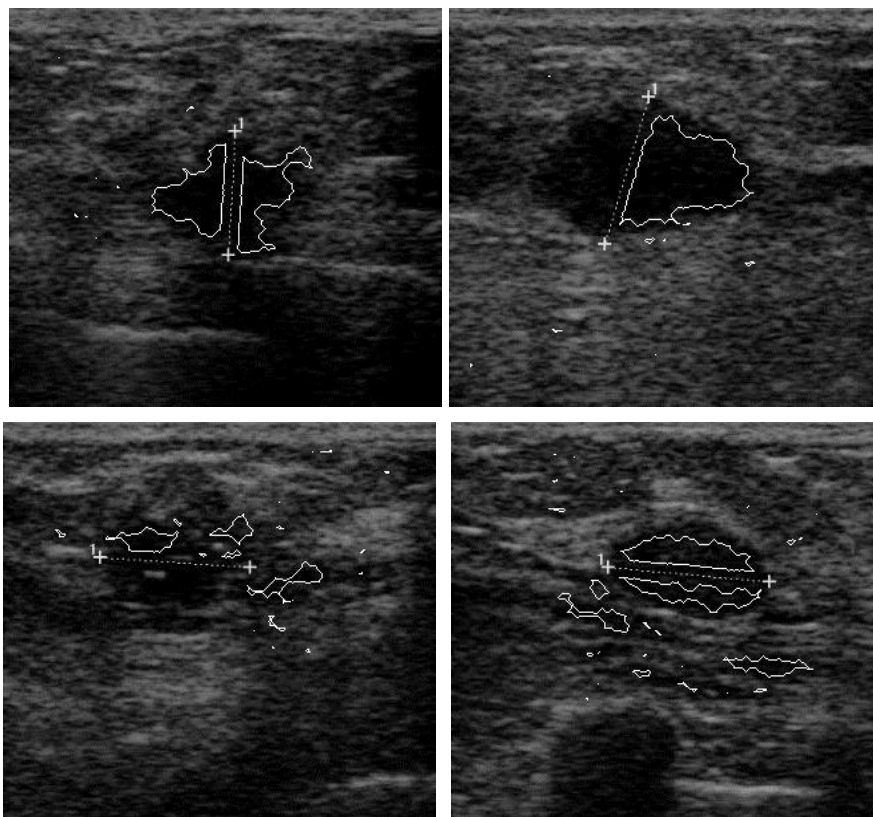


Figura 38 - Exemplo de quatro imagens de ultrassom mamário com marca, segmentadas

Desta forma as marcas tiveram que ser retiradas por um *software* open-source (GIMP) com uma determinada função ou filtro para que pudessem ser convenientemente processadas. Mesmo depois de terem sido retiradas essas marcas, este processo pode não ter deixado as imagens como se não tivessem qualquer marca inicial, podendo ter perturbado o resultado final. Todas as imagens foram ainda recortadas para apenas apresentarem a ecografia.

Com base nas imagens marcadas originalmente foi procedida uma segmentação manual destas, criando uma máscara da segmentação, que embora não tenha sido realizado por uma pessoa experiente, esta foi realizada por conhecimento empírico sendo marcados os nódulos visualmente aparentes, mas podendo não ter levado aos resultados mais precisos.

Posteriormente o algoritmo foi aplicado no conjunto de novas imagens já com as imagens sem as marcas. Certificamos que este não era de todo o melhor algoritmo a ser aplicado neste conjunto, ou na verdade, os parâmetros do algoritmo é que não eram os mais adequados, conforme mostrado na Figura 39. (Em alguns casos o resultado era de 0%)

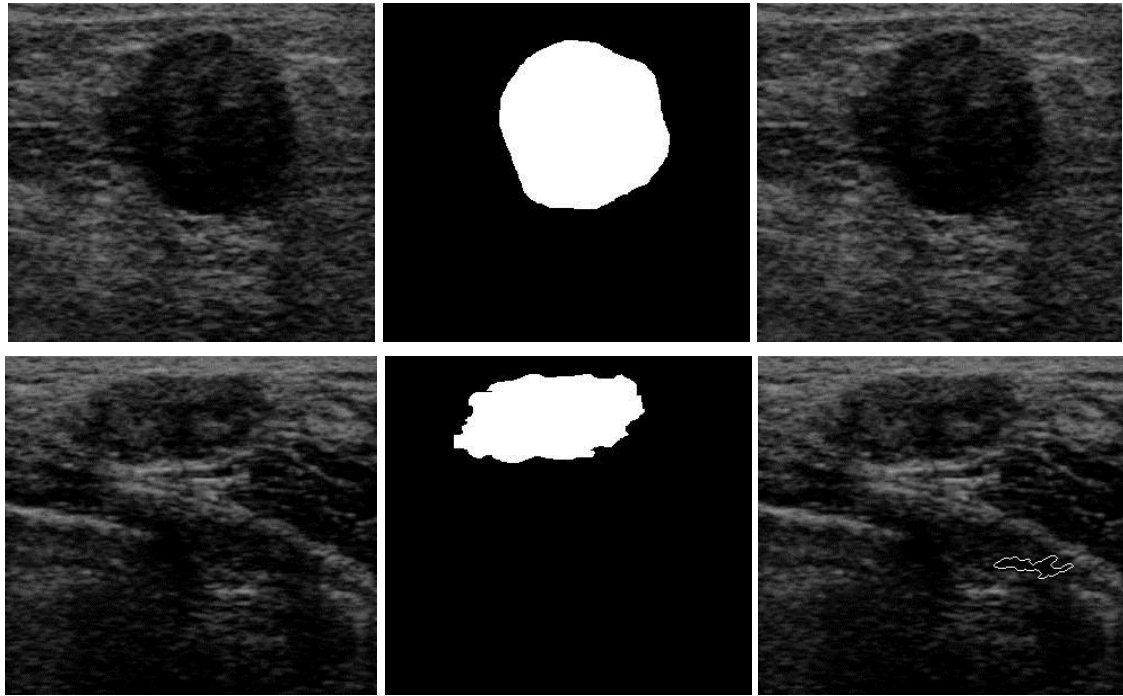


Figura 39 - Segmentação inicial

Com o objetivo de determinar o funcionamento ótimo do algoritmo anteriormente referido foram feitos diversos ensaios com várias ferramentas elegendo os Algoritmos Genéticos (AG), já resumido no capítulo 5. **O algoritmo até aqui utilizado torna-se a função fitness do AG.**

Diversos testes com os AG foram procedidos, várias definições experimentadas, como o tamanho da população, gama inicial de valores, taxas de mutação e crossover e algumas funções personalizadas.

Como podem ser utilizadas, por exemplo, na interface gráfica mostrada na figura.

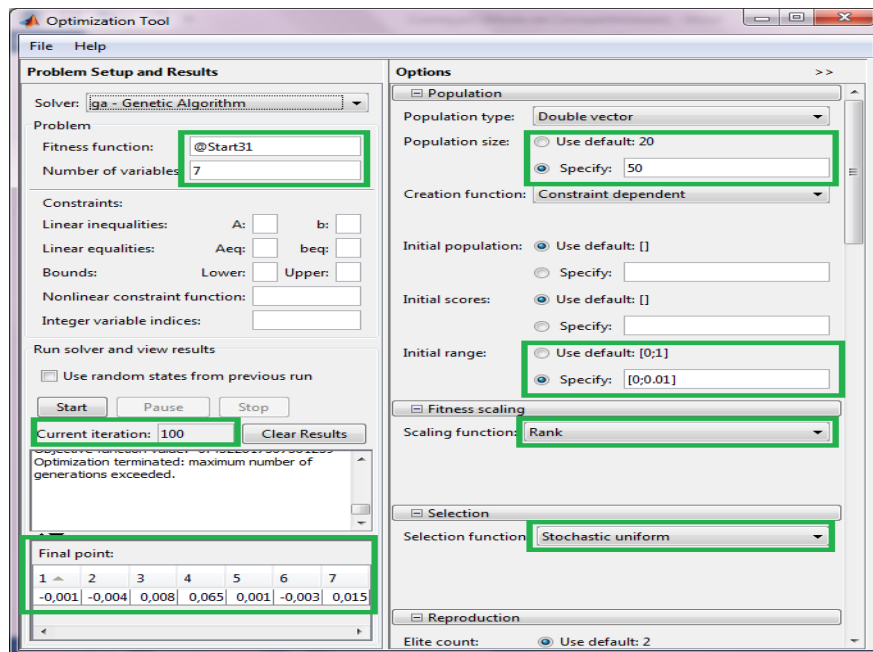


Figura 40- Interface gráfica

Como mencionado na secção 5.2. Optimization Toolbox™ - Toolbox Algoritmo Genético, para obter os melhores resultados do AG, pode ser necessário experimentar diferentes opções, envolvendo tentativa e erro.

A figura 41 ilustra um dos muitos testes que foram realizados.

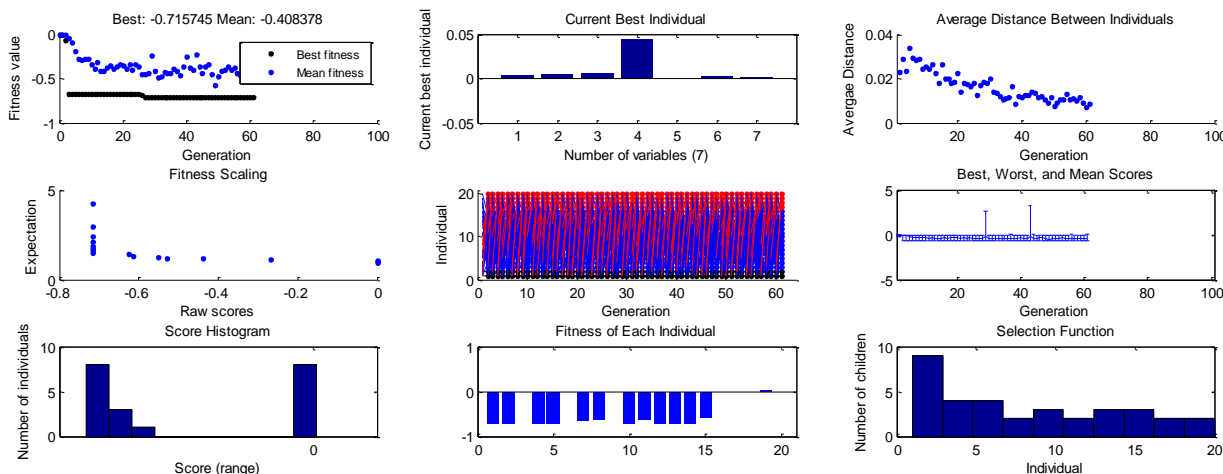


Figura 41 - Diversos gráficos que podem ser produzidos pela Toolbox

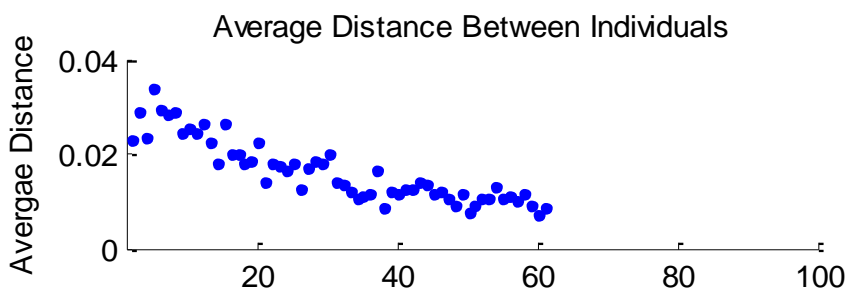


Figura 42 - “Average Distance Between Individuals”

Constata-se no gráfico “Average Distance Between Individuals” da figura 41 que aparece ampliado na figura 42 bem como nos gráficos das figuras 44 à 48 que a diversidade da população é um dos fatores mais importantes para determinar o quão bom é o AG. Se a distância média entre os indivíduos for grande, a diversidade será alta; se a distância média for pequena, a diversidade será baixa. A diversidade afeta o desempenho do AG. Se a diversidade for muito grande ou muito pequena, o AG poderá não ter um bom desempenho ou nem sequer ter capacidade para funcionar como ilustrado na figura 43.

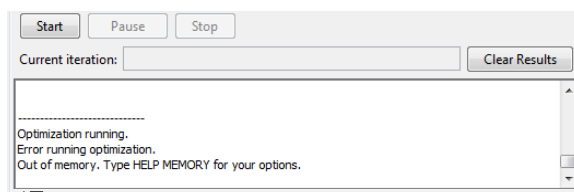


Figura 43 - Imagem ilustrativa mostrando que a memória foi excedida

Nesta opção a gama aqui considerada foi: [0:0,01] diminui o tamanho da amostra, mas mesmo assim ainda apresenta muitos valores de ordem de grandeza elevados, superiores a 10^{20} . Estas opções foram especificadas na gama inicial de modo que contivessem uma estimativa para a solução, aumentando desta forma a velocidade do AG.

A atribuição nas opções de dimensionamento da aptidão, fez com que a gama de valores calibrados tenha afetado o desempenho do AG. Quando os valores em escala variarem muito, os indivíduos com maiores valores reproduzem-se muito rapidamente, tomando conta da pool genética da população muito rapidamente, impedindo o AG de procurar outras áreas do espaço de solução. Por outro lado, se os valores em escala variarem pouco, todos os indivíduos têm, aproximadamente, a mesma possibilidade de reprodução e a pesquisa vai progredir muito lentamente.

Aumentando também o tamanho da população permitiu que o AG procurasse mais pontos e assim obter um melhor resultado. No entanto, quanto maior for o tamanho da população, mais tempo o AG irá demorar a calcular cada geração. Como pode ser comprovado na Tabela 2 e 3. Também o aumento do número de imagens a utilizar pelo algoritmo, permite obter valores de parâmetros mais eficazes, mas também aumenta o tempo de execução do AG, constatando na Tabela 5.

Nas tabelas 2 e 3, apresentam-se diversos testes para 5 imagens, com uma população 10, 25, 50, e 100 com Stall Time 50 e Stall Time 100 respectivamente em que T representa o Tempo de execução do GA em minutos, % representa a pontuação da função fitness em percentagem e n representa o número de gerações decorridas.

A exposição da diversidade de tempos de execução de qualquer execução de um AG pode ser visualizada na tabela 4, em que T representa o Tempo de execução do GA em segundos.

Já na tabela 5 são exibidos os testes de tempo para o número de algumas imagens utilizadas no AG 5, 10, 20 com Stall Time 50, População 25.

Tabela 2 - Testes usando Stall Time 50

N°Pop	10			25			50			100		
	T(min)	%	n	T(min)	%	n	T(min)	%	n	T(min)	%	n
1	0,43	78,2	51	1,32	80,62	55	2,37	82,61	52	4,53	85,17	51
2	0,45	77,54	51	1,1	81,54	51	2,27	81,39	51	4,28	82,77	51
3	0,46	78,29	51	1,13	80,27	51	2,18	80,95	51	4,4	81,54	51
4	0,54	77,93	51	1,26	81,4	57	2,24	81,78	51	5,05	82,21	60
5	0,52	77,49	51	1,26	80,29	51	2,27	80,55	51	4,42	81,62	51
Média	0,48	77,89	51	1,214	80,824	53	2,266	81,456	51,2	4,536	82,662	52,8

Tabela 3 - Testes usando Stall Time 100

N°Pop	10			25			50			100		
	T(min)	%	n	T(min)	%	n	T(min)	%	n	T(min)	%	n
1	1,2	80,53	100	2,61	80,11	100	4,73	81,42	100	9,26	82,61	100
2	1,11	80,27	100	2,7	82,77	100	4,95	81,15	100	9,78	82,61	100
3	1,13	79,03	100	2,58	82,61	100	4,71	82,61	100	9,34	83,32	100
4	1	80,17	100	2,61	80,45	100	4,64	82,13	100	9,58	82,17	100
5	1,05	80,34	100	2,43	80,27	100	4,68	81,54	100	10,37	82,17	100
Média	1,098	80,068	100	2,586	81,242	100	4,742	81,77	100	9,666	82,576	100

Tabela 4 - Diversidade dos Tempos de execução

Testes \ 10	Tempo (s)
1	31,64
2	32,58
3	29,16
4	31,29
5	34,3
6	32,0
7	39,31
8	37,33
9	36,19
10	33,16
Média	33,696

Tabela 5 - Testes de Tempo para o número de imagens 5, 10, 20 (Stall Time 50, População 25)

Testes \ N°Img	5 (2) T(s)	10(4) T(s)	20(6) T(s)
1	139,73	248,05	421,61
2	140,99	254,1	415,25
3	141,95	254,7	418,02
4	137,04	251,78	420,76
5	143,76	247,66	424,03
Média	140,694	251,258	419,934

Os gráficos da figura 44 à 48 foram obtidos com as diferentes opções de seleção da Toolbox

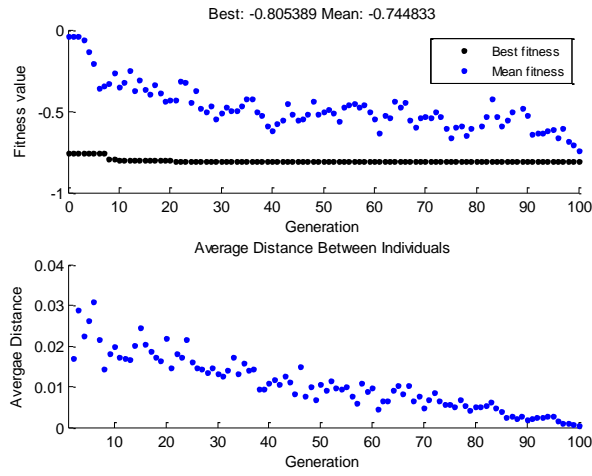


Figura 44 - Gráficos obtidos a partir da opção de seleção Stochastic uniform

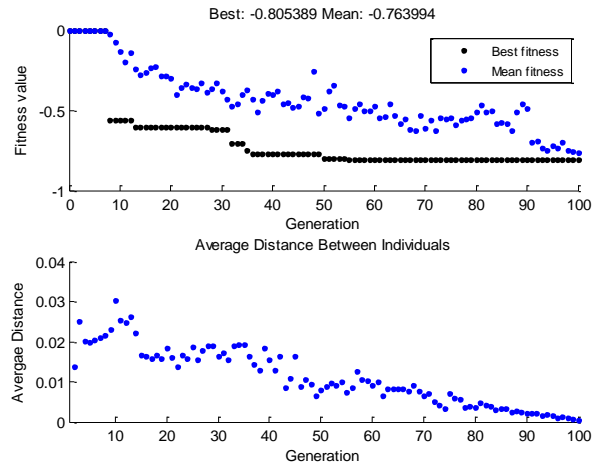


Figura 45 - Gráficos obtidos a partir da opção de seleção Reminder

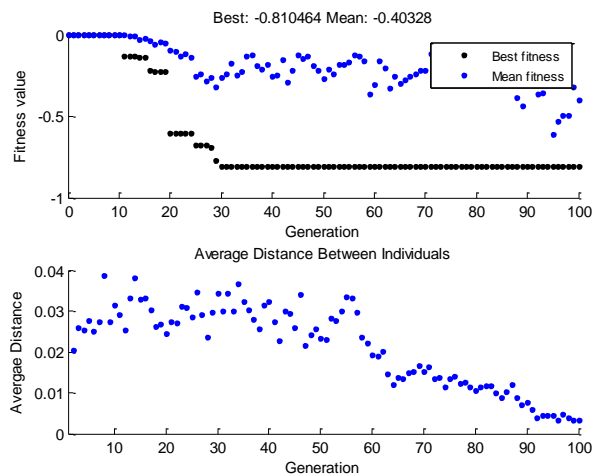


Figura 46 - Gráficos obtidos a partir da opção de seleção Uniform

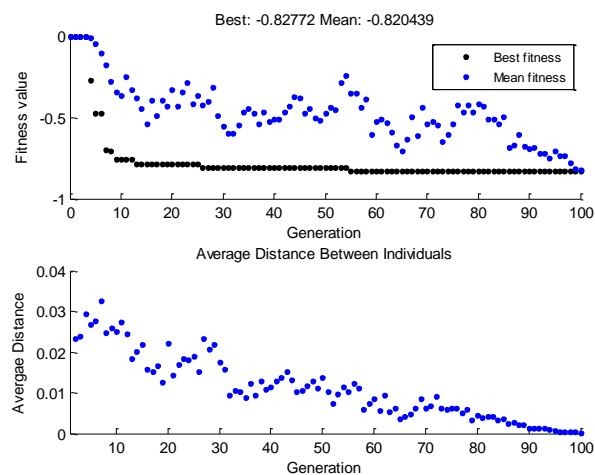


Figura 47 - Gráficos obtidos a partir da opção de seleção Roleta

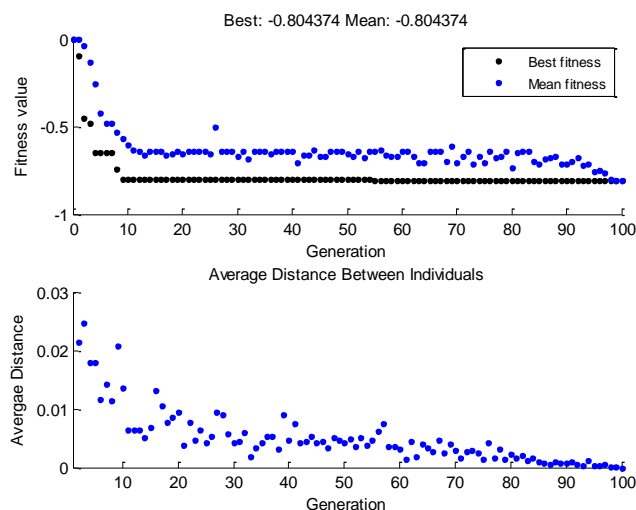


Figura 48 - Gráficos obtidos a partir da opção de seleção Torneio

Nas Opções de reprodução, “*Elite count*” - O número de indivíduos com os melhores valores de aptidão da geração atual são os mais aptos a sobreviver à próxima geração. “*Elite count*” um valor alto faz com que os indivíduos mais aptos dominem a população, tornando a busca menos eficaz.

A mutação aumenta a diversidade de uma população e, assim, aumenta a probabilidade de que o AG irá gerar indivíduos com melhores valores de aptidão.

Se a probabilidade de mutação ocorrer for muito elevada isto vai criar uma instabilidade na população.

Por exemplo, se o tamanho da população for 20, a contagem de Elite for 2, e a fração de Crossover for de 0.8, os números de cada tipo de “children”, a geração seguinte será:

- Existem 2 “*elite children*”
- 18 outros indivíduos da população para além desses dois “*elite children*”, o AG arredonda: $0.8 * 18 = 14,4$ para 14 a fim de obter o número “*children*” de crossover.
- Os restantes quatro indivíduos, são filhos de mutação.

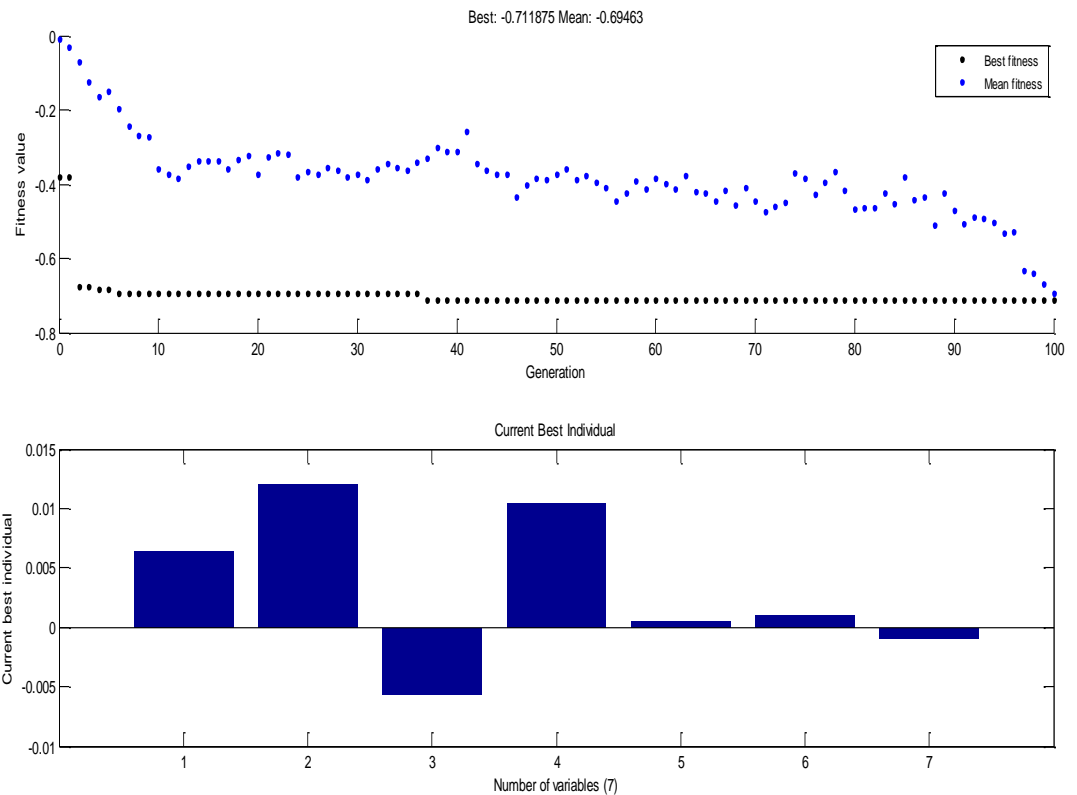


Figura 49 - Gráficos de Best Fitness e Current Best Individual

Após os parâmetros dos operadores terem sido determinados pelo AG que foram obtidos a partir do conjunto de treino, o conjunto de teste foi ensaiado, sendo possível constatar que nas imagens com nódulo, os resultados obtidos da segmentação automática foram na maioria visualmente superiores à segmentação manual - apresentando os contornos melhor delineados.

Já nas imagens sem nódulo não é segmentado qualquer objeto (ou um muito pequeno, que poderá ser considerada a desvalorização no algoritmo)

Todo o processo de segmentação foi elaborado, sem o auxílio de um médico/técnico especialista, que pudesse validar os reais contornos ou dar como referência os mesmos. Assim todo a segmentação foi guiada por um observador não treinado.

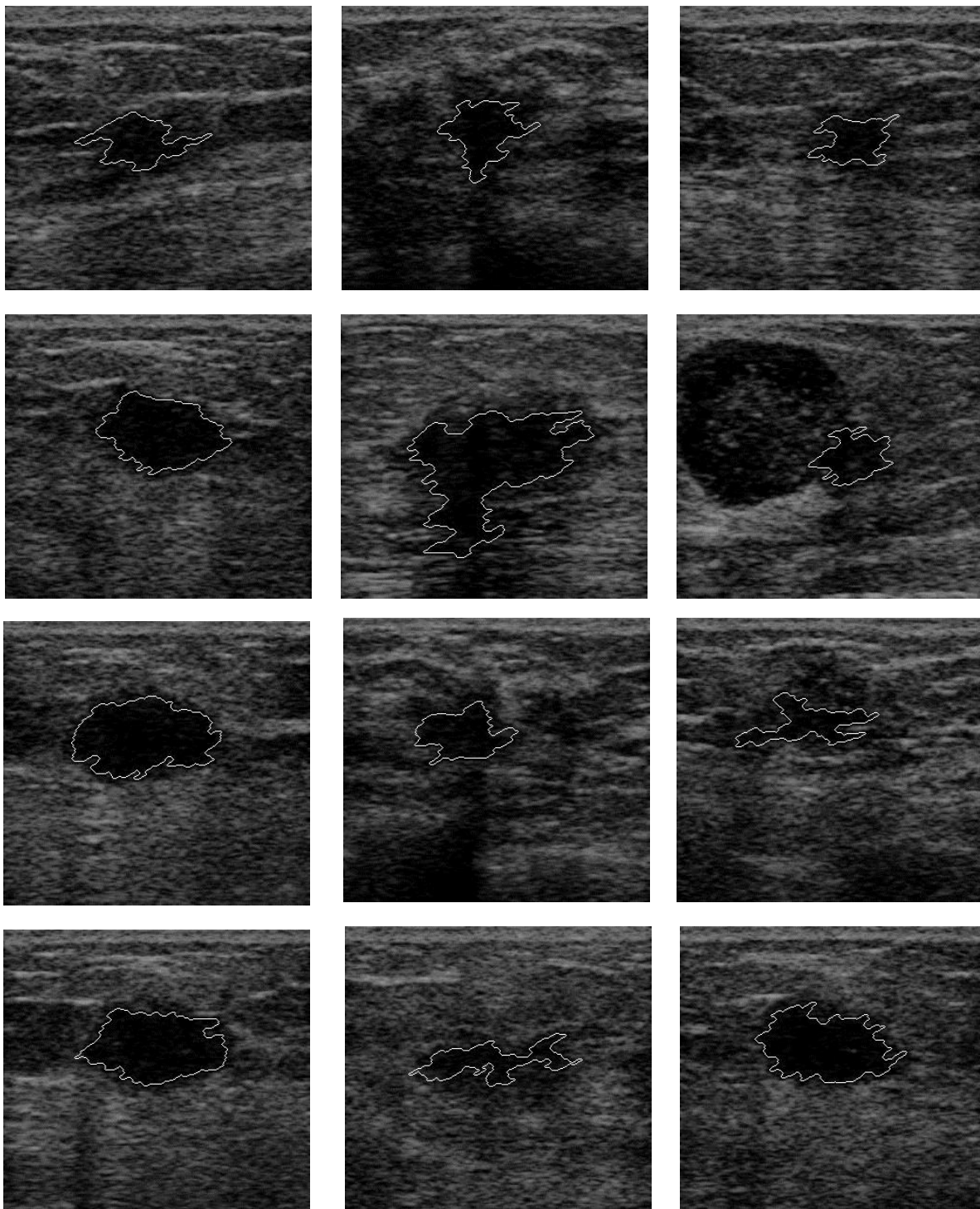


Figura 50 - Imagens segmentadas com o contorno à volta do nódulo

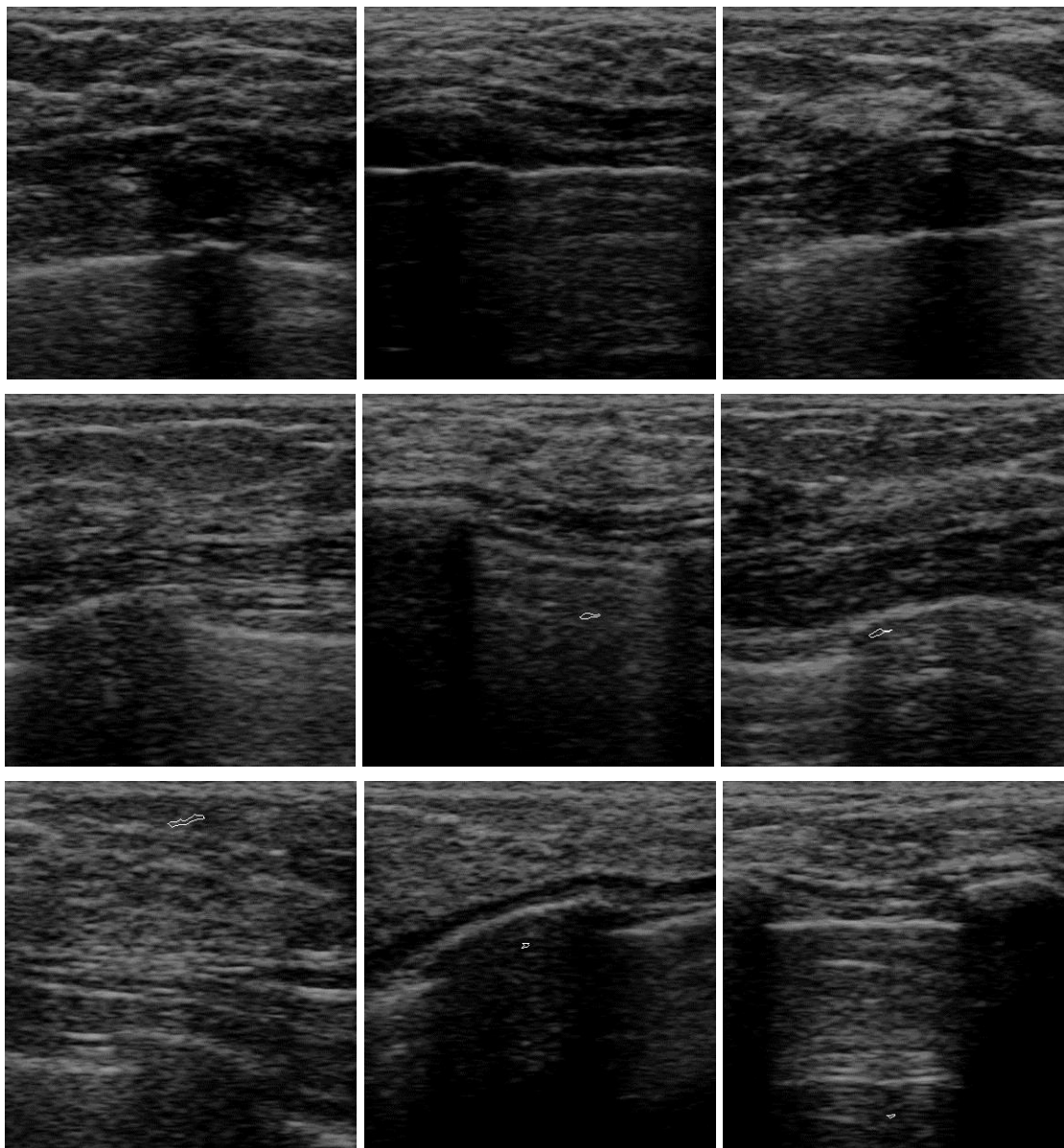


Figura 51 - Imagens segmentadas sem nódulo

Com referência à Análise Quantitativa os fatores que determinam a confiabilidade de um sistema CAD, em qualquer sistema de imagem médica, relacionam-se com a quantidade de casos Verdadeiro Positivos (VP) e Falso-Negativos (FN) que o CAD ajuda a detetar, e a quantidade de Falso Positivos (FP) e Verdadeiro-Negativos (VN) que ajuda a excluir. A análise quantitativa comparou a quantidade de informações apresentadas pelo médico no momento do teste com os resultados obtidos. Embora não seja aqui apresentada, a maioria dos casos obtidos eram Verdadeiros Positivos para as imagens com nódulo e Verdadeiro-Negativos nas imagens sem nódulo.

Capítulo 8

Conclusão

A mamografia é a técnica de diagnóstico que mais tem contribuído para a detecção e redução da mortalidade por cancro da mama. Estudos recentes têm mostrado que a probabilidade de mulheres morrerem de cancro da mama era igual à das mulheres que não fizeram o exame. Elevando deste modo a importância das ecografias como exame auxiliar das mamografias. O exame da mama através de imagens por ultrassom é o mais importante coadjuvante da mamografia, tornando-se deste modo um exame ainda mais importante, complementado por sistemas CAD, detetando e classificando nódulos mamários, limitando a maioria de fatores prejudiciais, evitando tratamentos e outras ações desnecessárias.

Esta dissertação tenta contribuir para melhorar a eficácia deste exame e para a procura de um diagnóstico cada vez mais precoce.

Foi apresentado um método de segmentação de nódulos mamários, baseado no uso de operadores morfológicos, detecção de contornos, métodos de thresholding e Algoritmos Genéticos, sendo aplicado a imagens reais de ultrassonografias da mama. Diversos testes/alterações foram feitos para que os AG determinassem os melhores parâmetros dos métodos enumerados. Aqueles que apresentaram melhores resultados foram os que a população dos AG era maior e mais vasta, bem como um número alargado de imagens, tendo demorado na verdade um período considerável de tempo a serem obtidos.

Neste trabalho fez-se um esforço de produzir uma nova metodologia, que embora não tenham obtido resultado excepcionais, foram alcançados resultados significativos/satisfatórios.

Contribuíram portanto os seguintes factos:

- É sabido que as imagens de ultrassom são de muito difícil trabalho, devido por exemplo ao ruído que contém mesmo assim não foi feito nenhum pré-processamento de imagem para não perder certas características;
- Não ter sido possível obter o auxílio de um médico/técnico especialista para realizar a segmentação manual;
- As imagens disponíveis não se encontrem no formato mais apropriado;
- A função de comparação muito minuciosa: qualquer pixel da imagem segmentada que se localize para fora e/ou para dentro da região segmentada manualmente era contado negativamente para o cálculo do fitness ou da aptidão;

- Algumas limitações por hardware: observou-se que a utilização do processador era bastante elevada, necessitando muitos recursos;
- O tempo de execução do Algoritmo Genético era muito elevado para populações grandes, bem como para um vasto número de imagens;

A maioria dos casos obtidos eram Verdadeiros Positivos nas imagens com nódulo e Verdeiro-Negativos nas imagens sem nódulo. Os contornos obtidos revelaram mais detalhes que os estabelecidos pela segmentação manual, sendo esta uma característica importante na avaliação da irregularidade do contorno dos nódulos. Demonstrando deste modo elevada potencialidade para o uso dos operadores e dos AG para segmentação e posterior classificação de nódulos da mama em imagens por ultrassom.

Porém a validação dos reais contornos ou uma referência dos mesmos, também não comprovados por um médico/técnico especialista. Assim todo o processamento do contorno foi guiado por um observador não treinado.

Em suma o trabalho realizado deixou boas perspectivas para o futuro, e com vários caminhos que podem ser percorridos para auxiliar o combate ao cancro da mama, bem como a possibilidade de adequação a outro tipo de diagnósticos.

8.1. Trabalho futuro

Como sugestão para uma próxima etapa, começamos por sugerir a implementação do uso da GPU, aproveitando os recursos da placa gráfica, bem como o uso da “computação paralela com otimização” tornando diversos processos mais rápidos e eficazes.

Aplicação do método a um maior número de imagens e avaliação/classificação de nódulos conforme a malignidade ou benignidade.

Como uma última sugestão, não podemos deixar de mencionar a criação de uma interface (portátil) o mais simples possível, quer do ponto de vista de manuseamento quer do ponto de vista da análise das imagens, por exemplo, para que seja possível o seu uso doméstico ou em localizações remotas..., com possível observação dos dados detalhados e com possibilidade de envio das imagens/informações para um médico ou observador experiente - telemedicina - Cloud.

Referências

- [1] Anthony B. Miller, MB, FRCP; Teresa To; Cornelia J. Baines, MD; and Claus Wall, Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. 11 February 2014).
- [2] Andy Milkowski, MS, Yadong Li, David Becker, MS, Syed O. Ishrak, Speckle Reduction Imaging, GE Medical Systems, Ultrasound, 2003.
- [3] J. Alison Noble, Djamel Boukerroui; Ultrasound Image Segmentation: A Survey, IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 25, NO. 8, AUGUST 2006
- [4]- Os cinco cancros mais frequentes nas mulheres <http://www.cancrodamama.com/index/os-cinco-cancros-mais-frequentes-nas-mulheres> revisitado em setembro 2014.
- [5] O CANCRO DA MAMA - <http://www.ligacontracancro.pt/gca/index.php?id=14> revisitado em setembro 2014.
- [6] E.D. Pisano, C. Gatsonis, E. Hendrick, M. Yaffe, J.K. Baum, S. Acharyya, E.F. Conant, L. L. Fajardo, L. Bassett, C. D'Orsi, R. Jong, M. Rebner and Digital Mammographic Imaging Screening Trial (DMIST) Investigators Group. Diagnostic Performance of Digital versus Film Mammography for Breast-Cancer Screening. New England Journal of Medicine, 2005
- [7] (KREMKAU, 1996) no jornal científico Reproductive Biology and Endocrinology A equipa da Escola de Medicina da Universidade da Carolina do Norte.
- [8] L. Birdwell, D. M. Ikeda, K. D. O'Shaughnessy, and E. A. Sickles. Mam-mographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. Radiology, 2001.
- [9] T. W. Freer and M. J. Ulissey. Screening mammography with computer-aided detection: Prospective study of 12860 patients in a community breast center. Radiology, 220:781-786, 2001.
- [10] P. Taylor, S. Hajnal, M. H. Dilhuydy, and B. Barreau. Measuring image texture to separate "difficult" from "easy" mammograms. British Journal of Radiology, 1994.
- [11] R2 Image Checker®, the digital detective. <http://www.pavilionimaging.com/mammography-center/r2-image-checker> revisitado em setembro 2014.
- [12] A. O. I. Malagelada. Automatic Mass Segmentation in Mammographic Images, Espanha, 2007.
- [13] ZONG, X.; LAINE, A. F.; GEISER, E. A. Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing. IEEE Transactions on Medical Imaging, 1998.
- [14] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, R. A. Borges, and A. F. Frère. Automatic identification of the pectoral muscle in mammograms. IEEE Transactions on Medical Imaging, 2004

- [15] GONZALEZ, R.C. Digital Image Processing. Prentice Hall, 2002.
- [16] L. G. Roberts. Machine perception of three-dimensional solids. In J. Tippet, D. Berkowitz, L. Clapp, C. Koester, and A. Vanderburgh, editors, Optical and Electro-Optical Information Processing, pages 159-197. MIT Press, Cambridge, Massachusetts, 196
- [17] N. Petrick, H. P. Chan, B. Sahiner, and D. Wei. An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection. IEEE Transactions on Medical Imaging, 1996.
- [18] N. Petrick, H. P. Chan, D. Wei, B. Sahiner, M. A. Helvie, and D. D. Adler. Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification. Medical Physics, 23(10). 1996.
- [19] DOUGHERTY, Geoff (2009) - Digital Image Processing for Medical Applications. Cambridge University PRESS. ISBN:9780511533433
- [20] T. Matsubara, H. Fujita, T. Hara, S. Kasai, O. Otsuka, Y. Hatanaka, and T. Endo. Development of a new algorithm for detection of mammographic masses. In Proc. International Workshop on Digital Mammography, 1998.
- [21] GONZALEZ, Rafael C.; WOODS, Richard E.- Digital image Processing (2001) - 2nd Edition. Prentice Hall. ISBN:0201180758
- [22] Z. Huo, M. L. Giger, C. J. Vyborny, U. Bick, P. Lu, D. E. Wolverton, and R. A. Schmidt. Analysis of spiculation in the computerized classification of mammographic masses. Medical Physics, 1995.
- [23] Chapter 5 - MATHEMATICAL MORPHOLOGY - PDF - www.bioss.ac.uk/people/chris/ch5.pdf
- [24] Morphological Operations - MATLAB - Image Processing Toolbox - Image Enhancement - <http://www.mathworks.com/help/images/morphological-filtering.html> revisitado em setembro 2014.
- [25] CAMPILHO, Aurélio - Morphological imaging Processing. Biomedical Image Analysis. Master Degree on Biongingineering.
- [26] Genetic Algorithms in Search, Optimization and Machine Learning 1st Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA ©1989 ISBN:0201157675
- [27] Pereira, F. (2012). Capítulo 2 Algoritmos Genéticos. In Gaspar-Cunha, A, Takahashi, R, Antunes, C, Manual de Computação Evolutiva e Meta Heurística (pp. 25-47). Imprensa da Universidade de Coimbra
- [28] An Introduction to Genetic Algorithms (Complex Adaptive Systems) Paperback - February 6, 1998, Melanie Mitchell, ISBN-13: 978-0262631853
- [29] Evolutionary Analysis, 3/E Jon C. Herron Scott Freeman, University of Washington, ISBN-10: 0321616677
- [30] Silva, A. Gramaxo, F. Santos, M. Mesquita, A. Baldaia, L, Félix, J. (2007). Capítulo 7 Evolução biológica e sistemática dos seres vivos. In Terra, Universo de Vida | 1ª Parte Biologia | Biologia e Geologia 11º ou 12º ano (ano 2) (pp. 115-145), Porto Editora
- [31] Genetic Algorithm and Direct Search Toolbox, <http://www.mathworks.com/products/gads/> revisitado em setembro 2014.

- [32] A Field Guide to Genetic Programming - Riccardo Poli William B. Langdon Nicholas F. McPhee - <http://cswww.essex.ac.uk/staff/rpoli/gp-field-guide/> revisitado em setembro 2014.
- [33] A.V. Alvarenga, A.F.C. Infantosi, C.M. Azevedo, W.C.A. Pereira. Applying Genetic Algorithms to Optimise Breast Ultrasound Images Segmentation Procedure Based on Morphological Operators.
- [34] A.V. Alvarenga, A.F.C. Infantosi, C.M. Azevedo, W.C.A. Pereira; Application of morphological operators on the segmentation and contour detection of ultrasound breast images. Revista Brasileira de Engenharia Biomédica, v. 19, n. 2, p. 91-101, agosto 2003
- [35] Dimitris K. Iakovidis · Michalis A. Savelonas · Stavros A. Karkanis · Dimitris E. Maroulis, A genetically optimized level set approach to segmentation of thyroid ultrasound images Published online: 16 June 2007 © Springer Science+Business Media, LLC 2007
- [36] Constituição do órgão - <http://www.pop.eu.com/portal/publico-geral/tipos-de-cancro/cancro-da-mama/o-cancro-da-mama.html> acedido em março 2014.
- [37] Irrigação sanguínea e drenagem linfática da mama <http://www.belezain.com.br/medestetica/mamoplastia.asp> acedido em março 2014.
- [38] NAZÁRIO, A.; REGO, M.; OLIVEIRA, V. Nódulos Benignos da Mama: Uma Revisão dos Diagnósticos Diferenciais e Conduta. Revista Brasileira Ginecologia Obstetrícia;v, 2007
- [11] E.D. Pisano, C. Gatsonis, E. Hendrick, M. Yaffe, J.K. Baum, S. Acharyya, E.F. Conant, L. L. Fajardo, L. Bassett, C. D'Orsi, R. Jong, M. Rebner and Digital Mammographic Imaging Screening Trial (DMIST) Investigators Group. Diagnostic Performance of Digital versus Film Mammography for Breast-Cancer Screening. New England Journal of Medicine, 2005
- [39] http://resources.edb.gov.hk/physics/articleIE/smartmaterials/SmartMaterials_e.htm
- [40] Diferentes tipos de transdutores Modos de visualização da informação obtida através dos ecos de ultrassons - <http://www.crcamp.com.br/> acedido em março 2014.
- [41] Biomedical Technology and Devices Handbook, CRC PRESS, 2004.
- [42] www.veterinaria.com.pt/media//DIR_27001/VCP1-1-e22.pdf.
- [43] Modos de visualização da informação obtida através dos ecos de ultrassons http://doctor-hill.com/iol-main/immersion_ascan.htm acedido em março 2014.
- [44] Modos de visualização da informação obtida através dos ecos de ultrassons <http://www.omjournal.org/> acedido em março 2014.
- [45] Modos de visualização da informação obtida através dos ecos de ultrassons <http://library.wolfram.com/examples/echocardiogram/echocardiography.html> acedido em março 2014.

Anexos

O cancro da mama

A investigação continua a esclarecer questões relacionadas com o cancro da mama: são descobertos novos dados acerca das suas causas e novos modos de prevenir, detetar e tratar esta doença. Assim, as pessoas com cancro da mama podem esperar uma melhor qualidade de vida e menor hipótese de morrer devido a esta doença.

Constituição do órgão: A mama

A mama é uma glândula que pode produzir leite. Cada mama assenta nos músculos peitorais que cobrem as costelas.

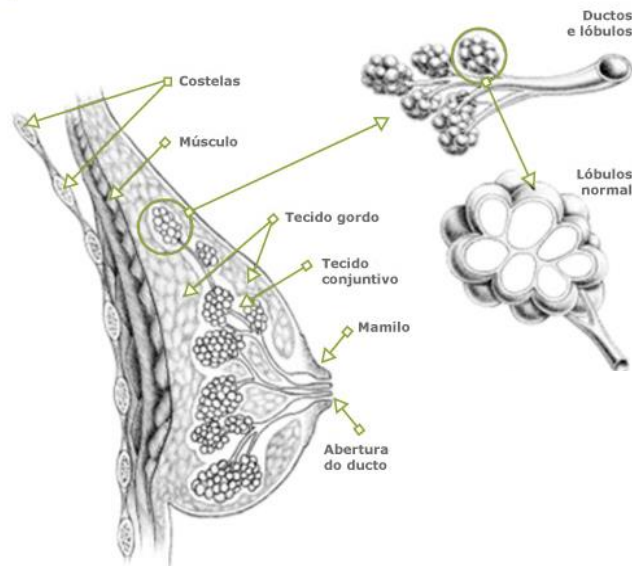


Figura 52 - Constituição de uma mama/secção transversal [36]

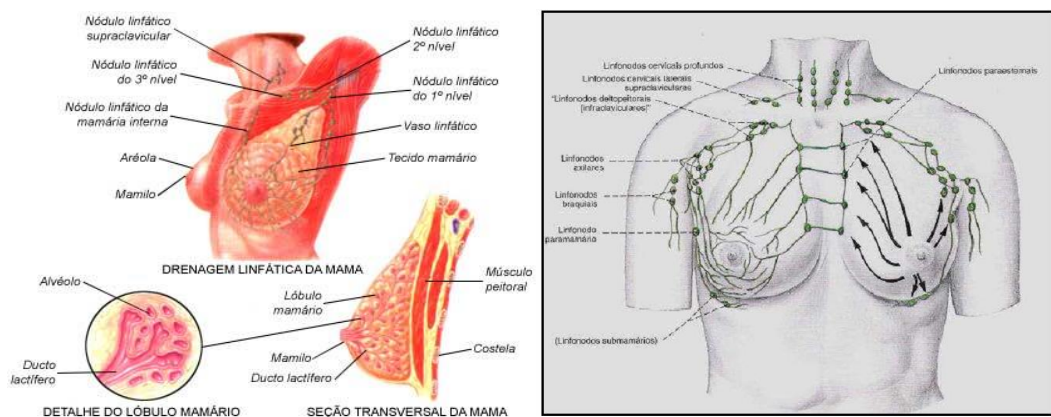


Figura 53 - Irrigação sanguínea e drenagem linfática da mama [37]

Cada mama encontra-se dividida em 15 a 20 secções, chamados de lobos. Estes contêm muitos lóbulos mais pequenos, que contêm grupos de pequenas glândulas que produzem leite.

O leite flui dos lóbulos através de uns tubos finos, os ductos, até ao mamilo. O mamilo é o centro de uma área escura de pele, a aréola. O espaço entre os lóbulos e os ductos é preenchido com gordura.

- Ácino - menor parte da glândula e responsável pela produção do leite durante a lactação;
- Lóbulo mamário - conjunto de ácinos;
- Lobo mamário - conjunto de lóbulos mamários que se liga à papila através de um ducto;
- Ductos mamários - em número de 15 a 20 canais, conduzem a secreção de (leite) até a papila;
- Tecido glandular - conjunto de lobos e ductos;
- Papila - protuberância elástica onde confluem os ductos mamários;
- Aréola - estrutura central da mama onde se projeta a papila;
- Tecido adiposo - todo o restante da mama é preenchido por tecido adiposo ou gorduroso, cuja quantidade varia com as características físicas, estado nutricional e idade da mulher.
- A mama também tem vasos linfáticos, transportando linfa. Os vasos linfáticos terminam nos gânglios linfáticos. Encontram-se grupos de gânglios linfáticos perto da mama, nas axilas (debaixo do braço), acima da clavícula, no peito (atrás do esterno), e em muitas outras partes do corpo. Os gânglios linfáticos "prendem" e retêm bactérias, células cancerosas, ou outras substâncias malignas, que se podem encontrar no sistema linfático. Quando as células de cancro da mama entram no sistema linfático, podem ser encontradas nos gânglios linfáticos próximo da mama (regionais).

Nódulos mamários

A deteção de um nódulo palpável na mama é, para a maioria das mulheres, um fator de impacto emocional, em parte pela conotação negativa associada ao tratamento, que altera a imagem e a sexualidade da paciente. Por esta razão, é importante ressaltar que cerca de 80% dos tumores mamários palpáveis são alterações benignas e que não aumentam, significativamente, o risco de desenvolvimento de cancro de mama. [38]

Segundo o Instituto Nacional Do Cancro o limite da normalidade no tecido mamário é difícil de ser definido devido às modificações dinâmicas verificadas nas mamas ao longo da vida das mulheres. Daí surgiu o conceito de alterações funcionais benignas da mama, antigamente denominadas doença fibrocística ou displasia mamária. Embora tais termos antigos ainda sejam utilizados pelos médicos, é importante entender que eles são sinónimos.

O nódulo mamário difere dependendo da idade e do sexo, uma vez que alguns nódulos e algumas alterações são considerados normais. Por exemplo, nódulos encontrados nos mamilos de crianças recém-nascidas é dilatação considerada normal e também esperada. Nas jovens com aproximadamente 6 anos de idade, geralmente, há uma "protuberância mamária" indicando o início da puberdade. Outra forma de dilatação comum ocorre em crianças durante a puberdade, conhecida como ginecomastia, onde a dilatação pode ser pequena ou bem significativa, sendo uma reação à alteração de hormonas e desaparece no decorrer de alguns meses.

Alterações fibrocísticas benignas que podem ocorrer numa ou nas duas mamas são observadas com frequência. Os nódulos benignos normalmente são arredondados com bordas lisas ou elásticos ou levemente flutuantes (capazes de sofrer depressões ou alterações na forma) e não são fixos. Há sempre uma secreção do mamilo que ocorre apenas após a sua manipulação. A secreção é leitosa e normalmente pode ser produzida por ambas as mamas.

O tumor benigno, também conhecido como fibroadenoma, é comumente encontrado em mamas femininas, ocorrendo mais frequentemente em mulheres férteis, sendo a neoplasia mamária mais comum em pacientes com menos de 35 anos. Os fibroadenomas não são sensíveis e normalmente não são malignos. Embora estes sejam inofensivos, a remoção cirúrgica ainda é recomendada.

Há duas razões principais para esta recomendação. A não ser que o tumor seja removido e estudado em um microscópio, não há como garantir que este não seja canceroso e, embora o nódulo não envolva o tecido mamário, ele pode crescer bastante e deformar a mama.

A doença fibrocística é caracterizada por um número de pequenos quistos (sacos cheios de líquido ao redor do tecido fibroso) também é comum e ocorre com mais frequência durante os anos reprodutivos, constituindo uma das causas mais frequentes de tumores mamários. A aspiração com agulha é utilizada para diagnosticar e tratar esta doença. Se o líquido que estiver a ser retirado for de uma cor amarelada e contiver sangue e os nódulos desaparecerem por completo após o procedimento, não há necessidade de estudos mais detalhados.

Um papiloma intraductal é caracterizado por um pequeno crescimento dentro do ducto mamário. É inofensivo e frequentemente pode não ser sentido, havendo em alguns casos uma secreção aquosa e rosa do mamilo. Uma vez que a secreção aquosa ou com sangue possa estar associada a cancro (muito raramente), um exame microscópico do líquido e normalmente uma mamografia são recomendados para avaliar as possibilidades. Neste caso a remoção cirúrgica é recomendada.

Potencialmente, os nódulos mamários malignos podem ser acompanhados de secreção espontânea dos mamilos em apenas um ducto em uma mama. A secreção ocorre em aproximadamente 4% dos cancros mamários.

Enquanto menos de 20% dos nódulos mamários representam um risco de vida, os sinais clínicos de cancro mamário não são facilmente distinguidos desta doença.

Deteção do cancro da mama

Para a deteção precoce do cancro da mama, é recomendado que:

- Mulheres com 40 anos ou mais, devem fazer uma mamografia (raio-X da mama) anualmente ou em cada dois anos.
- Mulheres que apresentem um risco aumentado (relativamente à média) de ter cancro da mama, devem falar com o seu médico acerca de fazer uma mamografia antes dos 40 anos, e saber qual a frequência para as próximas.

Exame clínico da mama

Durante um exame clínico da mama, o médico palpa as mamas em diferentes posições...

O médico procura quaisquer diferenças entre as mamas, incluindo diferenças invulgares de tamanho ou forma. Na pele, é verificada a presença de vermelhidão, depressões cutâneas ou outros sinais anormais. Os mamilos devem ser pressionadas para verificar se existe alguma secreção ou perda de líquido.

O médico poderá ainda examinar toda a mama, usando a ponta dos dedos para sentir quaisquer alterações e/ou nódulos, a área axilar e a área da clavícula, primeiro de um lado e depois do outro, Um nódulo apresenta, geralmente, o tamanho de uma ervilha, antes que alguém o consiga sentir ou palpar. Podem ser verificados os gânglios linfáticos perto da mama, para ver se estão inchados.

Se a mulher tiver um nódulo na mama, o médico deverá conseguir caracterizá-lo e determinar qual a sua causa (ou etiologia) através da palpação (pelo tato), sentindo o nódulo e a pele à sua volta. Na palpação, os nódulos benignos são diferentes dos cancerígenos. O médico pode verificar o tamanho, a forma, e a textura do nódulo, e sentir se este se move facilmente. Os gânglios que se apresentam macios, lisos, redondos e móveis são, provavelmente, benignos. Um nódulo duro, com forma estranha e irregular, que se sente bem preso (ou fixo) dentro da mama é, provavelmente, cancro.

Um exame clínico completo da mama, pode demorar cerca de 10 minutos a ser realizado.

Deverá fazer um exame físico. O médico irá fazer perguntas relacionadas com a história clínica e familiar. Pode precisar de fazer uma biópsia, para verificar se na zona suspeita existem células cancerosas.

Biópsia

Muitas vezes é necessário retirar tecido ou líquido, para ajudar o médico a perceber se é um cancro. Este procedimento é denominado por biópsia. Para efetuar uma biópsia, o médico pode recomendar uma consulta com um cirurgião ou um médico especializado em doenças da mama.

Por vezes, uma área suspeita, visível numa mamografia, não é palpável (sentida) no exame clínico da mama. O médico poderá usar aparelhos que fornecem imagens, para ver a área onde será retirado o tecido. Estes procedimentos incluem a biópsia guiada por ultrassons, com agulha localizada.

Os médicos podem retirar células ou tecido da mama, recorrendo a diferentes métodos:

- Aspiração com agulha fina: o médico usa uma agulha fina para remover líquido e/ou células de um nódulo na mama, que deverá ser analisado num laboratório, onde um patologista usa um microscópio para procurar células cancerosas.
- Biópsia "Core": o médico usa uma agulha para remover tecido mamário. Um patologista analisa o tecido, para ver se tem células cancerosas (microbiópsia);
- Biópsia cirúrgica: numa biópsia incisional, o cirurgião remove uma amostra de um nódulo ou de uma zona anormal. Se for uma biópsia excisional, o cirurgião remove completamente o nódulo ou a zona anormal. Posteriormente, um patologista analisa o tecido retirado, para ver se tem células cancerosas.

Se existirem células cancerosas, o patologista pode, caracterizar o tipo de cancro. O tipo de cancro da mama mais comum é o carcinoma ductal: tem início no interior dos ductos. Outro tipo de cancro da mama é o carcinoma lobular: tem início nos lóbulos.

Ultrassons

Os ultrassons podem ser usados em alta ou baixa intensidade. Quando usado em alta intensidade provocam modificações no meio atingido. Por outro lado, em baixa intensidade, é usado para obtenção de informações do meio no qual foi aplicado, como por exemplo, em diagnósticos mamários.

Transdutores de ultrassons -materiais piezoelétricos

Um equipamento de ultrassons é constituído por uma unidade de geração e transmissão de pulsos elétricos para excitação dos transdutores, uma unidade de receção e amplificação dos sinais captados, uma unidade de controlo e processamento que é utilizada para configurar os parâmetros das unidades de transmissão e receção, e uma unidade para visualização dos resultados do processamento.

Os equipamentos de ultrassons em geral utilizam uma frequência variada dependendo do tipo de transdutor, desde 2 até 14 Mhz, emitindo através de uma fonte de cristal piezoelétrico que fica em contacto com a pele e recebe os ecos gerados pelo mesmo.

O transdutor converte energia elétrica em vibração mecânica e vice-versa, e define a direção, frequência e geometria do feixe de som.

Os materiais piezoelétricos podem ser cristais de quartzo, que possuem uma elevada sensibilidade e reduzida largura de banda.

Quanto maior a frequência maior a resolução obtida.

O tipo de análise e de processamento efetuados nos sinais enviados e recebidos pelos transdutores dependem das informações requeridas pelo médico, que podem ser a apresentação de um órgão interno num monitor, a atenuação ou velocidade de propagação da onda ultrassónica num determinado órgão, a estimação de fluxo sanguíneo ou o deslocamento de determinadas estruturas em relação às suas posições normais.

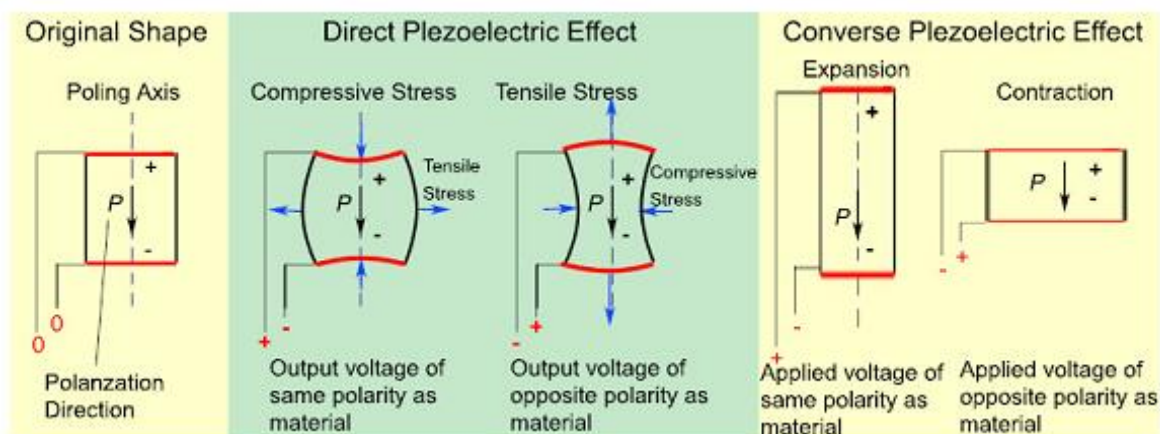


Figura 54 - Comportamento de um material piezoelétrico [39]

O transdutor emite ondas de ultrassom num corpo, estas são refletidas e captas pelo mesmo, transforma-as, por sua vez numa tensão, a qual pode ser lida e interpretada levando à obtenção de imagens. Neste tipo de exame não são usadas quaisquer formas de radiação ou de matérias contrastantes que possibilitem a formação de imagens.

Existem diversos tipos de transdutores, cada um adequado para um diferente tipo de exame. O transdutor convexo é o mais adequado para o exame de abdómen e pélvis, e o linear, para o exame da tiroide e mamas.



Figura 55 - Diferentes tipos de transdutores: Endoscópico, convexo, lineares, laparoscópicos...[40]

Na prática, os transdutores de menor frequência (de 3,5MHz) são utilizados para o exame de tecidos profundos, como, por exemplo, o abdómen (fígado, vesícula, baço, rins...), e a pélvis (bexiga, útero, ovários, próstata, obstétrico...). Os transdutores de frequência elevada (maiores que 7,5MHz) são utilizados para exame de tecidos superficiais, como a mama, tiroide, pele, testículo...

Quanto maior a frequência do transdutor, menor o comprimento da onda sonora e melhor a resolução espacial. Num meio homogêneo, o ultrassom propaga-se em linha reta, sendo que a velocidade de propagação depende da densidade desse meio. Num meio heterogêneo, as ondas sonoras são refletidas a cada densidade diferente, retornando ecos para o transdutor.

Quando o ultrassom passa de um meio para outro com densidade diferente, parte de sua energia é refletida e parte penetra no novo meio. Ambas as ondas ultrassônicas, refletida e transmitida, sofrem desvio de propagação.

O emissor de ultrassons ou transdutor é constituído por uma cerâmica piezoelétrica que responde a estímulos elétricos, emitindo uma onda curta que se propaga nos tecidos e é refletida nas diversas interfaces. O mesmo emissor regista o eco de retorno, de modo que é também um recetor. Apenas os ecos provenientes de interfaces perpendiculares às ondas sonoras são recebidos. Entre a emissão do ultrassom e a receção do seu eco correspondente há um intervalo de tempo que aumenta com a distância entre o transdutor e a interface refletora. Assim sendo, é possível estabelecer a profundidade desta interface de acordo com este intervalo.

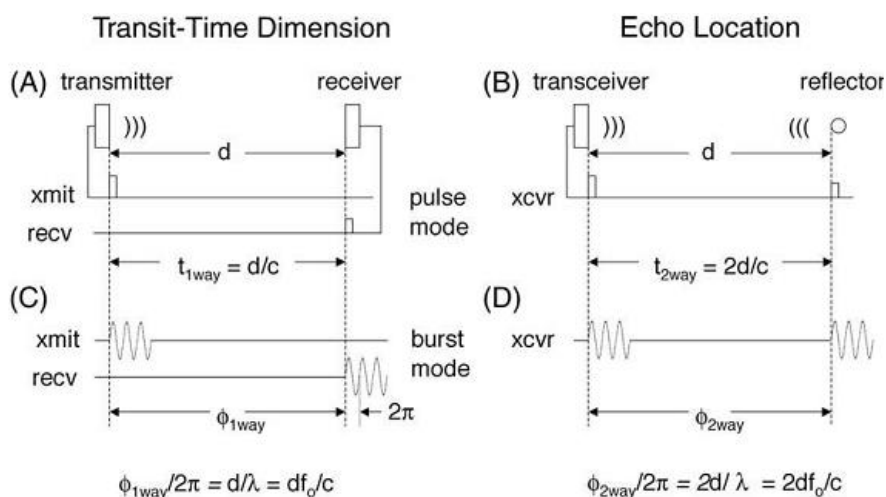


Figura 56 - Medição de distâncias por intermédio do princípio do tempo de trânsito [41]

Os transdutores de ultrassons são construídos a partir do material piezoelétrico e podem ter um único elemento ou ser multielementares. Os elementos básicos deste tipo de transdutor são: cerâmica piezoelétrica (transdutor), camada de retaguarda, camada frontal elementos de casamento de impedância elétrica, cabos e material de encapsulamento do transdutor.

Dependendo da aplicação, o elemento piezoelétrico é que determina a frequência de operação do transdutor. Em geral, os transdutores são acondicionados num suporte plástico para lhes dar proteção mecânica e elétrica. Na superfície por onde emergem as ondas ultrassônicas há uma camada especial para permitir o perfeito acoplamento acústico e também para dar proteção aos elementos piezoelétricos. Os transdutores são hermeticamente fechados e têm externamente o dispositivo para a conexão elétrica com o

módulo eletrónico do equipamento. As ondas ultrassónicas podem sofrer reflexão, refração ou difração quando o feixe encontra uma interface entre meios com características acústicas diferentes. Baseados nesse comportamento, dois métodos experimentais são geralmente utilizados para caracterização de meios biológicos e não biológicos por ultrassom: o método transmissão-recepção e o método pulso-eco (reflexão). A combinação desses dois métodos permite obter informações sobre distância entre objetos, velocidade de propagação e atenuação de ondas ultrassónicas nos meios.

As estruturas de densidade comparada à da água permitem propagar facilmente os ecos emitidos pelo ultrassom, ao contrário do que ocorre com estruturas ósseas ou gasosas, quando a variação de impedância acústica entre água e osso ou água e ar é tamanha que há reflexão total das ondas sonoras. Por isso o estudo ultrassonográfico de áreas com estas características (cérebro, pulmões) é muito pobre. É esta diferença de impedância acústica que obriga o contacto apropriado entre o transdutor e a pele do paciente.

O ultrassom, em geral, propaga-se através de líquidos, tecidos e sólidos.

Velocidade de propagação do som em diversos meios.
Speed of sound in different mediums

Meio / Medium	Velocidade / Speed (m/s)
Tecidos moles / Soft tissues	1540
Músculo / Muscle	1580
Água / Water	1497
Cérebro / Brain	1521
Osso / Bone	3500
Gordura / Fat	1460
Sangue / Blood	1570
Rim / Kidney	1561
Fígado / Liver	1570
Ar / Air	331

Figura 57 - Os princípios físicos da imagem e os artefactos na ecografia em modo B. [42]

A impedância acústica está relacionada com a resistência ou a dificuldade do meio à passagem do som, é o produto da densidade do material pela velocidade do som no mesmo. Quando o feixe sonoro atravessa uma interface entre dois meios com a mesma impedância acústica, não há reflexão e a onda é toda transmitida ao segundo meio. A diferença de impedância acústica entre dois tecidos define a quantidade de reflexão na interface, promovendo a identificação na imagem. Um nódulo no fígado por exemplo, será mais facilmente identificado se a impedância acústica for bastante diferente do parênquima hepático ao redor; ao contrário, quanto mais próxima a impedância acústica do parênquima hepático normal, mais dificuldade haverá em identificá-lo, porque pouca reflexão sonora irá

ocorrer. Ou seja, quanto maior for a diferença de impedância entre duas estruturas, maior será a intensidade de reflexão.

Modos de visualização da informação obtida através dos ecos de ultrassons

A técnica de ultrassons para diagnóstico compreende basicamente três modos: o Modo A (de Amplitude), Modo B (de Brilho) e o Modo M (de Movimento).

Os princípios básicos do Modo A (Amplitude) referem-se à simplicidade de operação de um sistema de ultrassons. Pulsos de ultrassons de curta duração são enviados por um único transdutor que também funciona como recetor dos ecos refletidos nas interfaces. A amplitude do eco retro espalhado é registada em função do tempo. Neste método de apresentação de informações, a distância alvo-transdutor é representada no eixo horizontal e a amplitude do eco, no eixo vertical na tela de um monitor de vídeo.

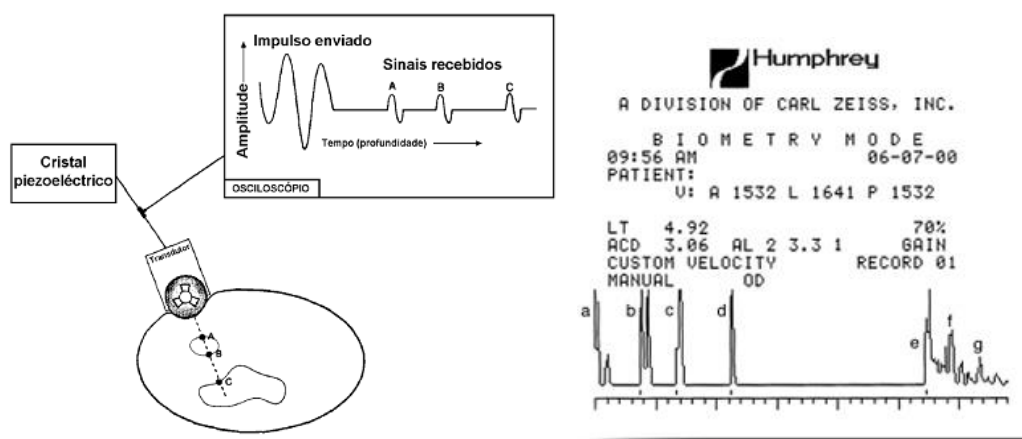


Figura 58 a) e b) - Sistema de ultrassons e leitura para operação no Modo A. [42] [43]

O Modo A foi bastante usado para detetar tumores cerebrais, localizando a linha mediana do cérebro, que normalmente separa os ventrículos simetricamente. O deslocamento dessa linha pode indicar presença de um tumor. O procedimento é chamado ecoencefalografia. Com o Modo A podem-se também distinguir diferentes tipos de tecido; por exemplo, o tecido mole produz ecos de baixa amplitude, enquanto o líquido de um quisto não produz eco.

Os Princípios básicos do Modo B (Brilho) baseiam-se num sistema que é mais usado para obtenção de imagens do interior do corpo humano ou de animais. Sistemas mais sofisticados permitem o uso de vários transdutores de multifrequência, e inúmeros recursos de pré-processamento na geração da imagem e de pós processamento da imagem congelada no monitor.

O princípio de funcionamento baseia-se na obtenção dos ecos como no Modo A, mas o sinal é tratado de maneira diferente. O Modo B pode ser estático ou dinâmico. No Modo B, os ecos

são representados por pontos brilhantes na tela. A intensidade do brilho (geralmente apresentada numa escala de cinza) corresponde à amplitude dos ecos do Modo A. O Modo B produz uma imagem bidimensional do meio sob estudo pela combinação dos sinais do Modo A em várias direções.

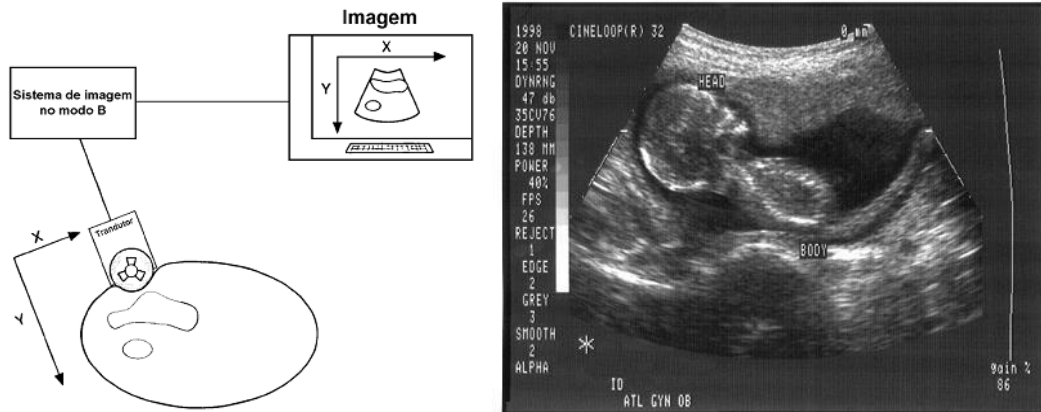


Figura 59 a) e b) - diagrama simplificado de um sistema de ultrassonografia típico [42] e ultrassom fetal Modo-B. [43]

O Modo B é amplamente utilizado em obstetrícia, onde a taxa de crescimento fetal, orientação de estruturas, e anormalidades podem ser mapeadas sem o risco da exposição do feto à radiação ionizante. Outras aplicações importantes: exames abdominais para localização de tumores e análise da imagem de órgãos internos; em ginecologia, para localização de quistos e tumores no ovário e mamas, no músculo-esquelético para observação dos membros inferiores e superiores; e em cardiologia, na visualização das paredes e válvulas cardíacas.

O Modo-M (Movimento) é utilizado para analisar qualitativa e quantitativamente o movimento de estruturas como válvulas cardíacas. Possui algumas características do Modo A e algumas do B. Como no Modo B, o brilho da linha mostrada é modulado de acordo com a amplitude do sinal recebido e é similar ao Modo A porque os ecos são coletados em apenas uma direção e apresentados na direção horizontal do monitor.

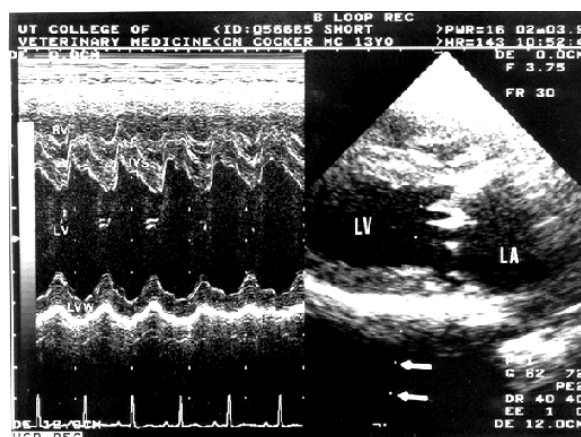


Figura 603 - Ultrassom Modo-M. [44]

Alcance e limitação dos Ultrassons

A ultrassonografia mamária diagnóstica e intervencionista tem grande participação como exame complementar à mamografia e à clínica, tornando-se um método valioso e bem estabelecido no diagnóstico das doenças mamárias

É um exame que permite obter informações da textura e conteúdo de nódulos mamários. Não é utilizado como método de rastreio do cancro da mama porque não tem condições para detetar microcalcificações que, muitas vezes, representam a única forma de expressão da doença. Outra limitação importante é a identificação de tumores menores que 1 cm, quando estes estão localizados profundamente em mamas volumosas e com grande quantidade de tecido adiposo, a identificação torna-se precária, porque pequenos tumores e porções de gordura geram imagens semelhantes.

Portanto, na maioria dos casos é usado como método complementar da mamografia.

A ultrassonografia tem grande aplicabilidade na diferenciação entre tumores císticos e sólidos e também é capaz de identificar lesões no interior de um quisto, indicando a remoção através de cirurgia.

Este exame tem melhores resultados quando feito em mamas densas, com tecido glandular exuberante, como as mamas das mulheres jovens.

Este método apresenta as vantagens de ser um procedimento não-invasivo, bem tolerado pelos pacientes e que pode trazer informações importantes, complementares ao exame físico e à mamografia. Amplamente usado como auxiliar no diagnóstico médico dentro da obstetrícia e ginecologia, oftalmologia, neurologia e cardiologia além de sua utilização como ferramenta comum em procedimentos terapêuticos.

Bibliografia adicional

- BYRNE, C.; SCHAIRER, C.; WOLFE, J.; PAREKH, N.; SALANE, M.; BRINTON, L. A.; HOOVER, R.; HAILE, R. Mammographic features and breast cancer risk: effects with time, age, and menopause status. *Journal of the National Cancer Institute*, 1995
- WOLFE, J.N. Breast patterns as an index of risk for developing breast cancer. *Am J Roentgenol*, 1976.
- BOYD, N. F.; LOCKWOOD, G. A.; BYNG, J. W.; TRITCHLER, D. L.; YAFFE, M. J. Mammographic densities and breast cancer risk. *Cancer Epidemiol Biomarkers*,
- VAN, G. C.H. Mammographic density and breast cancer risk. *Eur J Obstet Gynecol Reprod Biol*, , 1999.
- BRIGHT, R. A.; MORRISSON, A. S.; BRISSON, J.; BURSTEIN, N. A.; SADOWSKY, N. S.; KOPANS, D. B.; MEYER, J. E. Relationship between mammographic and histologic features of breast tissue in women with benign biopsies, Pubmed.
- L. W. Basset and R. H. Gold. Breast Cancer Detection: Mammograms and Other Methods in Breast Imaging. Grune & Stratton, New York, 1987.
- E. A. Sickles, W. N. Weber, and H. B. Galvin. Baseline screening mam-mography: one vs two views per breast. *American Journal of Roentgenology* 1986.
- W. T. Ho and P. W. T. Lam. Clinical performance of computer-assisted de-tection (CAD) system in detecting carcinoma in breasts of different densities. *Clinical Radiology* , 2003
- American College of Radiology. Illustrated Breast Imaging Reporting and Data
- System BIRADS. American College of Radiology, 3rd edition, 1998.
- S. E. Selvan, C. C. Xavier, N. Karssemeijer, J. Sequeira, R. A. Cherian, and B. Y. Dhala. Parameter estimation in stochastic mammogram model by heuristic optimization techniques. *IEEE Transactions on Information Technology in Biomedicine*, 10:685-695, 2006
- R. Sivaramakrishna, N. A. Obuchowski, W. A. Chilcote, and K. A. Powell. Automatic segmentation of mammographic density. *Academic Radiology*, 2001
- S. van Engeland, P. R. Snoeren, H. Huisman, C. Boetes, and N. Karssemeijer. Volumetric breast density estimation from full-field digital mammograms. *IEEE Transactions on Medical Imaging*, 2006.
- P. Miller and S. M. Astley. Classification of breast tissue by texture analysis. *Image and Vision Computing*, 1992.
- J. W. Byng, N. F. Boyd, E. Fishell, R. A. Jong, and M. J. Yaffe. Automated analysis of mammographic densities. *Physics in Medicine and Biology*, 1996.
- K. Bovis and S. Singh. Classification of mammographic breast density using a combined classifier paradigm. In *Proc. Medical Image Understanding and Analysis*, 2002
- D. Raba, J. Martí, R. Martí, and M. Peracaula. Breast mammography asym-metry estimation based on fractal and texture analysis. In *Proc. of the Com-puted Aided Radiology and Surgery*, 2005

- FRERY, A.C.; SANT'ANNA, S.J.S. Redução de ruído em imagens SAR pelo uso de filtros robustos, Simpósio Brasileiro de Sensoriamento Remoto, 1993.
- ROBINSON, E.A.; TREITEL, S. Geophysical signal analysis. Prentice-Hall, New Jersey, 1980.
- LEE, J. S. Speckle analysis and smoothing of synthetic aperture radar images. Computer Graphics and Image Processing, 1981.
- FROST, V.S.; STILES, J. A.; SHANMUGAN, K. S.; HOLTZMAN J.C. A model for radar images and its applications to adaptive digital filtering of multiplicative noise. IEE Trans. Pattern Analysis machine Intelligence, 1982.
- AMARA, G. (1995) 'An Introduction to Wavelet', IEEE Computational Science and Engineering, Summer 1995
- MINISTÉRIO DA CIÊNCIA E TECNOLOGIA.