

**Universidade de São Paulo
Escola de Engenharia de São Carlos
Departamento de Engenharia Elétrica**



**ESTUDO DA CARACTERIZAÇÃO DE
NÓDULOS EM MAMOGRAMAS ATRAVÉS DE
UMA CONFIGURAÇÃO DE REDE NEURAL
ARTIFICIAL**

Sérgio Koodi Kinoshita

Dissertação apresentada ao
Departamento de Engenharia
Elétrica, Escola de Engenharia de
São Carlos, da Universidade de
São Paulo, para obtenção do título
de mestre em Engenharia Elétrica.

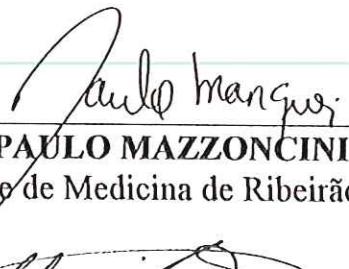
Orientador: **Dr. Paulo Mazzoncini de Azevedo Marques**

**São Carlos
1998**

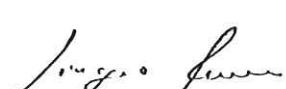
FOLHA DE APROVAÇÃO

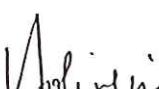
Candidato: Engenheiro **SÉRGIO KOODI KINOSHITA**

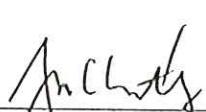
Dissertação defendida e aprovada em 27.10.98
pela Comissão Julgadora:


Prof. Dr. **PAULO MAZZONCINI DE AZEVEDO MARQUES** (Orientador)
(Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo)


Prof. Dr. **CLÓVIS SIMÃO TRAD**
(Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo)


Doutor **SÉRGIO SHIGUEMI FURUIE**
(InCor - Instituto do Coração)


Prof. Doutor **DENIS VINICIUS COURY**
Coordenador da Área de Engenharia Elétrica


JOSÉ CARLOS A. CINTRA
Presidente da Comissão de Pós-Graduação da EESC

*“ O que vale na vida não é o ponto de
partida e sim a caminhada.
Caminhando e semeando, no fim terás o
que colher.”*

Cora Coralina

Agradecimentos

À minha amada esposa, Angela, companheira e incentivadora deste trabalho.

Aos meus filhos, Daniele, Tiago e Cássia pelo amor singelo e sincero que só um coraçõozinho infantil pode proporcionar.

Aos meus pais que se foram deste mundo, mas deixaram em mim a vontade de viver e sorrir.

Aos meus queridos irmãos, cunhadas e cunhados que sempre me incentivaram e hoje elevo ainda mais o meu apreço.

À professora Annie que desde a graduação me ensinou a respeitá-la pela sua sabedoria e generosidade.

Ao professor Paulo, companheiro e amigo, que aceitou o desafio deste trabalho mesmo sabendo das dificuldades de tempo e espaço.

Aos amigos Ricardo (Jacaré), Ricardinho e Roberto (Chico Bento) pela amizade e companheirismo.

Aos colegas do laboratório, Henrique, Irita, Márcio, Márcia, Renatinha, Silvia e Lucas pelos bons momentos e acolhida dispensada.

A coordenação de pós-graduação por permitir a minha presença como aluno do programa de mestrado.

Aos professores e funcionários do Departamento de Engenharia Elétrica pelo apoio e presteza de seus serviços, em especial a Rosane pela competência e boa vontade.

A CAPES e à pró-reitoria de pesquisa e pós-graduação da Universidade de Marília pelo apoio financeiro.

Ao professor Sidnei Mariano pela sua equanimidade e amizade.

Aos professores do Departamento de Ginecologia e Obstetrícia da Faculdade de Medicina de Ribeirão Preto, da Universidade de São Paulo, em especial ao Dr. Heitor R. C. Marana e ao Dr. Luciano A. R. Spina.

E a todos que de forma direta ou indireta participaram deste trabalho.

Resumo

Este trabalho apresenta um estudo de classificação de nódulos em mamogramas digitalizados através de um classificador de rede neural artificial (RNA). O algoritmo de treinamento de “backpropagation” foi utilizado para ajustar os pesos da RNA. O objetivo principal deste trabalho foi determinar uma metodologia para analisar e selecionar a melhor configuração de atributos e topologia da RNA para classificar lesões mamárias do tipo nódulo. Foram escolhidas 118 imagens de regiões de interesse (ROI), sendo 68 benignas e 50 malignas de duas bases de imagens: uma do Hospital das Clínicas de Ribeirão Preto, da Universidade de São Paulo, e outra do MIAS-UK (Mammographic Image Analysis Society). O processo completo envolveu quatro etapas: detecção, extração e seleção de atributos, e classificação. Na etapa de detecção, as imagens foram submetidas ao processo combinado das técnicas segmentação de thresholding, morfologia matemática e crescimento de região. Foram extraídos 14 atributos de textura e 14 atributos de forma. Para selecionar os atributos mais discriminantes, foi utilizado o método de Jeffries-Matusita. Foram selecionados três grupos de atributos de forma, três de atributos de textura e três de atributos combinados. Análise pela curva ROC foram dirigidas para avaliar o desempenho do classificador de rede neural artificial (RNA). Os melhores resultados obtidos foram: para o grupo de atributos de forma com 5 unidades escondidas, a área dentro da curva ROC foi de 0.99, taxa de acerto de 98,21%, taxa de especificidade de 98,37% e taxa de sensibilidade de 98,00%; para o grupo de atributos de textura com 4 unidades escondidas, a área dentro da curva foi de 0.98, taxa de acerto de 97,08%, taxa de especificidade de 98,53% e taxa de sensibilidade de 95,11%; para o grupo de atributos combinados de textura e forma com 3 unidades escondidas, a área dentro da curva foi de 0.99, taxa de acerto de 98,21%, taxa de especificidade de 100,00% e taxa de sensibilidade de 95,78%.

Abstract

This work presents a study of masses classification in digitized mammograms by means of artificial neural network (ANN). The backpropagation training algorithm was used to adjust the weights of ANN. The aim of this work was to determine a methodology to analyze and selection of the best feature subset and ANN topology to classify masses lesions. A total of 118 regions of interest images were chosen (68 benign and 50 malignant lesions) from two image databases: one from "Hospital das Clínicas de Ribeirão Preto", at the University of São Paulo, and other from Mammographic Image Analysis Society (MIAS-UK). The whole process involved four steps: segmentation, feature extraction, selection, and classification. In the first step, the images were submitted to a combined process of thresholding, mathematical morphology, and region growing techniques. In the second step, fourteen texture features and fourteen shape features were extracted. The Jeffries-Matusita method was used to select the best features. The results of this stage were the selection of three shape feature sets, three texture feature sets, and three combined feature sets. The Receiver Operating Characteristic (ROC) analysis were conducted to evaluated the ANN classifier performance. The best result obtained for shape feature set was obtained using a ANN with 5 hidden units, the area under ROC curve was of 0.99, classification rate of 98.21%, specificity rate of 98.37% and sensitivity rate of 98.00%. For texture feature set, the best result was using a ANN with 4 hidden units, the area under ROC curve was of 0.98, classification rate of 97.08%, specificity rate of 98.53% and sensitivity rate of 95.11%. Finally; for the combined feature set (texture and shape) the best result obtained was using a ANN with 3 hidden units, the area under ROC curve was of 0.99, classification rate of 98.21%, specificity rate of 100.00% and sensitivity rate of 95.78%

índice

Capítulo 1

Introdução	1
------------------	---

Capítulo 2

Características de Nódulos em Mamogramas	4
--	---

Capítulo 3

Digitalização de Imagens Mamográficas	6
---	---

Capítulo 4

Detecção de Nódulos em Mamogramas Digitalizados	8
4.1 Método de Realce	8
4.1.1 Equalização de Histograma	10
4.1.2 Modificação do Contraste	12
4.2 Método de Limiarização	13
4.3 Método de Filtros Morfológicos	14
4.4 Método de Crescimento de Região	15

Capítulo 5

Extração e Seleção de Atributos de Textura e Forma em Nódulos em Mamograma	17
5.1 Atributos de Forma	17
5.1.1 Medida de Circularidade	18
5.1.2 Grau de Comapactação – Solidez	18
5.1.3 Descritores de Gupta	19
5.1.4 Descritores de Shen	21
5.1.5 Descritores de Hu	22
5.1.6 Descriptor de Fourier	23
5.1.7 Descritores de Wee	24
5.2 Atributos de Textura	24
5.2.1 Descritores de Haralick	25
5.3 Método de Seleção de Jeffries-Matusita	27

Capítulo 6

Classificação de Nódulos em Mamogramas	29
6.1 Técnicas Estatísticas	30
6.2 Redes Neurais Artificiais	31
6.2.1 Arquitetura de Redes Neurais Artificiais	32
6.2.2 Redes de Multi-Camadas com Algoritmo de Treinamento de “Backpropagation”.....	34

Capítulo 7

Estado da Arte de Classificação de Nódulos em Mamogramas	38
7.1 Classificação com Atributos de Textura	38
7.2 Classificação com Atributos de Forma	41
7.3 Classificação com Atributos de Textura e Forma	49
7.4 Classificação com Outros Tipos de Atributos	53

Capítulo 8

Materiais e Métodos	56
8.1 Base de Imagens e Aquisição das Imagens	57
8.2 Etapa de Detecção e Segmentação	59
8.3 Etapa de Extração e Seleção de Atributos	60
8.3.1 Atributos de Textura	60
8.3.2 Atributos de Forma	61
8.3.3 Seleção de Atributos	61
8.4 Etapa de Classificação através de Redes Neurais Artificiais	62
8.5 Etapa de Avaliação de Resultados	65

Capítulo 9

Resultados e Discussão	67
9.1 Detecção e Segmentação	67
9.2 Extração e Seleção de Atributos	68
9.3 Classificação através de Redes Neurais Artificiais	69
9.4 Avaliação de Resultados	70
9.4.1 Comparação com Outros Resultados	79

Capítulo 10

Conclusões	81
10.1 Conclusões Finais	81
10.2 Sugestões para Trabalhos Futuros	82

Capítulo 11

Anexos	83
---------------------	----

Capítulo 12

Referências Bibliográficas	106
---	-----

Listas de Figuras

Capítulo 2

- Figura 2.1 -** [Pinotti-91] Imagem de tumor benigno com contorno regular bem definido..... 5
- Figura 2.2 -** [Pinotti-91] Carcinoma mamário com bordas irregulares e espículas..... 5

Capítulo 4

- Figura 4.1 -** Histograma de uma imagem com separação em definida entre duas classes..... 13

Capítulo 6

- Figura 6.1 -** [Fausset-94] ilustra uma rede de camada única.; uma rede de multi-camadas; uma rede com camada competitiva 32

Capítulo 8

- Figura 8.1 -** Diagrama de blocos da metodologia utilizada para determinação da melhor configuração de um classificador 54
- Figura 8.2 -** Sequência de imagens em que foi aplicado os algoritmos de realce. 59
- Figura 8.3 -** Sequência de imagens em que não foi aplicado o algoritmo de realce..... 59

Capítulo 9

- Figura 9.1 -** Imagem de uma ROI em que o processo de segmentação da lesão não foi bem sucedida..... 68
- Figura 9.2 -** Apresentação das curvas ROC e das áreas A_z dentro das curvas em relação aos desempenhos obtidos pela rede neural artificial (RNA), de acordo com o número de unidades escondidas (u.e.). Na entrada da RNA foram aplicados os conjuntos de dados de atributos de textura selecionados pela distância de Jeffries-Matusita: $C1_{TF}$, $C2_{TF}$,e $C3_{TF}$; 70
- Figura 9.3 -** Apresentação das curvas ROC e das áreas A_z dentro das curvas em relação aos desempenhos obtidos pela rede neural artificial (RNA), de acordo com o número de unidades escondidas (u.e.). Na entrada da RNA foram aplicados os conjuntos de dados de atributos de textura selecionados pela distância de Jeffries-Matusita: $C1_T$, $C2_T$, e $C3_T$; 71
- Figura 9.4 -** Apresentação das curvas ROC e das áreas A_z dentro das curvas em relação aos desempenhos obtidos pela rede neural artificial (RNA), de acordo com o número de unidades escondidas (u.e.). Na entrada da RNA foram aplicados os conjuntos de dados de atributos de textura selecionados pela distância de Jeffries-Matusita: $C1_F$, ($C2_F$, e $C3_F$; 71
- Figura 9.5 –** Histogramas de distribuição de resultados na saída das redes neurais artificiais (RNA) de melhores desempenhos na classificação de casos benignos e malignos. Eles representam os conjuntos de atributos de: (a) textura e forma $C3_{TF}$ da rede com 3 unidades escondidas; (b) textura $C2_T$ com 4 unidades escondidas; (c) forma $C2_F$ com 5 unidades escondidas.....74,75

Lista de Tabelas

Capítulo 8

Tabela 8.1 -	Relação dos 28 atributos de Textura e Forma utilizados.....	63
Tabela 8.2 -	Relação dos conjuntos de atributos selecionados pela Distância de Jeffries-Matusita.....	64

Capítulo 9

Tabela 9.1 -	Cálculo da Distância de Jeffries-Matusita para os 28 atributos de textura e forma; 14 atributos de textura, e 14 atributos de forma.	69
Tabela 9.2 -	Resultados da Performance de Classificação da Rede Neural Artificial, de acordo com a Área na Curva ROC, Porcentagem de Acerto, Especificidade e Sensibilidade.....	73
Tabela 9.3 -	Melhores configurações de acordo com os resultados de performance de classificação da rede neural artificial, considerando a Área dentro da curva ROC, taxas médias de Acerto, Especificidade e Sensibilidade.....	74
Tabela 9.4 -	Valores de limiares fornecidos para auxiliar o radiologista na indicação de biópsia.....	76
Tabela 9.5 -	Resultados de classificação utilizando distância de Mahalanobis, base de imagens “The Mammographic Image Analysis Society” (MIAS), combinada com outra base de imagem do “Alberta Program for the Early Detection of Breast Cancer” APEDBC).....	80

Capítulo 11

Tabela 11.1 - Relação das regiões de interesse (ROI) obtidas das imagens mamográficas da base do Hospital de Clínicas de Ribeirão Preto (HCRP) e do “The Mammographic Images Analysis Society database” (MIAS).....	84
Tabela 11.2 - Resultados normalizados dos atributos obtidos da extração de atributos de textura das imagens das regiões de interesse originais.....	85
Tabela 11.3 - Resultados normalizados dos atributos obtidos da extração de atributos de forma das imagens das regiões de interesse originais.....	89
Tabela 11.4 - Um exemplo dos resultados obtidos na saída da RNA com o algoritmo de treinamento de “Backpropagation”, com aproximação de Levenberg-Marquardt, 3, 4, 5 e 6 unidades escondidas, treinamento e teste com o método “leaving-one-out” e entrada com o conjunto de atributos de Textura e Forma $C1_{TF}$	93
Tabela 11.5 - Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C1_{TF}$	96
Tabela 11.6 - Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C2_{TF}$	97
Tabela 11.7 - Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C3_{TF}$	98
Tabela 11.8 - Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C1_T$	99
Tabela 11.9 - Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C2_T$	100
Tabela 11.10- Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C3_T$	101

Tabela 11.11-Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C1_F$ 102

Tabela 11.12-Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C2_F$ 103

Tabela 11.13-Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C3_F$ 104

Capítulo 1

Introdução

Câncer de mama é a segunda causa de mortalidade feminina no mundo [Boring-92]. De acordo com o relatórios da American Cancer Society, nos países ocidentais, o câncer de mama é a segunda causa de mortalidade na população feminina. No Brasil, um estudo comparativo da Fundação Oncocentro entre os anos 70 e 90 mostra um aumento significativo: de 12,9 mulheres com a doença em grupos de 100 mil, a incidência passou para 15,3 por 100 mil. Estima-se que atualmente no Brasil o câncer de mama seja responsável pela morte de 10 mil mulheres por ano. Existe uma tendência do aumento da incidência de câncer de mama em mulheres mais jovens. Menarca precoce, menstruação tardia, nuliparidade, primiparidade tardia, dieta gordurosa e utilização de hormônio contraceptivo oral são alguns dos fatores de risco cogitados para esse aumento. A resistência ao auto-exame é um fator que contribui bastante para os índices de mortalidade.

Os Estados Unidos têm uma das maiores incidências mundiais desse tipo de câncer. Entre 1982 a 1987, a taxa de incidência de câncer nas mulheres aumentou em torno de 4% ao ano. Em 1992, a estimativa da American Cancer Society foi de 180.000 novos casos entre as mulheres e, em 1996, esta estimativa aumentou para 184.300. Deste número, de acordo com a American Câncer Society, a estimativa de mortalidade foi de 44.300 de mulheres. Dados recentes apontam que o índice de mortalidade continua a decair em mulheres brancas, e pela primeira vez, está decaindo entre as mulheres jovens negras americanas, porém os índices permanecem em patamares altos. Este pequeno decréscimo deve-se a detecção precoce e tratamento preventivo.

No Japão, a taxa de mulheres afetadas, na proporção de 4/100 mil até a década de 70 , cresceu 50% na última década. Na Inglaterra, o câncer mamário é também a causa de maior mortalidade feminina [Davies-90].

A influência da herança genética está presente entre 5% a 10% dos casos. Mulheres com história de câncer de mama na família são candidatas em potencial para o desenvolvimento da doença.

A idade média do surgimento do câncer de mama nas mulheres é de 50 anos. Quanto antes o câncer de mama for detectado, maiores são as chances de sobrevivência e cura. Quando o diagnóstico é feito em fase inicial (Estágio I e II) a chance de sobrevivência fica, em média, em torno de 85%. No caso específico das mulheres, se o diagnóstico ocorrer numa fase inicial, a extração da mama (mastectomia radical) pode ser evitada através de outros métodos cirúrgicos, que permite extrair apenas a porção da glândula mamária diretamente afetada.

O diagnóstico precoce só é possível, porém, através de procedimentos preventivos, os quais podem ser basicamente de dois tipos: mamografia e auto-exame. A detecção precoce por mamografia é considerada o método mais eficaz, permitindo a identificação do tumor quando ainda não é palpável.

A mamografia permite ainda obter a detecção de microcalcificações que mostram aspectos suspeitos de malignidade, quando agrupadas e de tamanhos e formatos variados. Porém, mesmo sendo a mamografia a melhor forma de detecção precoce de câncer mamário, entre 10% a 30% das mulheres que desenvolveram câncer de mama e passaram por mamografia tiveram resultados negativos nos mamogramas [Andersson-84] [Baines-86] [Basset-87] [Martin-79] [Moskowitz-83]. Por outro lado, aproximadamente 75% dos casos diagnosticados como suspeitos, e enviados para a biópsia, apresentaram diagnóstico benigno. Essa alta taxa de falso-positivo é atribuída a diversas dificuldades presentes na avaliação e formulação de um diagnóstico mais preciso [Andersson-84][Buchanan-83][Holland-82] [Martin-79].

Diversos trabalhos foram propostos, buscando essencialmente aprimorar a visualização de imagens radiográficas por meios computacionais, de forma a auxiliar o radiologista a elaborar um diagnóstico mais preciso. Alguns pesquisadores desenvolveram sistemas automáticos de detecção de nódulos em mamogramas digitalizados, diferenciando-os de tecidos normais presentes nas mamografias

[Winsberg-67] [Giger-90A] [Giger-90B] [Yin-93] [Yin-94A] [Petrosian-94] [Wei-95] [Sahiner-96A] [Petrick-96A] [Petrick-96B] [Wei-97] [Sahiner-98]. Outros pesquisadores desenvolveram trabalhos que procuraram, ainda, melhorar esta detecção, classificando as lesões em normais, espiculadas, ou circunscritas, sem porém caracterizá-las em malignas ou benignas [Kimme-75][Hand-79][Lai-89] [Li-93][Kegelmeyer-94][Li-95][Karssemeiger-96]. Muitos pesquisadores, têm dedicado atenção ao desenvolvimento de sistemas automáticos de caracterização de lesões em malignas ou benignas, auxiliando os radiologistas de forma mais completa no diagnóstico das lesões mamárias [Ackermann-72][Ackermann-73][Gale-87][Cook-89][Giger-90A] [Giger-90B] [Yin-91B] [Kilday-93] [Wu-93] [Huo-95A] [Huo-95B] [Velanovich-95] [Huo-96] [Highnam-96] [Pohlman-96] [Rangayyan-96] [El-Faramawy-96].

De modo geral, esses trabalhos têm o objetivo principal de auxiliar os radiologistas na detecção e caracterização das lesões mamárias, procurando reduzir a quantidade de falsos-positivo no diagnóstico da malignidade dos nódulos mamários, diminuindo assim a morbidade e também o número de biópsias desnecessárias.

Dentro desta perspectiva, desenvolvemos uma metodologia para determinar as melhores configurações do classificador para caracterizar nódulos mamários em malignos e benignos. O classificador utilizado foi a rede neural artificial com arquitetura do tipo multi-camadas e treinada com o algoritmo de “backpropagation”.

Utilizamos 118 imagens de regiões de interesse selecionadas de duas bases de imagens: uma do Hospital das Clínicas de Ribeirão Preto, da Universidade de São Paulo; e outra do MIAS (Mammographic Image Analysis Society) Extraímos 14 atributos de textura e 14 forma das imagens. Eles foram submetidos a um processo de seleção para escolher os 10 atributos mais discriminantes e, também, para eliminar os atributos redundantes.

Utilizamos o método “leaving one-out” para realizar o treinamento e teste do classificador, pois esse método fornece condições de avaliação da generalização do classificador. Para avaliar os resultados obtidos, calculamos a área dentro da curva ROC, as taxas de acerto, especificidade e sensibilidade.

No final, apresentamos as conclusões finais sobre os resultados obtidos e propomos alguns trabalhos futuros para aperfeiçoar algumas etapas de nosso trabalho.

Capítulo 2

Características de Nódulos em Mamogramas

A mamografia é parte importante dos métodos diagnósticos de doenças da mama. Tornou-se um procedimento diagnóstico padrão para clínicos, ginecologistas e muitos cirurgiões que lidam com a patologia mamária [Haagensen-89] [Pinotti-91] [Faria-94]. A mamografia é realizada em aparelhos dedicados, chamados de mamógrafos. Neste tipo de aparelho, o tubo gerador de raios-X pode ter o alvo constituído de tungstênio, molibdênio e rodio, com ânodo fixo ou giratório e janela de berílio. Possui dispositivos mecânicos de compressão e tração que possibilitam uma espessura uniforme da mama, sem necessidade de movimentar a paciente.

Normalmente, um par de incidências constituído pelas vistas crânio-caudal (frente) e médio-lateral obliqua (perfil), permite a correta interpretação e a localização topográfica da lesão. São radiografadas, sempre que possível, as mamas esquerda e direita para exames de comparações de similaridades geométricas e estruturais dos tecidos glandulares. Podem ser feitas incidências complementares como as de prolongamento axilar e as de cavidade axilar quando existe lesão prévia neste local.

O objetivo principal da mamografia é a detecção de lesões não palpáveis em estado pré-clínico (nódulos de tamanho menor que 2 cm e agrupamentos de microcalcificações). Os sinais mamográficos são classificados em dois tipos: diretos e indiretos [Faria-94]. Os sinais diretos são o nódulo, algumas formas de calcificações agrupadas, opacidades assimétricas acompanhadas de distorções da arquitetura, e o nódulo emergente. Os sinais indiretos são a dilatação monoductal, contração ou retração tecidual, vascularização aumentada regionalmente, padrão

ductal proeminente setorial, espessamento da pele e outros. Os sinais indiretos obviamente são menos fiéis e, quando utilizados isoladamente, são de baixa especificidade [Faria-94].

O nódulo representa a forma mais frequente do câncer da mama. Quando apresenta contornos irregulares, espiculados, serrilhadas, cheias de reentrâncias, contorno regular microlobulado, ou contorno regular maldefinido, é diagnosticado como altamente suspeito. A presença de microcalcificações pleomórficas, de densidade variável, reforça ainda mais a suspeita de malignidade. Quando o nódulo apresenta contorno regular bem definido, densidade assimétrica e sem distúrbio do parênquima, calcificações em casca de ovo com cisto oleoso, presença de hilo radiolucente parcial ou total, é considerada uma lesão com característica benigna [Haagensen-89] [Pinotti-91] [Faria-94]. A figura 2.1 e figura 2.2, apresentam exemplos de nódulos com características típicas de benignidade e malignidade, respectivamente.

Microcalcificações de forma variável (pleomórficas), mais ou menos densas, em elevado número, agrupadas, arborescentes, calcificações em forma das últimas letras do alfabeto (X, Y, Z), pontiformes, em forma de bastonetes e amorfas envolvem suspeita de malignidade [Faria-94].

A denominação “Nódulo Emergente”, consiste nos casos em que sinais suspeitos de nódulos inexistem em exames mamográficos anteriores, e surgem em mamografias mais recentes.

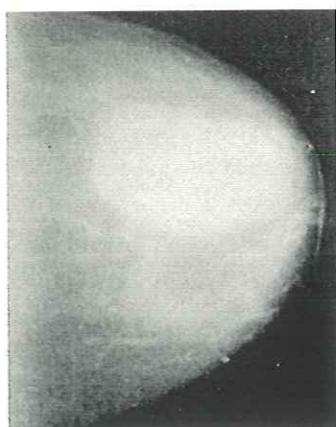


Figura 2.1 - [Pinotti-91] Imagem de tumor benigno com contorno regular bem definido.

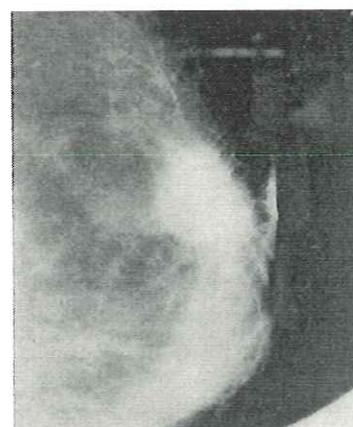


Figura 2.2 - [Pinotti-91] Carcinoma mamário com bordas irregulares e espícula

Capítulo 3

Digitalização de Imagens Mamográficas

O processo de digitalização de imagens mamográficas é uma etapa importante no processo de detecção e classificação de lesões em mamografia.

A imagem digital é uma imagem que sofreu o processo de digitalização tanto das coordenadas espaciais ou amostragem da imagem, quanto da intensidade ou quantificação dos níveis de cinza (quantização ou intensidade).

Em detecção de microcalcificações (tamanho em torno de 0.1mm) a qualidade da digitalização está diretamente associada à taxa de detecção dessas estruturas [Chan-94] [Ferrari-96]. Em detecção de nódulos, não existe nenhum estudo a respeito da resolução mínima de digitalização dos mamogramas, em função do tamanho das estruturas a serem analisadas. O número de pontos por polegada estabelece o limite de tamanho das estruturas que serão visíveis no mamograma. Quanto maior a resolução espacial, menores serão as estruturas identificáveis, porém, maior será a carga computacional devido ao tamanho da imagem, e maior será a presença de ruídos. Os níveis de cinza são responsáveis pelo grau de contraste, quanto maior o número de níveis de cinza maior será a possibilidade de se obter um melhor contraste, entretanto, o custo computacional de processamento também será maior. Estes fatores exigem um compromisso entre a resolução necessária ao processamento e o custo computacional, sem desprezo da presença de ruído. Em vários trabalhos apresentados não existe um padrão definido para a digitalização de mamogramas, embora alguns estudos tenham sido realizados [Nishikawa-94]. Pode-se verificar esta não padronização em diversos trabalhos apresentados de detecção e classificação de lesões do tipo massa. Em alguns artigos as informações de digitalização não existem ou estão incompletas, além disso, vários tipos de

equipamentos foram utilizados para a digitalização de mamogramas: como câmeras de vídeo e scanners óptico de média e alta resolução. Existem, porém, algumas configurações específicas que foram utilizadas por pesquisadores que investigaram a detecção e classificação de massas em mamogramas digitalizados, que serão citadas a seguir.

Giger e colaboradores, do “Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, University Chicago”, utilizaram tamanho de pixel de 100 μm e 10 bits de quantização de níveis de cinza para a classificação de lesões [Yin-91A] [Yin-91B] [Yin-93] [Yin-94A] [Yin-94B] [Giger-90A] [Giger-90B] [Giger-94] [Huo-95A] [Huo-95B] [Huo-96] [Bick-95].

Judy Kilday e colaboradores, digitalizaram os filmes à partir de uma câmera de vídeo com resolução espacial de 1.0 lpp/mm e 8 bits de quantização de níveis de cinza para a classificação e detecção das lesões nos mamogramas [Kilday-93].

R.M. Rangayyan e colaboradores, trabalharam com imagens digitalizadas com resolução espacial de 50 μm e 8 bits de quantização de níveis de cinza e, também, imagens obtidas do “Alberta Program for the Early Detection of Breast Cancer” que tinham resolução espacial de 62 μm e 8 bits de quantização de níveis de cinza [Rangayyan-96].

Scott Pohlman e colaboradores utilizaram quantização de 12 bits de níveis de cinza com uma câmera digital de sensor CCD de alta resolução. Utilizaram para uma subregião de cada mamograma contendo a lesão e tecidos normais, uma quantização de 10 bits de níveis de cinza e matriz de 4096x4096 pixels, resultando num tamanho médio de pixel de 12 μm [Pohlman-96].

Li, M. Kallergi, L.P e colaboradores, da “University of South Florida”, digitalizaram as imagens dos mamogramas de seu trabalho com resolução aproximada de 220 μm e 8 bits de quantização [Li-93].

H.P. Chan e colaboradores, do “Department of Radiology, University of Michigan”, digitalizaram os mamogramas de seus trabalhos com resolução de 100 μm e 12 bits de quantização [Petrosian-94] [Wei-95][Wei-97][Sahiner-96A] [Sahiner-96B][Petrick-96A][Petrick-96B][Chan-95][Sahiner-98].

Capítulo 4

Detecção de Nódulos em Mamogramas Digitalizados.

A detecção de nódulos computadorizada em mamogramas é um processo que se exprime pela dificuldade de distinção das características bastante similares entre as lesões e os tecidos glandulares (parênquima) que as cercam. Portanto, o processo de segmentação deve utilizar critérios bem definidos para diferenciar estes dois tipos de tecidos. De acordo com [Gonzalez-92], a segmentação é o processo pelo qual a imagem é dividida em suas partes constituintes, seguindo um ou mais critérios de similaridade. Ainda, de acordo com [Gonzalez-92], a descontinuidade diz respeito a mudanças abruptas, por exemplo, dos níveis de cinza, denominadas de borda. Em particular, para o estudo de nódulos em mamogramas digitalizados a borda fornece informações importantes como a forma espiculada e circunscrita de nódulos. A seguir apresentaremos os métodos utilizados para o processo de realce e segmentação utilizados em nosso trabalho.

4.1 Método de Realce

As técnicas de realce procuram manipular as imagens no sentido de melhorá-las para uma análise visual humana. Elas estão relacionadas com a alteração do contraste, destaque de bordas e detalhes, e suavização. As técnicas de realce podem ser divididas em duas categorias principais: métodos de vizinhança e métodos “ponto a ponto”. Os métodos de vizinhança são aplicados para suavização de ruídos e para realce de

contornos e de detalhes na imagem. Os métodos “ponto a ponto” procuram melhorar o contraste da imagem e foi utilizado em nosso trabalho. Vamos simplificadamente apresentar os métodos de vizinhança e detalharemos melhor o métodos “ponto a ponto” [Facon-93].

Os métodos de vizinhança mais utilizados para suavização de ruídos são os filtros passa-baixa, filtros passa-alta, filtros da média, filtros mediano, filtros da ordem, filtros de máscara de Nagao-Matsuyama e outros. Numa filtragem, o processamento de um nível de cinza de um pixel depende dos valores do pixel e dos seus vizinhos. Em geral, na vizinhança, os pixels mais próximos contribuem mais na definição do novo valor de seu nível de cinza do que dos pixels mais afastados. Por exemplo, no filtro mediano, toma-se os vizinhos mais próximos do pixel central e dentre eles assume-se o valor de nível de cinza mediano. No filtro da média, assume-se o valor da média dos níveis de cinza da vizinhança de um pixel central.

6	3	7
5	P_i	0
5	1	5

Dado o pixel P_i e sua vizinhança, o valor de nível de cinza mediano será igual a 5, pois a sequência em ordem crescente dos valores de nível de cinza dos pixels vizinhos é: 0, 1, 3, 5, 5, 5, 6, 7); e o valor médio dos níveis de cinza entre os pixels da vizinhança de P_i será igual a 4. Os outros métodos utilizam operações que manipulam os valores de nível de cinza dos pixels vizinhos de um pixel central.

Os métodos de realce de contornos e detalhes, utilizam operadores de diferenciação. Esses operadores são máscaras que aplicadas sobre as imagens realizam uma transformação em seus pixels. Os principais operadores são: o de Sobel, de Prewitt e Laplaciano. Por exemplo, o operador de Sobel é dado abaixo:

$$G_x = \begin{array}{|c|c|c|} \hline 1 & 0 & -1 \\ \hline 2 & 0 & -2 \\ \hline 1 & 0 & -1 \\ \hline \end{array} \quad G_y = \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 0 & 0 & 0 \\ \hline -1 & -2 & -1 \\ \hline \end{array}$$

A máscara é aplicada em todos os pixels da imagem. Num pixel dado, os valores da máscara são multiplicados pelos valores dos níveis de cinza dos pontos correspondentes vizinhos desse pixel. A soma dessas multiplicações representa o valor do operador aplicado ao pixel central dessa vizinhança.

Os métodos “ponto a ponto” aplicam operações pontuais sobre cada pixel procurando modificar a distribuição dos níveis de cinza na imagem preservando os contornos. Essa distribuição pode ser apresentada em um histograma que fornece a frequência de valores de níveis de cinza em uma imagem. Portanto, os métodos atuam no sentido de realizar alterações no histograma para um fim específico. Os métodos que apresentaremos é o da equalização do histograma, e da modificação do contraste [Facon-93].

4.1.1 Equalização do Histograma

Seja r , a variável representando o nível de cinza suposto normalizado entre 0 (preto) e 1 (branco), e uma transformação $T(\bullet)$:

$$s = T(r) \quad (4.1)$$

Supondo que a transformação $T(\bullet)$ satisfaz as condições de que a função é monotonicamente crescente no intervalo $[0,1]$; e $0 < T(r) < 1$, para $r \in [0,1]$.

Dada uma imagem original, as frequências de níveis de cinza são dadas por:

$$p_r(r_i) = \frac{n_i}{n} \quad (4.2)$$

onde i representa o nível de cinza e $i \in [0, L-1]$ e n_i corresponde ao número de pixels do nível de cinza i e n corresponde ao número total dos pixels.

A transformação associada à uniformização é dada por:

$$s_k = T(r_k) = \sum_{j=0}^k \frac{n_j}{n} = \sum_{j=0}^k p_k(r_k) \quad (4.3)$$

Ao processo geral denominamos de equalização do histograma. Por causa dos problemas de aproximação de densidade, o histograma obtido não corresponde exatamente ao histograma de uma distribuição uniforme.

Haralick propõe um outro algoritmo de equalização do histograma que tem a vantagem de fixar o número K de níveis de cinza. Fixando K menor que o valor de origem permitindo aproximar-se de uma distribuição uniforme. O algoritmo utiliza a filosofia que:

Sejam $F(\bullet)$ a função de repartição de r e $Q(\bullet)$ uma função de quantificação equiprovável de K níveis de cinza.

$$Q(r) = K \Leftrightarrow q_{k-1} < r < q_k \quad (4.4)$$

O problema é obter q_0, q_1, \dots, q_k . O algoritmo define estes valores de uma forma iterativa. Partimos de um valor inicial q arbitrário, por exemplo zero.

Suponhamos conhecidos q_0, q_1, \dots, q_k . Na k -ésima iteração, $F(q_{k-1})$ é distribuída com $(k-1)$ níveis de cinza. Falta então distribuir $[1-F(q_{k-1})]$.

Esse resto deve ser repartido de uma forma uniforme sobre os níveis de cinza restantes. Portanto, cada nível de cinza obterá:

$$\frac{[1 - F(q_{k-1})]}{[K - k + 1]} \quad (4.5)$$

Então:

$$F(q_k) = F(q_{k-1}) + \frac{[1 - F(q_{k-1})]}{[K - k + 1]} \quad (4.6)$$

$F(\bullet)$ é uma função em escada, não temos a certeza de encontrar um q_k que satisfaça a igualdade. Portanto, deve-se procurar o valor de q mais próximo de:

$$F(q_{k-1}) + \frac{[1 - F(q_{k-1})]}{[K - k + 1]} \quad (4.7)$$

4.1.2 Modificação do Contraste

Uma modificação do contraste é uma aplicação da função $T(\bullet)$ tal que:

$$z' = T(z), \text{ onde } z > z_{\min}, z' < z_{\max} \quad (4.8)$$

Para aumentar o contraste, com $z \in [a,b] \subset [z_{\min}, z_{\max}]$

$$z' = \left[\frac{(z_{\max} - z_{\min})}{(b - a)} \right] (z - a) + z_{\min} \quad (4.9)$$

Existe também uma outra possibilidade quando a maioria dos níveis de cinza estão entre $[a,b]$: se $z < a$, $z' = z_{\min}$; se $z > b$, $z' = z_{\max}$; se $a < z < b$ vale a equação 4.9.

Esta transformação quase linear comprime os intervalos $[z_{\min}, a]$ e $[b, z_{\max}]$, isto é aceitável quando temos poucos níveis cinza nesses intervalos.

Se considerarmos, por exemplo:

$z_{\max} = 255$; $z_{\min} = 0$, teremos:

$$z' = 255 \cdot \frac{(z - a)}{(b - a)} \quad (4.10)$$

4.2 Método de Limiarização

O princípio da limiarização (thresholding) consiste em separar as regiões de uma imagem quando esta apresenta duas classes (o fundo e o objeto). A técnica de limiarização pode ser baseada em operadores de limiarização ou no histograma da imagem. Na primeira técnica o valor de limiar é definida através de características vizinhas ao pixel. A segunda técnica baseia-se no histograma da imagem. A figura 4.1 mostra um exemplo de limiarização em que o histograma apresenta distintamente duas classes na forma de dois picos por um “vale”. Neste caso, a limiarização é trivial.

O processo de limiarização consiste, no caso de um objeto mais claro que o fundo, em escolher no histograma um valor T , tal que para um pixel $p(x,y)$:

$$\text{Se } I[p(x,y)] > T \Rightarrow p(x,y) \in \text{objeto}$$

$$\text{Senão} \Rightarrow p(x,y) \in \text{fundo}$$

Onde I é a intensidade e o valor de T separa as duas classes (objeto e fundo).

A escolha do limiar é muitas vezes empírica e depende muito da qualidade dos contrastes, da qualidade, da natureza da imagem e aquilo que se quer segmentar na imagem.

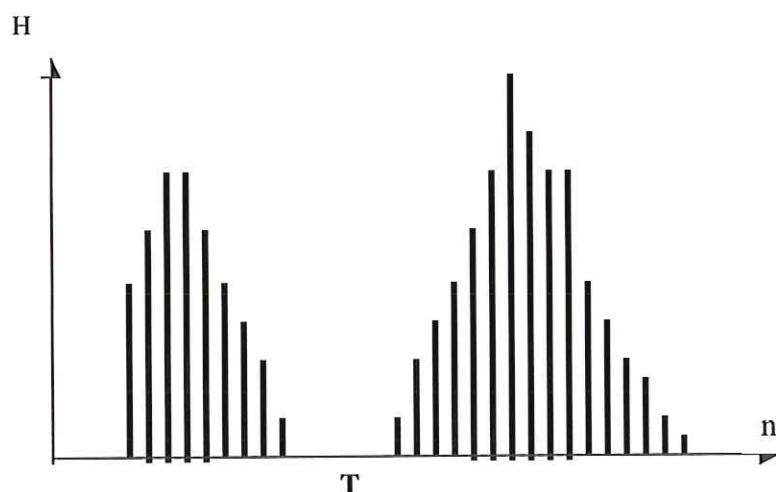


Figura 4.1 – Histograma de uma imagem com separação bem definida entre duas classes. O valor de limiar (T), neste caso, é trivialmente escolhido.

4.3 Método de Filtros Morfológicos.

A técnica de filtros morfológicos são baseadas na morfologia matemática [Facon-93]. Eles incluem o processo de erosão e dilatação, and modificações e combinações dessas operações. De forma geral, existem dois tipos de morfologia matemática: a morfologia binária que se aplica sobre imagens binárias e a morfologia cinzenta que se aplica sobre imagens em níveis de cinza [Facon-96] [Russ-94].

O método de erosão ou dilatação baseia-se em um *elemento estruturante* que é um conjunto conhecido (forma e tamanho), que é comparado, a partir de uma transformação, ao conjunto desconhecido da imagem. O resultado dessa transformação permite avaliar o conjunto desconhecido, testando e quantificando de que maneira o elemento estruturante *"está ou não contido"* na imagem [Facon-96].

A erosão segundo a definição de [Serra-82], os pixels da imagem a ser erodido, é deslocado em função de posições permitidas pelo elemento estruturante B. O deslocamento é realizado em relação ao ponto central dos elementos do conjunto estruturante. Uma operação de interseção (ou a operação matemática AND) entre os deslocamentos obtidos é realizada na operação de erosão. O exemplo abaixo ilustra uma operação de erosão, a partir de um elemento estruturante [Facon-96]:

$$B = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{ero} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Os efeitos obtidos pela erosão são [Facon-96]:

- Diminuição da partículas e eliminação de grãos de tamanho inferior ao tamanho do elemento estruturante;
- Aumento dos buracos e separação de grãos próximos.

Na dilatação, é utilizada a operação de união (ou a operação matemática OR) dos deslocamentos obtidos. Um exemplo é ilustrado abaixo com o mesmo elemento estruturante da erosão.

$$\begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \text{dil} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} Y \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} Y = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

Os efeitos principais obtidos são [Facon-96]:

- Preenchimento de pequenos buracos na imagem;
- Conexão de grãos próximos;
- Dilatação da imagem.

Duas operações morfológicas estendidas da erosão e dilatação são obtidas: a operação de fechamento e abertura. A operação de fechamento é o processo de aplicação da operação de dilatação seguida da operação de erosão. Pequenos buracos na imagem são preenchidos e pequenos detalhes da borda da imagem são conservados. A operação de abertura processa a operação de erosão seguida da operação de dilatação. Detalhes de borda e pequenas estruturas são removidas enquanto conserva estruturas grandes da imagem [Russ-94].

4.4 Método de Crescimento de Região.

Crescimento de região é o processo de agregação de pixels ou grupo de pixels de uma imagem de acordo com um critério de similaridade. É o mais conhecido método de segmentação em processamento de imagem. [Gonzalez-92]. O processo baseia-se em um pixel ou um grupo de pixels chamados de “pixels semente” que selecionam pixels vizinhos com as mesmas características similares. Os pixels sementes são obtidos de modo automático ou interativo e pertencem a região de interesse analisada. O resultado do processo de crescimento de região é uma simples região contendo pixels com

propriedades similares aos pixels sementes. Os pixels que não satisfazem a condição de similaridade são definidos como pixels de borda da região crescida.

Para definir um critério de similaridade, utiliza-se um valor de limiarização (thresholding) e o resultado de propriedades locais da região de interesse (textura, nível de cinza, cor). Desta forma, o critério é dependente não só de um valor limiar, mas também de propriedades que determinam a homogeneidade ou não da região de interesse.

Os valores de limiarização podem ser fixos ou adaptativos. Os valores fixos exigem conhecimento a priori das propriedades da imagem analisada. Os valores adaptativos variam durante o processamento dependendo de problemas de similaridade encontrados como, por exemplo, encontrar regiões com valores de pixels próximo aos valores dos pixels semente.

Capítulo 5

Extração e Seleção de Atributos de Textura e Forma de Nódulos em Mamogramas

O processo de extração de atributos é imprescindível para a classificação de nódulos mamários. São importantes para caracterizar e discriminar um objeto em uma imagem. São propriedades mensuráveis ou padrões dos objetos a serem reconhecidos na imagem ou parte dela. A área, perímetro, circularidade, orientação, e número de buracos são exemplos de atributos [Parker-94]. Nos trabalhos tanto de diferenciação do nódulo em relação ao tecido normal, como na sua classificação como maligno ou benigno, atributos de forma e textura são amplamente utilizados.

Os atributos de forma são importantes nos diagnósticos de doenças da mama, pois lesões com formas espiculadas ou estreladas no contorno são suspeitas de malignidade. Os atributos de textura são importantes para determinação de diferenças de densidade e distribuição de níveis de cinza no nódulo e ao redor dele. Dependendo da forma e estrutura dessa distribuição, podem representar propriedades importantes na caracterização de nódulos mamários do tipo circunscrito.

5.1 Atributos de Forma

Frequentemente, os objetos de uma classe podem ser distinguidos de outros objetos pela sua forma. Atributos de forma são amplamente utilizados para reconhecimento de padrões de nódulos em mamogramas. Apresentaremos resumidamente os atributos utilizados em nosso trabalho.

5.1.1 Medida de Circularidade:

De acordo com alguns autores a circularidade é definida de maneira diferente. Pohlman e colaboradores definem a circularidade como a razão entre a distância média radial do contorno do objeto (μ_R) e o desvio-padrão (σ_R) [Pohlman-96].

$$C_i = \frac{\mu_R}{\sigma_R} \quad (5.1)$$

Haralick demonstra que esta medida aumenta quanto mais circular for a forma do objeto independentemente da orientação ou da sua área [Haralick-74]. Desta forma, as lesões malignas, com formas espiculadas ou estreladas, tem valores menores de circularidade do que as lesões benignas, com formas regulares de contorno.

5.1.2 Grau de Compactação - Solidez:

Pohlman e colaboradores definem o grau de compactação ou solidez como razão do perímetro (P) e a área do objeto (A) [Pohlman-96].

$$C = \frac{P}{A} \quad (5.2)$$

Uma forma de contornar problemas computacionais, é restringir os valores de C entre [0,1], alterando-se a equação (5.3) [Rangayyan-96]:

$$C = 1 - \frac{4\pi A}{P^2} \quad (5.3)$$

Observe que para objetos com forma circular, o valor da solidez será de zero e aumenta de acordo com a complexidade ou rugosidade da forma.

5.1.3 Descritores de Gupta:

Os descritores de Gupta são fatores de forma baseados em momento e são poucos sensíveis a ruídos sendo indicados para formas mais grosseiras [El-Faramawy-96]. A ordem $(p+q)^{\text{ésimo}}$ de um momento bi-dimensional m_{pq} em um objeto é definida como:

$$m_{pq} = \sum_i \sum_j i^p j^q I(i, j) \quad p, q = 0, 1, \dots, N \quad (5.4)$$

Onde o somatório é sobre todos os pixels na fronteira e contorno da região de interesse (ROI) e $I(i,j)$ é a intensidade de imagem de cada pixel (i,j) . O centróide do objeto, de acordo com [Levine-85], é definido como

$$\bar{x} = m_{10}/m_{00} \quad (5.5)$$

$$\bar{y} = m_{01}/m_{00} \quad (5.6)$$

Uma variação dos momentos em função da diferença absoluta de nível de cinza de cada pixel na ROI em relação a média de nível de cinza (\bar{I}) é dada por:

$$m'_{pq} = \sum_i \sum_j i^p j^q |I(x, y) - \bar{I}| \quad p, q = 0, 1, \dots, N \quad (5.7)$$

Uma definição alternativa de momentos é usada para uma sequência de pixels de contorno [Shen-94]. Por esta definição, momentos estatísticos unidimensionais são extraídos de uma representação sequencial de fronteira de uma região. Levine indicou que o contorno em torno do centróide, é plotado à distância z_i desde o centróide até cada ponto da fronteira e deve ser normalizada pela máxima distância [Levine-85]. Então, o p -ésimo momento é definido como:

$$m_p = \frac{1}{N} \sum_{i=1}^N |z_i|^p \quad (5.8)$$

e o p -ésimo momento central é definido como:

$$M_p = \frac{1}{N} \sum_{i=1}^N |z_i - m_1|^p \quad (5.9)$$

Os momentos normalizados correspondentes são definidos como [Gupta-87]:

$$\overline{m_p} = \frac{m_p}{(M_2)^{p/2}} = \frac{\frac{1}{N} \sum_{i=1}^N |z_i|^p}{\left[\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^p \right]^{p/2}} \quad (5.10)$$

$$\overline{M_p} = \frac{M_p}{(M_2)^{p/2}} = \frac{\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^p}{\left[\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^p \right]^{p/2}} \quad (5.11)$$

Uma conclusão importante é que os momentos apresentados são invariantes quanto a rotação, translação e escala.

Momentos de altas ordens são mais sensíveis e torna os classificadores menos tolerantes a ruídos. Por causa disso, Gupta e Srinath selecionaram três momentos de baixa ordem para formar um conjunto de descritores de forma, que são os seguintes [Gupta-87]:

$$F_1 = \frac{(M_2)^{1/2}}{m_1} = \frac{\left[\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^2 \right]^{1/2}}{\frac{1}{N} \sum_{i=1}^N z_i} \quad (5.12)$$

$$F_2 = \frac{M_3}{(M_2)^{3/2}} = \frac{\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^3}{\left[\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^2 \right]^{3/2}} \quad (5.13)$$

$$F_3 = \frac{M_4}{(M_2)^2} = \frac{\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^4}{\left[\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^2 \right]^2} \quad (5.14)$$

5.1.4 Descritores de Shen:

Baseadas nas propostas de Gupta e Srinath em que apresentam os descritores de momento (equações 5.13 , 51.4 , e 5.15), Shen *et al.* [Shen-94] apresentam um novo conjunto de descritores de forma:

$$F_1' = F_1 = \frac{(M_2)^{1/2}}{m_1} = \frac{\left[\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^2 \right]^{1/2}}{\frac{1}{N} \sum_{i=1}^N z_i} \quad (5.15)$$

$$F_2' = \frac{(M_3)^{1/3}}{m_1} = \frac{\left[\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^3 \right]^{1/3}}{\frac{1}{N} \sum_{i=1}^N z_i} \quad (5.16)$$

$$F_3' = \frac{(M_4)^{1/4}}{m_1} = \frac{\left[\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^4 \right]^{1/4}}{\frac{1}{N} \sum_{i=1}^N z_i} \quad (5.17)$$

$$F_4' = F_3' - F_1' = \frac{\left[\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^4 \right]^{1/4}}{\frac{1}{N} \sum_{i=1}^N z_i} - \frac{\left[\frac{1}{N} \sum_{i=1}^N |z_i - m_1|^2 \right]^{1/2}}{\frac{1}{N} \sum_{i=1}^N z_i} \quad (5.18)$$

De acordo com [Shen-94], o descritor F_3' descreve a forma de rugosidade melhor do que F_3 de Gupta. Quanto maior o valor de F_3' , mais rugosa é a forma do contorno do objeto. O fator ($F_3' - F_1'$) fornece informação sobre a rugosidade da forma que individualmente F_3' e F_1' não informam.

5.1.5 Descritores de Hu:

Hu e Levine, apresentaram sete momentos invariantes de baixa-ordem, que são funções de segunda e terceira ordem de momentos centrais. Eles são invariantes a escala, posição e orientação. A seguir são apresentadas os descritores [Hu-62] [Levine-85]:

$$M_1 = \mu_{20} + \mu_{02} \quad (5.19)$$

$$M_2 = (\mu_{20} + \mu_{02})^2 + 4\mu_{11} \quad (5.20)$$

$$M_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \quad (5.21)$$

$$M_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \quad (5.22)$$

$$M_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + \\ + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[(3\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \quad (5.23)$$

$$M_6 = (\mu_{20} - \mu_{02}) [(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + \\ + 4\mu_{11} (\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03})(\mu_{21} + \mu_{03}) \quad (5.24)$$

$$M_7 = (3\mu_{31} - \mu_{03}) (\mu_{30} + \mu_{12})^2 [(\mu_{30} + \mu_{12})^2] - (\mu_{21} + \mu_{03})^2] - \\ - (\mu_{30} - 3\mu_{12}) (\mu_{21} + \mu_{03}) [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \quad (5.25)$$

Onde: $\mu_{00} = m_{00} = \mu$

$$\mu_{10} = \mu_{01} = 0$$

$$\mu_{20} = \mu_{20} - \mu_{10}$$

$$\mu_{11} = \mu_{11} - \mu \bar{X} \bar{Y}$$

$$\mu_{02} = \mu_{02} - \mu \bar{Y}^2$$

$$\mu_{30} = \mu_{30} - 3\mu_{20} \bar{X} + 2\mu \bar{X}^3$$

$$\mu_{21} = \mu_{21} - \mu_{20} \bar{Y} - 2\mu_{11} \bar{X} + 2\mu \bar{X}^2 \bar{Y}$$

$$\mu_{12} = \mu_{12} - \mu_{02} \bar{X} - 2\mu_{11} \bar{Y} + 2\mu \bar{X} \bar{Y}^2$$

$$\mu_{03} = \mu_{03} - 3\mu_{02} \bar{Y} + 2\mu_{11} \bar{Y}^3 + 2\mu \bar{X} \bar{Y}^2$$

Os descritores de Hu não são independentes do contraste do objeto, e são numericamente sensíveis [El-Faramawy-96].

5.1.6 Descritor de Fourier:

Os descritores de Fourier são definidos para cada ponto com coordenada (x,y) em uma matriz periódica de N pontos de fronteira de um objeto representados com valores complexos, $Z_i = x_i + j y_i$, onde $i = 0, 1, 2, \dots, N-1$, com a seguinte equação:

$$A(n) = \frac{1}{N} \sum_{i=0}^{N-1} Z_i \exp\left[\frac{-j2\pi ni}{N}\right] , n=0,1,2,\dots,N-1 \quad (5.26)$$

Usando somente a magnitude dos descritores de Fourier (FD) e fazendo sua normalização, pode-se obter:

$$NFD(k) = 0 \quad , \text{para } k = 0;$$

$$NFD(k) = \frac{A(k)}{A(1)} \quad , \text{para } k = 1, 2, \dots, N/2;$$

$$NFD(k) = \frac{A(k + N)}{A(1)} \quad , \text{para } k = -1, -2, \dots, -(N/2) + 1;$$

Enfatizando os componentes de baixa-frequência, podemos derivar o descritor de Fourier (FD) baseado em fator de forma (FF) como descrito em [Shen-94]:

$$FF = 1 - \frac{\sum_{k=N/2}^{N/2} \frac{\|NFD(k)\|}{\|k\|}}{\sum_{k=N/2}^{N/2} \|NFD(k)\|} \quad (5.27)$$

A vantagem desta medida é que não é sensível a ruídos e é invariante em relação a rotação, translação, ponto inicial, e tamanho do contorno.

5.1.7 Descritores de Wee:

Os descritores de Wee calculam a variação de borda em relação a centróide da lesão [Wee-75].

$$\text{Wee}_1 = \frac{1}{N} \left[\sum_{i=1}^N \sqrt{D(i)} \right] \quad (5.28)$$

onde:

$$D(i) = \sqrt{(X_i - C_x)^2 + (Y_i - C_y)^2}$$

C_x, C_y : Centróide da região da imagem do objeto;

X_i, Y_i : Fronteira da região do objeto;

$$\text{Wee}_2 = \frac{1}{N} \left[\sqrt{\sum_{i=1}^N (D(i) - \text{Wee}_1)^2} \right] \quad (5.29)$$

5.2 Atributos de Textura.

Textura é uma importante característica usada para identificar objetos ou regiões de interesse numa imagem. A textura contém informações sobre a distribuição espacial de variações de tonalidades dentro de uma banda. A tonalidade representa o nível de cinza da imagem. Logo, tonalidade e textura estão sempre presentes numa imagem.

Esses atributos tem sido utilizados para detectar nódulos em mamografias através da variação de sua densidade em relação aos tecidos circunvizinhos. Essa variação é detectada utilizando-se matrizes de co-ocorrência de níveis de cinza (SGLD) que calculam a probabilidade combinada entre direção e distância entre pares de pixels. Logo, essas variações podem indicar valores referenciais de malignidade ou normalidade de uma lesão numa mamografia.

Haralick *et al* [Haralich-73] sugerem 14 atributos de textura que podem ser extraídos das matrizes SGLD:

5.2.1 Descritores de Haralick:

Os atributos de textura sugeridos por Haralick *et al.* são: Variância, Correlação, Contraste, Entropia, Soma da Entropia, Diferença da Entropia, Momento da Diferença Inversa (MDI), Média da Soma, Variância da Soma, Média da Diferença, Variância da Diferença, Energia, Medida de Informação de Correlação 1 (MIC₁), e Medida de Informação de Correlação 2 (MIC₂):

$$\text{Variância} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i - \mu_x)(j - \mu_y)p(i,j) \quad (5.30)$$

Onde n é o número de níveis de cinza da imagem e $p(i,j)$ é a matriz de co-ocorrência de nível de cinza (SGLD).

$$\text{Correlação} = \frac{\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i - \mu_x)(j - \mu_y)p(i,j)}{\sigma_x \sigma_y} \quad (5.31)$$

Onde, as média e varianças são dadas por:

$$\begin{aligned} \mu_x &= \sum_{i=0}^{n-1} i p_x(i) & \sigma_x^2 &= \sum_{i=0}^{n-1} (i - \mu_x)^2 p_x(i) \\ \mu_y &= \sum_{j=0}^{n-1} j p_y(j) & \sigma_y^2 &= \sum_{j=0}^{n-1} (j - \mu_y)^2 p_y(j) \\ p_x(i) &= \sum_{j=0}^{n-1} p(i,j) & p_y(i) &= \sum_{i=0}^{n-1} p(i,j) \end{aligned}$$

$$\text{Contraste} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i - j)^2 p(i,j) \quad (5.32)$$

$$\text{Entropia} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} p(i,j) \log_2 p(i,j) \quad (5.33)$$

$$\text{Soma_Entropia} = \sum_{k=0}^{2n-2} p_{x+y}(k) \log_2[p_{x+y}(k)] \quad (5.34)$$

Onde:

$$p_{x+y}(k) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} p(i,j)$$

$$i+j = k, k = 0, \dots, 2n-2$$

$$\text{Diferença_Entropia} = \sum_{k=0}^{2n-2} p_{x-y}(k) \log_2[p_{x-y}(k)] \quad (5.35)$$

Onde:

$$p_{x-y}(k) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} p(i,j)$$

$$|i - j| = k, k = 0, \dots, n-1$$

$$\text{MDI} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \frac{1}{1 + (i-j)^2} p(i,j) \quad (5.36)$$

$$\text{Média_Soma} = \sum_{k=0}^{2n-2} k p_{x+y}(k) \quad (5.37)$$

$$\text{Variança_Soma} = \sum_{k=0}^{2n-2} (k - \text{média_soma})^2 p_{x+y}(k) \quad (5.38)$$

$$\text{Média_Diferença} = \sum_{k=0}^{n-1} k p_{x-y}(k) \quad (5.39)$$

$$\text{Variança_Diferença} = \sum_{k=0}^{n-1} (k - \text{média_diferença})^2 p_{x-y}(k) \quad (5.40)$$

$$\text{Energia} = \sum_{k=0}^{n-1} p(k)^2 \quad (5.41)$$

$$\text{MIC}_1 = \frac{\text{Entropia} - H_1}{\max[H_X, H_Y]} \quad (5.42)$$

$$\text{MIC}_2 = (1 - \exp[-2(H_2 - H_{XY})])^{1/2} \quad (5.43)$$

onde:

$$H_X = - \sum_{i=0}^{n-1} p_x(i, j) \log_2 [p_x(i)]$$

$$H_Y = - \sum_{j=0}^{n-1} p_y(i, j) \log_2 [p_y(j)]$$

$$H_1 = - \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} p(i, j) \log_2 [p_x(i) \cdot p_y(j)]$$

$$H_2 = - \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} p_x(i) \cdot p_y(j) \cdot \log_2 [p_x(i) \cdot p_y(j)]$$

5.3 Método Seleção de Jeffries-Matusita.

A definição da distância de Jeffrie –Matusita é dada por:

$$J_T = \left\{ \int [\sqrt{p(\xi / \omega_1)} - \sqrt{p(\xi / \omega_2)}]^2 d\xi \right\}^{1/2} \quad (5.44)$$

onde $p(\xi / \omega_1)$ e $p(\xi / \omega_2)$ são as funções de probabilidade condicional das classes ω_1 e ω_2 [Young-86] .

Quando as funções de densidade estão sobrepostas, a integral é igual a zero, logo a distância J_T entre as classes será igual a zero. Agora, se as funções de probabilidade não estão sobrepostas, a distância J_T entre as classes atinge o valor próximo de $\sqrt{2}$.

Quando supomos uma distribuição de probabilidade paramétrica e em particular uma distribuição Gaussiana, isto é, quando a função densidade é definida como:

$$p(\xi / \omega_i) = [(2\pi)^d |\Sigma_i|]^{-1/2} \exp\left\{-\frac{1}{2} (\xi - \mu_i)^T \Sigma_i^{-1} (\xi - \mu_i)\right\} \quad (5.45)$$

onde μ_i e Σ_i são os vetores de média e matriz de covariância da distribuição da i -ésima classe, respectivamente. desta forma obteremos para a distância de Matusita:

$$J_T = \{2[1-\exp(-J_B)]^{1/2} \quad (5.46)$$

onde :

$$J_B = \frac{1}{4} (\mu_2 - \mu_1)^T [\Sigma_1 + \Sigma_2]^{-1} (\mu_2 - \mu_1) + \frac{1}{2} \ln \left[\frac{\left| \frac{1}{2} (\Sigma_1 + \Sigma_2) \right|}{\sqrt{|\Sigma_1||\Sigma_2|}} \right] \quad (5.47)$$

Portanto, o valor máximo da distância de Jeffries-Matusita será igual a $\sqrt{2}$. isto significa que quando a distância atinge esse valor, temos uma separação máxima entre as classes [Young-86].

Capítulo 6

Classificação de Nódulos em Mamogramas

As técnicas de classificações recaem entre duas categorias: numéricas e não numéricas [Beale-90]. Técnicas numéricas incluem medidas determinísticas e estatísticas que podem ser consideradas como medidas originárias do espaço geométrico. Técnicas não numéricas são aquelas que trabalham no domínio do processamento simbólico.

A etapa de classificação efetua a identificação de atributos em classes. Essa classificação envolve a combinação de atributos para fornecer um único valor numérico indicando a classe. Normalmente, mais do que um atributo são necessários para classificar uma imagem [Parker-94]. Esses atributos são apresentados aos classificadores em forma de vetores, onde os elementos são valores numéricos representando medidas de forma, textura ou contraste da imagem. A dimensionalidade do vetor representa o número de elementos pertencentes a um vetor de atributos, e é chamado de espaço de atributos [Beale-90]. O sistema de classificação apresenta duas fases importantes: treinamento e teste.

A fase de treinamento agrupa as amostras de dados em classes e encontra a fronteira divisora que separa as classes, chamada de função discriminante. A função discriminante pode ser linear ou não-linear. O treinamento pode ser supervisionado ou não supervisionado. O processo supervisionado exige a disponibilidade de um tutor externo que forneça ao classificador a resposta desejada para um determinado conjunto de treinamento. No treinamento não supervisionado, o sistema se auto-

organiza e não existe um tutor externo; não existem exemplos específicos a serem aprendidos pelo sistema. Ele deve aprender com base em uma medida de qualidade.

Na fase de teste, apresenta-se um conjunto de padrões não utilizados no treinamento para verificar se o sistema assimilou conhecimento e a capacidade de generalização.

Os principais métodos de classificação utilizados na caracterização de nódulos mamários variam, principalmente, entre o enfoque estatístico e determinístico, divididos em paramétricos (regra de Bayes) e não-paramétricos (K-NN e LDA) e o de redes neurais artificiais (ANN) que é considerado um método semi-paramétrico [Beale-90].

6.1 Técnicas Estatísticas:

De acordo com Fukunaka, a proposta de classificação estatística de padrões, é determinar a que categoria ou classe pertence uma amostra [Young-86]. Através de observações e processos de medidas, obtemos um conjunto de números que compõem um vetor de medidas. O vetor é aleatório e sua função de densidade condicional depende de sua classe. A classificação bayesiana baseia-se na teoria estatística de probabilidade e probabilidade condicional [Beale-90]. Para a classificação de padrões são utilizadas medidas originárias de vetores de atributos para estimar a probabilidade de um determinado padrão pertence ou não a uma classe. O mérito do classificador bayesiano é que podemos assumí-lo como um classificador linear, realizando suposições em relação aos padrões fornecidos. A desvantagem dele é que erros podem ocorrer quando os desvios-padrões das distribuições de classes são elevados, podendo ocorrer o problema de sobreposição de classes em determinada região do espaço de atributos [Beale-90].

O K-NN é um método não-paramétrico e determinístico. A classificação é feita através da distância entre as amostras de uma classe em relação a outra. Por exemplo, temos duas classes representadas no espaço de padrões e queremos decidir a qual classe pertence um padrão X. Basta calcular a menor distância entre X e as amostras de cada classe. A classificação é baseada em função discriminante linear e utiliza fronteiras de decisão.

Diversas distâncias métricas podem ser utilizadas para medir a similaridade de amostras de padrões num espaço geométrico: distância de Hamming, distância Euclidiana, distância de Bloco de Cidade (Manhattan), e distância Quadrática. Pode-se utilizar a técnica em um número finito de vetores de atributos, isto é, em K vizinhos mais próximos (KNN). Cada elemento do vetor de atributo possui um fator de ponderação e a decisão é feita no máximo número de elementos de uma mesma classe dentre os K-elementos. A aplicação desta técnica é restrita, pois é necessário encontrar uma medida de distância confiável e várias medidas de distância com vários exemplos de cada classe [Beale-90].

A análise de discriminante linear tem a função de separar um conjunto de amostras em grupos ou classes, através de uma função discriminante linear que serve como fronteira de decisão no espaço de padrões. Normalmente, utiliza-se a distância de Mahalanobis para determinação da menor distância discriminante.

A distância de Mahalanobis é dada por vetores de médias para um número de amostras que são consideradas como estimativas dos verdadeiros vetores de média para as classes. A distância entre os componentes e as médias para cada classe são calculadas e o menor valor determina a associação de cada elemento a uma classe. O problema mais comum encontrado neste tipo de classificador é a implementação da fronteira de decisão e a definição desta fronteira no espaço de padrões [Beale-90].

6.2 Redes Neurais Artificiais.

As redes neurais artificiais - artificial neural networks (ANN) - é um sistema de processamento de informações que tem certas características em comum com redes neurais biológicas [Fausset-94]. As ANN tem sido desenvolvidas como generalização de modelos matemáticos de cognição humana ou biologia neural, baseada em áreas que: o processamento de informações ocorre em muitos elementos simples chamados de neurônios; os sinais são passados de um neurônio a outro através de ligações de conexões (os neurônio funcionam como integradores de sinais); cada conexão tem associado um peso, chamado de peso sináptico, numa rede neural típica, multiplicado pelo sinal transmitido (um neurônio coleta sinais de outros neurônios, integra os dados e realiza ajustes nos pesos); cada neurônio aplica uma função de ativação e um limiar (threshold ou bias) na sua entrada para determinar o sinal na sua saída na rede. Uma rede neural artificial (ANN) é caracterizada por seus

padrões de conexões entre os neurônios, definindo a sua arquitetura; por seus métodos de determinação de pesos sinápticos nas conexões, chamados de treinamento ou aprendizado; e por suas funções de ativação [Fausset-94].

6.2.1 Arquitetura de Redes Neurais Artificiais:

A arquitetura de uma rede neural artificial é determinada pelo arranjo de neurônios em camadas e as conexões entre os neurônios da mesma camada e entre o neurônio das outras camadas. Quando os neurônios são conectados entre os de mesma camada são chamados de camadas competitivas. Quando os neurônios são conectados somente com os de outras camadas e os sinais seguem em uma única direção, é chamada de feedforward. Existem redes que possuem conexões ligando os neurônios de saída aos neurônio de entrada e são chamadas de recorrentes. As redes são classificadas de acordo com o número camadas: rede de camada única (single-layer) ou de multi-camadas (multi-layers). Na determinação do número de camadas, normalmente, a camada de unidades de entrada não é contada como tal [Fausset-94]. Como na camada de entrada não são atribuídos pesos, o número de camadas pode ser determinado pelo número de camadas que possuem pesos. A rede de única camada é formada, então, pela camada de entrada e a camada de saída. A figura 6.1 ilustra uma rede de única camada; uma rede multi-camadas formada por uma camada de entrada, uma ou mais camadas intermediárias (camadas escondidas), e uma camada de saída; e uma rede com camada competitiva.

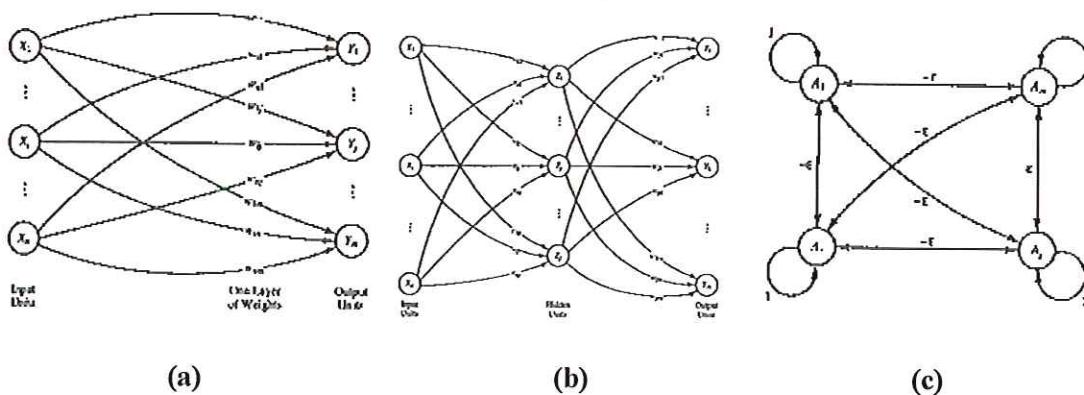


Figura 6.1 - [Fausset-94] (a) ilustra uma rede de camada única; (b) ilustra uma rede de multi-camadas; (c) ilustra uma rede com camada competitiva.

Uma rede neural de camada única tem uma camada de conexões com pesos. As unidades de entrada recebem os sinais externos (X_i , $i = 1, \dots, n$), e as unidades da camada de saída fornecem as respostas da rede (Y_j , $j = 1, \dots, m$). Exceção é a rede de Hopfield em que as unidades existentes tem ambas as funções de entrada e saída. Para a classificação de padrões, a entrada recebe os vetores de atributos e define para cada atributo uma unidade de entrada. Na camada de saída as unidades definem qual a classe em que os atributos de entrada pertencem.

Uma rede neural de multi-camadas possui um conjunto de pesos entre as camadas adjacentes (camada de entrada, escondida e de saída). Essa arquitetura pode resolver problemas mais complexos do que a de camada única, porém o treinamento é mais difícil.

Uma camada do tipo competitiva possui inter-conexões entre neurônios na mesma camada. Eles têm pesos iguais e é aplicada a operação “winner-take-all”, isto é, o vencedor-leva-tudo.

Os pesos são importantes numa rede pois a sua configuração determina o grau de aprendizado da rede. Esse processo de configuração dos pesos é realizado pelo treinamento da rede, através de um processo iterativo de ajustes. Considerando um par de neurônios x_i e x_j conectados por um peso sináptico w_{ji} (k) aplicado num tempo k, tem-se:

$$w_{ji}(k+1) = w_{ji}(k) + \Delta w_{ji}(k) \quad (6.1)$$

Onde w_{ji} (k+1) e w_{ji} (k) são os valores novo e antigo do peso sináptico e Δw_{ji} é o ajuste do peso.

Quanto ao tipo de conexões existentes em redes neurais temos: o excitatório ($w_{ji} > 0$), o inibitório ($w_{ji} < 0$), e o inexistente ($w_{ji} = 0$).

Uma característica que difere uma rede neural, além obviamente de sua arquitetura, é a função de ativação e a adição de parâmetros de limiar (threshold) e ganho. Portanto, na saída de cada neurônio teremos:

$$Y_j = f\left(\sum_{i=1}^n (w_{ij} \cdot x_i - \theta_j)\right) \quad (6.2)$$

Onde x_i é o estado do neurônio i da camada precedente, w_{ij} é o peso sináptico da conexão entre o neurônio i de uma camada k ao neurônio j da camada $k+1$; e θ_j é o limiar (threshold). A função de ativação f mapeia o resultado do somatório. As funções de ativação mais utilizadas são: a função sigmóide logarítmica (bipolar ou binária), e a sigmóide tangente hiperbólica, pois o cálculo de sua derivada de primeira ordem é numericamente fácil de ser determinado, o que simplifica computacionalmente a determinação do mínimo erro e a sua implementação. A função de ativação da saída da rede é geralmente a função identidade.

O treinamento da rede pode ser realizado de modo supervisionado, não-supervisionado e híbrido. Como já foi explicado anteriormente, o processo supervisionado exige a disponibilidade de um tutor externo, enquanto, no treinamento não supervisionado, o sistema se auto-organiza e não existe um tutor externo. O tipo híbrido combina os paradigmas supervisionados e não supervisionados. Como exemplo, temos o treinamento por reforço em que a rede recebe mensagem especificando se sua saída está correta ou errada.

O treinamento supervisionado pode ser utilizado através de vários algoritmos de treinamento dependendo da arquitetura da rede: algoritmo de Perceptron, Adaline, Madaline, “*backpropagation*”, função de base radial, etc.

Iremos apresentar a seguir o tipo mais utilizado de redes neurais artificiais em caracterização de massas em mamogramas digitalizados que é o modelo de multi-camada com algoritmo de treinamento de “*backpropagation*”.

6.2.2 Redes de Multi-Camadas com Algoritmo de Treinamento de “*backpropagation*”

Uma rede de multi-camadas apresenta arquitetura com uma camada de entrada, uma ou mais camadas intermediárias, e uma camada de saída.

O treinamento por “*backpropagation*” é a regra mais utilizada em redes de multi-camadas do tipo feedforward e envolve três estágios: o estágio de “feedforward” do treinamento dos padrões de entrada, o estágio de “*backpropagation*” do erro associado, e o estágio de ajustes dos pesos.

Durante o estágio de “feedforward”, cada unidade de entrada (X_i) recebe um sinal de entrada e transmite este sinal para cada uma das unidades escondidas (Z_i) da camada intermediária. Cada unidade escondida, então, computa sua ativação e envia o sinal (z_i) para a unidade de saída. Cada unidade de saída (Y_k), computa sua ativação (y_k) para formar a resposta da rede para uma dada entrada padrão.

Durante o treinamento, cada unidade da saída compara sua ativação y_k computada com os valores alvo t_k para determinar o erro associado para aquele padrão da unidade. Baseado neste erro, o fator δ_k ($k = 1, \dots, m$) é computado. O valor de δ_k é usado para distribuir o erro da unidade de saída Y_k para todas as unidades da camada anterior (as unidades escondidas que estão conectadas aos Y_k). É também usado posteriormente para atualizar os pesos da camada de saída e da camada escondida. De maneira similar, o fator δ_j ($j = 1, \dots, p$) é computado para cada unidade escondida Z_j . Não é necessário propagar o erro anterior para a camada de entrada, mas δ_j é usada para atualizar os pesos entre a camada de entrada e a camada escondida.

Após todos os fatores δ terem sido determinados, os pesos de todas as camadas são ajustados simultaneamente. Os ajustes dos pesos w_{jk} (da unidade escondida Z_i para a unidade de saída Y_k) é baseado no fator δ_k e na ativação z_j da unidade Z_j . O ajuste dos pesos v_{ij} (da unidade de entrada X_i para a unidade escondida Z_j) é baseado no fator δ_j e na ativação x_i da unidade de entrada.

A função de ativação para uma rede com “*backpropagation*” deve ter diversas características importantes: ser contínua, diferenciável, e monotonicamente não decrescente. Entretanto, para eficiência computacional, é desejável que sua derivada seja de fácil cálculo. A função de ativação mais comumente utilizada é a sigmóide binária (logarítmica) e a sigmóide tangente hiperbólica.

O algoritmo de treinamento por “*backpropagation*” é descrito a seguir [Fausett-94]:

Passo 0: Inicializa-se os pesos

Passo 1: Enquanto a condição de parada é falsa, faça os passos 2 à 9,

Passo 2: Para cada par de treinamento, faça os passos 3 à 8,

Passo 3: Cada unidade de entrada (X_i , $i=1,\dots,n$) recebe-se o sinal de entrada x_i e transmite-se este sinal para todas as unidades da camada posterior (camada escondida),

Passo 4: Para cada unidade escondida (Z_j , $j=1,\dots,p$) soma-se os seus sinais de parada ponderado,

$$z_{inj} = v_{0j} + \sum_{i=1}^n x_i v_{ij}, \quad (6.3)$$

Aplica-se a função de ativação para calcular o seu sinal de saída,

$$z_j = f(z_{inj}), \quad (6.4)$$

e envia-se este sinal para todas as unidades da camada posterior.

Passo 5: Para cada unidade de saída (Y_k , $k=1,\dots,m$) soma-se o seu sinal de entrada ponderado,

$$y_{in k} = w_{0k} + \sum_{j=1}^p z_j w_{jk}, \quad (6.5)$$

Aplica-se a função de ativação para calcular o seu sinal de saída,

$$y_k = f(y_{in k}), \quad (6.6)$$

Passo 6: Determina-se o erro de “backpropagation”:

Para cada unidade de saída (Y_k , $k=1,\dots,m$) recebe-se um valor alvo padrão correspondente ao valor padrão de treinamento de entrada e calcula-se os seus termos de informação de erro,

$$\delta_k = (t_k - y_k) f'(y_{in k}), \quad (6.7)$$

Calcula-se os termos de correção de pesos, usados para atualizar w_{jk} posteriormente,

$$\Delta w_{jk} = \alpha \delta_k z_j, \quad (6.8)$$

Calcula-se o termo de correção de bias, usado para atualizar w_{0k} posteriormente,

$$\Delta w_{0k} = \alpha \delta_k, \quad (6.9)$$

E emite-se δ_k para unidades da camada posterior.

Passo 7: Para cada unidade escondida (Z_j , $j=1,\dots,p$), soma-se suas entradas delta (das unidades da camada anterior),

$$\delta_{inj} = \sum_{k=1}^m \delta_k w_{jk}, \quad (6.10)$$

Multiplica-se pela derivada de sua função de ativação para calcular seu termo de informação de erro,

$$\delta_j = \delta_{inj} f'(z_{inj}), \quad (6.11)$$

Calcula-se o termo correção de peso, usado para atualizar v_{ij} posteriormente,

$$\Delta v_{ij} = \alpha \delta_j x_i, \quad (6.12)$$

E calcula-se os termos de correção de bias (usado para atualizar v_{0i} posteriormente),

$$\Delta v_{0j} = \alpha \delta_j \quad (6.13)$$

Passo 8: Atualizar os pesos e biases: Para cada unidade de saída (Y_k , $k=1,\dots,m$):

Atualiza-se os bias e pesos ($j=0,\dots,p$):

$$w_{jk} (\text{novo}) = w_{jk} (\text{velho}) + \Delta w_{jk}. \quad (6.14)$$

Para cada unidade escondida (Z_j , $j = 1,\dots,p$) atualiza-se os bias e pesos ($i = 0,\dots,n$):

$$v_{ij} (\text{novo}) = v_{ij} (\text{velho}) + \Delta v_{ij} \quad (6.15)$$

Passo 9: Testa-se a condição de parada.

A rede de multi-camadas com treinamento por “*backpropagation*” é o método mais utilizado, pois é de fácil implementação. De acordo com o teorema de Kolmogorov, utilizando a rede feedforward com três camadas (entrada, intermediária e saída), pode-se aproximar-a de qualquer função contínua ou booleana [Beale-90]. A rede possui capacidade de generalização de dados mesmo não apresentados a rede anteriormente, isto é, estados intermediários de padrões que possuem características semelhantes aos dados apresentados anteriormente, são classificados corretamente. Outra vantagem da rede é que ela é tolerante a falhas e ruídos. A rede pode executar a re-aprendizagem, onde novos caminhos para se atingir mínimos locais desejados serão alcançados; uma vez que a informação é processada por elementos de processamento paralelo e distribuído, então danos parciais podem não causar mudanças significativas nas respostas da rede.

Uma das desvantagens é que o treinamento por “*backpropagation*” é muito lento em superfícies de erros com formas complexas, por exemplo com muitos mínimos locais. Para resolver esse problema deve-se considerar os efeitos de segunda ordem para o gradiente descendente [Beale-90]

Outra desvantagem é que a convergência da rede pode direcionar para um mínimo local indesejado, apresentando resultados incorretos na saída da rede. Esses problemas podem ser evitados, segundo [Beale-90], por:

- * **Taxa de aprendizado decrescente:** faz-se a taxa inicialmente alta e progressivamente faz-se o seu decréscimo. Os pesos sinápticos serão alterados mais rapidamente nos estágios iniciais de aprendizagem e mais suavemente em estágios posteriores. A desvantagem é que torna mais lento o processo de aprendizagem;
- * **Adicionar nós intermediários:** esta opção cria mais limites para as possíveis soluções eliminando-se possibilidades não viáveis. Torna, também, o treinamento mais lento;
- * **Utilizar momento:** este termo aumentará a mudança dos pesos sinápticos proporcionalmente a variação destes e aumentará a velocidade de convergência. Esta opção pode não ser boa se a variação entre pesos sucessivos tiver direções opostas.
- * **Adicionar ruídos:** permite que os mínimos locais indesejáveis sejam evitados.

Capítulo 7

Estado da Arte de Classificação de Nódulos em Mamogramas

Apresentaremos nesta secção uma revisão de trabalhos desenvolvidos por pesquisadores que têm contribuído na investigação da classificação de nódulos mamários. Os trabalhos serão divididos em processos que utilizam atributos de textura, forma, e combinados, extraídos de imagens mamográficas através de processamento de imagem. Alguns trabalhos utilizam combinações de atributos extraídos visualmente com dados clínicos. Os métodos de detecção serão apresentados, também, conjuntamente com os tipos de classificadores adotados nos trabalhos.

7.1 Classificação com Atributos de Textura

O grupo de pesquisadores do Departamento de Radiologia da Universidade de Michigan, H.P. Chan, A. Petrosian, D. Wei, N. Petrick, B. Sahiner *et al* , são os que apresentaram os principais trabalhos utilizando atributos de textura.

Em 1994, Petrosian *et al.*, utilizaram atributos de textura proposta por Haralick [Haralick-73], e aplicaram os atributos na matriz de dependência espacial de níveis de cinza , spatial-grey-level dependence matrix (SGLD) [Petrosian-94]. A capacidade de classificação das características foram analisadas pelos seus coeficientes de correlação e pela distância da variança normalizada. Foi aplicado o método de teste “leaving-one-out” no treinamento. O processo foi testado em 45 mamogramas com nódulos, cuja malignidade foi comprovada por biópsia. O

resultado da melhor classificação, de acordo com os autores, apontou para uma sensibilidade de aproximadamente 89% e 76% de especificidade.

Em 1995, D. Wei *et al.*, do mesmo grupo de pesquisa da Universidade de Michigan, investigaram a possibilidade de utilização da análise de multiresolução de textura para diferenciar os nódulos dos tecidos normais em mamogramas [Wei-95]. A transformada de wavelet foi utilizada para decompor as regiões de interesse (ROI) em várias escalas extraídas das imagens dos mamogramas. As características de textura utilizadas foram as mesmas de [Petrosian-94]. Foi utilizada a análise de discriminante linear para maximizar a separação entre nódulos e tecidos normais, selecionando características que vieram dos três espaços de textura. A habilidade dos três vetores de características para a classificação entre nódulos e tecidos normais foi comparada através da curva ROC. O método foi aplicado em 168 ROIs contendo nódulos comprovados por laudos de biópsias e 504 ROIs contendo tecidos normais. Os resultados obtidos, de acordo com os autores, foram que para características de textura baseadas nos coeficientes de wavelet e distância variável, a análise de curva ROC apontou uma área $A_z = 0.89$ e $A_z = 0.86$ para o grupo de treinamento e para o grupo de teste, respectivamente.

Em 1997, os pesquisadores D. Wei *et al.*, apresentaram um trabalho o qual utilizaram o mesmo processo de [Wei-95] em 168 mamogramas digitalizados divididos igualmente e aleatoriamente em dois grupos de treinamento e teste [Wei-97]. Destes mamogramas, dois conjuntos de imagens, todos extraídos manualmente, foram formados. O primeiro conjunto continha quatro regiões de interesse (ROI), sendo uma região de nódulo, provado por biópsia, e três regiões de tecidos normais. O segundo conjunto continha ROIs extraídas manualmente, juntamente com ROIs de tecidos normais extraídas por um algoritmo de DWCE para detecção de falso-positivo. Uma transformada de wavelet foi usada para decompor uma ROI em diversas escalas. Características de textura global foram obtidas pelo cálculo de coeficientes passa-baixas na imagem transformada. Características de textura local foram calculadas para o objeto suspeito e regiões periféricas. Modelos de discriminante linear usando características globais, locais, ou espaço de características combinando as duas, foram estabelecidos para maximizar a separação entre as classes de nódulo e tecido normal. A classificação utilizando características globais foi comparada com a classificação por características locais. De acordo com

os autores, com as mesmas características globais e locais as classificações resultaram, para o conjunto manual de mamogramas, em uma área dentro da curva ROC, de $A_z=0.92$, e para o conjunto híbrido de $A_z=0.96$, demonstrando, segundo os autores, que a combinação de características globais e locais podem efetivamente reduzir os falsos-positivo na classificação de nódulos e tecidos normais.

Em 1996, Sahiner *et al.*, apresentaram outro trabalho de investigação de caracterização de nódulo [Sahiner-96A]. Neste trabalho, utilizaram redes neurais artificiais para realizar a classificação de características extraídas através de duas técnicas. A primeira técnica empregou a média e sub-amostragem (averaging and subsampling). A segunda técnica empregou métodos de extração de características de textura aplicados em pequenas subregiões dentro de uma área de interesse (ROI). A rede neural artificial utilizada neste trabalho foi a rede neural de convolução, convolution neural network (CNN), ao invés da rede neural artificial multilayer com algoritmo de treinamento por “*backpropagation*” (BPN). Segundo os autores, a diferença entre a BPN e a CNN, em classificação de imagens, é que a CNN opera com informações das imagens; em vez de características extraídas da imagem. Para testar o classificador CNN, um conjunto de 168 ROIs contendo nódulos e 504 ROIs contendo tecidos normais foi utilizado. Esses dados foram usados para treinar e testar a rede. Com a melhor combinação da arquitetura da rede CNN e os parâmetros de textura, segundo os autores de [Sahiner-96A], resultou numa área $Az=0.87$ da curva ROC, que correspondeu a uma fração de verdadeiro-positivo de 90% e de fração de falso-positivo de 31%. Demonstrando, segundo os autores de [Sahiner-96A] que a rede CNN pode ser utilizada na classificação de nódulos e tecidos normais em mamogramas.

Em 1994, W.P. Kegelmeyer Jr *et al.*, propuseram um método para detecção de lesões em mamogramas com aparência irregular e rodeadas por espículas radiais [Kegelmeyer-94]. O método utilizado foi baseado em análise estatística. Eles utilizaram análise de histograma através de janelas de orientações de borda que são transversalmente passadas na imagem. Após a conversão do mamograma em vetores de características, um classificador de árvore de decisão binária – “binary decision tree” (BDT) - foi utilizado para marcar os pixels da imagem como pertencentes a prováveis regiões suspeitas. A principal característica utilizada para a detecção de lesões espiculadas foi a análise da orientação de borda local – “analysis of local

oriented edges” (ALOE) - que mediou a textura de energia baseada em uma análise de histograma de orientações de bordas em uma janela local. O processo de detecção de lesões espiculadas foi testado em 85 casos com as 4 vistas (crânio-caudal e médio-lateral-obliqua), sendo 36 casos com carcinoma comprovado por biópsia e 49 casos negativos de carcinoma. A detecção ainda realizou a reprodução em filmes das lesões suspeitas. Quatro radiologistas analisaram os filmes com as lesões suspeitas destacadas e os filmes dos mamogramas originais. O algoritmo conseguiu uma sensibilidade de 100%, com uma especificidade de 82% para a detecção das lesões suspeitas nos 85 casos testados. O processo aumentou a sensibilidade dos radiologistas em 9.7% , passando a sua média de 80.6% para 90.3%, sem alteração de especificidade.

7.2 Classificação com Atributos de Forma

Laurens V. Ackermann *et al.*, foram os primeiros pesquisadores que desenvolveram medidas quantitativas para avaliar morfológicamente o tumor em mamas [Ackermann-72]. Eles investigaram quatro medidas de malignidade: formato espiculado do tumor, rugosidade de fronteira, forma, e presença de calcificações. Foram calculadas 20 medidas distribuídas da seguinte forma: 10 medidas de razão de perímetro pela área, variando-se os valores de threshold; 4 medidas de propriedade de espicularidade, variando-se a janela={2,4,8,16}; 4 medidas de propriedade de rugosidade, variando-se a janela={2,4,8,16}; e 2 propriedades da presença de calcificação. Para detectar a presença de calcificação, duas técnicas foram utilizadas. Uma análise de histograma foi construído sobre a imagem inteira. Ao eixo das abscissas foi atribuído a diferença de densidade absoluta entre pontos adjacentes, e ao eixo das ordenadas foi atribuído o número de ocorrências dessa particular diferença de densidade. Um valor de threshold foi definido como o ponto no histograma que o dividiu entre 85% para valores abaixo do threshold e 15% acima do valor de threshold. Três métodos de decisão foram usados para analisar os dados: Análise Gaussiana Multivariada e duas Técnicas do Vizinho Mais Próximo (KNN). Para cada método, diferentes combinações de propriedades foram experimentadas com diferentes parâmetros de controle da técnica KNN. Os resultados foram então plotados na forma de curvas ROC. Foram utilizadas 120 imagens xeroradiográficas

digitalizadas, sendo 60 com lesões malígnas e 60 com lesões benignas. Todas os mamogramas foram acompanhados de exames patológicos comprovando a malignidade ou não das lesões. Um radiologista determinou visualmente a área suspeita antes da realização da digitalização dos mamogramas utilizados na classificação. Dois radiologistas fizeram a classificação manual dos mesmos mamogramas para servirem de referência para o desempenho do sistema computadorizado. Os resultados obtidos, segundo os autores de [Ackermann-72], foram que os classificadores KNN tiveram melhor desempenho do que o classificador de Análise Gaussiana Multivariada.

Em 1990, D. Brzakovic *et al.*, apresentaram uma proposta utilizando classificação com análise bayesiana [Brzakovic-90]. Eles descreveram no trabalho um sistema de duas etapas: a primeira etapa detectou regiões de interesse com agrupamento de pixels com características suspeitas, usando o processamento de imagem de multiresolução baseada em “Fuzzy Pyramid Linking”; e a segunda etapa empregou a classificação hierárquica para identificar lesões benignas ou malignas. A etapa de detecção primeiro identificou o agrupamento de pixels que tinham diferentes propriedades de intensidades; o classificador extraiu regiões em três classes: não lesão, lesão benigna e lesão maligna. Ele foi aplicado dentro do processo hierárquico. Em cada nível extraiu as medidas de área, descritor de forma, descritor de variação de distância de borda, e variação de intensidade de borda. O processo de classificação requereu uma fase de aprendizagem do classificador de Bayes. O processo de aprendizagem envolveu a obtenção da média e do desvio-padrão para cada função de probabilidade de densidade (distribuição Gaussiana), através de uma amostra de imagens sendo previamente conhecidos os tipos de lesões. Foram processados para a avaliação do esquema computadorizado 25 mamogramas, sendo 5 sem lesão nenhuma e 20 com lesões de todos os tipos. Fora deste conjunto, mais 10 amostras de mamogramas (5 com lesão maligna e 5 com lesão benigna) foram utilizados para treinamento para obter a média e o desvio-padrão. Na primeira etapa, o sistema computadorizado conseguiu detectar 95% das lesões nas imagens mamográficas. Na segunda etapa, o classificador conseguiu diferenciar entre três classes: imagens sem lesão, imagens com lesão benigna e imagem com lesão maligna. O classificador alcançou 85% de acerto dos casos.

Em 1990, os pesquisadores M.L. Giger *et al.*, do “Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, University of Chicago”, apresentaram um trabalho de detecção e classificação de nódulos em mamogramas. A detecção utilizou um esquema computadorizado chamado de Técnica de Subtração Bilateral. As características foram extraídas através de um esquema que envolveu filtros morfológicos, teste de tamanho, teste de fronteira e teste de correlação. A classificação dos atributos foi realizada utilizando o grau de espicularidade exibida pelo nódulo [Giger-90A] [Giger-90B]. A localização dos nódulos mamográficas envolveu a ajuda de um radiologista. Feito isso, o nódulo foi então extraído automaticamente do fundo da imagem ao seu redor através da técnica de crescimento de região, resultando em uma imagem binária. Para quantificar o grau de espicularidade, informações de borda foram extraídas. Foram testados 36 pares de imagens de mamogramas, sendo empregadas as duas vistas crânio-caudal (CC) e médio-lateral-oblíqua (MLO) dos mamogramas de 18 pacientes. As técnicas de extração de características empregadas reduziram substancialmente o número de detecção de falsos-positivo. Esse estudo produziu uma taxa de verdadeiro-positivo de 95% (sensibilidade) com uma média de 3 falsos-positivo por imagem. O classificador foi aplicado em 35 nódulos mamográficas. De acordo com os autores, as duas medidas de espicularidade correlacionaram-se muito bem com as medidas obtidas por radiologistas. Os autores observaram que se o teste fosse realizado numa base maior de imagens poderiam melhorar a investigação desse processo de classificação.

Em 1991, F.F. Yin *et al.*, pesquisadores do mesmo laboratório do “Kurt Rossmann Laboratories for Radiologic Image Research”, apresentaram trabalhos semelhantes em dois artigos com a mesma técnica de detecção e classificação para 60 pares de mamogramas, sendo 18 de pacientes com carcinomas e 12 de pacientes com casos normais acompanhados de resultados de exames de biópsia[Yin-91A] [Yin-91B]. Os resultados da detecção foram iguais aos obtidos em [Giger-90A] e [Giger-90B], taxa de 95% de verdadeiro-positivo para uma média de 3 falsos-positivo por imagem. Os resultados da classificação apresentaram boa concordância entre a média de espicularidade calculada pelo esquema automático e a avaliação subjetiva de um radiologista. A análise de espicularidade de nódulo indicou uma

taxa de verdadeiro-positivo de 97% e uma taxa de falso-positivo de 79%, contra uma taxa de falso-positivo de 95% de classificação realizada pelo radiologista.

Em 1993, J. Kilday *et al.*, pesquisadores do “Department of Electrical and Systems Engineering, University of Connecticut”, desenvolveram um classificador para determinar três tipos mais comuns de lesões mamárias, fibroadenomas, cistos e carcinoma [Kilday-93]. As lesões foram segmentadas usando a técnica de limiarização simples para destacar os contornos da lesão. Foram extraídas sete descritores de forma da imagem mamográfica: circularidade, média de comprimento radial normalizado, desvio-padrão do comprimento radial, entropia do histograma de comprimento radial, razão de área, “contagem de cruzamento zero”, e rugosidade de fronteira. A idade da paciente foi a única informação clínica utilizada no estudo, como característica na classificação do processo. Para a seleção de características a técnica de análise de discriminante linear foi utilizada para distinguir entre diferentes classes de objetos baseados num conjunto pré-selecionado. O algoritmo de classificação usado foi o não paramétrico (LDA) e dependeu da distância entre os membros das classes em geral. Foram adotados os métodos de “leaving-one-out” e o “método aparente” (apparent method) para treinar e testar o algoritmo de classificação. Um total de 82 filmes mamográficos foram selecionados , sendo que 25 possuíam cistos, 38 cânceres e 19 fibroadenomas. O melhor resultado obtido pelo teste de método aparente resultou na classificação correta de 82% das lesões segmentadas usando todas as características. Para o teste utilizando o método “leaving-one-out” o melhor resultado de classificação correta foi de 69% usando como características a idade da paciente, a razão de área, a contagem de cruzamentos zero, e a medida de rugosidade de fronteira. O resultado da classificação com todas as características, excluindo a idade da paciente, foi de 72% utilizando o teste do “método aparente” e de 46% utilizando o método “leaving-one-out”.

Em 1995, Z. Huo *et al.*, do “Kurt Rossmann Laboratories for Radiologic Image Research”, apresentaram um trabalho de classificação de nódulo mamário computadorizada baseada na análise de espicularidade chamada de análise de gradiente radial de borda [Huo-95B]. O esquema consiste de duas etapas: um método automático de segmentação da lesão mamográfica do parênquima circunvizinho e a extração de características usando uma nova técnica desenvolvida. Duas medidas foram obtidas utilizando a técnica: FWHM (full width a half maximum) e o gradiente

radial normalizado. A performance de cada medida de espicularidade foi testada usando a curva ROC que avaliou as suas habilidades para determinar a probabilidade de malignidade de uma nódulo. Analisaram, também, a dependência da performance dessas medidas na escolha das vizinhanças. Para realizar a análise desse esquema de classificação, os pesquisadores utilizaram uma base de dados de 95 mamogramas clínicos contendo lesões, sendo 38 com lesões benignas e 57 com lesões malignas. Todos os casos estavam acompanhados de seus respectivos exames de biópsia. Um radiologista experiente foi convocado para classificar manualmente os graus de espiculação das nódulos nos mamogramas para comparação com o sistema computadorizado. Os resultados obtidos, de acordo com os autores do [Huo-95B], foram que: (a) o esquema computadorizado produziu, para a extração de características, valores mais altos usando a vizinhança da lesão. Obtiveram valores de área A_z na curva ROC de 0.83 e 0.85, respectivamente, para vizinhança sem correção e com correção de FWHM. Esse valor foi semelhante ao valor obtido por um radiologista experiente que obteve valor de A_z na curva ROC de 0.85; (b) quando foram combinados os quatro tipos de vizinhanças com o FWHM corrigido, obtiveram um resultado de 0.88 da área A_z na curva ROC.

No mesmo ano, em 1995, Z. Huo *et al.*, em outro trabalho, apresentaram um artigo em que relataram o desenvolvimento de um esquema computadorizado para extrair características relacionadas com as margens, formas e densidades de lesões mamárias, similares às características utilizadas pelos radiologistas na interpretação de mamogramas [Huo-95A]. O método computadorizado caracterizou as nódulos encontradas em mamogramas pela espicularidade. Os resultados obtidos pelo esquema automático foram comparados com os resultados obtidos por um radiologista experiente. Para melhor distinguir as nódulos benignos de malignos com bordas levemente espiculadas, características de densidade relativa foram usadas tanto para nódulos de alta densidade e baixa densidade, como para nódulos com densidade uniforme ou não uniforme. Uma rede neural artificial foi utilizada como meio para combinar as múltiplas características obtidas pelo sistema automático e na saída determinou a probabilidade de malignidade. Para analisar o método apresentado, 95 mamogramas foram testados, sendo 57 com lesões malignas e 38 com lesões benignas. O resultado apresentado pelo autores foi que o método computadorizado apresentou sensibilidade de 100% para uma especificidade de 58%,

dando uma biópsia positiva de 75%. A rede neural artificial teve performance avaliada pela curva ROC, obtendo-se, de acordo com os autores, valor de área $A_z = 0.90$, o que foi melhor do que o valor obtido pelo radiologista que conseguiu um valor de $A_z = 0.83$.

Em 1995, Vic Velanovich, apresentou uma forma de caracterizar lesões malignas e benignas através da dimensão fractal [Velanovich-95]. Foram utilizados mamogramas de 10 pacientes com doenças benignas (fibroadenoma, adenose esclerosada, doenças fibrocística não proliferativa, etc); e mamogramas de 10 pacientes com carcinoma ductal. Todos acompanhados com laudos de exames de biópsia. Foram selecionados somente nódulos mamográficos discretos e lesões que apresentavam misturas com microcalcificações foram rejeitadas. A média da dimensão fractal composta para lesões benignas foi de 1.831 (taxa de 1.359 a 2.009) e para lesões malignas foi de 2.477 (taxa de 2.084 a 3.158). Nestes 20 mamogramas analisados, as lesões malignas tinham dimensão fractal maiores do que 2.084; e as lesões benignas tinham dimensão fractal menores do que 2.009. De acordo com o autor do [Velanovich-95], a dimensão fractal foi perfeitamente discriminatória entre as lesões malignas e benignas.

Em 1996, os pesquisadores R.M. Rangayyan, N. El-Faramawy, *et al*, apresentaram um outro trabalho de classificação de nódulos mamários [Rangayyan-96]. O algoritmo de classificação foi baseado na distância de Mahalanobis e com treinamento da técnica “leaving-one-out”. A diferença é que apresentaram uma nova medida de espicularidade chamada de “measure of edge profile acutance” (medida de borda com contorno pontiagudo). Foram testados 39 imagens, sendo 16 CB, 4 CM, 12 SB, e 7 SM, todos comprovados por biópsia. As imagens foram selecionadas da mesma base de dados da “Mammographic Image Analysis Society” (MIAS, UK). Para aumentar o número de CM e SM, 15 imagens foram combinadas ao conjunto de testes, imagens obtidas do “Alberta Program for the Early Detection of Breast Cancer” (APEDBC). De acordo com os autores, a medida de borda para contorno pontiagudo obteve taxa de 94.9% de acerto de classificação para o conjunto de imagens do MIAS (39 imagens) e 92.6% para o conjunto de imagens MIAS+APEDBC. Uma taxa de verdadeiro-negativo de 92.9% e uma taxa de verdadeiro-positivo de 100% para a base do MIAS, e 92,9% e 92,3% para MIAS+APEDBC. Para a característica de solidez, de acordo com os autores, apresentou um resultado

de taxa de acerto de 61.5% para o MIAS e 72.2% para MIAS+APEDBC. Combinando a medida de borda com contorno pontiagudo (A) com as características de fator de forma solidez (C), medidas baseadas em descritores de Fourier (FF), e medidas baseadas em momento (MF), os autores obtiveram os resultados: combinando (A e C) , obtiveram taxa de acerto de 84.6% para o MIAS e 83.3% para o MIAS+APEDBC; e combinando (A, FF e MF), obtiveram taxa de acerto de 94.9% para o MIAS e 87% para o MIAS+APEDBC .

Em 1996, os pesquisadores S. Pohlman *et al*, apresentaram um trabalho que segmentou as lesões com a técnica de crescimento de região, extraiu e selecionou os atributos através de seis descritores morfológicos: circularidade, solidez, classificador de momento normalizado, flutuação de margem, dimensão fractal e rugosidade de contorno do tumor (TBR - tumor boundary roughness) [Pohlman-96]. A dimensão fractal, segundo os autores, foi baseada no conceito desenvolvido por Mandelbrot e tem sido usada para descrever a rugosidade de superfícies [Mandelbrot-84]. A dimensão fractal é a quantidade que pode ser usada para descrever a irregularidade da forma ou textura de um objeto. Os autores usaram uma variação do método de Mandelbrot, proposto por Dubuc [Dubuc-87] [Dubuc-89], chamada de rugosidade de contorno do tumor (TBR). Utilizaram mamogramas de 48 pacientes, resultando em 51 lesões, sendo 25 benignas e 26 malignas, todos acompanhados de laudo de exame de biópsia. Todos os mamogramas foram selecionados do registro de pacientes do “Cleveland Clinic Foundation, Cleveland”, Ohio. A seleção de cada atributo foi analisada através da curva ROC e comparada com os outros atributos, e também comparada com resultados obtidos com dois radiologistas experientes. Na etapa de segmentação, de acordo com os autores de [Pohlman-96], 97% das lesões foram segmentadas. Na etapa de classificação, os resultados da análise indicaram que as seis características poderiam servir como parâmetros de caracterização de malignidade com áreas A_z da curva ROC variando de 0.759 a 0.928. O diagnóstico de classificação do modelo baseado nestas características foram similares aos obtidos pelos dois radiologistas (A_z de TBR = 0.928; A_z do 1º radiologista = 0.969; e A_z do 2º radiologista = 0.932)

Em 1996, os pesquisadores Petrick *et al.*, da Universidade de Michigan, apresentaram um trabalho de investigação de segmentação de áreas suspeitas em mamogramas digitalizados usando um filtro adaptativo de realce de densidade ponderada, “adaptative density-weighted contrast enhancement” (DWCE), em conjunção com o detector de borda Laplaciano-Gaussiano (LG) [Petrick-96A]. O DWCE realçou estruturas dentro do mamograma digitalizado tal que um simples algoritmo de detecção de borda poderia ser utilizado para definir os contornos do nódulo no mamograma. Uma vez conhecido o contorno do nódulo, características morfológicas puderam ser extraídas e usadas na classificação de diferentes regiões dentro da imagem. Foram utilizados classificadores baseados em limiarização (thresholding) sequencial, análise de discriminante linear (LDA), e uma rede neural artificial do tipo multilayer-backpropagation (BPN) para reduzir a detecção de falso-positivo. O processo foi aplicado em 25 sub-amostras de mamogramas digitalizados acompanhados de laudos de biópsia. Após o estágio de segmentação, 461 objetos separados foram obtidos e suas características morfológicas calculadas. Os conjuntos de características foram usados para classificar os objetos como nódulos potenciais, utilizando os esquemas de classificação com Limiarização, LDA, e BPN. O melhor resultado obtido, segundo os autores, foi a detecção da taxa de verdadeiro-positivo de 96% com 4.5 falsos-positivo por imagem para o classificador LDA, porém com diferenças desprezíveis em relação aos outros classificadores Limiarização e BPN. De acordo com os autores, o método por DWCE mostrou-se eficiente para detectar nódulos mamários em mamogramas digitalizados.

7.3 Classificação com Atributos de Textura e Forma

Em 1995, os pesquisadores H.D. Li et al., da Universidade do Sul da Flórida, propuseram uma técnica para detecção de tumores suspeitos em mamogramas digitalizados [Li-95]. Aproveitaram as pesquisas de detecção de nódulos realizadas em 1993 [Li-93], que segmentavam as regiões de interesse através do processo de campo aleatório de Markov – “Markov random fields” (MRF), porém sem classificá-las em nódulo suspeita ou não. Neste trabalho, o processo de segmentação foi aperfeiçoado com a introdução da análise de multiresolução, o que aumentou a velocidade de segmentação da área de interesse. Esta área de interesse segmentada

foi , então, classificada em regiões suspeitas e normais, através da aplicação de uma árvore de decisão binária fuzzy (fuzzy binary decision tree). Os atributos de forma utilizados são área, gradiente médio da fronteira da região, solidez e variação de distância de borda; e os atributos de textura utilizados são diferença de intensidade média e variação de intensidade O processo foi aplicado em 95 mamogramas com variações na densidade e estrutura parenquimal, sendo 48 com carcinomas comprovados por biópsias. O treinamento do classificador foi realizado em 20 mamogramas selecionados, sendo 13 com carcinomas. O desempenho de detecção foi avaliada pela curva ROC sendo explicitado o relacionamento entre a taxa de detecção de nódulos de verdadeiro-positivo e o número de falso-positivo por imagem. Segundo os autores, os resultados indicaram uma sensibilidade de 95% na detecção de diferentes tipos de nódulos com dois falsos-positivo por imagem. O algoritmo teve um desempenho de 94% de sensibilidade na detecção de lesões com tamanho menores que 10 mm, com 1.5 de falsos-positivo por imagem.

Em 1996, os pesquisadores N. El-Faramawy *et al.*, apresentaram um trabalho de investigação de atributos de fator de forma usando momentos, descritores de Fourier, descritores de Hu, estatística de comprimento de corda (chord length statistics), e grau de solidez; e atributos de textura utilizando momentos e descritores de Hu, em níveis de cinza da imagem [El-Faramawy-96]. A classificação de nódulos mamográficas foi realizada através de um algoritmo que calcula a distância de Mahalanobis e o treinamento foi realizada com a técnica “leaving-one-out”. Eles classificaram as lesões em quatro grupos: lesões circunscritas benignas (CB), lesões circunscritas malignas (CM), lesões espiculadas benignas (SB), e lesões espiculadas malignas (SM). Eles propuseram um novo fator de forma baseado em momento de característica de textura, que indicou, de acordo com os autores, uma excelente correlação destas características com a medida de rugosidade de calcificações e invariante em relação a variação de tamanho, rotação, e translação do objeto. Aplicaram o esquema em 54 tumores acompanhados de laudo de biópsia, sendo 16 CB, 7 CM, 12 SB, e 19 SM. As imagens foram obtidas da base de dados da “Mammographic Image Analysis Society” (MIAS, UK) e da base do “Alberta Program for the Early Detection of Breast Cancer” (APEDBC). Os melhores resultados obtidos, segundo os autores, para classificação de lesões como circunscritas/espiculadas, a taxa de acerto foi de 94,4% para a base de dados do



MIAS e 95% para a base combinada do MIAS+APEDBC; para a classificação de lesão maligna/benigna, a taxa de acerto foi de 77% para a base do MIAS e 69% para a base combinada MIAS+APEDBC. Sendo que o primeiro resultado de circunscrita/espiculada, o grupo de imagens era específico para cada grupo analisado, enquanto para os resultados de maligna/benigna, foram aplicados os quatro grupos combinados.

Em 1996, Z. Huo *et al.*, pesquisadores do “Kurt Rossmann Laboratories for Radiologic Image Research”, investigaram um esquema computadorizado incluindo segmentação e extração de quatro características automáticas (espicularidade, rugosidade, densidade e textura) e três classificadores automáticos de estimativa de probabilidade de malignidade do nódulo (método de regra-básica, uma rede neural artificial, e um sistema híbrido) [Huo-96]. A segmentação da lesão mamária foi obtida utilizando pontos de múltipla transição e a técnica de crescimento de região, já explicadas no artigo [Huo-95B]. Após a segmentação da lesão, as quatro características foram extraídas automaticamente das suas vizinhanças. Uma medida quantificou o grau de espicularidade do nódulo e foi obtida utilizando a técnica de gradiente radial de borda, detalhada no artigo [Huo-95B]. A medida de rugosidade de margem foi calculada pelo gradiente médio ao longo da margem identificada pelo computador. As medidas de densidade do nódulo e a textura foram quantificadas pela média de nível de cinza dentro do nódulo. A medida de textura foi obtida pelo desvio-padrão do gradiente dentro do nódulo. A avaliação da habilidade destas características individuais em distinguirem entre nódulos malignos e benignos foi realizada através de curvas ROC. As quatro características foram combinadas e aplicadas em três diferentes classificadores. Em um classificador com método de duas etapas de regra-básica, a característica de medida de espicularidade foi aplicada na primeira etapa para classificar as lesões em espiculadas e não espiculadas, então uma segunda característica, escolhida entre as três características remanescentes, foi aplicada para melhorar a classificação de nódulos não espiculadas. Num segundo classificador, uma rede neural artificial, do tipo multilayer com três camadas, com algoritmo “*backpropagation*” de treinamento, foi empregado. No terceiro classificador, um sistema híbrido foi utilizado aplicando primeiramente uma regra-básica na medida de espicularidade, e as outras três características remanescentes foram aplicadas na rede neural artificial, com três unidades de entrada, duas unidades

escondidas e uma unidade de saída, para melhorar a classificação de nódulos não espiculadas. Foram analisados 95 mamogramas contendo nódulos de 65 pacientes, sendo 26 pacientes tinham doenças mamárias benignas e 39 pacientes tinham diagnósticos de carcinomas mamários, acompanhados de resultados de laudos de biópsia. Um radiologista experiente e 5 radiologistas residentes, realizaram os testes visualmente para que fosse possível a comparação com o esquema computadorizado. De acordo com os autores, a análise de características mostrou que a medida de espiculação, produziu um valor da área na curva ROC de $A_z = 0.88$ comparativamente às outras características que obtiveram valores de área A_z da curva ROC entre 0.54 a 0.65. Com isso, concluímos que a medida de espicularidade é a mais importante característica, tanto para classificação visual quanto para a automática, pois correlacionou-se bem com o valor de área da curva ROC de $A_z = 0.85$, obtido pelo radiologista experiente. A performance dos classificadores foi avaliada com a análise da curva ROC em termos de área A_z e sua performance de 100% de sensibilidade em comparação com o radiologista experiente e com os 5 radiologistas residentes. O classificador com sistema híbrido desempenhou melhor que a média dos 5 radiologistas residentes e similarmente ao radiologista experiente. O sistema híbrido produziu um A_z de 0.94 e uma especificidade de 69.2% e 100% de sensibilidade, enquanto o radiologista experiente produziu $A_z=0.91$ e a média dos 5 radiologistas com pouca experiência produziram valores de $A_z= 0.81$.

Em 1996, Petrick *et al.* apresentaram um outro trabalho em que utilizaram a mesma técnica de segmentação por DWCE - density-weighted contrast enhancement aplicada em uma única vista mamográfica para realçar as estruturas importantes e um detector de borda Laplaciano-Gaussiano (LG) combinado para realizar a identificação da região de interesse [Petrick-96B]. Uma vez que a região potencial do nódulo foi identificada, foi realizada a extração de características por multiresolução de textura por coeficientes de wavelet, como já investigado em [Wei-95]; e para a classificação foi utilizada também análise de discriminante linear (LDA) para caracterizar as nódulos mamárias de tecidos normais [Petrick-96A]. Neste estudo utilizaram 168 mamogramas de uma única vista, sendo 85 com nódulos malignas e 83 com nódulos benignas. As 168 imagens foram divididas igualmente em dois grupos de 84 imagens, sendo um grupo utilizado para treinamento e outro para teste, e os resultados avaliados por análise da curva ROC. A seguir os grupos

foram invertidos, segundo os autores, passando o grupo que era de teste para o grupo de treinamento e vice-versa. Segundo os autores, os testes resultaram na produção de aproximadamente 4.4 falsos-positivo por imagem para uma taxa de verdadeiro-positivo de 90% no primeiro treinamento e teste; e 2.3 falsos-positivos por imagem para uma taxa de verdadeiro-positivo de 80% no outro treinamento e teste com os grupos invertidos.

7.4 Classificação com Outros Tipos de Atributos

Em um artigo de 1973, L.V. Ackerman *et al.*, propõem a classificação de lesões mamárias suspeitas em malignas ou benignas, através do processamento computacional de informações obtidas de um questionário especialmente respondido por um médico especialista que analisou 102 imagens mamográficas, divididas em 51 com lesões malignas e 51 com lesões benignas, escolhidas aleatoriamente do banco de imagens do Rush-Presbyterian-St Lukes Medical Center, Chicago, Illinois, nos EUA [Ackerman-73]. Pelas propriedades obtidas, uma probabilidade de malignidade foi determinada para cada caso, usando um algoritmo de análise de clusters com distribuição gaussiana para cada cluster de categorias malignas e benignas. O algoritmo foi treinado com metade de cada caso (30 imagens com lesões malignas e 30 com lesões benignas) e a outra metade foi utilizada como teste dos parâmetros obtidos com os valores treinados. Cada mamograma utilizado no estudo tinha um laudo previamente diagnosticado patologicamente e os resultados das biópsias realizadas para os casos suspeitos. Usando as propriedades e o método de decisão descrita no estudo de [Ackerman-73], a taxa de falso-positivo foi de 55%, para uma taxa de falso-negativo igual a zero. De acordo com os autores, isso sugere que a metodologia descrita no trabalho pode auxiliar no diagnóstico da malignidade de doenças da mama.

Em 1987 G. Gale *et al.*, pesquisadores do “Department of Radiology, University of Nottingham”, realizaram um trabalho para melhorar a interpretação das especificidades de uma mamografia, melhorando os detalhes de sinais importantes na determinação da malignidade das lesões encontradas nos mamogramas [Gale-87]. Para isso, primeiramente foi criada uma base de dados de 500 pacientes, sendo 260 com resultados de lesões malignas e 240 com resultados de lesões benignas, todos

comprovados através de exames de biópsia. Foram selecionadas 39 características referentes às propriedades padrões de lesões. Através de um algoritmo de análise de função discriminante, implementado em computador, estudaram os sinais característicos e seus relacionamentos com as doenças malignas e benignas da mama. A análise discriminante inicial identificou 9 características importantes na distinção entre lesões malignas e benignas. Através de estudo estatístico de validação-cruzada, identificaram mais 3 características importantes, resultando, então, na redução do número de características de 39 para 12 no total. Essas 12 características foram novamente aplicadas em cada caso individualmente, no sentido de estabelecer a probabilidade de prognóstico correto entre lesões malignas e benignas. Com isso, os detalhes dos sinais foram melhorados e obtiveram um prognóstico correto de 88% de casos benignos e 79% de casos malignos. Compararam esses resultados com os realizados pelos radiologistas de 49% para os casos benignos e 87,2% para os casos malignos. Apesar do resultado dos casos malignos apresentarem pior desempenho (78% contra 87,2%) , os autores enfatizaram o resultado dos casos benignos por terem alcançado melhor desempenho (88% contra 49%). Isso se deve, segundo os autores, à diminuição da taxa de falso-positivo de 21% alcançado, sendo 33,5% obtidos pelos radiologistas e somente 12,4% para o método empregado pelos pesquisadores. Desta forma, os autores argumentaram que, com a utilização do método apresentado, as biópsias de menos 107 casos estudados poderiam ter sido evitadas.

Em 1993, Y. Wu *et al* , pesquisadores do “Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, University of Chicago”, apresentaram dois trabalhos utilizando redes neurais artificiais (ANN) para auxiliar na tomada de decisão no diagnóstico de lesões em mamografia, através de características extraídas manualmente[Wu-93A][Wu-93B]. Foram selecionadas 43 características radiográficas distintivas entre lesões malignas e benignas. A arquitetura da rede utilizada foi a de três camadas, com algoritmo “*backpropagation*” de treinamento para interpretação das características extraídas dos mamogramas. A performance da rede neural foi avaliada pela análise da curva ROC. Foi aplicado, também, o método de “leaving-one-out” para testar a habilidade da rede neural. Os pesquisadores verificaram, também, que algumas características produziam pequenas ou nenhuma diferença nos valores médios entre casos malignos

e benignos, devido a seleção original não ter levado em consideração a redundância de características e pelo fato de que algumas características não serem discriminantes para a análise pretendida. Com isso, o número reduziu de 43 para 14 características. Esse novo conjunto de características foi assim distribuído: 5 características relacionadas com nódulos, 6 características relacionadas com as microcalcificações encontradas, e 3 características relacionadas com anormalidades secundárias. De acordo com os autores, a performance da rede neural, $A_z = 0.89$, foi mais alta do que a média de performance dos radiologistas especialistas, $A_z = 0.84$, e radiologistas residentes, $A_z = 0.80$, na classificação das lesões. Eles consideraram que a diferença entre os radiologistas atendentes e os residentes foi substancial, mas estatisticamente não significante.

Em 1997, J.Y. Lo et al, da “University of Durham”, apresentaram um trabalho utilizando a rede neural artificial (ANN) para a classificação de lesões em mammogramas [Lo-97]. Eles selecionaram 10 características extraídas dos mamogramas e das fichas cadastrais das pacientes. Foram selecionadas 266 lesões, sendo 170 benignas e 96 malignas, acompanhadas dos laudos de biópsias. A arquitetura da rede neural artificial utilizada foi a multilayer Perceptron (MLP) de três camadas, com a camada de entrada com 10 unidades correspondentes a cada característica selecionada; uma camada intermediária com 15 unidades escondidas; e uma camada de saída com uma unidade de saída. O treinamento da rede foi do tipo “*backpropagation*” e foi aplicado o método de “leaving-one-out” para testar a habilidade da rede neural. Para análise da performance foi utilizada a curva ROC. Eles treinaram a rede, primeiramente, somente para os casos de lesões malignas e obtiveram resultados de sensibilidade de $91\% \pm 3\%$ para as lesões com carcinoma *in situ* e invasivo. A partir dos resultados obtidos na saída da rede neural, um histograma foi plotado. Analisando o histograma, selecionaram um valor ótimo de threshold igual a 0.825, pois para este valor os carcinomas do tipo invasivo podem ser separados dos carcinomas *in situ* e consequentemente de lesões benignas. Neste caso, a sensibilidade foi de 71% e especificidade de 100% para os carcinomas invasivos. Com o valor de threshold de 0.825, a rede foi treinada com todos os 96 casos de lesões malignas e benignas e, de acordo com os autores, obtiveram uma sensibilidade de 97% e especificidade de 100%.

Capítulo 8

Materiais e Métodos

A seguir, iremos apresentar a metodologia utilizada para realizar as investigações e determinação das melhores configurações de um classificador de rede neural artificial com treinamento “backpropagation”. Essa configuração envolve desde a seleção dos melhores atributos, até a determinação das melhores características do classificador. A figura 8.1 ilustra o diagrama de blocos da metodologia utilizada neste trabalho.

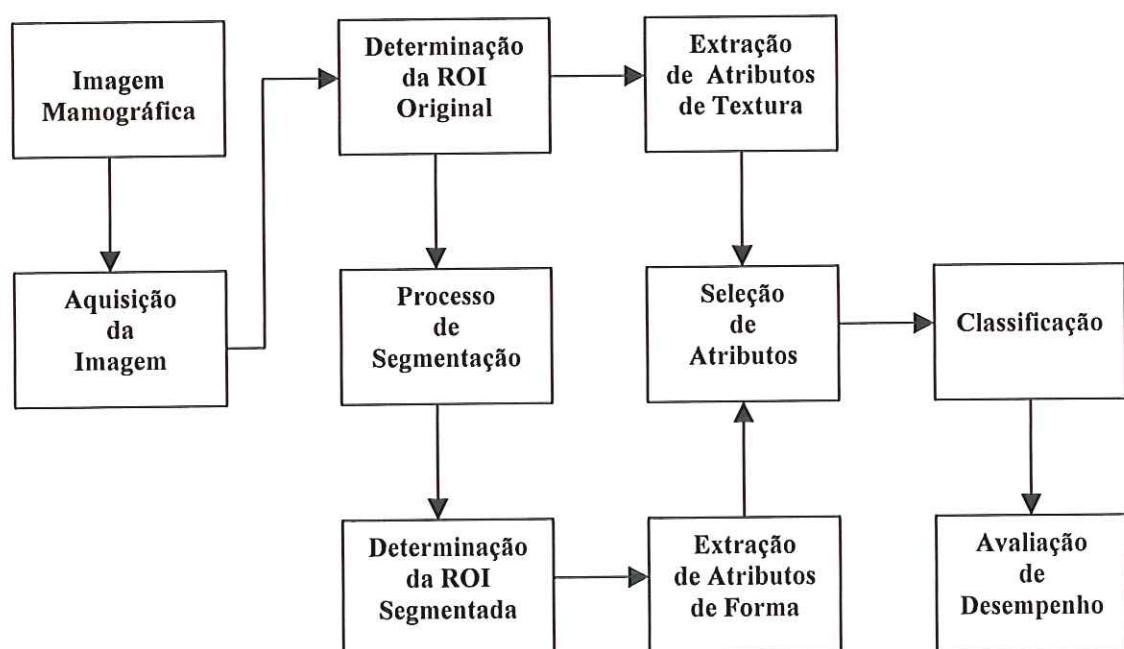


Figura 8.1 - Diagrama de blocos da metodologia utilizada para determinação da melhor configuração do classificador

As imagens foram apresentadas a especialistas que indicaram as regiões de interesse das lesões. Em seguida as imagens foram submetidas ao processo de segmentação que utilizou um processo interativo, de acordo com as dificuldades encontradas nas imagens. As regiões de interesse contendo os nódulos foram denominados de ROI original. Essa ROI foi, então submetida, conforme a necessidade, ao processo de melhoria de contraste (realce). A seguir, nessa ROI realçada foram aplicados os algoritmos de segmentação isolando a região da lesão e transformando essa imagem em binária (branco e preto). Essa imagem foi denominada de ROI segmentada. Na etapa seguinte, foram extraídos 28 atributos: 14 de textura e 14 de forma. Os atributos de textura foram extraídos das ROIs originais, e os atributos de forma, das ROIs segmentadas. Com todos os valores de atributos definidos, aplicou-se um processo de seleção de atributos para escolher os atributos mais discriminantes. Foram escolhidos conjuntos com 10 atributos: três conjuntos somente com atributos de forma, três conjuntos somente com atributos de textura, e três conjuntos com a combinação dos dois tipos de atributos. Selecionados os 9 conjuntos de atributos, realizamos o processo de classificação, utilizando uma rede neural artificial com algoritmo de treinamento por “*backpropagation*”. Foram investigadas várias configurações, variando-se o número de unidades escondidas da rede e o conjunto de atributos selecionados. Para determinar as melhores configurações, foram utilizados o método de avaliação de desempenho pela curva ROC (“*Receiver Operating Characteristic*”) e as taxas médias de acerto, sensibilidade e especificidade do classificador.

8.1 Base de Imagens e Aquisição das Imagens

Foram utilizadas duas bases de imagens neste trabalho de pesquisa: uma base desenvolvida junto ao Hospital de Clínicas de Ribeirão Preto (HCRP), com o apoio do grupo de serviço de Mastologia do Departamento de Ginecologia e Obstetrícia da Faculdade de Medicina de Ribeirão Preto, da Universidade de São Paulo; a outra base utilizada foi a do MIAS (The Mammographic Images Analysis Society database), disponível para livre utilização [Suckling-94].

Da base de imagens do HCRP foram selecionadas 60 imagens contendo nódulos (30 malignas e 30 benignas), representadas pelas vistas crânio-caudal e

médio-lateral-oblíqua, acompanhadas de respectivos resultados dos laudos de biópsias porém sem identificação dos tipos de nódulos (espiculadas ou circunscritas) presentes nas imagens.

A base de imagem do MIAS foi escolhida, pois possui uma boa representatividade dos tipos de lesões investigadas neste trabalho (lesões espiculadas e circunscritas). Foram selecionadas 57 imagens com um único nódulo e uma imagem contendo dois nódulos.

As imagens obtidas no HCRP, foram digitalizadas no digitalizador UMAX modelo UC 1260 com resolução espacial de 600 dpi, tamanho de pixel igual a 42 μ m e 8 bits de quantização (256 níveis de cinza). As imagens do MIAS já estavam digitalizadas e tinham tamanho de pixel de 50 μ m e, também, 8 bits de quantização.

Foram selecionadas manualmente regiões de interesse (ROI) nas imagens das duas bases (HCRP e MIAS) para os processos de segmentação e extração de atributos. Da base de imagens do HCRP, foram obtidas 60 regiões de interesse, sendo 30 com lesões malignas e 30 com lesões benignas. Da base de imagens do MIAS, foram obtidas 58 regiões de interesse, sendo 20 com lesões malignas e 38 com lesões benignas. As lesões estavam caracterizadas a priori e distribuídas da seguinte forma: 20 com lesões circunscritas benignas e 4 malignas; 11 com lesões espiculadas consideradas benignas e 8 malignas; 7 com lesões combinadas (espiculadas e circunscritas) benignas e 8 malignas. No total obtivemos 118 ROIs, sendo 68 com lesões benignas e 50 com lesões malignas. A região de interesse foi adquirida com tamanhos variados, dependendo do tamanho do nódulo, porém o tamanho da ROI foi padronizado para 512x512 pixels. A lista de todas as regiões de interesse adquiridas estão relacionadas na tabela 11.1, no capítulo anexos. Na tabela são fornecidas informações de identificação, base de imagem de origem, classificação histopatológica de cada região de interesse.

8.2 Etapa de Detecção e Segmentação

Após definidas as regiões de interesse, realizamos o tratamento das imagens através do processo de melhoria do contraste (realce) e de segmentação das imagens para destacar a lesão suspeita. O algoritmo utilizado para melhorar o realce foi o método de modificação do contraste (equação 4.8 e equação 4.9).

Nem todas as imagens necessitaram da aplicação do processo de realce, já que possuíam as regiões de interesse suficientemente contrastadas. A figura 8.2, ilustra uma sequência de imagens de uma ROI submetida ao processo de realce até a segmentação.

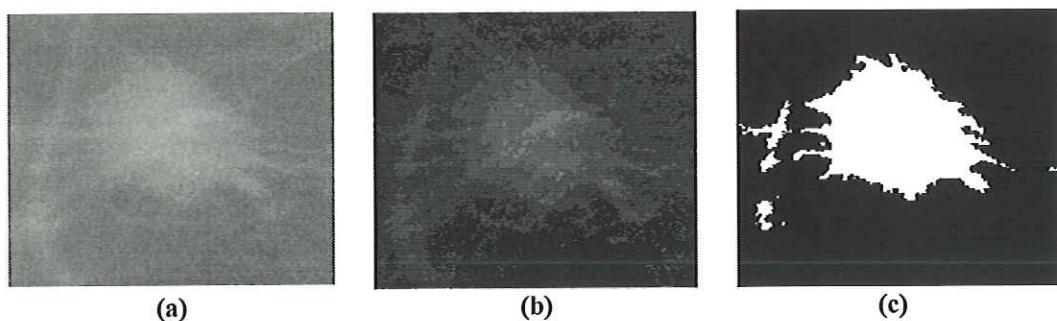


Figura 8.2 – Sequência de imagens em que foi aplicado os algoritmos de realce. (a) imagem original; (b) imagem após aplicada a técnica de modificação de contraste; (c) imagem posteriormente segmentada.

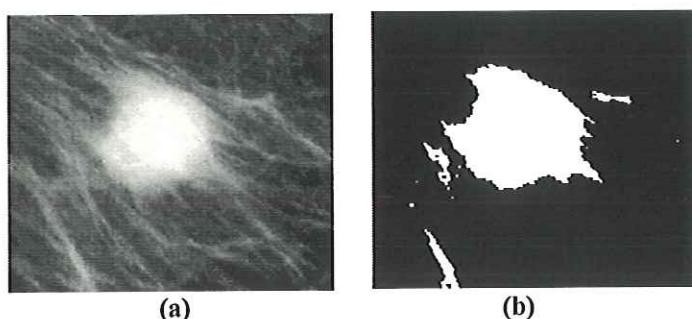


Figura 8.3 – Sequência de imagens em que não foi aplicado o algoritmo de realce. (a) imagem original; (b) imagem posteriormente segmentada.

e a figura 8.3, mostra um exemplo em que não foi aplicado o processo de realce de imagem para a segmentação.

Para a segmentação dos nódulos contidos nas regiões de interesse (ROIs originais), utilizou-se as técnicas investigadas em [Villela-96], que combinam métodos de crescimento de região [Shen-93] [Shen-94], limiarização (thresholding)

[Chan-87] [Chan-88A] [Chan-88B], e filtros morfológicos matemáticos [Dengler-93].

8.3 Etapa de Extração e Seleção de Atributos

Após a etapa de melhoria de contraste das imagens e de sua segmentação, realizamos a extração de atributos de textura e forma. Os atributos de textura foram calculados à partir das imagens das ROIs originais, pois de acordo com a metodologia proposta por Haralick *et al* [Haralick-73], utilizam-se imagens em níveis de cinza. Os atributos de forma foram calculados à partir das ROIs segmentadas, cujas imagens foram binarizadas. Alguns desses atributos de textura e forma foram já investigados por [Kilday-93] [Petrosian-94] [Wei-95] [Li-95] [El-Faramawy-96][Pohlman-96] [Sahiner-96A] [Sahiner-96B] [Sahiner-98] [Petrick-96A] [Wei-97] que comprovaram, em seus trabalhos, o alto grau de discriminação desses atributos.

8.3.1 Atributos de Textura

Os atributos de textura derivaram de cálculos realizados através da matriz de co-ocorrência, também conhecida como matriz de dependência espacial de nível de cinza (SGLD: spatial grey-level dependence matrix). Os atributos foram obtidos variando-se os ângulos de direção de 0° , 45° , 90° e 135° , e a distância entre pares de pixels de $d=1$. Desta forma, foram criadas 4 matrizes de SGLD, uma para cada ângulo de direção, com informações da probabilidade de pixels vizinhos com distância igual a um terem o mesmo valor de nível de cinza. Foram extraídos 14 atributos de textura das imagens que estão relacionados na tabela 8.1 e as equações para determinação de seus valores são apresentadas na secção 5.2 , do capítulo 5.

Os valores desses atributos foram normalizados para serem implementados na entrada da rede neural artificial (RNA), e estando relacionados na tabela 11.2, no capítulo anexos.

8.3.2 Atributos de Forma

Extraímos 14 atributos de forma das ROIs segmentadas. Aproveitamos o processo de crescimento de região, durante a etapa de segmentação do nódulo, para determinar os pixels de contorno e com isso possibilitar os cálculos de alguns atributos que utilizam essa informação. Os atributos de forma extraídos das ROIs estão relacionados, também, na tabela 8.1 e as equações para determinação de seus valores são apresentadas na seção 5.1 , do capítulo 5.

Os valores dos atributos de forma foram, também, normalizados para serem implementados na entrada da rede neural artificial (RNA) e estão relacionados na tabela 11.3, no capítulo anexos.

8.3.3 Seleção de Atributos

Em nosso estudo, utilizamos atributos de textura, atributos de forma, e a combinação dos dois tipos de atributos. Foram utilizadas combinações de atributos apresentados por pesquisadores diferentes (descritores de Hu, descritores de Shen, descriptor de Fourier, atributos de Wee, atributos de Haralick, etc).

Para definir o número de unidades da camada escondida e o número de amostras de imagens, aplicamos a regra apresentada por Timothy Masters [Masters-93]. De acordo com o autor, o número de amostras deve ser no mínimo igual ao dobro de $(n+1) \times m$, onde n é o número de entradas da rede (no nosso caso o número de atributos) e m é igual ao número de unidades da camada escondida. Se utilizássemos 28 atributos e 5 unidades na camada intermediária, teríamos que ter aproximadamente uma amostra de 290 imagens. A quantidade $(n+1) \times m$ determina o número de **parâmetros livres** da rede neural.

Levando-se em conta o número de parâmetros livres da rede neural, aplicamos um método de seleção para diminuir o número de atributos a ser aplicado na entrada da rede, e variamos o número de unidades da camada intermediária entre 3, 4, 5, e 6 . O método de seleção escolhido foi a distância de Jeffries-Matusita, cujo valor de melhor discriminação é o mais próximo de $\sqrt{2}$ ou seja, aproximadamente, 1,4142. Veja equações 5.47 e 5.48, do capítulo 5. Fixamos o número de elementos do conjunto igual a 10, e o método de seleção escolheu os 10 atributos mais

discriminantes para cada tipo de combinação. A tabela 8.2, mostra os três melhores conjuntos de atributos selecionados para cada tipo de combinação.

8.4 Etapa de Classificação através de Redes Neurais Artificiais

Para a classificação, escolhemos o método de redes neurais artificiais do tipo multi-camadas com algoritmo de treinamento por “*backpropagation*” (BPN). Ele foi escolhido pela facilidade de implementação e, de acordo com o teorema de Kolmogorov, utilizando-se esta rede com três camadas (entrada, intermediária e saída), pode-se aproximar qualquer função contínua ou booleana [Beale-90]. Outro fator importante é que a rede BPN possui alto grau de generalização e é tolerante a falhas e ruídos. A função de ativação escolhida foi a sigmóide logarítmica, com os valores de treinamento na saída da rede fixados em 0.02 para a classe benigna e 0.98 para a classe maligna.

Utilizamos apenas 3, 4, 5 e 6 unidades na camada intermediária e 10 unidades de entrada (número de atributos). O número de camadas intermediárias, também, ficou limitada devido ao número de parâmetros livres e o número de amostra de imagens, sendo utilizada a topologia com somente uma camada intermediária.

Alguns trabalhos utilizaram a classificação com redes neurais [Wu-93] [Sahiner-96] [Huo-95A][Huo96] [Lo-97], porém nenhum dos trabalhos utilizou a mesma combinação deste trabalho.

Tabela 8.1– Relação dos 28 atributos de Textura e Forma utilizados. A primeira coluna indica uma referência para cada atributo; na segunda coluna o tipo de atributo.

Código de Referência	Tipo de Atributo	Atributo
1	Textura	Variância
2	Textura	Entropia
3	Textura	Energia
4	Textura	Contraste
5	Textura	Correlação
6	Textura	Momento da Diferença Inversa (MDI)
7	Textura	Média da Soma
8	Textura	Variância da Soma
9	Textura	Entropia da Soma
10	Textura	Média da Diferença
11	Textura	Variância da Diferença
12	Textura	Entropia da Diferença
13	Textura	Medida de Informação de Correlação 1 (MC_1)
14	Textura	Medida de Informação de Correlação 2 (MC_2)
15	Forma	Descriptor de Hu 1
16	Forma	Descriptor de Hu 2
17	Forma	Descriptor de Hu 3
18	Forma	Descriptor de Gupta F1 ($F1'$)
19	Forma	Descriptor de Gupta F2
20	Forma	Descriptor de Gupta F3
21	Forma	Descriptor de Shen F2'
22	Forma	Descriptor de Shen F3'
23	Forma	Descriptor de Shen F3'-F1'
24	Forma	Descriptor de Fourier
25	Forma	Circularidade
26	Forma	Solidez
27	Forma	Descriptor de Wee 1
28	Forma	Descriptor de Wee 2

Tabela 8.2– Relação dos conjuntos de atributos selecionados pela Distância de Jeffries-Matusita. A primeira coluna indica uma identificação para cada conjunto de atributos indicados na tabela 8.1; na segunda coluna o tipo de atributo, na terceira coluna indica os códigos de referência dos atributos selecionados; e na última coluna o valor da distância de JM calculados.

Identificação do Conjunto	Tipo de Atributo	Atributos Selecionados	Distância de Jeffries-Matusita
$C1_{TF}$	Textura e Forma	3, 7, 9, 15, 18, 20, 23, 24, 25, 26	1,4126
$C2_{TF}$	Textura e Forma	3, 7, 9, 12, 15, 18, 20, 23, 24, 25	1,4126
$C3_{TF}$	Textura e Forma	1, 3, 7, 9, 12, 15, 18, 20, 23, 24, 25	1,4125
$C1_T$	Textura	1, 3, 4, 7, 8, 9, 10, 12, 13, 14	1,4097
$C2_T$	Textura	1, 2, 3, 4, 7, 8, 9, 10, 12, 13	1,4088
$C3_T$	Textura	1, 2, 3, 4, 8, 9, 10, 12, 13, 14,	1,4088
$C1_F$	Forma	15, 16, 17, 20, 21, 22, 24, 25, 26, 27	1,4079
$C2_F$	Forma	15, 16, 17, 20, 21, 24, 25, 26, 27, 28	1,4077
$C3_F$	Forma	15, 16, 17, 20, 21, 22, 24, 25, 27, 28	1,4076

A quantidade de imagens em nossa amostra não permitiu, também, uma flexibilidade na escolha dos grupos de imagens de treinamento e teste. Escolhemos, então, o método “leaving-one-out”, também conhecido como “jackknife procedure”. Neste método numa amostra com n elementos, deixa-se um elemento de fora do conjunto de treinamento e treina-se o classificador com os $(n-1)$ elementos restantes. O elemento deixado de fora é, então, utilizado para testar o classificador. Repete-se esse procedimento para todos os elementos da amostra. Com isso, os n elementos serão testados e participarão na amostra de treinamento dos $(n-1)$ elementos restantes. A avaliação do grau de generalização do classificador ficará, desta forma, garantida com este método. A única desvantagem, é o tempo gasto para o treinamento das n configurações de rede neural necessária. Para tanto, utilizamos uma variação no algoritmo de treinamento por “backpropagation” conhecido como “Aproximação de Levenberg-Marquardt” que aumenta a velocidade de treinamento sem perder a qualidade de desempenho da rede. Este método utiliza uma aproximação do método de Newton, cuja regra de atualização dos pesos é dada por:

$$\Delta \mathbf{W} = (\mathbf{J}^T \mathbf{J} + \alpha \mathbf{I})^{-1} \mathbf{J}^T \mathbf{E}$$

onde \mathbf{J} é a matriz Jacobiana da derivada de cada erro para cada peso. α é um escalar, e \mathbf{E} é vetor de erro. Se o escalar α é muito grande, a expressão acima aproxima-se do gradiente descendente. Se α for pequeno, a expressão acima transforma-se no próximo do erro mínimo, logo deve-se tentar manter a expressão mais próximo desse valor. Para isso, o algoritmo decresce o valor de α após cada etapa com sucesso e aumenta este valor somente quando em uma etapa aumenta o erro. O tempo de parada do treinamento da rede se dá quando um erro mínimo for atingido, ou um valor máximo de α for atingido, ou um número limite de ciclos for alcançado. Em nosso estudo, fixamos estes valores para todas as configurações de rede investigadas.

Realizamos o treinamento e teste do classificador com os nove conjuntos de atributos selecionados ($C1_{TF}, C2_{TF}, C3_{TF}, C1_T, C2_T, C3_T, C1_F, C2_F, C3_F$), e para cada conjunto configuramos a camada intermediária da rede para 3, 4, 5 e 6 unidades. Com isso, foram treinadas e testadas 36 configurações de RNA. A tabela 11.4, no capítulo anexos, exemplifica os resultados obtidos na saída de uma das configurações investigadas da classificação através de rede neural artificial.

8.5 Etapa de Avaliação de Resultados

Com os resultados obtidos de cada configuração, envolvendo o conjunto de atributos selecionados e o número de unidades da camada intermediária da rede, obtivemos as taxas de classificação da rede, variando-se o limiar na saída entre 0.1 à 0.9. Para cada limiar, calculamos as taxas de acerto, especificidade e sensibilidade. A taxa de acerto, refere-se a porcentagem do número de acertos da classe benigna e maligna juntas. A taxa de especificidade, também conhecida como taxa de verdadeiro-negativo, refere-se a porcentagem do número de acertos dentro da classe benigna; e a taxa de sensibilidade, conhecida como taxa de verdadeiro-positivo, refere-se a porcentagem de acertos dentro da classe maligna. Existe, ainda, a taxa de falso-positivo que corresponde a porcentagem de erro dentro da classe benigna, e a taxa de falso-negativo que corresponde a porcentagem de erro dentro da classe

maligna. Todas as taxas obtidas estão relacionadas nas tabelas 11.5, 11.6, 11.7, 11.8, 11.9, 11.10, 11.11, 11.12, 11.13, no capítulo anexos.

Para avaliar o desempenho do classificador utilizamos a curva ROC (Receiver Operating Characteristics curve) para cada uma das configurações e calculamos a área dentro da curva. A curva ROC foi traçada de acordo com dois índices: fração verdadeiro-positivo e fração falso-positivo. O valor da área forneceu um índice de desempenho do classificador [Evans-81][Metz-86]. Calculamos, também, a média e desvio-padrão das taxas de acerto, especificidade e sensibilidade de cada configuração da rede neural. De posse desses valores, área dentro da curva ROC e taxa de acerto, tornou-se mais fácil e eficaz a avaliação de desempenho do classificador. Em caso de coincidência de valores da área dentro da curva de duas ou mais configurações, consultamos as taxas média de acerto, especificidade e sensibilidade, respectivamente, como medidas de desempate. A avaliação foi realizada individualmente para cada uma das combinações (textura, forma, e os dois tipos juntos). Desta forma, obtivemos as melhores configurações do classificador, fornecendo o conjunto de atributo ($C1_{TF}$, $C2_{TF}$, $C3_{TF}$, $C1_F$, $C2_F$, $C3_F$, $C1_T$, $C2_T$, $C3_T$) e o número de unidades escondidas (3, 4, 5 ou 6).

Capítulo 9

Resultados e Discussão

A partir da metodologia definida, os resultados serão apresentados e analisados em cada etapa investigada. Iniciaremos a apresentação pela etapa de detecção e segmentação das imagens, a seguir a etapa de extração e seleção de atributos, classificação e avaliação.

9.1 Detecção e Segmentação

Durante a fase de detecção e segmentação, verificamos que algumas imagens não obtiveram resultados satisfatórios com os métodos empregados. Das 118 imagens processadas 14 imagens (11,2%) obtiveram visualmente resultados imprecisos. A causa principal é a própria característica de alguns tipos de nódulos, principalmente do tipo circunscrito, que possuem densidade na lesão similar a do tecido normal da mama. A figura 9.1, ilustra um exemplo de uma ROI em que a segmentação não foi bem sucedida. Note que a lesão possui parcialmente contorno bem definido. Este problema na segmentação poderia ser resolvido se utilizássemos técnicas que definissem o contorno da lesão. Esta definição do contorno da lesão é de difícil solução devido as características de similaridade de densidade entre os tecidos da mama e da lesão. Novas técnicas devem ser investigadas, porém, em nosso trabalho corrente, priorizamos as investigações do processo de extração de atributos e classificação para determinação das configurações ótimas do classificador.

As outras 104 imagens (88,8%) foram segmentadas satisfatoriamente, pois o processo interativo, permitiu determinação de contorno mais definidos do que num processo automático.

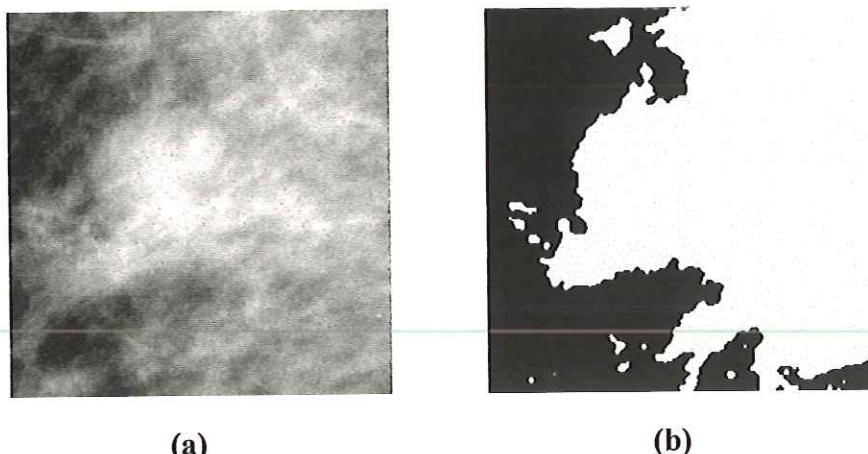


Figura 9.1 - Imagem de uma ROI em que o processo de segmentação da lesão não foi bem sucedida. Note que o nódulo se confunde com o tecido do parêquima (a) ROI original. (b) ROI segmentada.

9.2 Extração e Seleção de Atributos

Na secção 7.3, a extração dos atributos foi descrita e os resultados foram apresentados. A tabela 7.1, relaciona os atributos de textura e forma empregados, e as tabelas 11.2 e 11.3, do capítulo anexos, apresentam os valores normalizados dos resultados da extração dos atributos.

A etapa de seleção de atributos, como foi descrito na secção 7.3, foi necessária para diminuir a quantidade de atributos para atender as exigências impostas pelas condições dos parâmetros livres da rede neural e pelo número de amostras das imagens disponível. O método de Jeffries-Matusita forneceu um índice de discriminação dos atributos muito próximo do valor de $\sqrt{2}$, que indicou um alto grau de discriminação dos conjuntos de atributos selecionados. A tabela 7.2, mostra os três melhores conjuntos de atributos selecionados para cada tipo de combinação e os valores das distâncias calculados. Essa seleção é importante, também, para evitar atributos redundantes. Os 28 atributos escolhidos mostraram-se altamente discriminantes pelas investigações realizadas por vários pesquisadores que estudaram os atributos individualmente ou em conjunto [Kilday-93][Kegelmeyer-94] [Yin-91A] [Yin-91B][Yin-94A][Yin-94B][Huo-95][Rangayyan-96][Pohlman-96][Petrosian-94]

[Wei-95][Wei-97][Sahiner-96A][Sahiner-96B][Petrick-96A][Petrick-96B] [Chan-95] [Sahiner-98]. Para comprovar esse alto grau de discriminação dos atributos, calculamos a distância de Jeffries-Matusita para os 28 atributos de textura e forma, para os 14 atributos de textura, e para os 14 atributos de forma. A tabela 9.1, mostra esses valores de distância. Verifica-se que os valores são muito próximo do valor de $\sqrt{2}$, comprovando o alto grau de discriminação dos atributos escolhidos para o nosso estudo.

Tabela 9.1– Cálculo da Distância de Jeffries-Matusita para os 28 atributos de textura e forma; 14 atributos de textura, e 14 atributos de forma.

Tipo de Atributo	Quantidade de Atributos	Distância de Jeffries-Matusita
Textura e Forma	28	1,4142
Textura	14	1,4138
Forma	14	1,4136

9.3 Classificação através de Redes Neurais Artificiais

Na secção 7.4, descrevemos todo o processo de escolha das possíveis configurações da rede neural artificial a serem utilizadas. O algoritmo de treinamento foi o “*backpropagation*” com adaptação da “Aproximação de Levenberg-Marquardt”. A função de ativação escolhida foi a sigmóide logarítmica, com os valores de treinamento na saída da rede fixados em 0,02 para a classe benigna e 0,98 para a classe maligna, o número de unidades de escondidas ficou em 3, 4, 5 e 6 numa única camada intermediária, e o método de treinamento e teste foi o “*leaving-one-out*” aplicados sobre as 118 imagens.

A rede neural mostrou-se estável em todo o processo, convergindo em todos os treinamentos realizados. Para testar a habilidade de aprendizagem da rede neural, treinamos o classificador com os atributos das 118 imagens, tomando-se todas as configurações adotadas. A seguir testamos as configurações da rede neural com os mesmos valores de atributos. O classificador obteve uma performance de 100% de

acerto em todas as configurações, demonstrando que a rede neural foi capaz de aprender todos os padrões impostos [Lo-97]

Uma desvantagem observada no método de “leaving-one-out” é o tempo gasto para o treinamento da rede. O algoritmo de treinamento de “Aproximação de Levenberg-Marquardt” aumentou de modo significativo a velocidade de treinamento e, desta forma, que minimizou este problema.

9.4 Avaliação de Resultados

Os resultados obtidos das configurações, envolvendo o conjunto de atributos selecionados, o número de unidades da camada intermediária de acordo com a variação de limiar dos resultados, estão ilustrados nas tabelas 11.5, 11.6, 11.7, 11.8, 11.9, 11.10, 11.11, 11.12, 11.13, no capítulo anexos.

Para realizar a avaliação da melhor configuração envolvendo o conjunto de atributos e o número de unidades escondidas, utilizamos o critério que envolveu a área dentro da curva ROC e as taxas médias de acertos, especificidade e sensibilidade.

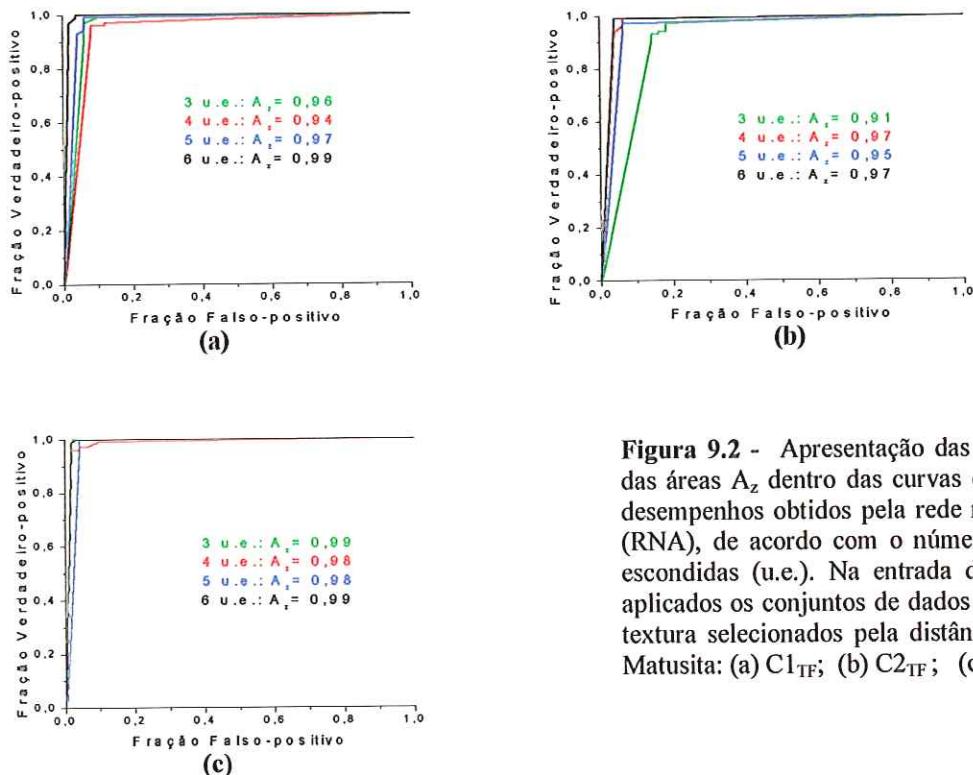
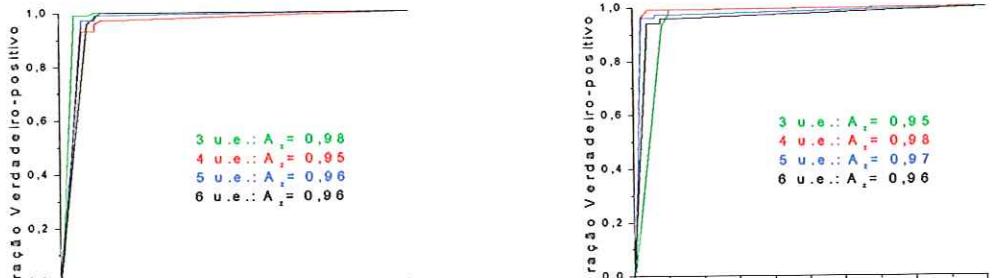
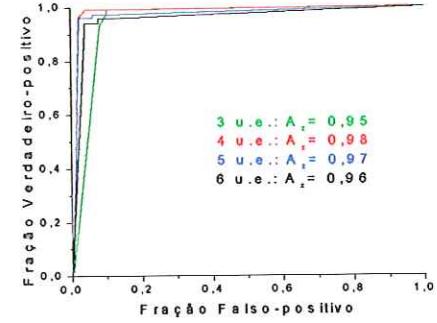


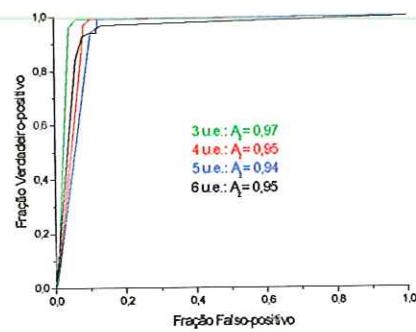
Figura 9.2 - Apresentação das curvas ROC e das áreas A_z dentro das curvas em relação aos desempenhos obtidos pela rede neural artificial (RNA), de acordo com o número de unidades escondidas (u.e.). Na entrada da RNA foram aplicados os conjuntos de dados de atributos de textura selecionados pela distância de Jeffries-Matusita: (a) C1_{TF}; (b) C2_{TF}; (c) C3_{TF}.



(a)

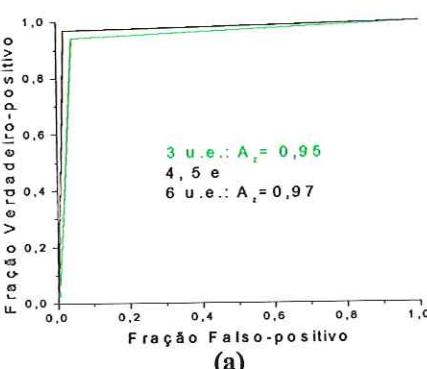


(b)

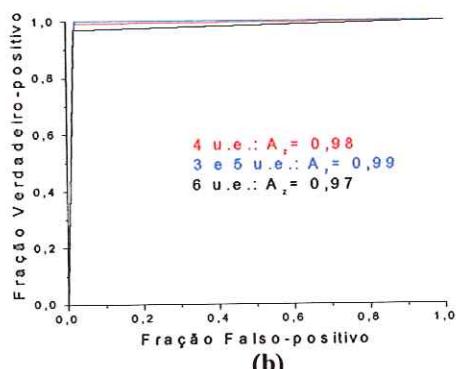


(c)

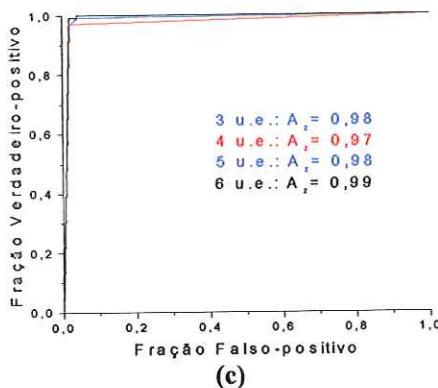
Figura 9.3 - Apresentação das curvas ROC e das áreas A_z dentro das curvas em relação aos desempenhos obtidos pela rede neural artificial (RNA), de acordo com o número de unidades escondidas (u.e.). Na entrada da RNA foram aplicados os conjuntos de dados de atributos de textura selecionados pela distância de Jeffries-Matusita: (a) $C1_T$; (b) $C2_T$; (c) $C3_T$.



(a)



(b)



(c)

Figura 9.4 - Apresentação das curvas ROC e das áreas A_z dentro das curvas em relação aos desempenhos obtidos pela rede neural artificial (RNA), de acordo com o número de unidades escondidas (u.e.). Na entrada da RNA foram aplicados os conjuntos de dados de atributos de textura selecionados pela distância de Jeffries-Matusita: (a) $C1_F$; (b) $C2_F$; (c) $C3_F$.

A figuras 9.2, 9.3, e 9.4, apresentam as curvas ROC das 9 configurações estudadas. A área dentro da curva é, também, apresentada em cada curva. A figura 9.2, refere-se aos conjuntos de atributos de textura e forma $C1_{TF}$, $C2_{TF}$, e $C3_{TF}$, variando-se o número de unidades escondidas. A figura 8.3, refere-se aos conjuntos de atributos de textura $C1_T$, $C2_T$, e $C3_T$, variando-se o número de unidades escondidas. A figura 9.3, refere-se aos conjuntos de atributos de forma $C1_F$, $C2_F$, e $C3_F$, variando-se o número de unidades escondidas.

Com os valores das áreas dentro da curva ROC calculadas e as médias das taxas de acerto, especificidade e sensibilidade, também, definidas, realizamos a escolha das melhores configurações. A tabela 9.2, apresenta os resultados da performance da classificação da rede neural artificial.

Observamos que as referências para a escolha das melhores configurações estão bem claras e podem ser facilmente selecionadas. Comparando estas tabelas com as tabelas 11.5, 11.6, 11.7, 11.8, 11.9, 11.10, 11.11, 11.12, 11.13, no capítulo anexos, observamos que a avaliação através das taxas de acerto, sensibilidade e especificidade destas tabelas é difícil, imprecisa e cansativa, devido a grande variações apresentadas de número de camadas escondidas e valor limiar. Com isso, obtivemos as seguintes configurações:

- Para o conjunto de atributos de textura e forma, foi selecionada a configuração com o conjunto $C3_{TF}$ e a rede com 3 unidades escondidas que apresentou área dentro da curva ROC igual a 0,99 e taxa de acerto de 98,21% de média e 1,67% de desvio-padrão;
- Para o conjunto de atributos de textura, foi selecionada a configuração com o conjunto $C2_T$ e a rede com 4 unidades escondidas que apresentou área dentro da curva ROC igual a 0,98 e taxa de acerto de 97,08% de média e 1,13% de desvio-padrão;
- Para o conjunto de atributos de forma, foi selecionada a configuração com o conjunto $C2_F$ e a rede com 5 unidades escondidas que apresentou área dentro da curva ROC igual a 0,99 e taxa de acerto de 98,21% de média e 0,28% de desvio-padrão;

A tabela 9.3, apresenta as melhores configurações baseadas na performance de classificação da rede neural artificial.

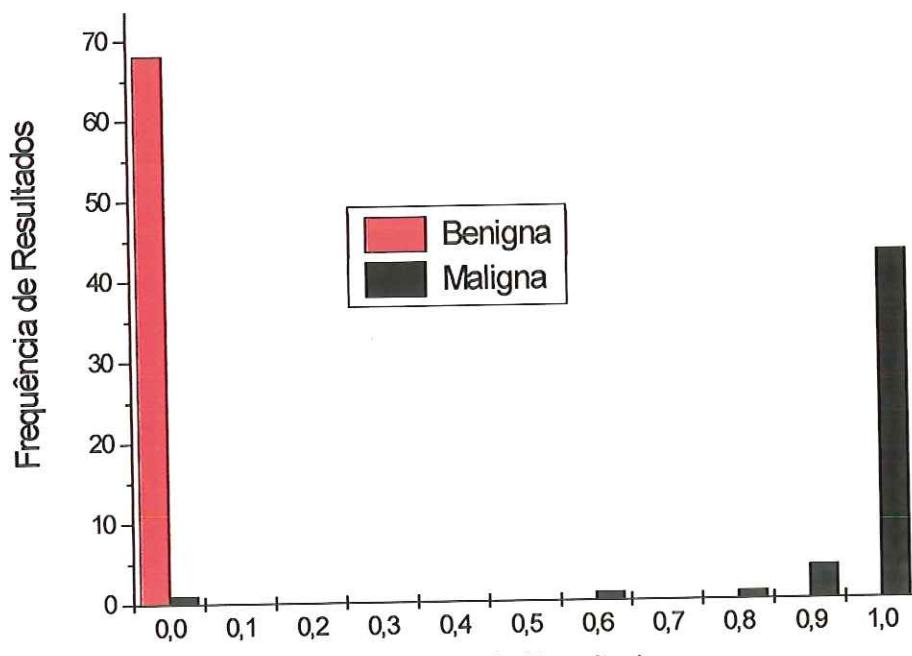
Tabela 9.2 - Resultados da Performance de Classificação da Rede Neural Artificial, de acordo com a Área na Curva ROC, Porcentagem de Acerto, Especificidade e Sensibilidade.

Conjunto de Atributos	Unidades Escondidas	Área dentro da Curva ROC	% Acertos $\mu \pm \sigma$	% Especificidade $\mu \pm \sigma$	% Sensibilidade $\mu \pm \sigma$
C1 _{TF}	3	0,96	94,07 ± 1,41	94,44 ± 2,83	93,56 ± 1,33
	4	0,94	95,10 ± 0,93	96,41 ± 1,82	93,33 ± 3,00
	5	0,97	95,10 ± 0,93	96,57 ± 1,04	93,11 ± 2,03
	6	0,99	97,18 ± 1,34	98,37 ± 1,87	95,56 ± 3,71
C2 _{TF}	3	0,91	89,45 ± 0,75	93,30 ± 2,09	84,22 ± 1,56
	4	0,97	93,22 ± 0,42	95,75 ± 0,49	89,78 ± 1,20
	5	0,95	94,73 ± 1,10	95,26 ± 2,92	94,00 ± 2,00
	6	0,97	96,42 ± 0,57	97,71 ± 1,66	94,67 ± 2,24
C3 _{TF}	3	0,99	98,21 ± 1,67	100,00 ± 0,00	95,78 ± 3,93
	4	0,98	95,95 ± 0,57	96,57 ± 1,04	95,11 ± 2,47
	5	0,98	96,61 ± 1,80	99,02 ± 1,64	93,33 ± 4,58
	6	0,99	97,46 ± 0,95	98,20 ± 1,91	96,44 ± 2,60
C1 _T	3	0,98	96,14 ± 0,86	99,02 ± 0,74	92,22 ± 2,54
	4	0,95	92,37 ± 0,60	94,77 ± 1,49	89,11 ± 2,03
	5	0,96	94,07 ± 1,12	97,88 ± 1,30	88,89 ± 4,01
	6	0,96	94,35 ± 1,12	100,00 ± 0,00	86,67 ± 2,65
C2 _T	3	0,95	93,88 ± 0,57	97,71 ± 0,78	88,67 ± 2,00
	4	0,98	97,08 ± 1,13	98,53 ± 0,00	95,11 ± 2,67
	5	0,97	95,29 ± 0,62	96,41 ± 0,78	93,78 ± 1,56
	6	0,96	93,69 ± 0,86	94,61 ± 0,74	92,44 ± 2,40
C3 _T	3	0,97	94,92 ± 0,95	98,53 ± 0,00	90,00 ± 2,24
	4	0,95	94,82 ± 0,28	97,39 ± 0,65	91,33 ± 1,41
	5	0,94	93,13 ± 0,51	97,22 ± 0,88	87,56 ± 1,33
	6	0,95	92,18 ± 0,57	94,77 ± 1,30	88,67 ± 1,41
C1 _F	3	0,95	94,92 ± 0,00	94,12 ± 0,00	96,00 ± 0,00
	4	0,97	96,80 ± 0,56	96,57 ± 0,74	97,11 ± 1,45
	5	0,97	96,61 ± 0,95	96,08 ± 1,27	97,33 ± 2,00
	6	0,97	97,36 ± 0,28	96,90 ± 0,49	98,00 ± 0,00
C2 _F	3	0,99	97,83 ± 1,13	98,04 ± 2,08	97,56 ± 1,33
	4	0,98	97,18 ± 0,73	96,73 ± 1,43	97,78 ± 0,67
	5	0,99	98,21 ± 0,28	98,37 ± 0,49	98,00 ± 0,00
	6	0,97	97,18 ± 0,42	96,73 ± 0,65	97,78 ± 0,67
C3 _F	3	0,98	98,49 ± 0,37	98,86 ± 0,65	98,00 ± 0,00
	4	0,97	97,18 ± 0,42	97,06 ± 0,00	97,33 ± 1,00
	5	0,98	97,46 ± 0,00	97,22 ± 0,49	97,78 ± 0,67
	6	0,99	97,65 ± 0,37	97,71 ± 1,07	97,56 ± 0,88

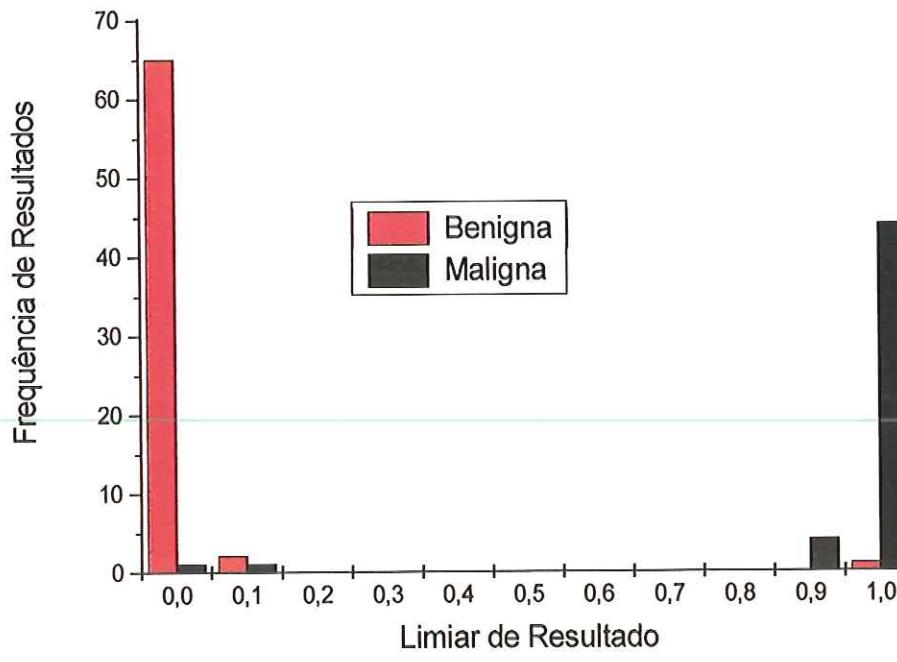
Tabela 9.3 - Melhores configurações de acordo com os resultados de performance de classificação da rede neural artificial, considerando a Área dentro da curva ROC, taxas médias de Acerto, Especificidade e Sensibilidade.

Conjunto de Atributos	Unidades Escondidas	Área dentro da Curva ROC	% Acertos $\mu \pm \sigma$	% Especificidade $\mu \pm \sigma$	% Sensibilidade $\mu \pm \sigma$
C3 _{TF}	3	0,99	98,21 ± 1,67	100,00 ± 0,00	95,78 ± 3,93
C2 _T	4	0,98	97,08 ± 1,13	98,53 ± 0,00	95,11 ± 2,67
C2 _F	5	0,99	98,21 ± 0,28	98,37 ± 0,49	98,00 ± 0,00

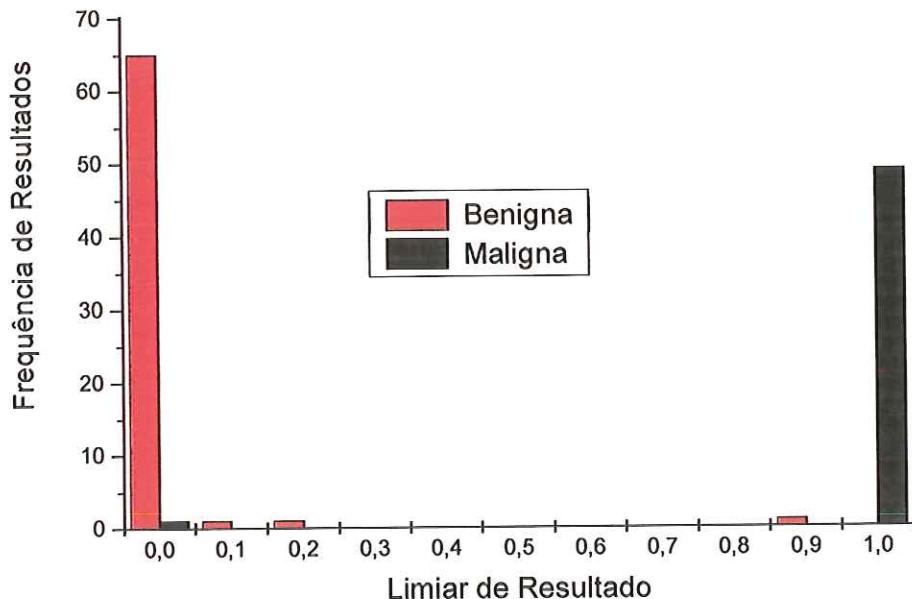
Definidas as melhores configurações de acordo com o conjunto de atributos e o número de unidades escondidas da rede neural, traçamos o histograma da distribuição de resultados da saída do classificador (RNA), em relação ao valor de limiar entre 0,0 e 1,0. A figura 9.5, apresenta os histogramas para as configurações selecionadas: conjunto de atributos de textura e forma C3_{TF}, com 3 unidades escondidas; conjunto de atributos de textura C2_T, com 4 unidades escondidas; e conjunto de atributos de forma C2_F, com 5 unidades escondidas, respectivamente.



(a)



(b)



(c)

Figura 9.5 – Histogramas de distribuição de resultados na saída das redes neurais artificiais (RNA) de melhores desempenhos na classificação de casos benignos e malignos. Eles representam os conjuntos de atributos de: (a) textura e forma C₃_{TF} da rede com 3 unidades escondidas; (b) textura C₂_T com 4 unidades escondidas; (c) forma C₂_F com 5 unidades escondidas.

Em situações clínicas, médicos radiologistas não diagnosticam simplesmente a lesão em maligna ou benigna. Normalmente, quando a lesão é suspeita, ele recomenda o encaminhamento para a biópsia. Além do auxílio ao diagnóstico, o classificador pode ser útil na tomada de decisão do radiologista em relação ao envio de casos para a biópsia. Para isso, pode-se utilizar o valor de limiar da saída da rede para tomar essa decisão. Analisamos qual deve ser o valor mais apropriado para obter uma alta sensibilidade ou especificidade, considerando, também, o custo de falso-positivo, isto é, o custo de não ter enviado um caso positivo para a biópsia. A alta especificidade é importante para não enviar para a biópsia casos benignos desnecessariamente.

Analisando os resultados obtidos, verificamos que podemos determinar um valor de limiar para que o classificador determine se um caso estudado deve ser recomendado para a biópsia. Tomando-se os valores de limiar igual a 0,1; 0,5; e 0,9, teremos as taxas de falso-negativo, falso-positivo, verdadeiro-negativo (especificidade), e verdadeiro-positivo (sensibilidade) dentro da amostra tratada. A tabela 9.4, apresenta esses valores de acordo com a configuração selecionada.

Tabela 9.4 – Valores de limiares são fornecidos, também, para auxiliar o radiologista para na indicação de biópsia. Com elas apresentamos as taxas de verdadeiro-positivo, falso-negativo, falso-positivo e verdadeiro-negativo.

Conjunto de Atributos	Unidades Escondidas	Limiar	Verdadeiro Negativo (Especificidade)			Falso Positivo		Verdadeiro Positivo (Sensibilidade)		Falso Negativo	
C3 _{TF}	3	0,1	68	100,00%	0	0,00%	49	98,00%	1	2,00%	
		0,5	68	100,00%	0	0,00%	49	98,00%	1	2,00%	
		0,9	68	100,00%	0	0,00%	43	86,00%	7	14,00%	
C2 _T	4	0,1	67	98,53%	1	1,47%	48	96,00%	2	4,00%	
		0,5	67	98,53%	1	1,47%	48	96,00%	2	4,00%	
		0,9	67	98,53%	1	1,47%	44	88,00%	6	12,00%	
C2 _F	5	0,1	66	97,06%	2	2,94%	49	98,00%	1	2,00%	
		0,5	67	98,53%	1	1,47%	49	98,00%	1	2,00%	
		0,9	67	98,53%	1	1,47%	49	98,00%	1	2,00%	

Da tabela, podemos observar que:

- Para a configuração com o conjunto de atributos C3_{TF}, com 3 unidades escondidas, o custo de falso-negativo é em torno de 2.0% para limiar de 0.1 e 0.5; e 14% para limiar igual a 0.9. A especificidade ficou em 100% para os três valores de limiares, enquanto a sensibilidade permaneceu igual a 98% para o limiar de 0.1 e 0.5, e baixou para 86% no limiar de 0.9.
- Para a configuração com o conjunto de atributos C2_T, com 4 unidades escondidas, o custo de falso-negativo é em torno de 4.0% para limiar de 0.1 e 0.5; e 12% para limiar igual a 0.9. A especificidade ficou em 98.53% para os três valores de limiar, enquanto a sensibilidade permaneceu igual a 96% para o limiar de 0.1 e 0.5, e baixou para 88% no limiar de 0.9.
- Para a configuração com o conjunto de atributos C2_F, com 5 unidades escondidas, o custo de falso-negativo é em torno de 2.0% para os três valores de limiar. A especificidade ficou em 97.06% para o limiar de 0.1, e 98.53% para os limiares de 0.5 e 0.9, enquanto a sensibilidade permaneceu igual a 98% para os três valores de limiar.

Analizando o histograma, verificamos que é possível selecionar um valor de limiar para a configuração do classificador, levando-se em consideração o custo de falso-negativo (o não envio de um caso positivo para a biópsia), a taxa de falso-positivo (envio de um caso negativo para a biópsia desnecessariamente), a taxa de especificidade, e a taxa de sensibilidade. Obviamente que a busca deve ser pela alta especificidade e sensibilidade, porém em alguns casos devem ser analisados para obter-se um equilíbrio entre essas taxas. Nestas configurações de classificadores, a recomendação mais apropriada seria o limiar de 0.5, pois possui a menor taxa de falso-positivo, para taxas mais altas de especificidade e sensibilidade.

Em relação as altas taxas de acerto, especificidade e sensibilidade obtidas para as três combinações de atributos, podemos creditá-las as algumas características observadas durante esse processo:

- O processo interativo da segmentação realizado em nosso trabalho permitiu, na maioria dos casos, a manutenção das bordas das lesões investigadas. Como os atributos de forma estão intrinsecamente relacionadas com esta característica, isto contribuiu com um alto índice na extração de atributos e na classificação. Se fosse utilizado um processo automático de segmentação, certamente teríamos resultados não tão bons de segmentação e, consequentemente, no processo de extração e de classificação.
- O alto grau de discriminação dos atributos utilizados influenciou de modo importante na classificação. As pesquisas realizadas por outros autores e a investigação realizada em nosso trabalho, comprovam estas altas taxas para as três combinações de atributos.
- O classificador de RNA utilizado não apresentou problemas de convergência e nem de desempenho durante o processo de treinamento e teste, porém, existem situações em que uma rede mesmo tendo bom desempenho, sofre o processo de “overfitting”. Esse processo, de acordo com [Masters-93], ocorre quando utilizamos uma quantidade pequena de amostras, ou quando aplicamos um número de entradas muito grande (centenas, milhares, etc). Logo, está relacionado com o número de parâmetros livres. Quando o “overfitting” acontece, a rede ao invés de “aprender”, ela “decora” os padrões impostos. Com isso, podemos obter resultados altos ou baixos, sempre constante, mesmo variando os parâmetros de treinamento (grupo de treinamento e teste, número de unidades da camada intermediária, número de camadas intermediária, etc). Em nosso trabalho, as altas taxas obtidas de acerto, especificidade, e sensibilidade, apontam um suposto processo de “overfitting”, porém atendemos a exigência primordial da rede que a quantidade de amostra

deve ser o dobro do número de parâmetros livres. Porém, esta hipótese deve ser levada em consideração.

9.4.1 Comparação com Outros Resultados

Comparação com outros trabalhos não foi possível, pois as bases de imagens utilizadas eram diferentes; trabalhos que utilizaram os atributos de textura de Haralick, aplicaram esses atributos para detecção de nódulos e não para classificação entre malignos e benignos [Petrosian-94] [Wei-95][Wei-97][Sahiner-96A][Petrick-96A]. Os trabalhos que utilizaram atributos de forma, como circularidade, solidez, descritores de Fourier, etc, utilizaram somente alguns dos atributos. Os únicos trabalhos que tiveram algumas características semelhantes ao nosso trabalho foram os de [El-Faramawy-96] e [Rangayyan-96], que utilizaram as bases de imagens do MIAS – “The Mammographic Image Analysis Society”- combinada com outra base de imagem do “Alberta Program for the Early Detection of Breast Cancer”. O método de treinamento e teste foi, também, o “leaving-one-out”. A tabela 9.5, apresenta os resultados resumidos da taxa de acerto, especificidade (verdadeiro-negativo), e sensibilidade (verdadeiro-positivo) desses trabalhos.

Observe que em [Rangayyan-96] os resultados foram superiores aos apresentados em [El-Faramawy-96]. Os atributos de forma foram mais discriminantes do que os atributos de textura utilizados e combinados com atributos de forma em [El-Faramawy-96]. Mesmo esses resultados não podem ser diretamente comparados com os resultados obtidos em nosso trabalho, pois o grupo de atributos selecionados foi muito diferente do utilizado, principalmente os atributos de textura, que em nosso trabalho mostraram-se superiores quanto ao grau de discriminação.

Tabela 9.5 - Resultados de classificação utilizando distância de Mahalanobis, base de imagens “The Mammographic Image Analysis Society” (MIAS), combinada com outra base de imagem do “Alberta Program for the Early Detection of Breast Cancer” APEDBC). Os atributos utilizados são circularidade (C); medida de borda pontiaguda (A); Descritor de Fourier (FF); Descritor de Shen F3’-F1’ (MF); Fator de forma de distribuição de corda ($M_{1,4}$); e descriptores de Hu e momentos utilizando os pixels de borda em níveis de cinza (M^{bd}), e pixels dentro da ROI em níveis de cinza (M^{rd}) e (M'^{rd}).

Referência de Trabalho	Base de Imagem	Atributo	% Acerto	% Verdadeiro Negativo	% Verdadeiro Positivo
[Rangayyan-96]	MIAS	A	94,9	92,9	100
	MIAS +APEDBC	A	92,6	92,9	92,3
	MIAS	C	61,5	—	—
	MIAS +APEDBC	C	72,2	—	—
	MIAS	A, FF, MF	94,9	—	—
	MIAS +APEDBC	A, FF, MF	87,0	—	—
	MIAS	A, C	84,6	—	—
	MIAS +APEDBC	A, C	83,3	—	—
[EL-Faramawy-96]	MIAS	M_{bd}, M_{rd}, M'^{rd}	74,5	—	—
	MIAS +APEDBC	M_{bd}, M_{rd}, M'^{rd}	70,4	—	—
	MIAS	$M_{1,4}, C, FF, MF$	76,9	—	—
	MIAS +APEDBC	$M_{1,4}, C, FF, MF$	68,5	—	—

Capítulo 10

Conclusões

10.1 Conclusões Finais:

- 1) Este estudo demonstrou que o esquema proposto pode ser utilizado para determinar a melhor configuração de um classificador de rede neural artificial para a caracterização de nódulos em mamogramas digitalizados. A rede estudada foi de arquitetura multi-camadas, com algoritmo de treinamento de “*backpropagation*” com adaptação da “aproximação de Levenberg-Marquardt”. Os parâmetros de configuração investigados foram os atributos de forma, textura, e os dois combinados; o parâmetro da rede neural artificial foi o número de unidades escondidas.
- 2) Os 28 atributos escolhidos mostraram um alto grau de discriminação. O método de seleção de Jeffries-Matusita selecionou eficientemente os atributos sem perder o grau de discriminação, com a vantagem da eliminação de atributos redundantes. O esquema proposto permitiu, também, investigar o grau de discriminação dos dois tipos de atributos individualmente (textura ou forma) e combinados. As altas taxas de acerto, especificidade, e sensibilidade deve-se a esse alto grau de discriminação dos atributos, e ao processo de segmentação que, maioria dos casos, isolou eficientemente o nódulo sem perder informações de borda. A vantagem dos atributos de textura foi excluir a etapa de segmentação.

- 3) O esquema proposto pode ser utilizado, acrescido de um processo automático de detecção e segmentação, no auxílio ao diagnóstico de nódulos em mamografia. O esquema permite definir um índice para minimizar as taxas de falso-positivo e falso-negativo, e buscar um equilíbrio entre as taxas de especificidade e sensibilidade.

10.2 Sugestões para Trabalhos Futuros:

Apesar de nosso esquema ter demonstrado várias qualidades importantes para a caracterização de nódulos em mamogramas digitalizados, podemos aperfeiçoá-lo acrescentando ou substituindo algumas características:

- Deve-se buscar uma maior quantidade de amostras, procurando, também, um alto grau de representatividade dessas amostras.
- Deve-se investigar processos automáticos de detecção e segmentação que, conjuntamente com um número maior de amostras, possibilitaria a sua utilização clínica.
- O processo de seleção de atributos da distância de Jeffries-Matusita, obteve um bom resultado, mas tem algumas restrições como: assumir uma distribuição Gaussiana para os atributos; não fornecer automaticamente uma quantidade ótima de atributos mais discriminantes. Em nosso trabalho, fixamos o número de atributos desejados (10 atributos) e o esquema calculou a distância de cada combinação. Por isso, consideramos importante a investigação de outros métodos de seleção, de preferência automáticos, como, por exemplo, a Análise de Componentes Principais (PCA).
- Deve-se fazer um estudo estatístico da representatividade de uma amostra de imagens dentro de uma população de casos. Isto permitirá determinar uma quantidade mínima que seja representativa num processo de classificação.

Capítulo 11

Anexos

As tabelas que apresentaremos a seguir são importantes para a compreensão do nosso trabalho. Foram relacionadas aqui porque a sua extensão tornaria cansativa a leitura do texto. As tabelas relacionadas são: a tabela com a relação de imagens das regiões de interesse (ROI) obtidas das imagens mamográficas da base do Hospital das Clínicas de Ribeirão Preto (HCRP) e do “The Mammographic Images Analysis Society database” (MIAS); tabela dos resultados normalizados dos atributos obtidos da extração de textura e forma das ROIs; tabela de um exemplo de resultados obtidos na saída da RNA com o algoritmo de treinamento de “Backpropagation”, com aproximação de Levenberg-Marquardt”, com 3, 4, 5, e 6 unidades escondidas e com um dos conjuntos de atributos; e finalmente a tabela de resultados de performance do classificador, tendo como entrada um dos conjuntos de atributos.

Tabela 11.1– Relação das regiões de interesse (ROI) obtidas das imagens mamográficas da base do Hospital de Clínicas de Ribeirão Preto (HCRP) e do “The Mammographic Images Analysis Society database” (MIAS).

Identificação da Região de Interesse	Identificação da Imagem Original	Classificação Clínica	Base da Imagem
ROI001	Imag001rp	Benigna	HCRP
ROI002	mdb207	Benigna	MIAS
ROI003	Imag002rp	Benigna	HCRP
ROI004	mdb001	Benigna	MIAS
ROI005	Imag003rp	Benigna	HCRP
ROI006	mdb199	Benigna	MIAS
ROI007	mdb005	Benigna	MIAS
ROI008	mdb010	Benigna	MIAS
ROI009	mdb012	Benigna	MIAS
ROI010	mdb015	Benigna	MIAS
ROI011	mdb017	Benigna	MIAS
ROI012	mdb019	Benigna	MIAS
ROI013	mdb021	Benigna	MIAS
ROI014	mdb025	Benigna	MIAS
ROI015	mdb069	Benigna	MIAS
ROI016	mdb080	Benigna	MIAS
ROI017	mdb091	Benigna	MIAS
ROI018	mdb132	Benigna	MIAS
ROI019	mdb132	Benigna	MIAS
ROI020	mdb142	Benigna	MIAS
ROI021	mdb244	Benigna	MIAS
ROI022	mdb290	Benigna	MIAS
ROI023	mdb315	Benigna	MIAS
ROI024	mdb145	Benigna	MIAS
ROI025	mdb175	Benigna	MIAS
ROI026	mdb188	Benigna	MIAS
ROI027	mdb190	Benigna	MIAS
ROI028	mdb191	Benigna	MIAS
ROI029	mdb204	Benigna	MIAS
ROI030	Imag004rp	Benigna	HCRP
ROI031	mdb198	Benigna	MIAS
ROI032	mdb005	Benigna	MIAS
ROI033	mdb002	Benigna	MIAS
ROI034	mdb195	Benigna	MIAS
ROI035	mdb193	Benigna	MIAS
ROI036	mdb199	Benigna	MIAS
ROI037	mdb013	Benigna	MIAS
ROI038	mdb030	Benigna	MIAS
ROI039	mdb032	Benigna	MIAS
ROI040	mdb063	Benigna	MIAS
ROI041	mdb144	Benigna	MIAS
ROI042	mdb312	Benigna	HCRP
ROI043	mdb314	Benigna	MIAS
ROI044	Imag005rp	Benigna	HCRP
ROI045	Imag006rp	Benigna	HCRP
ROI046	Imag012rp	Benigna	HCRP
ROI047	Imag013rp	Benigna	HCRP
ROI048	Imag014rp	Benigna	HCRP
ROI049	Imag015rp	Benigna	HCRP
ROI050	Imag016rp	Benigna	HCRP
ROI051	Imag017rp	Benigna	HCRP
ROI052	Imag018rp	Benigna	HCRP
ROI053	Imag019rp	Benigna	HCRP
ROI054	Imag020rp	Benigna	HCRP
ROI055	Imag021rp	Benigna	HCRP
ROI056	Imag022rp	Benigna	HCRP
ROI057	Imag023rp	Benigna	HCRP
ROI058	Imag024rp	Benigna	HCRP
ROI059	Imag025rp	Benigna	HCRP
ROI060	Imag026rp	Benigna	HCRP

Tabela 11.1– Relação das regiões de interesse (ROI) obtidas das imagens mamográficas da base do Hospital de Clínicas de Ribeirão Preto (HCRP) e do “The Mammographic Images Analysis Society database” (MIAS) (*continuação*).

Identificação da Região de Interesse	Identificação da Imagem Original	Classificação Clínica	Base da Imagem
ROI061	Imag027rp	Benigna	HCRP
ROI062	Imag028rp	Benigna	HCRP
ROI063	Imag033rp	Benigna	HCRP
ROI064	Imag033rp	Benigna	HCRP
ROI065	Imag033rp	Benigna	HCRP
ROI066	Imag034rp	Benigna	HCRP
ROI067	Imag034rp	Benigna	HCRP
ROI068	Imag034rp	Benigna	HCRP
ROI101	mdb023	Maligna	MIAS
ROI102	mdb028	Maligna	MIAS
ROI103	mdb041	Maligna	MIAS
ROI104	mdb270	Maligna	MIAS
ROI105	Imag035rp	Maligna	HCRP
ROI106	mdb178	Maligna	MIAS
ROI107	mdb181	Maligna	MIAS
ROI108	mdb184	Maligna	MIAS
ROI109	mdb186	Maligna	MIAS
ROI110	mdb202	Maligna	MIAS
ROI111	mdb206	Maligna	MIAS
ROI112	mdb058	Maligna	MIAS
ROI113	mdb134	Maligna	MIAS
ROI114	mdb264	Maligna	MIAS
ROI115	mdb265	Maligna	MIAS
ROI116	mdb267	Maligna	MIAS
ROI117	mdb271	Maligna	MIAS
ROI118	mdb274	Maligna	MIAS
ROI119	Imag036rp	Maligna	HCRP
ROI120	Imag037rp	Maligna	HCRP
ROI121	Imag038rp	Maligna	HCRP
ROI122	Imag039rp	Maligna	HCRP
ROI123	Imag040rp	Maligna	HCRP
ROI124	Imag041rp	Maligna	HCRP
ROI125	Imag042rp	Maligna	HCRP
ROI126	Imag043rp	Maligna	HCRP
ROI127	Imag044rp	Maligna	HCRP
ROI128	Imag045rp	Maligna	HCRP
ROI129	Imag046rp	Maligna	HCRP
ROI130	Imag047rp	Maligna	HCRP
ROI131	Imag048rp	Maligna	HCRP
ROI132	Imag049rp	Maligna	HCRP
ROI133	mdb148	Maligna	MIAS
ROI134	Imag050rp	Maligna	HCRP
ROI135	Imag051rp	Maligna	HCRP
ROI136	mdb144	Maligna	MIAS
ROI137	Imag005rp	Maligna	HCRP
ROI138	Imag006rp	Maligna	HCRP
ROI139	Imag007rp	Maligna	HCRP
ROI140	Imag007rp	Maligna	HCRP
ROI141	Imag007rp	Maligna	HCRP
ROI142	Imag008rp	Maligna	HCRP
ROI143	Imag009rp	Maligna	HCRP
ROI144	Imag010rp	Maligna	HCRP
ROI145	Imag011rp	Maligna	HCRP
ROI146	Imag029rp	Maligna	HCRP
ROI147	Imag030rp	Maligna	HCRP
ROI148	Imag031rp	Maligna	HCRP
ROI149	Imag032rp	Maligna	HCRP
ROI150	mdb179	Maligna	MIAS

Tabela 11.2 – Resultados normalizados dos atributos obtidos da extração de atributos de textura das imagens das regiões de interesse originais.

ROI	Variância	Entropia	Energia	Contraste	Correlação	Momento Diferença Inversa	Média da Soma
ROI001	0,2853	0,4708	0,1993	0,0503	0,9679	0,3889	0,0000
ROI002	0,5621	0,6498	0,0299	0,2063	0,9601	0,2529	0,7581
ROI003	0,7505	0,5062	0,1429	0,3505	0,9601	0,4771	0,2257
ROI004	0,9224	0,9339	0,0100	0,1592	0,9913	0,2656	0,5875
ROI005	0,4477	0,7876	0,3322	0,2155	0,9367	0,1706	0,2555
ROI006	0,4277	0,6608	0,0332	0,2238	0,9289	0,2525	0,7405
ROI007	0,0274	0,1458	0,3953	0,2051	0,0000	0,4143	0,1019
ROI008	0,4006	0,4226	0,1561	0,4178	0,8501	0,4427	0,5853
ROI009	0,3064	0,5262	0,0963	0,1993	0,8908	0,3329	0,8764
ROI010	0,1348	0,4473	0,1163	0,0200	0,9532	0,7619	0,8299
ROI011	0,5506	0,6772	0,0498	0,0437	0,9939	0,6455	0,7138
ROI012	0,6234	0,3095	0,2359	0,3713	0,9385	0,7619	0,6410
ROI013	0,7164	0,5648	0,1063	0,0952	0,9913	0,8249	0,8804
ROI014	0,1852	0,7132	0,0100	0,0855	0,9003	0,2167	0,6851
ROI015	0,5724	0,6086	0,0399	0,5308	0,8951	0,1991	0,5478
ROI016	0,9369	0,6802	0,0997	0,1047	0,9965	0,4898	0,5064
ROI017	0,5153	0,3205	0,2159	1,0000	0,7608	0,4754	0,5248
ROI018	0,4748	0,3987	0,1894	0,3432	0,9081	0,5160	0,2588
ROI019	0,4743	0,4652	0,1296	0,1904	0,9497	0,6510	0,5351
ROI020	0,3289	0,3355	0,1728	0,5743	0,7201	0,2520	0,2171
ROI021	0,7129	0,6783	0,0465	0,1348	0,9861	0,4045	0,8192
ROI022	0,1204	0,4778	0,1229	0,0366	0,9177	0,7662	0,7103
ROI023	0,6229	0,8229	0,0066	0,1291	0,9809	0,2412	0,9155
ROI024	0,6295	0,7899	0,0033	0,3155	0,9497	0,0501	0,4514
ROI025	0,4230	0,5896	0,0631	0,2970	0,9029	0,2962	0,7418
ROI026	0,3379	0,7691	0,0066	0,0655	0,9688	0,2774	0,7973
ROI027	0,0517	0,0000	1,0000	0,0494	0,7886	1,0000	0,6616
ROI028	0,4900	0,3066	0,1794	0,7115	0,8154	0,3089	0,6237
ROI029	0,4856	0,2789	0,2591	0,6704	0,8241	0,4763	0,3167
ROI030	0,5016	0,8753	0,0000	0,5919	0,8536	0,0043	0,7313
ROI031	1,0000	0,7788	0,1296	0,0885	0,9991	0,4486	0,6175
ROI032	0,4385	0,7245	0,0233	0,0363	0,9905	0,5591	0,4485
ROI033	0,7824	0,6530	0,0797	0,0391	1,0000	0,7308	0,8014
ROI034	0,1877	0,4444	0,1229	0,0096	0,9809	0,9536	0,2525
ROI035	0,1608	0,4443	0,1429	0,0204	0,9619	0,8476	0,8256
ROI036	0,1816	0,5767	0,0698	0,0303	0,9575	0,6866	0,7093
ROI037	0,0558	0,3957	0,1296	0,0175	0,9012	0,6815	0,9436
ROI038	0,0179	0,2227	0,3189	0,0068	0,8899	0,9315	1,0000
ROI039	0,3106	0,8206	0,0000	0,0861	0,9532	0,2203	0,7616
ROI040	0,0811	0,4389	0,1096	0,0181	0,9246	0,7003	0,9632
ROI041	0,0000	0,1580	0,5150	0,0124	0,8128	0,8666	0,7422
ROI042	0,0719	0,3131	0,2857	0,0105	0,9367	0,9308	0,8952
ROI043	0,0890	0,4106	0,1694	0,0241	0,9168	0,8419	0,7520
ROI044	0,1962	0,7988	0,0066	0,1793	0,8154	0,0591	0,6858
ROI045	0,1584	0,7398	0,0066	0,1297	0,8224	0,1142	0,4120
ROI046	0,1422	0,5329	0,1063	0,0613	0,8969	0,6021	0,1984
ROI047	0,2058	0,5531	0,1163	0,0551	0,9411	0,4497	0,2927
ROI048	0,1776	0,4828	0,1395	0,0490	0,9350	0,4538	0,0651
ROI049	0,0092	0,0536	0,8605	0,0000	0,9133	0,8541	0,1463
ROI050	0,3736	0,8171	0,0100	0,1915	0,9237	0,1242	0,2808
ROI051	0,6194	0,7169	0,1661	0,2610	0,9575	0,1436	0,2224
ROI052	0,2069	0,4205	0,3920	0,0696	0,9289	0,4810	0,0543
ROI053	0,4510	0,8258	0,0332	0,2613	0,9237	0,1491	0,2274
ROI054	0,5510	0,8184	0,0432	0,4476	0,9055	0,1481	0,2691
ROI055	0,3510	0,8823	0,0066	0,2256	0,9003	0,0415	0,9067
ROI056	0,1678	0,3934	0,1794	0,4764	0,4203	0,3691	0,0371
ROI057	0,2657	0,6143	0,0664	0,1724	0,8830	0,2848	0,1020
ROI058	0,3500	0,8509	0,0000	0,2416	0,8925	0,1150	0,3046
ROI059	0,9352	1,0000	0,0399	0,3974	0,9705	0,0673	0,6576
ROI060	0,8043	0,9754	0,0399	0,5775	0,9393	0,0092	0,5432
ROI061	0,0780	0,6896	0,0266	0,3195	0,1629	0,0000	0,3983

Tabela 11.2 – Resultados normalizados dos atributos obtidos da extração de atributos de textura das imagens das regiões de interesse originais(continuação).

ROI	Variância	Entropia	Energia	Contraste	Correlação	Momento Diferença Inversa	Média da Soma
ROI062	0,1447	0,6373	0,0365	0,1208	0,8154	0,1688	0,2181
ROI063	0,2076	0,6725	0,0233	0,0712	0,9272	0,2797	0,3524
ROI064	0,1892	0,7394	0,0100	0,0974	0,8908	0,1809	0,5875
ROI065	0,6213	0,8204	0,0332	0,1196	0,9827	0,2072	0,3346
ROI066	0,2731	0,7721	0,0066	0,0890	0,9402	0,2105	0,4310
ROI067	0,1389	0,6523	0,0266	0,0719	0,8787	0,2522	0,5364
ROI068	0,5048	0,8213	0,0066	0,1096	0,9757	0,2115	0,3649
ROI101	0,4397	0,8197	0,0066	0,0150	1,0000	0,3627	0,4397
ROI102	0,2058	0,2727	0,1978	0,3834	0,8420	0,6771	0,2058
ROI103	0,4570	0,7892	0,0086	0,0140	0,9919	0,5621	0,4570
ROI104	0,5693	0,6010	0,0283	0,0000	0,9675	0,7668	0,5693
ROI105	0,1506	0,7629	0,0056	0,0922	0,9408	0,2075	0,1506
ROI106	0,5679	0,3739	0,0984	0,1308	0,9721	0,6743	0,5679
ROI107	0,2808	0,8894	0,0174	0,0596	0,9907	0,3075	0,2808
ROI108	0,7363	0,7767	0,0056	0,0210	0,9361	0,3455	0,7363
ROI109	0,6732	0,7230	0,0102	0,0341	0,5738	0,3721	0,6732
ROI110	0,3045	0,8966	0,0043	0,0313	1,0000	0,3075	0,3045
ROI111	0,3742	0,7545	0,0112	0,0109	0,9733	0,5988	0,3742
ROI112	0,6095	0,8083	0,0039	0,0419	0,7909	0,2514	0,6095
ROI113	0,4970	0,7456	0,0092	0,0207	0,8711	0,3321	0,4970
ROI114	0,2000	0,8708	0,0020	0,0232	0,9861	0,3819	0,2000
ROI115	0,6807	0,8065	0,0039	0,0171	0,9628	0,3882	0,6807
ROI116	0,6510	0,6331	0,0217	0,0061	0,8269	0,5087	0,6510
ROI117	0,6335	0,7495	0,0086	0,0129	0,9350	0,4300	0,6335
ROI118	0,4592	0,8089	0,0039	0,0361	0,8595	0,2864	0,4592
ROI119	0,3670	0,7973	0,0049	0,0853	0,3926	0,1304	0,3670
ROI120	0,4978	0,7671	0,0063	0,0453	0,6098	0,2162	0,4978
ROI121	0,3962	0,7714	0,0072	0,0448	0,7573	0,2278	0,3962
ROI122	0,1778	0,4882	0,0477	1,0000	0,2288	0,2053	0,1778
ROI123	0,1391	0,7601	0,0072	0,0199	0,9117	0,3357	0,1391
ROI124	0,6294	0,9254	0,0030	0,1047	0,9803	0,1998	0,6294
ROI125	0,5644	0,9383	0,0026	0,1094	0,9861	0,2120	0,5644
ROI126	0,0182	0,5502	0,0342	0,0470	0,9129	0,3332	0,0182
ROI127	0,3435	0,7406	0,0217	0,2205	0,9210	0,1955	0,3435
ROI128	0,4026	0,6778	0,0118	0,3968	0,6167	0,0859	0,4026
ROI129	0,1365	0,7571	0,0059	0,7653	0,1510	0,0000	0,1365
ROI130	0,3231	0,8676	0,0023	0,2025	0,9524	0,0812	0,3231
ROI131	0,3024	0,8382	0,0030	0,4048	0,8455	0,0300	0,3024
ROI132	0,3458	0,8224	0,0036	0,0412	0,8711	0,2192	0,3458
ROI133	0,0673	0,6258	0,2163	0,3657	0,9071	0,2979	0,0673
ROI134	0,1739	0,7635	0,0056	0,0155	0,9454	0,3578	0,1739
ROI135	0,2393	0,7732	0,0063	0,0955	0,0000	0,1114	0,2393
ROI136	0,5820	0,6528	0,0135	0,0117	0,6307	0,3838	0,5820
ROI137	0,2070	0,8959	0,0026	0,0715	0,9663	0,1847	0,2070
ROI138	0,2148	0,8853	0,0013	0,0520	0,9559	0,2044	0,2148
ROI139	0,3851	1,0000	0,0000	0,0981	0,9698	0,1410	0,3851
ROI140	0,2624	0,9554	0,0023	0,0819	0,9803	0,1864	0,2624
ROI141	0,5337	0,9542	0,0237	0,1202	0,9895	0,2255	0,5337
ROI142	0,1508	0,8091	0,0043	0,0509	0,7642	0,2038	0,1508
ROI143	0,1599	0,9360	0,0086	0,2128	0,9280	0,1318	0,1599
ROI144	0,3233	0,9802	0,0013	0,1195	0,9710	0,1384	0,3233
ROI145	0,0000	0,8362	0,0122	0,1385	0,8931	0,1913	0,0000
ROI146	0,3864	0,7278	0,0079	0,0137	0,9013	0,3411	0,3864
ROI147	0,2440	0,6689	0,0132	0,0065	0,8862	0,4153	0,2440
ROI148	0,2854	0,6902	0,0145	0,0045	0,9744	0,6111	0,2854
ROI149	0,3917	0,7482	0,0089	0,0140	0,9489	0,5102	0,3917
ROI150	1,0000	0,0000	1,0000	0,0203	0,5226	1,0000	1,0000

Tabela 11.2 – Resultados normalizados dos atributos obtidos da extração de atributos de textura das imagens das regiões de interesse originais(continuação).

ROI	Variância da Soma	Entropia da Soma	Média da Diferença	Variância da Diferença	Entropia da Diferença	Informação Correlação 1	Informação Correlação 2
ROI001	0,1204	0,4209	0,1895	0,0380	0,3607	0,3602	0,9720
ROI002	0,3624	0,6083	0,4565	0,1685	0,3772	0,3538	0,9860
ROI003	0,5979	0,6558	0,5102	0,3748	0,4941	0,2680	0,9840
ROI004	0,8642	1,0000	0,3608	0,1469	0,6163	0,2874	0,9980
ROI005	0,2471	0,6249	0,4930	0,1650	0,7401	0,5711	0,9540
ROI006	0,2290	0,7469	0,4800	0,1821	0,6446	0,4379	0,9720
ROI007	0,0050	0,1220	0,4287	0,1839	0,6063	0,8419	0,0000
ROI008	0,2048	0,3593	0,6210	0,4034	0,3816	0,4120	0,9400
ROI009	0,1339	0,5636	0,4253	0,1753	0,5985	0,5076	0,9280
ROI010	0,0405	0,4119	0,0671	0,0205	0,1794	0,2444	0,9840
ROI011	0,3506	0,7173	0,1237	0,0461	0,2922	0,1793	0,9980
ROI012	0,4322	0,3577	0,4659	0,4381	0,1416	0,1046	0,9880
ROI013	0,5520	0,7929	0,1690	0,1187	0,2522	0,0181	1,0000
ROI014	0,0629	0,5527	0,2877	0,0573	0,4867	0,5675	0,9420
ROI015	0,3724	0,6802	0,7935	0,4413	0,6346	0,4769	0,9540
ROI016	0,8890	0,8673	0,2520	0,1047	0,4690	0,1740	0,9980
ROI017	0,3091	0,1773	1,0000	1,0000	0,2113	0,4510	0,8940
ROI018	0,2719	0,4802	0,5088	0,3617	0,4230	0,3680	0,9500
ROI019	0,2721	0,6335	0,3278	0,2128	0,3330	0,1700	0,9920
ROI020	0,1481	0,3341	0,8180	0,4929	0,5137	0,5873	0,8100
ROI021	0,5472	0,8545	0,3150	0,1270	0,5399	0,2492	0,9940
ROI022	0,0347	0,4314	0,0942	0,0420	0,2223	0,2542	0,9820
ROI023	0,4327	0,8765	0,3410	0,1055	0,5850	0,3450	0,9940
ROI024	0,4398	0,8723	0,6343	0,2252	0,7812	0,4919	0,9740
ROI025	0,2245	0,6555	0,5528	0,2541	0,6105	0,4553	0,9580
ROI026	0,1569	0,7016	0,2390	0,0445	0,4240	0,4018	0,9860
ROI027	0,0119	0,0104	0,0725	0,0690	0,0382	0,1764	0,9240
ROI028	0,2850	0,1827	0,9161	0,6196	0,1660	0,4515	0,9040
ROI029	0,2809	0,3281	0,8062	0,6611	0,3588	0,4225	0,9020
ROI030	0,2970	0,8527	0,9110	0,4331	0,9884	0,6764	0,9220
ROI031	1,0000	0,8740	0,2328	0,0852	0,4671	0,2057	0,9980
ROI032	0,2395	0,7597	0,1319	0,0314	0,3003	0,2092	0,9980
ROI033	0,6449	0,7538	0,0920	0,0472	0,2346	0,0787	1,0000
ROI034	0,0645	0,5207	0,0109	0,0140	0,0388	0,0000	0,9980
ROI035	0,0517	0,4403	0,0516	0,0245	0,1386	0,1574	0,9920
ROI036	0,0614	0,5610	0,1022	0,0292	0,2477	0,2418	0,9900
ROI037	0,0132	0,3100	0,0740	0,0147	0,1884	0,3647	0,9560
ROI038	0,0037	0,1601	0,0086	0,0095	0,0386	0,2114	0,9640
ROI039	0,1372	0,7288	0,2878	0,0580	0,4865	0,4477	0,9840
ROI040	0,0209	0,3888	0,0715	0,0164	0,1861	0,2955	0,9760
ROI041	0,0000	0,0123	0,0325	0,0148	0,0979	0,3947	0,8680
ROI042	0,0180	0,2913	0,0175	0,0144	0,0548	0,1367	0,9860
ROI043	0,0234	0,3791	0,0576	0,0294	0,1518	0,2030	0,9840
ROI044	0,0680	0,6054	0,4780	0,1128	0,6890	0,6950	0,8940
ROI045	0,0501	0,5353	0,3905	0,0819	0,6046	0,6786	0,8920
ROI046	0,0434	0,4626	0,1686	0,0625	0,3664	0,3790	0,9700
ROI047	0,0735	0,5056	0,1876	0,0469	0,3972	0,3628	0,9780
ROI048	0,0594	0,4022	0,1720	0,0418	0,3616	0,4144	0,9600
ROI049	0,0019	0,0000	0,0000	0,0000	0,0000	0,2370	0,9340
ROI050	0,1837	0,7209	0,4585	0,1466	0,7009	0,5314	0,9680
ROI051	0,4279	0,6472	0,5374	0,2118	0,7598	0,4748	0,9720
ROI052	0,0740	0,3045	0,1959	0,0688	0,4076	0,4608	0,9360
ROI053	0,2499	0,7487	0,4832	0,2504	0,7266	0,4803	0,9800
ROI054	0,3492	0,7646	0,5089	0,5519	0,7184	0,4570	0,9820
ROI055	0,1660	0,6864	0,5411	0,1493	0,7559	0,6631	0,9320
ROI056	0,0529	0,3041	0,3263	0,7110	0,4804	0,5666	0,8680
ROI057	0,1074	0,5507	0,3217	0,1916	0,5573	0,4636	0,9640
ROI058	0,1652	0,7179	0,4828	0,2165	0,7198	0,5668	0,9640
ROI059	0,8847	0,9229	0,6927	0,3211	0,8948	0,4829	0,9880
ROI060	0,6748	0,8166	0,8619	0,4640	1,0000	0,6224	0,9600
ROI061	0,0185	0,3521	0,5945	0,2702	0,7875	1,0000	0,2020

Tabela 11.2 – Resultados normalizados dos atributos obtidos da extração de atributos de textura das imagens das regiões de interesse originais(continuação).

ROI	Variância da Soma	Entropia da Soma	Média da Diferença	Variância da Diferença	Entropia da Diferença	Informação Correlação 1	Informação Correlação 2
ROI062	0,0442	0,4595	0,3547	0,0864	0,5824	0,6750	0,8660
ROI063	0,0744	0,5653	0,2470	0,0514	0,4435	0,4917	0,9640
ROI064	0,0649	0,5893	0,3235	0,0612	0,5319	0,5796	0,9440
ROI065	0,4308	0,8222	0,3531	0,0850	0,5914	0,3612	0,9920
ROI066	0,1124	0,6735	0,3007	0,0580	0,5104	0,4867	0,9740
ROI067	0,0420	0,5145	0,2567	0,0481	0,4455	0,5497	0,9400
ROI068	0,3024	0,7888	0,3329	0,0783	0,5673	0,4051	0,9900
ROI101	0,3367	0,8539	0,1246	0,0078	0,5862	0,4828	0,9949
ROI102	0,4561	0,3452	0,4405	0,5832	0,1621	0,2752	0,9012
ROI103	0,2115	0,8594	0,0882	0,0173	0,5283	0,3639	1,0000
ROI104	0,0392	0,6915	0,0262	0,0013	0,3409	0,2217	1,0000
ROI105	0,2751	0,8472	0,2753	0,0964	0,7338	0,5129	0,9864
ROI106	0,6606	0,4989	0,2378	0,2051	0,3369	0,1565	0,9881
ROI107	0,5566	0,8998	0,2010	0,0702	0,7261	0,5167	0,9949
ROI108	0,0728	0,7755	0,1397	0,0151	0,6188	0,5815	0,9710
ROI109	0,0087	0,6791	0,1516	0,0383	0,6463	0,7076	0,8620
ROI110	0,5595	0,9393	0,1632	0,0276	0,6572	0,4612	1,0000
ROI111	0,0970	0,8242	0,0780	0,0135	0,4981	0,3495	1,0000
ROI112	0,0367	0,7588	0,1954	0,0342	0,6984	0,7075	0,9131
ROI113	0,0341	0,7211	0,1423	0,0134	0,6187	0,6513	0,9319
ROI114	0,2326	0,9197	0,1357	0,0212	0,6213	0,4503	0,9983
ROI115	0,0997	0,8289	0,1241	0,0125	0,5954	0,5124	0,9898
ROI116	0,0072	0,6318	0,0813	0,0024	0,5130	0,5829	0,9233
ROI117	0,0515	0,7634	0,1094	0,0084	0,5695	0,5358	0,9779
ROI118	0,0508	0,7740	0,1790	0,0296	0,6794	0,6676	0,9387
ROI119	0,0189	0,6973	0,2990	0,0642	0,8146	0,9038	0,5451
ROI120	0,0162	0,7090	0,2126	0,0331	0,7173	0,7822	0,8092
ROI121	0,0324	0,7288	0,2066	0,0355	0,7120	0,7401	0,8705
ROI122	0,2452	0,4978	1,0000	1,0000	0,5781	0,7554	0,5162
ROI123	0,0503	0,7735	0,1381	0,0138	0,6110	0,5665	0,9744
ROI124	0,6691	0,9178	0,2918	0,1114	0,8315	0,6014	0,9864
ROI125	0,8150	0,9440	0,2909	0,1226	0,8323	0,5709	0,9932
ROI126	0,1055	0,6552	0,1760	0,0566	0,5607	0,4834	0,9523
ROI127	0,4930	0,8422	0,4093	0,2640	0,8069	0,5360	0,9796
ROI128	0,1962	0,7561	0,6440	0,3557	1,0000	0,7277	0,8228
ROI129	0,1667	0,7620	0,9223	0,6628	0,9587	0,9022	0,4940
ROI130	0,6828	0,9199	0,4587	0,1726	0,8911	0,6221	0,9761
ROI131	0,4886	0,8791	0,6627	0,3456	0,9432	0,7389	0,9046
ROI132	0,0631	0,8010	0,2046	0,0288	0,7035	0,6630	0,9506
ROI133	0,6994	0,7029	0,5220	0,4499	0,8228	0,5665	0,9370
ROI134	0,0681	0,7975	0,1245	0,0094	0,5848	0,5146	0,9864
ROI135	0,0084	0,6436	0,3203	0,0704	0,8331	1,0000	0,0000
ROI136	0,0000	0,6195	0,1135	0,0046	0,5627	0,7037	0,8296
ROI137	0,3361	0,9148	0,2623	0,0601	0,7861	0,5665	0,9898
ROI138	0,2075	0,8826	0,2269	0,0396	0,7391	0,5986	0,9830
ROI139	0,4791	1,0000	0,3117	0,0825	0,8379	0,5849	0,9949
ROI140	0,5249	0,9781	0,2747	0,0743	0,8075	0,5462	0,9966
ROI141	1,0000	0,9808	0,3023	0,1382	0,8473	0,5200	0,9983
ROI142	0,0386	0,7811	0,2209	0,0407	0,7259	0,7002	0,9250
ROI143	0,5146	0,9237	0,4355	0,2209	0,9529	0,6325	0,9813
ROI144	0,5927	0,9890	0,3388	0,1072	0,8691	0,5755	0,9949
ROI145	0,2406	0,8516	0,3321	0,1548	0,8660	0,6027	0,9761
ROI146	0,0338	0,7384	0,1255	0,0052	0,5786	0,5819	0,9625
ROI147	0,0170	0,6951	0,0945	0,0000	0,5194	0,5409	0,9642
ROI148	0,0707	0,7697	0,0590	0,0047	0,4658	0,3517	0,9966
ROI149	0,0681	0,7924	0,0990	0,0143	0,5550	0,4543	0,9915
ROI150	0,0007	0,0000	0,0000	0,0496	0,0000	0,0000	0,5111

Tabela 11.3 – Resultados normalizados dos atributos obtidos da extração de atributos de forma das imagens das regiões de interesse originais.

ROI	Descriptor de Hu 1	Descriptor de Hu 2	Descriptor de Hu 3	Descriptor de Gupta F1 (F1')	Descriptor de Gupta F2	Descriptor de Gupta F3	Descriptor de Shen F2'
ROI001	0,7532	0,7477	0,5798	0,4546	0,4668	0,3555	0,4300
ROI002	0,9823	0,9741	0,8559	0,6803	0,5141	0,4024	0,6210
ROI003	0,6820	0,6801	0,5364	0,4679	0,4897	0,3780	0,4460
ROI004	0,7738	0,7741	0,6353	0,5356	0,5407	0,4297	0,5125
ROI005	0,6645	0,6638	0,4685	0,4747	0,3842	0,2783	0,4250
ROI006	0,1871	0,1495	0,1744	0,1684	0,3899	0,2837	0,1916
ROI007	0,6610	0,6504	0,0000	0,0000	0,0000	0,0000	0,0000
ROI008	0,7346	0,7335	0,5166	0,4011	0,3655	0,2617	0,3641
ROI009	0,5360	0,5363	0,2665	0,2252	0,1859	0,1182	0,1924
ROI010	0,9166	0,9186	0,8655	0,8202	0,7209	0,6309	0,7950
ROI011	0,8997	0,9003	0,7961	0,7342	0,6262	0,5215	0,6965
ROI012	0,9667	0,9673	0,9693	0,8369	0,8138	0,7461	0,8353
ROI013	0,9637	0,9642	0,9534	0,8709	0,8009	0,7295	0,8610
ROI014	0,6939	0,6886	0,5053	0,4091	0,4237	0,3145	0,3842
ROI015	0,7579	0,7546	0,4746	0,3144	0,2868	0,1953	0,2804
ROI016	0,9026	0,8822	0,0168	0,0570	0,0843	0,0498	0,0541
ROI017	0,8464	0,8479	0,7250	0,6570	0,5834	0,4746	0,6216
ROI018	0,7501	0,7494	0,5277	0,4162	0,3688	0,2647	0,3764
ROI019	0,6751	0,6739	0,2856	0,2290	0,1267	0,0772	0,1813
ROI020	0,6789	0,6775	0,3311	0,2529	0,1708	0,1074	0,2086
ROI021	0,8829	0,8821	0,7387	0,5455	0,5237	0,4121	0,5161
ROI022	0,9340	0,9339	0,8245	0,6913	0,5797	0,4707	0,6485
ROI023	0,8838	0,8850	0,7872	0,7498	0,6358	0,5323	0,7121
ROI024	0,4938	0,4885	0,5536	0,4700	0,8428	0,7837	0,5250
ROI025	0,7316	0,7254	0,4520	0,3144	0,2831	0,1924	0,2795
ROI026	0,8620	0,8583	0,7469	0,6472	0,6165	0,5107	0,6225
ROI027	0,6563	0,6536	0,5302	0,4835	0,5083	0,3965	0,4629
ROI028	0,6475	0,6470	0,6286	0,7213	0,7776	0,7002	0,7264
ROI029	0,0563	0,0435	0,3652	0,3267	0,9542	0,9349	0,4195
ROI030	0,9365	0,9348	0,8776	0,7744	0,7143	0,6231	0,7545
ROI031	0,6618	0,6620	0,6814	0,7305	0,8837	0,8378	0,7612
ROI032	0,5398	0,5390	0,4956	0,5969	0,6297	0,5254	0,5845
ROI033	0,7378	0,7249	0,7496	0,6564	0,9704	0,9577	0,7162
ROI034	0,0000	0,0000	0,1369	0,3030	0,4173	0,3086	0,3007
ROI035	0,9014	0,8965	0,8597	0,8026	0,7783	0,7011	0,7961
ROI036	0,6498	0,6300	0,6089	0,4329	0,7582	0,6764	0,4762
ROI037	0,6017	0,5868	0,5146	0,4632	0,5950	0,4873	0,4667
ROI038	0,9690	0,9681	0,9405	0,8162	0,7346	0,6475	0,7955
ROI039	0,6827	0,6612	0,4992	0,3768	0,4545	0,3437	0,3661
ROI040	0,6418	0,6413	0,6825	0,7112	0,9499	0,9287	0,7602
ROI041	0,4859	0,4826	0,3876	0,3924	0,5130	0,4013	0,3914
ROI042	0,5150	0,5117	0,3966	0,3956	0,4698	0,3584	0,3843
ROI043	0,6963	0,6877	0,6653	0,6250	0,7923	0,7187	0,6477
ROI044	0,7640	0,7597	0,5258	0,3724	0,4013	0,2939	0,3506
ROI045	0,7117	0,7076	0,5816	0,5152	0,5741	0,4648	0,5042
ROI046	0,7209	0,7215	0,5739	0,5080	0,5265	0,4150	0,4868
ROI047	0,6337	0,6351	0,4621	0,4290	0,4403	0,3301	0,4036
ROI048	0,7073	0,7029	0,2469	0,1756	0,1296	0,0791	0,1450
ROI049	0,6810	0,6747	0,4026	0,3905	0,3095	0,2139	0,3423
ROI050	0,7802	0,7723	0,5989	0,5011	0,4767	0,3652	0,4690
ROI051	0,5568	0,5545	0,2707	0,2233	0,1859	0,1182	0,1911
ROI052	0,3183	0,2744	0,3318	0,2197	0,5615	0,4516	0,2624
ROI053	0,5945	0,5792	0,5588	0,4565	0,7009	0,6072	0,4843
ROI054	0,6110	0,6021	0,5169	0,3822	0,5822	0,4735	0,3980
ROI055	0,6636	0,6466	0,5038	0,4033	0,4696	0,3582	0,3903
ROI056	0,4388	0,4375	0,3457	0,4513	0,4320	0,3223	0,4189
ROI057	0,5198	0,5175	0,4549	0,5753	0,5604	0,4502	0,5493
ROI058	0,4871	0,4775	0,4490	0,5142	0,5868	0,4784	0,5063
ROI059	0,4707	0,4700	0,3307	0,3341	0,3721	0,2676	0,3147
ROI060	0,3687	0,3155	0,0480	0,0146	0,1012	0,0605	0,0283
ROI061	0,6085	0,5962	0,3303	0,2550	0,2255	0,1474	0,2228

Tabela 11.3 – Resultados normalizados dos atributos obtidos da extração de atributos de forma das imagens das regiões de interesse originais(continuação).

ROI	Descriptor de Hu 1	Descriptor de Hu 2	Descriptor de Hu 3	Descriptor de Gupta F1(F1')	Descriptor de Gupta F2	Descriptor de Gupta F3	Descriptor de Shen F2'
ROI062	0,7372	0,7348	0,5500	0,4434	0,4247	0,3154	0,4110
ROI063	1,0000	1,0000	0,9726	0,7912	0,6778	0,5801	0,7583
ROI064	0,9006	0,8953	0,8251	0,6812	0,6795	0,5821	0,6671
ROI065	0,8651	0,8643	0,7230	0,5459	0,5351	0,4238	0,5192
ROI066	0,8979	0,8937	0,9339	0,8612	1,0000	1,0000	0,9065
ROI067	0,9723	0,9713	1,0000	1,0000	0,8932	0,8506	1,0000
ROI068	0,9062	0,9053	0,8688	0,8559	0,7713	0,6924	0,8397
ROI101	0,3319	0,3520	0,2105	0,3841	0,2616	0,1798	0,3206
ROI102	0,5534	0,5814	0,4742	0,5546	0,3632	0,2652	0,4745
ROI103	0,4710	0,4925	0,2627	0,3818	0,1762	0,1147	0,3011
ROI104	0,7317	0,7648	0,9568	1,0000	0,9608	0,9449	1,0000
ROI105	0,7015	0,7322	0,8041	0,7967	0,6830	0,5912	0,7502
ROI106	0,6806	0,7110	0,8169	0,8401	0,7645	0,6882	0,8078
ROI107	0,1861	0,1489	0,2780	0,2939	0,6116	0,5115	0,3082
ROI108	0,6812	0,7127	0,6889	0,7568	0,4711	0,3655	0,6630
ROI109	0,1897	0,2029	0,0845	0,3445	0,2510	0,1714	0,2885
ROI110	0,4765	0,4854	0,0234	0,2425	0,0354	0,0209	0,1724
ROI111	0,4664	0,4858	0,4480	0,5699	0,4674	0,3620	0,5104
ROI112	0,0881	0,0989	0,0694	0,3848	0,3165	0,2249	0,3321
ROI113	0,4888	0,5123	0,3195	0,4143	0,2265	0,1523	0,3361
ROI114	0,4400	0,4586	0,5230	0,6239	0,6698	0,5761	0,5991
ROI115	0,4754	0,5008	0,6664	0,8301	0,9530	0,9340	0,8429
ROI116	0,5468	0,5620	0,5942	0,6265	0,6115	0,5109	0,5888
ROI117	0,6816	0,7045	0,7295	0,7290	0,5649	0,4608	0,6640
ROI118	0,4727	0,5013	0,6719	0,9849	1,0000	1,0000	0,9959
ROI119	0,4142	0,4374	0,4696	0,6173	0,5986	0,4968	0,5783
ROI120	0,2895	0,2988	0,2193	0,3443	0,3321	0,2382	0,3037
ROI121	0,4252	0,4422	0,1966	0,3495	0,1527	0,0979	0,2724
ROI122	0,0478	0,0519	0,1511	0,3438	0,4975	0,3918	0,3320
ROI123	0,3122	0,3302	0,2100	0,4140	0,2845	0,1983	0,3482
ROI124	0,5250	0,5461	0,7201	0,8195	0,9871	0,9816	0,8408
ROI125	0,5223	0,5412	0,4750	0,5084	0,4229	0,3195	0,4512
ROI126	0,0000	0,0000	0,1032	0,4819	0,4790	0,3735	0,4414
ROI127	0,5884	0,6186	0,5623	0,6556	0,4530	0,3480	0,5764
ROI128	0,5716	0,6003	0,6823	0,9189	0,7453	0,6648	0,8719
ROI129	0,1994	0,2165	0,2608	0,5292	0,5088	0,4030	0,4861
ROI130	0,5857	0,6106	0,6018	0,6676	0,5425	0,4373	0,6076
ROI131	0,4053	0,4185	0,4401	0,5842	0,5686	0,4648	0,5441
ROI132	0,7137	0,7448	0,7843	0,7457	0,5759	0,4725	0,6807
ROI133	0,4518	0,4727	0,4496	0,5812	0,4889	0,3831	0,5244
ROI134	0,5574	0,5848	0,7238	0,7632	0,9199	0,8887	0,7749
ROI135	0,3754	0,3886	0,3601	0,4830	0,4423	0,3378	0,4348
ROI136	1,0000	1,0000	0,0000	0,0000	0,0000	0,0000	0,0000
ROI137	0,6548	0,6856	0,6752	0,6740	0,5189	0,4131	0,6073
ROI138	0,6589	0,6896	0,6547	0,7267	0,4668	0,3613	0,6374
ROI139	0,5915	0,6066	0,7461	0,7384	0,8869	0,8445	0,7455
ROI140	0,6126	0,6353	0,7483	0,7933	0,8155	0,7517	0,7787
ROI141	0,5632	0,5646	0,6192	0,5528	0,6632	0,5686	0,5369
ROI142	0,3820	0,4008	0,3634	0,5330	0,4370	0,3328	0,4739
ROI143	0,3808	0,4012	0,4333	0,6533	0,5806	0,4775	0,6045
ROI144	0,2281	0,2149	0,2785	0,3881	0,5130	0,4072	0,3709
ROI145	0,5933	0,6144	0,5889	0,6550	0,4947	0,3888	0,5859
ROI146	0,8149	0,8487	1,0000	0,8911	0,6404	0,5428	0,8200
ROI147	0,7532	0,7867	0,8471	0,7756	0,5634	0,4592	0,7025
ROI148	0,7221	0,7538	0,9153	0,9306	0,8813	0,8370	0,9171
ROI149	0,7480	0,7789	0,9434	0,9564	0,8332	0,7743	0,9279
ROI150	0,7441	0,7750	0,9387	0,9401	0,8597	0,8086	0,9201

Tabela 11.3 – Resultados normalizados dos atributos obtidos da extração de atributos de forma das imagens das regiões de interesse originais(continuação).

ROI	Descriptor de Shen F3'	Descriptor de Shen F3'-F1'	Descriptor de Fourier	Circularidade	Solidez	Descriptor de Wee 1	Descriptor de Wee 2
ROI001	0,4164	0,5902	0,7315	0,2975	0,9126	0,3114	0,4759
ROI002	0,5951	0,4196	0,7315	0,1429	0,9306	0,8347	0,5385
ROI003	0,4333	0,5726	0,7315	0,2872	0,9229	0,2502	0,4925
ROI004	0,4994	0,5068	0,7315	0,2342	0,9383	0,3400	0,5483
ROI005	0,4027	0,6097	0,7315	0,2813	0,8715	0,2332	0,3910
ROI006	0,1954	0,8000	0,7315	0,6362	0,8766	0,0515	0,3595
ROI007	0,0000	1,0000	0,0000	1,0000	0,0000	0,0440	0,0000
ROI008	0,3465	0,6629	0,7315	0,3461	0,8612	0,2969	0,3786
ROI009	0,1781	0,8300	0,4047	0,5493	0,6607	0,1600	0,1966
ROI010	0,7827	0,2238	1,0000	0,0722	0,9743	0,6218	0,7335
ROI011	0,6788	0,3308	0,7315	0,1134	0,9589	0,5513	0,6401
ROI012	0,8328	0,1679	1,0000	0,0648	0,9871	0,8097	0,8279
ROI013	0,8553	0,1474	1,0000	0,0501	0,9846	0,7759	0,8146
ROI014	0,3706	0,6361	0,7315	0,3387	0,8946	0,2486	0,4298
ROI015	0,2644	0,7442	0,7315	0,4359	0,7969	0,2929	0,3036
ROI016	0,0516	0,9493	0,4047	0,8542	0,4062	0,0000	0,0455
ROI017	0,6043	0,4048	0,7315	0,1561	0,9486	0,4479	0,5949
ROI018	0,3578	0,6523	0,7315	0,3314	0,8638	0,3058	0,3823
ROI019	0,1628	0,8486	0,4047	0,5434	0,5321	0,1954	0,1447
ROI020	0,1905	0,8202	0,4047	0,5110	0,6324	0,2129	0,1879
ROI021	0,5008	0,5069	0,7315	0,2283	0,9332	0,5235	0,5407
ROI022	0,6286	0,3823	0,7315	0,1370	0,9486	0,6630	0,5985
ROI023	0,6944	0,3152	0,7315	0,1060	0,9814	0,5231	0,6484
ROI024	0,5423	0,4452	1,0000	0,2857	0,9897	0,1464	0,8086
ROI025	0,2633	0,7456	0,7315	0,4359	0,7918	0,2696	0,2986
ROI026	0,6092	0,3974	0,7315	0,1620	0,9563	0,4612	0,6273
ROI027	0,4506	0,5550	0,7315	0,2739	0,9280	0,2374	0,5088
ROI028	0,7256	0,2737	1,0000	0,1208	0,9820	0,2239	0,7616
ROI029	0,4512	0,5273	1,0000	0,4227	0,9974	0,0414	0,8655
ROI030	0,7440	0,2612	1,0000	0,0943	0,9743	0,6569	0,7282
ROI031	0,7712	0,2218	1,0000	0,1163	0,9923	0,2384	0,8642
ROI032	0,5757	0,4280	0,7315	0,1929	0,9589	0,1608	0,6118
ROI033	0,7378	0,2481	1,0000	0,1561	1,0000	0,2789	0,9519
ROI034	0,2951	0,7063	0,7315	0,4492	0,8895	0,0392	0,3787
ROI035	0,7912	0,2107	1,0000	0,0810	0,9820	0,5582	0,7868
ROI036	0,4883	0,5021	1,0000	0,3166	0,9794	0,2116	0,7419
ROI037	0,4634	0,5366	0,7315	0,2916	0,9537	0,1856	0,5837
ROI038	0,7849	0,2204	1,0000	0,0736	0,9769	0,8109	0,7513
ROI039	0,3575	0,6458	0,7315	0,3697	0,9075	0,2255	0,4567
ROI040	0,7779	0,2106	1,0000	0,1267	0,9974	0,2177	0,9235
ROI041	0,3862	0,6148	0,7315	0,3535	0,9306	0,1207	0,4941
ROI042	0,3755	0,6280	0,7315	0,3505	0,9152	0,1386	0,4578
ROI043	0,6531	0,3420	1,0000	0,1753	0,9846	0,2540	0,7795
ROI044	0,3383	0,6676	0,7315	0,3741	0,8817	0,2806	0,4115
ROI045	0,4955	0,5078	0,7315	0,2504	0,9486	0,2556	0,5730
ROI046	0,4745	0,5313	0,7315	0,2548	0,9332	0,2692	0,5292
ROI047	0,3898	0,6170	0,7315	0,3196	0,9023	0,1987	0,4399
ROI048	0,1323	0,8752	0,4047	0,6244	0,5398	0,1458	0,1412
ROI049	0,3212	0,6908	0,7315	0,3549	0,8175	0,2014	0,3170
ROI050	0,4527	0,5556	0,7315	0,2607	0,9177	0,3272	0,4865
ROI051	0,1770	0,8310	0,4047	0,5508	0,6607	0,1595	0,1964
ROI052	0,2729	0,7179	0,7315	0,5582	0,9434	0,0848	0,5296
ROI053	0,4904	0,5038	1,0000	0,2960	0,9717	0,1892	0,6842
ROI054	0,3990	0,5981	0,7315	0,3638	0,9486	0,1949	0,5734
ROI055	0,3809	0,6230	0,7315	0,3432	0,9152	0,2274	0,4713
ROI056	0,4027	0,6057	0,7315	0,3004	0,8972	0,1227	0,4196
ROI057	0,5354	0,4715	0,7315	0,2077	0,9434	0,1550	0,5459
ROI058	0,4989	0,5037	0,7315	0,2504	0,9512	0,1434	0,5684
ROI059	0,3033	0,7021	0,7315	0,4138	0,8638	0,1311	0,3656
ROI060	0,0312	0,9659	0,4047	0,9602	0,4602	0,0591	0,0986
ROI061	0,2082	0,7999	0,4047	0,5081	0,7249	0,1875	0,2370
ROI062	0,3948	0,6136	0,7315	0,3078	0,8946	0,2999	0,4351

Tabela 11.3 – Resultados normalizados dos atributos obtidos da extração de atributos de forma das imagens das regiões de interesse originais(continuação).

ROI	Descriptor de Shen F3'	Descriptor de Shen F3'-F1'	Descriptor de Fourier	Circularidade	Solidez	Descriptor de Wee 1	Descriptor de Wee 2
ROI063	0,7428	0,2656	1,0000	0,0854	0,9692	1,0000	0,6992
ROI064	0,6581	0,3458	1,0000	0,1429	0,9692	0,5673	0,6919
ROI065	0,5050	0,5021	0,7315	0,2268	0,9357	0,4834	0,5502
ROI066	0,9246	0,0645	1,0000	0,0545	1,0000	0,5550	1,0000
ROI067	1,0000	0,0000	1,0000	0,0000	0,9949	0,8185	0,9049
ROI068	0,8313	0,1729	1,0000	0,0574	0,9820	0,5850	0,7809
ROI101	0,2941	0,7219	0,5490	0,3021	0,7454	0,0993	0,2918
ROI102	0,4406	0,5798	0,5490	0,1786	0,8380	0,2431	0,4099
ROI103	0,2693	0,7505	0,5490	0,3041	0,6111	0,1565	0,2243
ROI104	0,9992	0,0035	1,0000	0,0000	0,9954	0,5608	1,0000
ROI105	0,7285	0,2850	1,0000	0,0647	0,9630	0,4714	0,7298
ROI106	0,7921	0,2181	1,0000	0,0483	0,9769	0,4398	0,8058
ROI107	0,3098	0,6883	1,0000	0,3938	0,9491	0,0606	0,5977
ROI108	0,6230	0,4013	0,5490	0,0801	0,8981	0,4377	0,5260
ROI109	0,2649	0,7492	0,5490	0,3398	0,7315	0,0486	0,2646
ROI110	0,1469	0,8696	0,0000	0,4575	0,1852	0,0555	0,0712
ROI111	0,4837	0,5322	0,5490	0,1689	0,8981	0,1798	0,4993
ROI112	0,3091	0,7045	0,5490	0,3021	0,8009	0,0345	0,3164
ROI113	0,3045	0,7149	0,5490	0,2761	0,6991	0,1768	0,2741
ROI114	0,5856	0,4224	1,0000	0,1400	0,9583	0,1701	0,6859
ROI115	0,8466	0,1528	1,0000	0,0521	0,9954	0,1972	0,9561
ROI116	0,5702	0,4408	1,0000	0,1390	0,9491	0,2462	0,6439
ROI117	0,6349	0,3827	1,0000	0,0907	0,9352	0,4335	0,6153
ROI118	1,0000	0,0000	1,0000	0,0039	1,0000	0,1897	0,9984
ROI119	0,5592	0,4521	1,0000	0,1438	0,9444	0,1546	0,6162
ROI120	0,2851	0,7257	0,5490	0,3398	0,8148	0,0872	0,3531
ROI121	0,2424	0,7764	0,0000	0,3349	0,5648	0,1234	0,1970
ROI122	0,3236	0,6807	0,5490	0,3407	0,9120	0,0352	0,4794
ROI123	0,3205	0,6962	0,5490	0,2770	0,7685	0,0935	0,3110
ROI124	0,8482	0,1491	1,0000	0,0560	1,0000	0,2314	0,9947
ROI125	0,4256	0,5896	0,5490	0,2075	0,8750	0,2171	0,4630
ROI126	0,4219	0,5893	0,5490	0,2249	0,9028	0,0243	0,4551
ROI127	0,5422	0,4784	0,5490	0,1245	0,8935	0,2855	0,4986
ROI128	0,8503	0,1636	1,0000	0,0232	0,9722	0,2785	0,7738
ROI129	0,4655	0,5465	0,5490	0,1940	0,9167	0,0711	0,5082
ROI130	0,5805	0,4358	1,0000	0,1187	0,9259	0,2832	0,5830
ROI131	0,5247	0,4868	1,0000	0,1612	0,9352	0,1448	0,5865
ROI132	0,6515	0,3661	1,0000	0,0840	0,9352	0,5173	0,6298
ROI133	0,4986	0,5167	0,5490	0,1631	0,9074	0,1738	0,5181
ROI134	0,7776	0,2220	1,0000	0,0772	0,9954	0,2554	0,9356
ROI135	0,4126	0,6004	0,5490	0,2239	0,8843	0,1317	0,4663
ROI136	0,0000	1,0000	0,0000	1,0000	0,0000	0,0000	0,0000
ROI137	0,5776	0,4402	1,0000	0,1158	0,9213	0,3842	0,5686
ROI138	0,5992	0,4239	0,5490	0,0917	0,8981	0,3906	0,5194
ROI139	0,7460	0,2547	1,0000	0,0869	0,9907	0,2899	0,9091
ROI140	0,7701	0,2357	1,0000	0,0656	0,9815	0,3211	0,8449
ROI141	0,5268	0,4789	1,0000	0,1795	0,9583	0,2492	0,6929
ROI142	0,4475	0,5681	0,5490	0,1911	0,8843	0,1358	0,4624
ROI143	0,5816	0,4320	1,0000	0,1255	0,9398	0,1377	0,5957
ROI144	0,3604	0,6452	1,0000	0,2992	0,9167	0,0772	0,5145
ROI145	0,5554	0,4629	0,5490	0,1245	0,9120	0,2929	0,5389
ROI146	0,7885	0,2309	1,0000	0,0319	0,9537	1,0000	0,7032
ROI147	0,6702	0,3494	1,0000	0,0724	0,9352	0,6590	0,6224
ROI148	0,9098	0,0961	1,0000	0,0193	0,9907	0,5343	0,9226
ROI149	0,9143	0,0952	1,0000	0,0116	0,9861	0,6273	0,8804
ROI150	0,9101	0,0973	1,0000	0,0164	0,9861	0,5966	0,9046

Tabela 11.4 – Um exemplo dos resultados obtidos na saída da RNA com o algoritmo de treinamento de “Backpropagation”, com aproximação de Levenberg-Marquardt, 3, 4, 5 e 6 unidades escondidas, treinamento e teste com o método “leaving-one-out” e entrada com o conjunto de atributos de Textura e Forma CI_{TF} .

Identificação da ROI	Classificação Clínica	Valor de Saída 3 unidades	Valor de Saída 4 unidades	Valor de Saída 5 unidades	Valor de Saída 6 unidades
ROI001	Benigna	0.0198	0.0215	0.0186	0.0199
ROI002	Benigna	0.0191	0.0471	0.0199	0.0201
ROI003	Benigna	0.0205	0.0177	0.0197	0.0153
ROI004	Benigna	0.0175	0.0136	0.0181	0.0124
ROI005	Benigna	0.0200	0.0237	0.0256	0.0634
ROI006	Benigna	0.0186	0.0202	0.5532	0.0000
ROI007	Benigna	0.5432	0.9921	0.9987	0.2829
ROI008	Benigna	0.0175	0.0199	0.0248	0.0189
ROI009	Benigna	0.0194	0.0205	0.0163	0.0167
ROI010	Benigna	0.0222	0.0200	0.0212	0.0201
ROI011	Benigna	0.0215	0.0046	0.0181	0.0201
ROI012	Benigna	0.0161	0.0077	0.0244	0.0158
ROI013	Benigna	0.0286	0.0201	0.0197	0.0187
ROI014	Benigna	0.0190	0.0191	0.0193	0.0194
ROI015	Benigna	0.0150	0.0199	0.0171	0.0188
ROI016	Benigna	0.6498	0.1464	0.0708	0.0208
ROI017	Benigna	0.0194	0.0425	0.0104	0.0195
ROI018	Benigna	0.0200	0.0205	0.0195	0.0198
ROI019	Benigna	0.2224	0.0194	0.0215	0.0209
ROI020	Benigna	0.0205	0.0202	0.0205	0.0184
ROI021	Benigna	0.0141	0.0283	0.0202	0.0202
ROI022	Benigna	0.0180	0.0235	0.0126	0.0195
ROI023	Benigna	0.0190	0.0074	0.0172	0.0115
ROI024	Benigna	0.0278	0.0683	0.6193	0.0006
ROI025	Benigna	0.0234	0.0207	0.0192	0.0195
ROI026	Benigna	0.0161	0.0064	0.0113	0.0174
ROI027	Benigna	0.0223	0.0187	0.0207	0.0202
ROI028	Benigna	0.0195	0.0122	0.0010	0.0261
ROI029	Benigna	0.0020	0.0296	0.0172	0.0385
ROI030	Benigna	0.0053	0.0760	0.0197	0.0240
ROI031	Benigna	0.7252	0.0103	0.0198	0.4098
ROI032	Benigna	0.0239	0.0856	0.0292	0.0297
ROI033	Benigna	0.0235	0.0845	0.0166	0.0201
ROI034	Benigna	0.9738	0.8709	0.4327	0.0415
ROI035	Benigna	0.0133	0.0059	0.0127	0.0204
ROI036	Benigna	0.0275	0.0292	0.0351	0.0193
ROI037	Benigna	0.0200	0.0202	0.0166	0.0193
ROI038	Benigna	0.0198	0.0258	0.0549	0.0192
ROI039	Benigna	0.0199	0.0172	0.0202	0.0175
ROI040	Benigna	0.0207	0.0133	0.0352	0.1181

Tabela 11.4 – Um exemplo dos resultados obtidos na saída da RNA com o algoritmo de treinamento de “Backpropagation”, com aproximação de Levenberg-Marquardt, 3, 4, 5 e 6 unidades escondidas, treinamento e teste com o método “leaving-one-out” e entrada com o conjunto de atributos de Textura e Forma $C1_{TF}$ (continuação).

Identificação da ROI	Classificação Clínica	Valor de Saída 3 unidades	Valor de Saída 4 unidades	Valor de Saída 5 unidades	Valor de Saída 6 unidades
ROI041	Benigna	0.0200	0.0201	0.0200	0.0200
ROI042	Benigna	0.0196	0.0182	0.0203	0.0213
ROI043	Benigna	0.0142	0.0100	0.0139	0.0184
ROI044	Benigna	0.0190	0.0212	0.0221	0.0208
ROI045	Benigna	0.0202	0.0188	0.0149	0.0198
ROI046	Benigna	0.0200	0.0188	0.0192	0.0207
ROI047	Benigna	0.0199	0.0194	0.0150	0.0212
ROI048	Benigna	0.0230	0.0136	0.0212	0.0197
ROI049	Benigna	0.0200	0.0200	0.0200	0.0202
ROI050	Benigna	0.0187	0.0153	0.0195	0.0162
ROI051	Benigna	0.0148	0.0197	0.0112	0.0188
ROI052	Benigna	0.0200	0.0208	0.0002	0.0181
ROI053	Benigna	0.0200	0.0174	0.0001	0.0278
ROI054	Benigna	0.0187	0.0208	0.0278	0.0165
ROI055	Benigna	0.0256	0.0202	0.0172	0.0207
ROI056	Benigna	0.0193	0.0162	0.1943	0.0283
ROI057	Benigna	0.0240	0.0432	0.7208	0.0370
ROI058	Benigna	0.0200	0.0120	0.0045	0.0113
ROI059	Benigna	0.0221	0.0161	0.0157	0.0178
ROI060	Benigna	0.0196	0.0208	0.1402	0.0211
ROI061	Benigna	0.0307	0.0210	0.0207	0.0265
ROI062	Benigna	0.0199	0.0201	0.0140	0.0190
ROI063	Benigna	0.0233	0.0199	0.0161	0.0200
ROI064	Benigna	0.0149	0.0063	0.0102	0.0218
ROI065	Benigna	0.0178	0.0292	0.0189	0.0218
ROI066	Benigna	0.0198	0.0212	0.0014	0.0087
ROI067	Benigna	0.2269	0.9894	0.0194	0.6110
ROI068	Benigna	0.8581	0.0397	0.0203	0.0131
ROI101	Maligna	0.9818	0.9811	0.9843	0.9900
ROI102	Maligna	0.9715	0.1941	0.9569	0.9704
ROI103	Maligna	0.9805	0.9804	0.9870	0.9481
ROI104	Maligna	0.9800	0.8109	0.9700	0.9586
ROI105	Maligna	0.9769	0.9831	0.9696	0.9760
ROI106	Maligna	0.9801	0.9927	0.9712	0.8643
ROI107	Maligna	0.0338	0.0431	0.9517	0.9398
ROI108	Maligna	0.8879	0.9801	0.9823	0.9598
ROI109	Maligna	0.9801	0.9804	0.9905	0.9908
ROI110	Maligna	0.9814	0.9801	1.0000	0.9819
ROI111	Maligna	0.9809	0.9798	0.9777	0.9682
ROI112	Maligna	0.9807	0.9805	0.9889	0.9676
ROI113	Maligna	0.9801	0.9842	0.9972	0.9885
ROI114	Maligna	0.9812	0.9803	0.9812	0.9834

Tabela 11.4 – Um exemplo dos resultados obtidos na saída da RNA com o algoritmo de treinamento de “Backpropagation”, com aproximação de Levenberg-Marquardt, 3, 4, 5 e 6 unidades escondidas, treinamento e teste com o método “leaving-one-out” e entrada com o conjunto de atributos de Textura e Forma CI_{TF} (*continuação*).

Identificação da ROI	Classificação Clínica	Valor de Saída 3 unidades	Valor de Saída 4 unidades	Valor de Saída 5 unidades	Valor de Saída 6 unidades
ROI115	Maligna	0.9808	0.9798	0.9916	0.9826
ROI116	Maligna	0.9808	0.9827	0.9827	0.9875
ROI117	Maligna	0.9795	0.9971	0.9803	0.9783
ROI118	Maligna	0.9801	0.9794	0.9882	0.9966
ROI119	Maligna	0.9796	0.9797	0.9772	0.9813
ROI120	Maligna	0.9916	0.0201	0.9647	0.9176
ROI121	Maligna	0.9739	1.0000	0.4220	0.9921
ROI122	Maligna	0.9307	0.9768	0.9946	0.9998
ROI123	Maligna	0.9796	0.9812	0.9991	0.9860
ROI124	Maligna	0.9778	0.9812	0.9797	0.9801
ROI125	Maligna	0.9806	0.9801	0.9856	0.9665
ROI126	Maligna	0.9805	0.9800	0.9808	0.9989
ROI127	Maligna	0.9791	0.9808	0.9795	0.9702
ROI128	Maligna	0.9801	0.9801	0.9803	0.9922
ROI129	Maligna	0.9768	0.9796	0.9798	0.9819
ROI130	Maligna	0.9809	0.9798	0.9788	0.9795
ROI131	Maligna	0.9801	0.9797	0.9915	0.9686
ROI132	Maligna	0.9797	0.9800	0.9756	0.9800
ROI133	Maligna	0.9752	0.9785	0.9384	0.9859
ROI134	Maligna	0.9866	0.9805	0.9805	0.9898
ROI135	Maligna	0.9753	0.9795	0.9815	0.9823
ROI136	Maligna	0.0198	0.0204	0.0120	0.0168
ROI137	Maligna	0.9800	0.9806	0.9831	0.9812
ROI138	Maligna	0.9803	0.9802	0.9832	0.9870
ROI139	Maligna	0.9784	0.9808	0.9844	0.9794
ROI140	Maligna	0.9799	0.9798	0.9798	0.9855
ROI141	Maligna	0.9799	0.9800	0.9699	0.9166
ROI142	Maligna	0.9832	0.9821	0.9808	0.9854
ROI143	Maligna	0.9826	0.9798	0.9697	0.8706
ROI144	Maligna	0.9770	1.0000	0.9931	0.9474
ROI145	Maligna	0.9790	0.9766	0.7545	0.9664
ROI146	Maligna	0.9802	0.9789	0.9839	0.9838
ROI147	Maligna	0.9829	0.9809	0.9886	0.9852
ROI148	Maligna	0.9792	0.9800	0.9698	0.9799
ROI149	Maligna	0.9732	0.9757	0.9815	0.9782
ROI150	Maligna	0.0031	0.0193	0.0196	0.3933

Tabela 11.5 – Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C1_{TF}$.

Unidades Escondidas	Limiar	Benigna e Maligna				Benigna				Maligna			
		Acerto		Erro		Acerto		Erro		Acerto		Erro	
3	0,1	108	91,53%	10	8,47%	61	89,71%	7	10,29%	47	94,00%	3	6,00%
	0,2	110	93,22%	8	6,78%	63	92,65%	5	7,35%	47	94,00%	3	6,00%
	0,3	110	93,22%	8	6,78%	63	92,65%	5	7,35%	47	94,00%	3	6,00%
	0,4	110	93,22%	8	6,78%	63	92,65%	5	7,35%	47	94,00%	3	6,00%
	0,5	111	94,07%	7	5,93%	64	94,12%	4	5,88%	47	94,00%	3	6,00%
	0,6	112	94,92%	6	5,08%	65	95,59%	3	4,41%	47	94,00%	3	6,00%
	0,7	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,8	113	95,76%	5	3,39%	66	97,06%	2	1,47%	47	94,00%	3	6,00%
	0,9	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
4	0,1	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,2	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,3	112	94,92%	6	5,08%	65	95,59%	3	4,41%	47	94,00%	3	6,00%
	0,4	112	94,92%	6	5,08%	65	95,59%	3	4,41%	47	94,00%	3	6,00%
	0,5	112	94,92%	6	5,08%	65	95,59%	3	4,41%	47	94,00%	3	6,00%
	0,6	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,7	114	96,61%	4	3,39%	67	98,53%	1	1,47%	47	94,00%	3	6,00%
	0,8	113	95,76%	5	4,24%	67	98,53%	1	1,47%	46	92,00%	4	8,00%
	0,9	110	93,22%	8	6,78%	67	98,53%	1	1,47%	43	86,00%	7	14,00%
5	0,1	111	94,07%	7	5,93%	64	94,12%	4	5,88%	47	94,00%	3	6,00%
	0,2	112	94,92%	6	5,08%	65	95,59%	3	4,41%	47	94,00%	3	6,00%
	0,3	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,4	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,5	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,6	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,7	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,8	112	94,92%	6	5,08%	66	97,06%	2	2,94%	46	92,00%	4	8,00%
	0,9	110	93,22%	8	6,78%	66	97,06%	2	2,94%	44	88,00%	6	12,00%
6	0,1	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,2	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,3	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,4	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,5	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,6	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%
	0,7	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%
	0,8	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%
	0,9	111	94,07%	7	5,93%	68	100,00%	0	0,00%	43	86,00%	7	14,00%

Tabela 11.6 – Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C2_{TF}$.

Unidades Escondidas	Limiar	Benigna e Maligna				Benigna				Maligna			
		Acerto		Erro		Acerto		Erro		Acerto		Erro	
		104	88,14%	14	11,86%	61	89,71%	7	10,29%	43	86,00%	7	14,00%
3	0,1	104	88,98%	13	11,02%	62	91,18%	6	8,82%	43	86,00%	7	14,00%
	0,2	106	89,83%	12	10,17%	63	92,65%	5	7,35%	43	86,00%	7	14,00%
	0,3	105	88,98%	13	11,02%	63	92,65%	5	7,35%	42	84,00%	8	16,00%
	0,4	106	89,83%	12	10,17%	64	94,12%	4	5,88%	42	84,00%	8	16,00%
	0,5	106	89,83%	12	10,17%	64	94,12%	4	5,88%	42	84,00%	8	16,00%
	0,6	106	89,83%	12	10,17%	64	94,12%	4	5,88%	42	84,00%	8	16,00%
	0,7	106	89,83%	12	10,17%	64	94,12%	4	5,88%	42	84,00%	8	16,00%
	0,8	105	88,98%	13	11,02%	64	94,12%	4	5,88%	41	82,00%	9	18,00%
	0,9	107	90,68%	11	9,32%	66	97,06%	2	2,94%	41	82,00%	9	18,00%
4	0,1	111	94,07%	7	5,93%	65	95,59%	3	4,41%	46	92,00%	4	8,00%
	0,2	110	93,22%	8	6,78%	65	95,59%	3	4,41%	45	90,00%	5	10,00%
	0,3	110	93,22%	8	6,78%	65	95,59%	3	4,41%	45	90,00%	5	10,00%
	0,4	110	93,22%	8	6,78%	65	95,59%	3	4,41%	45	90,00%	5	10,00%
	0,5	110	93,22%	8	6,78%	65	95,59%	3	4,41%	45	90,00%	5	10,00%
	0,6	110	93,22%	8	6,78%	65	95,59%	3	4,41%	45	90,00%	5	10,00%
	0,7	110	93,22%	8	6,78%	65	95,59%	3	4,41%	45	90,00%	5	10,00%
	0,8	109	92,37%	9	7,63%	65	95,59%	3	4,41%	44	88,00%	6	12,00%
	0,9	110	93,22%	8	6,78%	66	97,06%	2	2,94%	44	88,00%	6	12,00%
5	0,1	110	93,22%	8	6,78%	62	91,18%	6	8,82%	48	96,00%	2	4,00%
	0,2	111	94,07%	7	5,93%	63	92,65%	5	7,35%	48	96,00%	2	4,00%
	0,3	111	94,07%	7	5,93%	63	92,65%	5	7,35%	48	96,00%	2	4,00%
	0,4	111	94,07%	7	5,93%	64	94,12%	4	5,88%	47	94,00%	3	6,00%
	0,5	111	94,07%	7	5,93%	64	94,12%	4	5,88%	47	94,00%	3	6,00%
	0,6	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,7	114	96,61%	4	3,39%	67	98,53%	1	1,47%	47	94,00%	3	6,00%
	0,8	113	95,76%	5	4,24%	67	98,53%	1	1,47%	46	92,00%	4	8,00%
	0,9	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
6	0,1	113	95,76%	5	4,24%	65	95,59%	3	4,41%	48	96,00%	2	4,00%
	0,2	113	95,76%	5	4,24%	65	95,59%	3	4,41%	48	96,00%	2	4,00%
	0,3	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%
	0,4	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%
	0,5	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%
	0,6	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,7	114	96,61%	4	3,39%	67	98,53%	1	1,47%	47	94,00%	3	6,00%
	0,8	114	96,61%	4	3,39%	68	100,00%	0	0,00%	46	92,00%	4	8,00%
	0,9	113	95,76%	5	4,24%	68	100,00%	0	0,00%	45	90,00%	5	10,00%

Tabela 11.7 – Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura e forma $C3_{TF}$.

Unidades Escondidas	Limiar	Benigna e Maligna				Benigna				Maligna			
		Acerto		Erro		Acerto		Erro		Acerto		Erro	
		1	0,85%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
3	0,1	117	99,15%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
	0,2	117	99,15%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
	0,3	117	99,15%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
	0,4	117	99,15%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
	0,5	117	99,15%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
	0,6	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%
	0,7	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%
	0,8	115	97,46%	3	2,54%	68	100,00%	0	0,00%	47	94,00%	3	6,00%
	0,9	111	94,07%	7	5,93%	68	100,00%	0	0,00%	43	86,00%	7	14,00%
4	0,1	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,2	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,3	113	95,76%	5	4,24%	65	95,59%	3	4,41%	48	96,00%	2	4,00%
	0,4	113	95,76%	5	4,24%	65	95,59%	3	4,41%	48	96,00%	2	4,00%
	0,5	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%
	0,6	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,7	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,8	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,9	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
5	0,1	113	95,76%	5	4,24%	65	95,59%	3	4,41%	48	96,00%	2	4,00%
	0,2	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%
	0,3	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,4	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%
	0,5	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%
	0,6	115	97,46%	3	2,54%	68	100,00%	0	0,00%	47	94,00%	3	6,00%
	0,7	114	96,61%	4	3,39%	68	100,00%	0	0,00%	46	92,00%	4	8,00%
	0,8	114	96,61%	4	3,39%	68	100,00%	0	0,00%	46	92,00%	4	8,00%
	0,9	109	92,37%	9	7,63%	68	100,00%	0	0,00%	41	82,00%	9	18,00%
6	0,1	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,2	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,3	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,4	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,5	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,6	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%
	0,7	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%
	0,8	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%
	0,9	113	95,76%	5	4,24%	68	100,00%	0	0,00%	45	90,00%	5	10,00%

Tabela 11.8 – Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura CI_T .

Unidades Escondidas	Limiar	Benigna e Maligna				Benigna				Maligna			
		Acerto		Erro		Acerto		Erro		Acerto		Erro	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
3	0,1	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,2	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,3	113	95,76%	5	4,24%	67	98,53%	1	1,47%	46	92,00%	4	8,00%
	0,4	113	95,76%	5	4,24%	67	98,53%	1	1,47%	46	92,00%	4	8,00%
	0,5	113	95,76%	5	4,24%	67	98,53%	1	1,47%	46	92,00%	4	8,00%
	0,6	113	95,76%	5	4,24%	67	98,53%	1	1,47%	46	92,00%	4	8,00%
	0,7	114	96,61%	4	3,39%	68	100,00%	0	0,00%	46	92,00%	4	8,00%
	0,8	113	95,76%	5	4,24%	68	100,00%	0	0,00%	45	90,00%	5	10,00%
	0,9	112	94,92%	6	5,08%	68	100,00%	0	0,00%	44	88,00%	6	12,00%
4	0,1	108	91,53%	10	8,47%	63	92,65%	5	7,35%	45	90,00%	5	10,00%
	0,2	109	92,37%	9	7,63%	64	94,12%	4	5,88%	45	90,00%	5	10,00%
	0,3	109	92,37%	9	7,63%	64	94,12%	4	5,88%	45	90,00%	5	10,00%
	0,4	109	92,37%	9	7,63%	64	94,12%	4	5,88%	45	90,00%	5	10,00%
	0,5	109	92,37%	9	7,63%	64	94,12%	4	5,88%	45	90,00%	5	10,00%
	0,6	109	92,37%	9	7,63%	64	94,12%	4	5,88%	45	90,00%	5	10,00%
	0,7	110	93,22%	8	6,78%	65	95,59%	3	4,41%	45	90,00%	5	10,00%
	0,8	110	93,22%	8	6,78%	66	97,06%	2	2,94%	44	88,00%	6	12,00%
	0,9	108	91,53%	10	8,47%	66	97,06%	2	2,94%	42	84,00%	8	16,00%
5	0,1	112	94,92%	6	5,08%	65	95,59%	3	4,41%	47	94,00%	3	6,00%
	0,2	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,3	111	94,07%	7	5,93%	66	97,06%	2	2,94%	45	90,00%	5	10,00%
	0,4	111	94,07%	7	5,93%	66	97,06%	2	2,94%	45	90,00%	5	10,00%
	0,5	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
	0,6	111	94,07%	7	5,93%	67	98,53%	1	1,47%	44	88,00%	6	12,00%
	0,7	111	94,07%	7	5,93%	67	98,53%	1	1,47%	44	88,00%	6	12,00%
	0,8	109	92,37%	9	7,63%	67	98,53%	1	1,47%	42	84,00%	8	16,00%
	0,9	109	92,37%	9	7,63%	68	100,00%	0	0,00%	41	82,00%	9	18,00%
6	0,1	112	94,92%	6	5,08%	68	100,00%	0	0,00%	44	88,00%	6	12,00%
	0,2	112	94,92%	6	5,08%	68	100,00%	0	0,00%	44	88,00%	6	12,00%
	0,3	112	94,92%	6	5,08%	68	100,00%	0	0,00%	44	88,00%	6	12,00%
	0,4	112	94,92%	6	5,08%	68	100,00%	0	0,00%	44	88,00%	6	12,00%
	0,5	112	94,92%	6	5,08%	68	100,00%	0	0,00%	44	88,00%	6	12,00%
	0,6	112	94,92%	6	5,08%	68	100,00%	0	0,00%	44	88,00%	6	12,00%
	0,7	112	94,92%	6	5,08%	68	100,00%	0	0,00%	44	88,00%	6	12,00%
	0,8	109	92,37%	9	7,63%	68	100,00%	0	0,00%	41	82,00%	9	18,00%
	0,9	109	92,37%	9	7,63%	68	100,00%	0	0,00%	41	82,00%	9	18,00%



Tabela 11.9 – Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura $C2_T$.

Unidades Escondidas	Limiar	Benigna e Maligna				Benigna				Maligna			
		Acerto		Erro		Acerto		Erro		Acerto		Erro	
3	0,1	111	94,07%	7	5,93%	66	97,06%	2	2,94%	45	90,00%	5	10,00%
	0,2	111	94,07%	7	5,93%	66	97,06%	2	2,94%	45	90,00%	5	10,00%
	0,3	111	94,07%	7	5,93%	66	97,06%	2	2,94%	45	90,00%	5	10,00%
	0,4	111	94,07%	7	5,93%	66	97,06%	2	2,94%	45	90,00%	5	10,00%
	0,5	111	94,07%	7	5,93%	66	97,06%	2	2,94%	45	90,00%	5	10,00%
	0,6	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
	0,7	110	93,22%	8	6,78%	67	98,53%	1	1,47%	43	86,00%	7	14,00%
	0,8	110	93,22%	8	6,78%	67	98,53%	1	1,47%	43	86,00%	7	14,00%
	0,9	110	93,22%	8	6,78%	67	98,53%	1	1,47%	43	86,00%	7	14,00%
4	0,1	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,2	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,3	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,4	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,5	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,6	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,7	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,8	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,9	111	94,07%	7	5,93%	67	98,53%	1	1,47%	44	88,00%	6	12,00%
5	0,1	113	95,76%	5	4,24%	65	95,59%	3	4,41%	48	96,00%	2	4,00%
	0,2	112	94,92%	6	5,08%	65	95,59%	3	4,41%	47	94,00%	3	6,00%
	0,3	112	94,92%	6	5,08%	65	95,59%	3	4,41%	47	94,00%	3	6,00%
	0,4	112	94,92%	6	5,08%	65	95,59%	3	4,41%	47	94,00%	3	6,00%
	0,5	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,6	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,7	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,8	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
	0,9	111	94,07%	7	5,93%	66	97,06%	2	2,94%	45	90,00%	5	10,00%
6	0,1	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,2	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,3	110	93,22%	8	6,78%	64	94,12%	4	5,88%	46	92,00%	4	8,00%
	0,4	110	93,22%	8	6,78%	64	94,12%	4	5,88%	46	92,00%	4	8,00%
	0,5	110	93,22%	8	6,78%	64	94,12%	4	5,88%	46	92,00%	4	8,00%
	0,6	110	93,22%	8	6,78%	64	94,12%	4	5,88%	46	92,00%	4	8,00%
	0,7	111	94,07%	7	5,93%	65	95,59%	3	4,41%	46	92,00%	4	8,00%
	0,8	111	94,07%	7	5,93%	65	95,59%	3	4,41%	46	92,00%	4	8,00%
	0,9	109	92,37%	9	7,63%	65	95,59%	3	4,41%	44	88,00%	6	12,00%

Tabela 11.10 – Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura $C3_T$.

Unidades Escondidas	Limiar	Benigna e Maligna				Benigna				Maligna			
		Acerto		Erro		Acerto		Erro		Acerto		Erro	
		Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
3	0,1	114	96,61%	4	3,39%	67	98,53%	1	1,47%	47	94,00%	3	6,00%
	0,2	113	95,76%	5	4,24%	67	98,53%	1	1,47%	46	92,00%	4	8,00%
	0,3	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
	0,4	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
	0,5	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
	0,6	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
	0,7	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
	0,8	111	94,07%	7	5,93%	67	98,53%	1	1,47%	44	88,00%	6	12,00%
	0,9	110	93,22%	8	6,78%	67	98,53%	1	1,47%	43	86,00%	7	14,00%
4	0,1	112	94,92%	6	5,08%	66	97,06%	2	2,94%	46	92,00%	4	8,00%
	0,2	112	94,92%	6	5,08%	66	97,06%	2	2,94%	46	92,00%	4	8,00%
	0,3	112	94,92%	6	5,08%	66	97,06%	2	2,94%	46	92,00%	4	8,00%
	0,4	112	94,92%	6	5,08%	66	97,06%	2	2,94%	46	92,00%	4	8,00%
	0,5	112	94,92%	6	5,08%	66	97,06%	2	2,94%	46	92,00%	4	8,00%
	0,6	112	94,92%	6	5,08%	66	97,06%	2	2,94%	46	92,00%	4	8,00%
	0,7	112	94,92%	6	5,08%	66	97,06%	2	2,94%	46	92,00%	4	8,00%
	0,8	112	94,92%	6	5,08%	67	98,53%	1	1,47%	45	90,00%	5	10,00%
	0,9	111	94,07%	7	5,93%	67	98,53%	1	1,47%	44	88,00%	6	12,00%
5	0,1	109	92,37%	9	7,63%	65	95,59%	3	4,41%	44	88,00%	6	12,00%
	0,2	110	93,22%	8	6,78%	66	97,06%	2	2,94%	44	88,00%	6	12,00%
	0,3	110	93,22%	8	6,78%	66	97,06%	2	2,94%	44	88,00%	6	12,00%
	0,4	110	93,22%	8	6,78%	66	97,06%	2	2,94%	44	88,00%	6	12,00%
	0,5	110	93,22%	8	6,78%	66	97,06%	2	2,94%	44	88,00%	6	12,00%
	0,6	110	93,22%	8	6,78%	66	97,06%	2	2,94%	44	88,00%	6	12,00%
	0,7	110	93,22%	8	6,78%	66	97,06%	2	2,94%	44	88,00%	6	12,00%
	0,8	111	94,07%	7	5,93%	67	98,53%	1	1,47%	44	88,00%	6	12,00%
	0,9	109	92,37%	9	7,63%	67	98,53%	1	1,47%	42	84,00%	8	16,00%
6	0,1	109	92,37%	9	7,63%	63	92,65%	5	7,35%	46	92,00%	4	8,00%
	0,2	109	92,37%	9	7,63%	64	94,12%	4	5,88%	45	90,00%	5	10,00%
	0,3	108	91,53%	10	8,47%	64	94,12%	4	5,88%	44	88,00%	6	12,00%
	0,4	108	91,53%	10	8,47%	64	94,12%	4	5,88%	44	88,00%	6	12,00%
	0,5	108	91,53%	10	8,47%	64	94,12%	4	5,88%	44	88,00%	6	12,00%
	0,6	109	92,37%	9	7,63%	65	95,59%	3	4,41%	44	88,00%	6	12,00%
	0,7	109	92,37%	9	7,63%	65	95,59%	3	4,41%	44	88,00%	6	12,00%
	0,8	109	92,37%	9	7,63%	65	95,59%	3	4,41%	44	88,00%	6	12,00%
	0,9	110	93,22%	8	6,78%	66	97,06%	2	2,94%	44	88,00%	6	12,00%

Tabela 11.11 – Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura CI_F .

Unidades Escondidas	Limiar	Benigna e Maligna				Benigna				Maligna			
		Acerto		Erro		Acerto		Erro		Acerto		Erro	
3	0,1	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,2	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,3	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,4	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,5	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,6	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,7	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,8	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
	0,9	112	94,92%	6	5,08%	64	94,12%	4	5,88%	48	96,00%	2	4,00%
4	0,1	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,2	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,3	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,4	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,5	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,6	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,7	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%
	0,8	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%
	0,9	113	95,76%	5	4,24%	66	97,06%	2	2,94%	47	94,00%	3	6,00%
5	0,1	113	95,76%	5	4,24%	64	94,12%	4	5,88%	49	98,00%	1	2,00%
	0,2	113	95,76%	5	4,24%	64	94,12%	4	5,88%	49	98,00%	1	2,00%
	0,3	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,4	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,5	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,6	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,7	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,8	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,9	112	94,92%	6	5,08%	66	97,06%	2	2,94%	46	92,00%	4	8,00%
6	0,1	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,2	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,3	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,4	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,5	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,6	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,7	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,8	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,9	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%

Tabela 11.12 – Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura $C2_F$.

Unidades Escondidas	Limiar	Benigna e Maligna				Benigna				Maligna			
		Acerto		Erro		Acerto		Erro		Acerto		Erro	
3	0,1	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	3,00%
	0,2	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,3	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,4	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,5	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,6	117	99,15%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
	0,7	117	99,15%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
	0,8	117	99,15%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
	0,9	115	97,46%	3	2,54%	68	100,00%	0	0,00%	47	94,00%	3	6,00%
4	0,1	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,2	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,3	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,4	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,5	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,6	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,7	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,8	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,9	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
5	0,1	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,2	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,3	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,4	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,5	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,6	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,7	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,8	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,9	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
6	0,1	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,2	114	96,61%	4	3,39%	65	95,59%	3	4,41%	49	98,00%	1	2,00%
	0,3	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,4	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,5	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,6	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,7	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,8	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,9	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%

Tabela 11.13 – Resultados da performance do classificador (RNA) tendo como entrada o conjunto de atributos de textura $C3_F$.

Unidades Escondidas	Limiar	Benigna e Maligna				Benigna				Maligna			
		Acerto		Erro		Acerto		Erro		Acerto		Erro	
		116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	3,00%
3	0,1	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,2	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,3	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,4	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,5	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,6	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,7	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,8	117	99,15%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
	0,9	117	99,15%	1	0,85%	68	100,00%	0	0,00%	49	98,00%	1	2,00%
4	0,1	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,2	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,3	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,4	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,5	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,6	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,7	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%
	0,8	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%
	0,9	114	96,61%	4	3,39%	66	97,06%	2	2,94%	48	96,00%	2	4,00%
5	0,1	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,2	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,3	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,4	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,5	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,6	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,7	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,8	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,9	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
6	0,1	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,2	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,3	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,4	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,5	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,6	115	97,46%	3	2,54%	66	97,06%	2	2,94%	49	98,00%	1	2,00%
	0,7	116	98,31%	2	1,69%	67	98,53%	1	1,47%	49	98,00%	1	2,00%
	0,8	115	97,46%	3	2,54%	67	98,53%	1	1,47%	48	96,00%	2	4,00%
	0,9	116	98,31%	2	1,69%	68	100,00%	0	0,00%	48	96,00%	2	4,00%

Capítulo 12

Referências Bibliográficas

[ACKERMAN-72]

ACKERMAN, L. V. & GOSE, E.E.; **Breast Lesion Classification By Computer and Xeroradiograph**, *Cancer*, v.30, n.4, p.1025-1035, 1972.

[ACKERMAN-73]

ACKERMAN, L.V.; MUCCIARDI, A.N.; GOSE, E.E. & ALCORN, F.S.; **Classification of benign and malignant breast tumors on the basis of 36 radiographic properties**. *Cancer*, v.31, n.2, p.342-352, 1973.

[ANDERSSON-84]

ANDERSSON, I.; **What can we learn from interval carcinomas?**. *Recent Results Cancer Res.*, v.90, p.161-163, 1984.

[BAINES-86]

BAINES, C.J.; MILLER, A.B.; WALL, C. ; ET AL.; **Sensitivity and specificity of first screen mammography in the Canadian National Breast Screening Study: A preliminary report from five centers**. *Radiology*, v.160, p.295-298, 1986.

[BASSET-87]

BASSET, L.W & GOLD, R.H.; **Breast cancer detection: mammography and other methods in breast imaging**. Grune & Stratton, New York, 1987.

[BEALE-90]

BEALE,R & JACKSON, T.; **Neural computing: an introduction**. Adam-Hilger, Inst. of Physics Publishing Ltd, Bristol, 1990.

[BICK-95]

BICK, U.; GIGER, M.L.; SCHMIDT, A.; NISHIKAWA, R.M.; WOLVERTON, D.E. & DOI, K.; **Automated segmentation of digitized mammograms.** Acad. Radiol., v.2, n.1, p.1-9, 1994

[BORING-92]

BORING,C.C; SQUIRES,T.S & TONG, T; **Cancer Statistics, 1992.** CA 1992, v.42, p.19-38, 1992.

[BRZAKOVIC-90]

BRZAKOVIC, D.; LUO, X.M. & BRZAKOVIC, P.; **An approach to automated detection on tumors in mammograms.** IEEE Trans. on Medical Imaging, v.9, n.3, p.233-241, 1990.

[BUCHANANN-83]

BUCHANANN, J.R.; SPRATT, J.S. & HEUSER, L.S.; **Tumor growth, doubling times, and the inability of the radiologist to diagnose certain cancers.** Radiol. Clin. North Am., v.21, p.115, 1983.

[CHAN-87]

CHAN, H.P; DOI, K.; GALHOTRA, S.;ET AL.; **Image feature analysis and computer-aided diagnosis in digital radiography. I. Automated detection of microcalcifications in mammography.** Medical Physics, v.14, p.538-548, 1987.

[CHAN-88A]

CHAN, H.P.; DOI, K.; VYBORMY, C.J.; SCHMIDT, R.A. & METZ, C.E.; **Digital characterization of clinical mammographic microcalcifications: applications in computer-aided detection.** Proc. SPIE: Med. Imag. II p. Image Formations Detection Proc. and Interpret. , v.914, p.591-593, 1988.

[CHAN-88B]

CHAN, H.P.; DOI, K.; VYBORMY, C.J.; SCHMIDT, R.A. & LAM, K.L; **Computer-aided detection of microcalcifications in mammograms: methodology and preliminary clinical study.** Original Investigations, 1988.

[CHAN-94]

CHAN, H.P.; NIKLADSON, L.T.; IKEDA, D.M.; LAM, K.L. & ADLER, D.D.; **Digitalization requirements in mammography: effects on computer-aided detection of microcalcifications.** Medical Physics, v.21, n.7, p.1203-1211, 1994.

[CHAN-95]

CHAN, H.P.; WEI, D.; HELVIE, M.A.; ET AL.; **Computer-aided classification of mammographic masses and normal tissue: Linear discriminant analysis in texture feature space.**, Phys. Med. Biol, v.40, p.857-876, 1995.

[COOK-89]

COOK, H.M. & FOX, M.D.; **Application of expert systems to mammographic image analysis.** American Journal Phisiology Imaging, v.4, p.16-22, 1989.

[DAVIES-90]

DAVIES, D.H. & DANCE, D.R.; **Automatic computer detection of agrupamentoed calcification in digital mammograms.** Physics in Med. and Biol., v.35, n.8, p.1111-1118, 1990.

[DUBUC-87]

DUBUC, B.; CARMES, C.R.; TRICOT, C. & ZUCKER, S.W.; **The variation method: a technique to estimate the fractal dimension of surfaces.** Visual Commun. Image Processing II, SPIE 845, p.241-248, 1987.

[DUBUC-89]

DUBUC, B.; TRICOT, C. ; ZUCKER, S.W.; QUINIOU, J.F. & WEHBI, D.; **Evaluation the fractal dimension of surfaces.** Proc. R. Soc. London 425, p.113-127, 1989.

[EL-FARAMAWY-96]

EL-FARAMAWY, N.M.; RANGAYYAN, R.M.; DESAUTELS, J.E.L. & ALIM, O.A.; **Shape factors for analysis of breast tumors in mammograms.** IEEE Proc. Canadian Conf. on Electrical and Computer Eng., p.355-358, 1996.

[FACON-93]

FACON, J.; **Processamento e Análise de Imagens.** VI EBAI, 1993.

[FACON-96]

FACON, J.; **Morfologia Matemática: Teoria e Exemplos**, Editora Universitária Champagnat da PUC-PR, Curitiba, 1996.

[FARIA-94]

FARIA, S.L; LEME, L.H.S. & OLIVEIRA FILHO, J.A.; **Câncer da Mama : Diagnóstico e Tratamento**. Editora Médica e Científica Ltda, Rio de Janeiro, 1994.

[FAUSETT-94]

FAUSETT, L.; **Fundamentals of neural networks. Architectures, algorithms, and applications**. Prentice Hall Int., Inc., 1994.

[FERRARI-96]

FERRARI, R.J.; KINOSHITA, S.K.; SLAETS, A.F.F.; MARQUES, P.M.A. & SCHIABEL, H.; **Estudo comparativo entre tipos e tecnologias de scanners para a digitalização de mamogramas**. Anais do III Fórum Nacional de Ciência e Tecnologia em Saúde p.605-606, 1996.

[GALE-87]

GALE, A.G.; ROEBUCK, E.J.; RILEY, P. & WORTHINGTON, B.S. , **Computer aids to mammographic diagnosis**. The British Journal of Radiology, v.60, n.717, p.887-891, 1987.

[GIGER-90A]

GIGER, M.L.; YIN, F.F.; DOI, K.; METZ, C.E., SCHMIDT, R.A. & VYBORNY, C.J.; **Investigation of methods for the computerized detection and analysis of mammographic masses**. Proc. SPIE. v.1233, p.183-184, 1990.

[GIGER-90B]

GIGER, M.L.; YIN, F.F.; DOI, K.; METZ, C.E., SCHMIDT, R.A. & VYBORNY, C.J.; **Image features of mammographic masses used in the development of computerized schemes**. Proc. SCAR90, p.471-473, 1990.

[GIGER-94]

GIGER, M.L.; YIN, SCHMIDT, R.A. & VYBORNY, C.J.; **Computerized characterization of mammographic masses: analysis of speculation**. Cancer Lett., v.77, p.201-211, 1994.

[GONZALEZ-92]

GONZALEZ, R.C. & WOOD, R.E.; **Digital Image Processing.** Addison-Wesley Reading, 1992.

[GUPTA-87]

GUPTA, L. & SRINATH, M.D.; **Contour sequence moments for the classification of closed planar shapes.** Pattern Recognition, v.20, n.3, p.267-272, 1987.

[HAAGENSEN-89]

HAAGENSEN, C.D.; **Doenças da Mama.** 3^a edição, Tradução de Akemi Ishikawa, et al., Libraria Roca Ltda, São Paulo, SP, 1989.

[HAND-79]

HAND, W.; SEMMLOW, J.L.; ACKERMANN, L.V. & ALCORN, F.S.; **Computer screening of xeromammograms: a technique for defining suspicious areas of the breast.** Comput. Biomed. Res., v.12, p.445-460, 1979.

[HARALICK-73]

HARALICK, R.M.; SHANMUGAN, K. & DINSTEIN, I. ; **Textural features of images classification.** IEEE Trans. Systems Man. Cybernetics SMC-3, p.610-621, 1973.

[HARALICK-74]

HARALICK, R.M.; **A measure of circularity of digital figures.** IEEE Trans. Systems Man. Cybernetics SMC-3, p.394-396, 1974.

[HIGHNAM-96]

HIGHNAM, R.P.; BRADY, J.M. & SHEPSTONE, B.J.; **A quantitative feature to aid diagnosis in mammography.** Proc. of the 3rd Intern. Workshop on Digital Mamography, p.201-206, 1996.

[HOLLAND-82]

HOLLAND, R.; MRVUNAC, M.; HENDRIKS, J.H.C.L.; ET AL.; **So-called interval cancers of breast. Pathologic and radiographic analysis.** Cancer, v.49, p.2527-2533, 1982.

[HU-62]

HU, M.K.; **Visual pattern recognition by moment invariants.** IRE Trans. on Inf. Theory IT, v.8, n.2, p.179-187, 1962.

[HUO-95A]

HUO, Z.; GIGER, M.L.; VYBORNY, C.J.; WOLVERTON, D.E.; SCHMIDT, R.A. & DOI, K.; **Feature analysis and classification of mass lesions in digital mammograms.** Medical Physics, v.22, n.6, p.937-938, 1995.

[HUO-95B]

HUO, Z.; GIGER, M.L.; VYBORNY, C.J.; BICK, U.; LU, P.; WOLVERTON, D.E. & SCHMIDT, R.A.; **Analysis of spiculation in the computerized classification of mammographic masses.** Medical Physics, v.22, n.10, p.1569-1579, 1995.

[HUO-96]

HUO, Z.; GIGER, M.L.; VYBORNY, C.J.; BICK, U.; LU, P.; WOLVERTON, D.E. & SCHMIDT, R.A.; **Computer-aided diagnosis: automated classification of mammographic mass lesions.** Proc. of the 3rd Intern. Workshop on Digital Mammography, p.207-211, 1996.

[KARSSEMEIJER-96]

KARSSEMEIJER, N. & TE BRAKE, G.M.; **Detection of stellate distortions in mammograms.** IEEE Trans. on Medical Imaging, v.15, n.5, p.611-619, 1996.

[KEGELMEYER-94]

KEGELMEYER JR, W.P.; PRUNEDA, J.M.; BOURLAND, P.D.; HILLIS, A.; RIGGS, M.W. & NIPPER, M.L.; **Computer-aided mammographic screening for spiculated lesions.** Radiology, v.191, n.2, p.331-337, 1994.

[KILDAY-93]

KILDAY, J. ; PALMIERI, F. & FOX, M.D; **Classifying mammographic lesions using computerized image analysis.** IEEE Trans. on Medical Imag., v.12, n.4, p.664-669, 1993.

[KIMME-75]

KIMME, O.; O'LOUGHLIN, B.J. & SKLANSKY, J.; **Automatic detection of suspicious abnormalities in breast radiographs.** Data structures, computer graphics, and pattern recognition, New York: Academic Press, p.427-447, 1975.

[LAI-89]

LAI, S.M.; LI, X. & BISCHOF, W.F.; **On techniques for detection circumscribed masses in mammograms.** IEEE Trans. on Medical Imaging, v.8, n.4, p.377-386, 1989.

[LEVINE-85]

LEVINE, M.D.; **Vision in man and machine.** McGrawHill, New York, N.Y., 1985.

[LI-93]

LI, H.D.; KALLERGI, M.; CLARKE, L.P. ; Qian, W.; & Clark, R.A.; **Markov Random Field model for mammogram segmentation.** Proc. IEEE EMB, p.54-55, 1993.

[LI-95]

LI, H.D.; KALLERGI, M.; CLARKE, L.P. & JAIN, V.K., **Markov Random Field for tumor detection in digital mammography.** IEEE Trans. on Med. Imag., v.14, n.3, p.565-576, 1995.

[LO-97]

LO, Y.L.; BAKER, J.A.; KORNGUTH, P.J.; IGLEHART, J.D. & FLOYD JR, C.E.; **Predicting breast cancer invasion with artificial neural networks on the basis of mammographic features.** Radiology, v.203, n.1, p.159-163, 1997.

[MARTIN-79]

MARTIN, J.E.; MOSKOWITZ, M.; MILBRATH, J.R.; **Breast cancer by mammography,** AJR, v.132, p.737-739, 1979.

[MASTERS-93]

MASTERS, T.; **Practical neural network. Recipes in C++.** Academic Press, Inc., 1993.

[METZ-86]

METZ, C.E.; **ROC methodology in radiologic imaging.** Invest. Radiol., v.21, p.720-733, 1986.

[MOSKOWITZ-83]

MOSKOWITZ, M.; Screening for breast cancer: How effective are our tests? A critical review. CA Cancer J. Clin., v.33, p.25-39, 1983.

[NISHIKAWA-94]

NISHIKAWA, R.M.; GIGER, M.L.; DOI, K.; METZ, C.E.; VYBORNY, C.J.; SCHMIDT, R.A.; Effect of case selection on the performance of computer-aided detection schemes. Medical Physics, v.21, p.265-269, 1994.

[PARKER-94]

PARKER, J.R; Practical computer vision using C. John Wiley & Sons, Inc., 1994.

[PETRICK-96A]

PETRICK, N.; CHAN, H.P.; SAHINER, B. & WEI, D.; An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection. IEEE Trans. on Medical Imaging, v.15, n.1, p.59-67, 1996.

[PETRICK-96B]

PETRICK, N.; CHAN, H.P.; SAHINER, B.; WEI, D.; HELVIE, M.A. & ADLER, D.D.; Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification. Medical Physics, v.23, n.10, p.1685-1696, 1996.

[PETROSIAN-94]

PETROSIAN, A.; CHAN, H.P.; PETRICK, N.; WEI, D.; HELVIE, M.A.; ADLER, D.D. & GOODSITT, M.M.; Computer-aided diagnosis in mammography: classification of mass and normal tissue by texture analysis. Phys. Med. Biol., v.39, p.2273-2288, 1994.

[PINOTTI-91]

PINOTTI, J.A.; ET AL.; Compêndio de Mastologia. Editora Manole Ltda, São Paulo, 1991.

[POHLMAN-96]

POHLMAN, S; POWELL, K.A.; OBUCHOWSKI, N.A.; CHILCOTE, W.A. & BRONIATOWSKI, S.G.; Quantitative classification of breast tumors in digitized mammograms. Medical Physics, v.23, n.8, p.1337-1345, 1996.

[RANGAYYAN-96]

RANGAYYAN, R.M.; EL-FARAMAWY, N.M.; DESAUTELS, J.E.L. & ALIM, O.A.; **Discrimination between benign and malignant breast tumors using a region-based measure of edge profile acutance.** Proc. of the 3rd Intern. Workshop on Digital Mamography, p.213- 218, 1996.

[RUSS-94]

RUSS, J.C.; **The image processing handbook.** 2nd edition. CRC Press Inc., USA, 1994.

[SAHINER-96A]

SAHINER, B.; CHAN, H.P.; PETRICK, N.; WEI, D.; HELVIE, M.A.; ADLER, D.D. & GOODSITT, M.M.; **Classification of mass and normal breast tissue: a convolution neural networks classifier with spatial domain and texture images.** IEEE Trans. on Medical Imaging, v.15, n.5, p.598-610, 1996.

[SAHINER-96B]

SAHINER, B.; CHAN, H.P.; PETRICK, N.; WEI, D.; HELVIE, M.A.; ADLER, D.D. & GOODSITT, M.M.; **Image feature selection by a genetic algorithm: Application of mass and normal breast tissue.** Medical Physics, v.23, n.10, p.1671-1684, 1996.

[SAHINER-98]

SAHINER, B.; CHAN, H.P.; PETRICK, N.; HELVIE, M.A. & GOODSITT, M.M.; **Computerized characterization of masses on mammograms: the rubber band straightening transform and texture analysis.** Medical Physics, v.25, n.4, p.516-526, 1998.

[SERRA-82]

SERRA, J.; **Image analysis and mathematical morphology.** Academic Press, London, 1982.

[SHEN-93]

SHEN, L.; RANGAYYAN, R.M. & DESAUTELS, J.E.L.; **Detection and classification of mammographic calcifications.** International Journal of Pattern Recognition and Artificial Intelligence, v.7, n.6, p.1403-1416, 1993.

[SHEN-94]

SHEN, L.; RANGAYYAN, R.M. & DESAUTELS, J.E.L; **Application of shape analysis to mammographic calcification.** IEEE Trans. Medical Imaging, v.13, n.2, p.263-274, 1994.

[SUCKLING-94]

SUCKLING, J.; PARKER, J.; DANCE, D.R.; ASTLEY, S.; HUTT, I.; BOGGIS, C.R.M.; RICKETSS, I.; STAMATAKIS, E.; CERNEAZ, N.; KOK, S.L., TAYLOR, P.; BETAL, D. & SAVAGE, J.; **The mammographic image analysis society digital mammogram database.** Proc. of the 2nd Intern. Workshop on Digital Mammography, p.375-378, 1994.

[VELANOVICH-95]

VELANOVICH, V.; **Fractal analysis of mammographic lesion: a feasibility study quantifying the difference between benign and malignant masses.** The American Journal of The Medical Sciences, v.311, n.5, p.211-214, 1995.

[VILLELA-96]

VILLELA, R.L.; SLAETS, A.F.F; MARQUES, P.M.A.; SCHIABEL, H.; FLORIAN, R.V. & FERRARI, R.J.; **Estudo comparativo das técnicas de segmentação de microcalcificações em imagens mamográficas.** Anais do III Fórum Nacional de Ciências e Tecnologia em Saúde, v.2, p.575-576, 1996.

[WEE-75]

WEE, W.G., MOSKOWITZ, M.; CHANG, G.C; TING, Y.C. & PEMMERAJU, S.; **Evaluation of mammographic calcifications using a computer program.** Radiology, v.110, p. 717-720, 1975.

[WEI-95]

WEI, D.; H.P. CHAN; HELVIE, M.A.; SAHINER, B.; PETRICK, N., ADLER, D.D. & GOODSON, M.M.; **Classification of mass and normal tissue on digital mammograms: Multiresolution texture analysis.** Medical Physics, v.22, n.9, p.1501-1513, 1995.

[WEI-97]

WEI, D.; H.P. CHAN; HELVIE, M.A.; SAHINER, B.; PETRICK, N., ADLER, D.D. & GOODSON, M.M.; **False-positive reduction technique for detection of masses on digital mammograms: global and local multiresolution texture analysis.** Medical Physics, v.24, n.6, p.903-914, 1997.

[WINSBERG-67]

WINSBERG, F.; ELKIN, M.; MACY, J.; BORDAZ, V. & WEYMOUTH, W.; **Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis.** Radiology, v.89, p.211-215, 1967.

[WU-93A]

WU, Y.; GIGER, M.L.; DOI, K.; VYBORNY, C.J.; SCHMIDT, R.A. & METZ, C.E.; **Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer.** Radiology, v.187, n.1, p.81-87, 1993.

**[WU-93B]**

WU, Y.; GIGER, M.L.; DOI, K.; VYBORNY, C.J.; SCHMIDT, R.A. & METZ, C.E.; **Simulation studies of data classification by artificial neural networks: potential applications in medical imaging and decision making.** Journal of Digital Imaging, v.6, n.2, p.117-125, 1993.

[YIN-91A]

YIN, F.F.; GIGER, M.L.; DOI, K.; METZ, C.E.; VYBORNY, C.J. & SCHMIDT, R.A.; **Computerized detection of masses in digital mammograms: analysis of bilateral subtraction images.** Medical Physics, v.18, n.5, p.955-963, 1991.

[YIN-91B]

YIN, F.F.; GIGER, M.L.; DOI, K.; METZ, C.E.; VYBORNY, C.J.; SCHMIDT, R.A. & METZ, C.E.; **Computer-vision system for the detection and characterization of masses for use in mammographic screening programs.** Proc. SPIE, v.1396, p.2-4, 1991.

[YIN-93]

YIN, F.F.; GIGER, M.L.; DOI, K.; METZ, C.E.; VYBORNY, C.J. & SCHMIDT, R.A.; **Comparison of bilateral-subtraction and single-Image processing techniques in the computerized detection of mammographic masses.** Invest. Radiology, v.28, n.6, p.473-481, 1993.

[YIN-94A]

YIN, F.F.; GIGER, M.L.; DOI, K.; METZ, C.E.; VYBORNY, C.J. & SCHMIDT, R.A.; **Computerized detection of masses in digital mammograms: investigation of feature-analysis techniques.** J. of Digital Imaging, v.7, n.1, p.18-26, 1994.

[YIN-94B]

YIN, F.F.; GIGER, M.L.; DOI, K.; METZ, C.E.; VYBORNY, C.J. & SCHMIDT, R.A.; **Computerized detection of masses in digital mammograms: automated alignment of breast images and its effect on bilateral-subtraction technique.** Medical Physics, v.21, n.3. p.445-452, 1994.

[YOUNG-86]

YOUNG, T.Y. & FU, K.S.; **Handbook of pattern recognition and image processing.** Academic Press, USA, 1986.