

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA BIOMÉDICA

ISABELLA MARIA MOURA FERNANDES

**SISTEMA DE APOIO À CLASSIFICAÇÃO DE LESÕES
EM MAMOGRAFIAS CONSIDERANDO A DENSIDADE
MAMÁRIA**

Recife

2015

ISABELLA MARIA MOURA FERNANDES

**SISTEMA DE APOIO À CLASSIFICAÇÃO DE LESÕES
EM MAMOGRAFIAS CONSIDERANDO A DENSIDADE
MAMÁRIA**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Biomédica, da Universidade Federal de Pernambuco, como parte dos requisitos para a obtenção do título de Mestre em Engenharia Biomédica.

Orientador: Wellington Pinheiro dos Santos

Recife

2015

Catálogo na fonte
Bibliotecária Margareth Malta, CRB-4 / 1198

F363s Fernandes, Isabella Maria Moura.
 Sistema de apoio à classificação de lesões em mamografias
 considerando a densidade mamária / Isabella Maria Moura Fernandes. -
 Recife: O Autor, 2015.
 89 folhas, il., gráfs., tabs.

 Orientador: Prof. Dr. Wellington Pinheiro dos Santos.
 Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG.
 Programa de Pós-Graduação em Engenharia Biomédica, 2015.
 Inclui Referências e Apêndices.

 1. Engenharia Biomédica. 2. Mamografia. 3. Sistema CAD. 4.
 Classificação. 5. Redes Neurais Artificiais. I. Santos, Wellington Pinheiro
 dos. (Orientador). II. Título.

UFPE

610.28 CDD (22. ed.)

BCTG/2015-127

ATA DA DEFESA DE DISSERTAÇÃO DE MESTRADO, REALIZADA NO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA BIOMÉDICA DO CENTRO DE TECNOLOGIA E GEOCIÊNCIAS DA UNIVERSIDADE FEDERAL DE PERNAMBUCO, NO DIA 25 DE FEVEREIRO 2015.

Aos 25 (vinte e cinco) dias do mês de fevereiro de dois mil e quinze (2015), às 14 horas, no Laboratório da Computação Biomédica do Programa de Pós-graduação em Engenharia Biomédica da Universidade Federal de Pernambuco, reunida a Banca Examinadora, em sessão pública, teve início a defesa da Dissertação intitulada “*SISTEMA DE APOIO À CLASSIFICAÇÃO DE LESÕES EM MAMOGRAFIAS CONSIDERANDO A DENSIDADE MAMÁRIA*” da aluna **ISABELLA MARIA MOURA FERNANDES**, que havia cumprido anteriormente todos os demais requisitos regimentais para a obtenção do grau de Mestre em Engenharia Biomédica. A Banca Examinadora foi composta pelo Prof. Dr. Ricardo Yara, do departamento de Engenharia Biomédica da Universidade Federal de Pernambuco – UFPE, Prof. Dr. Adriano Lorena Inácio de Oliveira, do Departamento de Ciência da Computação da UFPE e o Prof. Dr. Wellington Pinheiro dos Santos, do Departamento de Engenharia Biomédica – UFPE, na qualidade de presidente. Depois de cumpridas as formalidades, o candidato foi convidado a discorrer sobre o conteúdo da Dissertação. Concluída a explanação, o candidato foi arguido pela Banca Examinadora que, em seguida, reuniu-se para deliberar e conceder ao mesmo a menção **APROVADO** da referida Dissertação. E, para constar, lavrei a presente Ata que vai por mim assinada, Secretaria de Pós-graduação, e pelos membros da Banca Examinadora e demais presentes.

Recife, 25 de fevereiro de 2015.

BANCA EXAMINADORA

Prof. Dr. Wellington Pinheiro dos Santos

Prof. Dr. Ricardo Yara

Prof. Dr. Adriano Lorena Inácio de Oliveira

RESUMO

O câncer de mama é uma preocupação mundial, pois é o segundo mais frequente no mundo e o mais comum entre as mulheres. Até o momento, o único método para rastreamento de câncer de mama provado ser efetivo é a mamografia, sendo o exame padrão ouro para tal finalidade. No entanto, este exame é de difícil leitura e interpretação, principalmente pelo baixo contraste da imagem gerada, inclusive devido à semelhança da densidade radiológica entre o tecido mamário normal e o patológico, principalmente para mamas densas. Estudamos neste trabalho diversos sistemas de apoio ao diagnóstico médico, estruturados a partir da classificação de imagens de mamografia através da descrição da região de interesse da imagem em relação às suas características de forma e de textura e da utilização de classificadores baseados em redes neurais artificiais de treinamento supervisionado. Os sistemas CAD estudados foram aplicados à base de dados IRMA e verificamos que ao organizarmos e separarmos previamente a base de dados utilizada em relação à característica tecidual da mama e sua densidade obtemos melhores resultados do que quando a consideramos de forma completa. Verificamos também que maior quantidade de características utilizadas para descrever a imagem possibilita melhores resultados no desempenho das redes neurais utilizadas para classificação. Obtivemos uma taxa de classificação correta de 94% para mamas pouco densas e de 89% para mamas extremamente densas, sendo esta última de maior dificuldade de classificação devido intrinsecamente às características do exame de mamografia e, mais especificamente, ao baixo contraste entre tecido patológico e tecido normal.

Palavras-chaves: Mamografia; Sistema CAD; Classificação; Redes Neurais Artificiais.

ABSTRACT

Breast cancer is a global concern and it is the second most common in the world and the most common among women. So far the only breast cancer screening method that has proved to be effective is mammography screening, being the gold standard for this purpose. However, this exam is difficult to read and to interpret, especially because the low contrast image generated due to the radiographic density similarity between normal breast tissue and the pathological tissue, especially for dense breasts. In this work we study different computer-aided diagnosis systems (CAD) applied to medical support, based on mammography image classification through the description of texture and shape feature of the image's region of interest. We used classifiers based on artificial neural network with supervised training. The CAD systems studied were applied to the IRMA database. We verified that the organization of the database according to the breast density and tissue characteristics we obtained better lesion classification in mammograms than when we use the entire IRMA database. Also we were able to demonstrate that as must characteristics used to describe the image the better is the performance of the artificial neural network in the classification task. As a result we obtained a percentage of instances correctly classified over 94% to slightly dense breasts and 89% to extreme dense breasts. This last is more difficult to be classified because the characteristic of the mammography exam that uses x ray radiation, which reduces the contrast between pathological and normal breast tissue.

Keywords: Mammography; CAD System; Classification; Artificial Neural Network.

LISTA DE ILUSTRAÇÕES

FIGURA 1 - PROCESSO DE ANGIOGÊNESE.	18
FIGURA 2 - COMPONENTES DO MAMÓGRAFO.	21
FIGURA 3 - A) COMPRESSÃO DA MAMA. B) MODIFICAÇÕES DA CONFORMAÇÃO DA MAMA DURANTE COMPRESSÃO.	21
FIGURA 4 - EXEMPLOS DE MAMOGRAFIAS DE MAMA Densa.	23
FIGURA 5 - MAMA PREDOMINANTEMENTE Densa.	23
FIGURA 6 - MAMA PREDOMINANTEMENTE ADIPOSA.	23
FIGURA 7 - MAMA ADIPOSA.	24
FIGURA 8 - ASPECTOS MORFOLÓGICOS DE MASSAS MAMÁRIAS.	25
FIGURA 9 - AMOSTRA DE IMAGENS DAS 12 CLASSES DA BASE IRMA. COLUNAS, DA ESQUERDA PARA DIREITA, TEMOS OS TIPOS DE MAMA ADIPOSA, FIBROGLANDULAR, Densa E EXTREMAMENTE Densa, RESPECTIVAMENTE. LINHAS DE CIMA PARA BAIXO REPRESENTAM O ESTÁGIO DA LESÃO INICIANDO POR REPRESENTAÇÃO DE IMAGEM SEM TUMOR, COM TUMOR BENIGNO E COM TUMOR MALIGNO, RESPECTIVAMENTE.	27
FIGURA 10 - REPRESENTAÇÃO DE UMA IMAGEM AMPLIADA DE 10x10 COM 256 TONS DE CINZA.	30
FIGURA 11 - ESQUEMA DE UM SISTEMA CAD.	33
FIGURA 12 - MATRIZ DE CO-OCORRÊNCIA DE NÍVEIS DE CINZA.	38
FIGURA 13 - ARQUITETURA DE UMA REDE RBF.	40
FIGURA 14 - ARQUITETURA DE REDES MLP.	41
FIGURA 15 - PADRÕES LINEARMENTE SEPARÁVEIS.	42
FIGURA 16 - DADOS NÃO LINEARMENTE SEPARÁVEIS.	42
FIGURA 17 - MARGEM DO HIPERPLANO QUE SEPARA DOIS PADRÕES. AMOSTRAS LOCALIZADAS NA MARGEM SÃO CHAMADAS DE VETORES DE SUPORTE.	43
FIGURA 18 - REPRESENTAÇÃO DE CURVAS ROC.	46
FIGURA 19 - TELA INICIAL DO APLICATIVO CONVIRMA.	49
FIGURA 20 - SELEÇÃO DE INFORMAÇÃO PARA GERAÇÃO DA BASE DE DADOS.	50
FIGURA 21 - SELEÇÃO DAS CONFIGURAÇÕES DE PROCESSAMENTO DA BASE DE DADOS.	50
FIGURA 22 - TELA INICIAL DO WEKA 3.6.11.	51
FIGURA 23 - ABA DE CLASSIFICAÇÃO DO WEKA.	52
FIGURA 24 - MELHORES CONFIGURAÇÕES DO SISTEMA CAD PARA CLASSIFICAÇÃO DE ATRIBUTOS DE TEXTURA.	57
FIGURA 25 - PERCENTUAL DE CLASSIFICAÇÕES CORRETAS DE LESÕES UTILIZANDO DESCRITORES DE FORMA.	58
FIGURA 26 - PERCENTUAL DE CLASSIFICAÇÕES CORRETAS DE LESÕES UTILIZANDO DESCRITORES DE FORMA E DE TEXTURA.	59
FIGURA 27 - CURVA ROC PARA MAMAS ADIPOSAS. CURVA 1: MAMAS SEM LESÃO. CURVA 2: LESÕES BENIGNAS. CURVA 3: LESÕES MALIGNAS.	63

FIGURA 28 - CURVA ROC PARA MAMAS FIBROGLANDULARES. CURVA 1: MAMAS SEM LESÃO. CURVA 2: LESÕES BENIGNAS. CURVA 3: LESÕES MALIGNAS.....	64
FIGURA 29 - CURVA ROC PARA MAMAS POUCO DENSAS. CURVA 1: MAMAS SEM LESÃO. CURVA 2: LESÕES BENIGNAS. CURVA 3: LESÕES MALIGNAS.....	65
FIGURA 30 - CURVA ROC PARA MAMAS MUITO DENSAS. CURVA 1: MAMAS SEM LESÃO. CURVA 2: LESÕES BENIGNAS. CURVA 3: LESÕES MALIGNAS.....	66

LISTA DE TABELAS

TABELA 1: DIVISÃO DA BASE DE DADOS IRMA UTILIZADA NO TRABALHO ESTUDADO CONFORME TIPO DE TECIDO MAMÁRIO E LESÃO..	28
TABELA 2: MATRIZ CONFUSÃO GENÉRICA.....	45
TABELA 3: QUANTIDADE DE INSTÂNCIAS POR ABORDAGEM.	55
TABELA 4: CONFIGURAÇÃO DOS MELHORES CLASSIFICADORES.	61
TABELA 5: MATRIZ CONFUSÃO E ÁREA SOB A CURVA ROC PARA CLASSIFICAÇÃO I-WA-2-HA-ZE-MLP-A.	62
TABELA 6: MATRIZ CONFUSÃO E ÁREA SOB CURVA ROC PARA CLASSIFICAÇÃO II-SOM-HA-ZE-MLP-A.	63
TABELA 7: MATRIZ CONFUSÃO E ÁREA SOB CURVA ROC PARA CLASSIFICAÇÃO III-WA-2-HA-ZE-MLP-I.....	64
TABELA 8: MATRIZ CONFUSÃO E ÁREA SOB CURVA ROC PARA CLASSIFICAÇÃO IV-WA-2-HA-ZE-MLP-A.	65

LISTA DE ABREVIATURAS E SIGLAS

ACR	American College of Radiology
ALARA	As Low As Reasonably Achievable
ANN	Artificial Neural Network
ARFF	Attribute-Relation File Format
BI-RADS	Breast Imaging Reporting and Data System
CDC	Centers for Disease Control and Prevention
CSV	Comma-separated values
DDSM	Digital Database for Screening Mammography
FDA	Food and Drug Administration
GLCM	Gray Level Co-Occurrence Matrix
IRMA	Image Retrieval in Medical Applications
LLNL	Lawrence Livermore National Laboratory
MIAS	Mammographic Image Analysis Society
MLP	MultiLayer Perceptron
NCI	National Cancer Institute
RBF	Radial Basis Function
ROC	Receiver Operating Characteristic
ROI	Regio of Interest
RWTH	Rheinisch-Westf`alische Technische Hochschule
SMO	Sequential Minimal Optimization
SOM	Self-Organizing Maps
SVM	Support Vector Machine
WEKA	Waikato Environment for Knowledge Analysis

SUMÁRIO

1 Introdução.....	11
1.1 Justificativa e Motivação.....	11
1.2 Objetivos	15
1.3 Organização da Dissertação	16
2 Análise de mamografias	18
2.1 Câncer de mama.....	18
2.2 Mamografia.....	20
<i>2.2.1 Tipos de lesões e classificação BI-RADS</i>	<i>24</i>
2.3 Base de dados	26
2.4 Sistemas CAD.....	28
<i>2.4.1 Análise digital de imagens</i>	<i>30</i>
<i>2.4.2 Segmentação de imagens.....</i>	<i>34</i>
<i>2.4.3 Extração de características</i>	<i>36</i>
<i>2.4.4 Classificação dos atributos.....</i>	<i>38</i>
2.5 Métricas de desempenho em classificação.....	43
3 Estudo proposto	48
3.1 Softwares utilizados	49
<i>3.1.1 CONVIRMA</i>	<i>49</i>
<i>3.1.2 WEKA.....</i>	<i>50</i>
3.2 Experimentos.....	53
<i>3.2.1 Experimento 1: Análise da textura.....</i>	<i>56</i>
<i>3.2.2 Experimento 2: Análise da forma</i>	<i>57</i>
<i>3.2.3 Experimento 3: Análise da forma e da textura.....</i>	<i>58</i>
3.3 Resultados	60
4 Conclusões e Discussões	67
4.1 Trabalhos futuros.....	68
Referências	70
APÊNDICE A – RESULTADO DAS SIMULAÇÕES DOS SISTEMAS CAD ESTUDADOS	74
APÊNDICE B – ARTIGO PUBLICADO NO XXIV CONGRESSO BRASILEIRO DE ENGENHARIA BIOMÉDICA	84

Capítulo 1

Introdução

Existem diversas técnicas de detecção do câncer de mama, tais como: técnicas óticas, de ultrassom, de termografia, raios X, impedância elétrica, entre outros. A mamografia, no entanto, é utilizada como ferramenta de triagem para detecção precoce do câncer de mama, sendo considerada como a técnica padrão ouro para tal, principalmente por ser capaz de revelar alterações na mama até dois anos antes de um paciente ou o médico poder senti-las.

A mamografia é um exame de difícil interpretação, pois não possui um bom contraste, devido principalmente à semelhança da densidade radiológica entre o tecido mamário normal e o patológico. Além disso, a predominância do tipo de tecido constituinte da mama altera sua densidade e, conseqüentemente, o contraste da imagem. Mulheres jovens tendem a possuir mamas mais densas, em que há pouca substituição do parênquima mamário por tecido adiposo. A substituição do tecido ocorre com a idade, de forma que mulheres mais velhas possuem mamas com predominância de tecido adiposo. A mamografia é capaz de detectar grande parte das lesões mamárias, mesmo em casos assintomáticos, possuindo sensibilidade mais baixa para mulheres mais jovens ou com mamas densas.

A mamografia possui sensibilidade variável entre 88% e 93,1%, e especificidade entre 85% e 94,2% (INCA,2007). Visto o alto valor de falsos positivos em casos em que o exame é avaliado por um único especialista e ao alto custo e escassez de profissionais capacitados para implementação de um sistema de dupla leitura para o diagnóstico do câncer de mama através da mamografia, acredita-se que a utilização de sistemas computacionais de apoio médico possa reduzir a quantidade de diagnósticos falsos positivos e, conseqüentemente, reduzir o número de biópsias desnecessárias em pacientes sadias.

1.1 Justificativa e Motivação

O câncer de mama é uma preocupação mundial. De acordo com o Instituto Nacional do Câncer (INCA), o câncer de mama é o segundo mais frequente no mundo e o mais comum entre as mulheres. O diagnóstico precoce de tal enfermidade, que atualmente vem sendo realizado pelo exame de mamografia, revela um prognóstico relativamente bom e pode aumentar as chances de cura.

A cada ano há, em média, 22% de novos casos, contabilizando-se para o Brasil 52.680 novos casos em 2012, tendo sido registrado uma taxa de mortalidade de 12.852 óbitos em 2010, sendo esperados em 2014, 57.120 casos novos de câncer de mama no Brasil. A mamografia e o exame clínico das mamas (ECM) são os métodos preconizados para o rastreamento na rotina da atenção integral à saúde da mulher (INCA, 2014).

Mamografias são imagens médicas de difícil leitura, devido principalmente ao baixo contraste e diferenças no tipo de tecido. As diferentes formas que podem assumir as lesões mamárias e as características das mesmas, que muitas vezes se assemelha ao tecido normal, torna a decisão da avaliação e classificação dessas imagens ainda mais árdua. Além da dificuldade intrínseca à técnica de imagem utilizada para o diagnóstico do câncer de mama e a característica dessas lesões, a subjetividade entre radiologistas na interpretação de mamografias resulta em maiores porcentagens de erros de diagnósticos (AZAR e EL-SAID, 2013).

Diagnósticos falso-positivos podem levar a biópsias desnecessárias. Dados estatísticos mostram que apenas 20 a 30% das biópsias de mamas são provadas serem de fato casos de câncer (AZAR e EL-SAID, 2013). A fim de evitar erros de diagnóstico, é possível a implementação de um sistema de dupla leitura, em que é necessário que o exame seja analisado por dois especialistas antes do diagnóstico final. Apesar das vantagens na utilização da dupla leitura para o diagnóstico do câncer de mama através da mamografia, essa técnica não é economicamente viável devido à sobrecarga aplicada aos especialistas e ao custo de tal sistema (RAPELYEA, 2015).

A interpretação de exames de mamografia por dois especialistas ao invés da análise por um único especialista aumenta a detecção de câncer de 4 a 14% e é uma prática comum em alguns países europeus (GILBERT et al, 2008). O custo desse benefício está geralmente associado ao aumento na taxa de *recall* e maior quantidade de biópsias negativas. Além disso, a dupla leitura também é dispendiosa no que diz respeito a quantidade de mão de obra especializada necessário para sua implementação. Devido ao custo e por ser uma tarefa dispendiosa, a dupla leitura não é empregada em alguns países, tendo crescido o interesse na utilização dos sistemas CAD para aumento da sensibilidade do diagnóstico, reduzindo descuidos ao mesmo tempo em que o radiologista continua responsável pela análise da lesão e interpretação e diagnóstico final (GROMET, 2008).

O uso de sistemas de apoio ao diagnóstico auxiliado por computador (CAD) tem o potencial de aumentar a taxa de detecção de câncer. Desta forma, acredita-se que um sistema de

detecção utilizando a análise de um especialista juntamente com um sistema CAD possa alcançar performances semelhantes ao da dupla leitura (GILBERT et al, 2008).

Segundo estudos, sistemas CAD melhoram o desempenho do diagnóstico de um único especialista, aumentando a sensibilidade com um pequeno aumento da taxa de *recall*. Conforme estudo de Gromet et al. (2008), a utilização do método de dupla leitura aumentou a sensibilidade em 6,6% (de 81,4% a 88%), com aumento de 1,7% na taxa de *recall* (de 10,2% para 11,9%), de forma que um sistema utilizando única leitura em conjunto com sistema CAD revelou um aumento na sensibilidade para 90,4%, com apenas 0,4% de aumento na taxa de *recall* (de 10,2% para 10,6%), permitindo concluir que tanto o sistema de dupla leitura como a utilização do sistema CAD podem aumentar a sensibilidade do diagnóstico de câncer de mama através da mamografia (GROMET,2008).

A utilização de técnicas de inteligência artificial, como redes neurais artificiais, mostra grande potencial em aplicações de leitura automática de mamografias. A associação dessas técnicas com reconhecimento de padrões, classificação e algoritmos de detecção pode ser aplicada no desenvolvimento de sistemas CAD (AZAR e EL-SAID, 2013). O estudo realizado por Azar e El-Said propôs a classificação de lesões em duas classes, benigna e maligna, de forma que foram realizados testes com redes neurais artificiais perceptron de múltiplas camadas (MLP) com uma camada oculta contendo 7 neurônios e 9 neurônios na camada de entrada e redes de função de base radial (RBF), com 23 neurônios na camada oculta. A base de dados utilizada nesse estudo foi disponibilizada pelo Hospital da Universidade de Wisconsin e consiste de 9 variáveis de entrada, como uniformidade do tamanho da célula, uniformidade do formato da célula e tamanho do tecido epitelial, 2 variáveis de saída (benigno ou maligno) e 699 casos, dos quais 458 foram diagnosticados como benigno e os demais 241, diagnosticados como maligno. No total, foram retirados do estudo 16 casos por falta de informações, de forma que foram analisados 683 casos, dos quais 65% representativos de lesões benignas e 35% de lesões malignas. Os resultados mostraram acurácia de 97,80% e área sob a curva ROC de 0.9956 utilizando redes MLP e acurácia de 98,10% e área sob a curva ROC de 0.9988 com redes RBF.

No trabalho de Eddaoudi et al. (2011) foram utilizados atributos de textura extraídos do vetor de Haralick calculados a partir da matriz de co-ocorrência de níveis de cinza para alimentação de um classificador baseado em máquinas de vetores de suporte. Utilizando uma janela com dimensão 9x9, orientação de 90° e deslocamento 4, aplicado à base de dados DDSM, os autores obtiveram uma média de classificação correta de 77%, obtendo uma me-

hora para 95% de classificações corretas ao utilizar métodos de pré segmentação por limiar máximo.

A segmentação de imagens baseia-se no princípio de que os pixels pertencentes ao tecido patológico possuem características diferentes dos pixels de regiões com tecido normal. A segmentação e a extração de características da imagem podem também ser realizadas através da transformada de wavelets. Utilizando a segunda decomposição da transformada de wavelet de Daubechies-4, com 4 coeficientes (aproximação, vertical, horizontal e diagonal), Kilic et al. (2010) alimentaram um classificador baseado em redes MLP com algoritmo de aprendizado por retropropagação e 20 neurônios na camada oculta e obtiveram uma classificação de 45 imagens de mamografia com sensibilidade da classificação de 73.3 a 89.2%.

Islam et al. (2010) propõem a classificação de imagens de mamografia da base de dados MiniMIAS utilizando atributos estatísticos e de textura e redes neurais artificiais. Os autores propuseram um classificador utilizando redes MLP com uma camada oculta contendo 5 neurônios. Foi verificado um resultado de classificações corretas de lesões benignas de 83,87% e de 90,81% para lesões malignas, utilizando um conjunto de treino composto por 25% das imagens e um conjunto de teste com 75% das imagens utilizadas.

Uma importante característica a ser analisada numa imagem de mamografia é densidade do parênquima mamário no que diz respeito à quantidade de tecido fibroglandular presente. Há uma relação entre os níveis de densidade do parênquima mamário e o risco de câncer de mama, de forma que mamografias com maior densidade mamária devem ser examinadas mais cuidadosamente (TZIKOPOULOSA, S.D. et al., 2011).

Tahmasbi et al. (2011) utilizaram a base de dados MIAS em seu estudo, que contém 322 mamografias digitais com resolução de 200nm pixels armazenadas em imagens de 1024 x 1024 pixels. Essas imagens foram categorizadas em quatro diferentes classes referentes à constituição tecidual da mama e em cinco classes do ponto de vista do tipo de lesão. Neste estudo, cada região de interesse foi segmentada por dois especialistas e, em seguida, realizada a equalização do histograma para aumentar o contraste. Foram utilizados os momentos de Zernike para extração de 32 atributos de forma das lesões segmentadas. Para classificação, foram utilizadas redes MLP, com 32 neurônios de entrada, função de ativação sigmoidal com experimento utilizando validação cruzada com 10 *folds*. Esse estudo resultou numa taxa de acertos de 6,43%, com área sob a curva ROC de 0,976.

No trabalho de Tzikopoulou et al. (2011), testado com as 322 imagens disponibilizadas na base de dados miniMIAS, foram utilizadas a extração de bordas seguido de uma etapa

para estimação e detecção do músculo pectoral conjuntamente com a identificação do mamilo para ser usado como ponto de referência. Além dessas etapas de segmentação, foi aplicado também um algoritmo para estimar a densidade do parênquima mamário, usando segmentação de primeira ordem, classificando-o em 3 classes. Para a classificação dos dados extraídos da imagem foi utilizado máquinas de vetores de suporte (SVM), alcançando acurácia de 85,7%.

Objetivando a redução da carga de trabalho depositada nos especialistas para a análise de mamografias, a redução dos diagnósticos falso-positivos e o aumento da sensibilidade e acurácia no diagnóstico do câncer de mama, tem-se optado pela utilização de sistemas com dupla leitura. Uma alternativa menos dispendiosa é a utilização de sistemas CAD, que são geralmente utilizados como segunda análise, de forma que a interpretação final continua sendo responsabilidade do médico especialista.

1.2 Objetivos

Este trabalho visa testar diversas configurações de sistemas computacionais de apoio ao diagnóstico médico para classificação de câncer de mama em imagens de mamografia. Acredita-se que um melhor desempenho de classificação pode ser obtido quando se considera informações não apenas diretamente ligadas à lesão a ser classificada, como tamanho, localização, formato e textura, mas também características do tecido constituinte da mama, que está indiretamente relacionada a fatores clínicos como idade da paciente. Dessa forma, previamente à utilização de ferramentas computacionais, é realizada uma organização das imagens conforme as características teciduais da mama e sua densidade.

Neste trabalho, objetiva-se o teste e análise dos resultados de diversas configurações de sistema de detecção e classificação de imagens auxiliadas por computador para apoio ao diagnóstico médico (Computer-aided diagnosis, CAD) utilizando classificadores baseados em redes neurais artificiais com treinamento supervisionado, que são modelos computacionais capazes de processar uma grande quantidade de informação simultaneamente através do aprendizado de casos anteriores. Anteriormente à etapa de classificação, aplicamos técnicas de segmentação e extração de características da imagem. Visa-se a verificação de qual configuração do sistema proporciona uma melhor taxa de acertos, considerando, além dos parâmetros do sistema CAD, as características do tecido constituinte da mama.

Foram testados dois métodos para segmentação de imagem, mapas auto organizáveis (SOM) e transformada de wavelets, e três métodos para extração de características da imagem (descrição das características de forma através dos momentos de Zernike, textura através do

descriptor de Haralick, baseado na matriz de co-ocorrência de níveis de cinza e uma abordagem que considerou ambas as informações). Em seguida, o vetor de característica gerado para descrição da imagem foi usado como entrada redes neurais artificiais utilizadas como classificadores. Foram testadas três tipos de redes neurais, com diferentes configurações. O sistema foi aplicado à base de dados IRMA, organizada previamente conforme as características da mama. Dessa forma, o sistema foi aplicado em quatro conjuntos de imagens conforme a densidade da mama, extremamente densa, predominantemente densa, fibroglandular e adiposa.

Este trabalho tem por objetivo principal verificar qual a melhor configuração dentre as testadas para os sistemas CAD na classificação de mamografias sem lesão e com lesão do tipo benigna e maligna para cada um dos quatro tipos de mamas definidos. Através da análise dos resultados, visamos igualmente verificar a viabilidade da utilização de um sistema computacional para auxílio ao diagnóstico do câncer de mama, de forma que possa ser usado como uma segunda opinião para o médico especialista e possibilite a redução de diagnósticos falso-positivos e um aumento na precisão de classificação.

1.3 Organização da Dissertação

O capítulo inicial apresenta a caracterização do problema, a motivação e os objetivos do trabalho. No segundo capítulo, será realizada a fundamentação teórica de conceitos importantes sobre as características teciduais da mama e sobre o câncer de mama, sua formação e a progressão de células cancerígenas, prevenção, e o diagnóstico através do exame de mamografia, ressaltando as vantagens e desvantagens e detalhes sobre análise desse tipo de imagem. Além disso, serão apresentados conceitos sobre processamento digital de imagens, particularizando sua aplicação em mamografias. Neste capítulo também será realizada a descrição da base de dados e das técnicas para segmentação, descrição da imagem por meio de suas características de forma e textura, e de redes neurais artificiais utilizadas como classificadores.

Posteriormente, no terceiro capítulo, serão descritas com maiores detalhes as características do sistema de apoio ao diagnóstico estudado, além dos resultados obtidos com as configurações testadas para o sistema, aplicado a cada conjunto de imagem da base IRMA formado conforme à constituição da mama e suas características.

O quarto capítulo expô-lo-á os resultados das simulações realizadas para classificação utilizando três modelos de redes neurais artificiais, verificando qual a melhor configuração para a proposta do trabalho, com melhor desempenho em relação às métricas utilizadas e maior poder de classificação. Este trabalho será encerrado por uma discussão de tais resultados, de

forma que serão propostas melhorias e diferentes perspectivas a serem abordadas em trabalhos futuros.

Capítulo 2

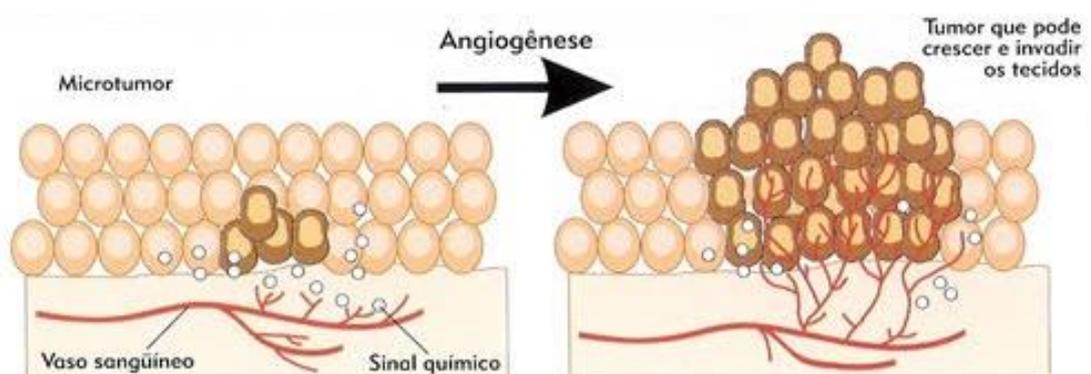
Análise de mamografias

Nas seções seguintes será realizada uma introdução a alguns conceitos importantes referentes à mama e a análise da mesma por imagem. Serão descritas as características da mama e dos tumores que podem lhes acometer, bem como as indicações, princípios de funcionamento e eficiência da mamografia, principal técnica de análise das mamas por imagem. Serão descritos igualmente algumas ferramentas de processamento de sinais, visão computacional e reconhecimento de padrões utilizadas na elaboração do sistema CAD estudado, bem como as ferramentas de análise da eficiência do sistema e os softwares utilizados na elaboração do projeto.

2.1 Câncer de mama

A unidade básica da vida são as células. Nosso corpo é constituído de diferentes tipos de células. Em condições normais, estas crescem e, quando necessário, dividem-se de forma ordenada para produzir mais células. No entanto, as células podem continuar dividindo-se além da necessidade do corpo e de forma desorganizada e descontrolada (Figura 1). Essas células extras formam uma massa de tecido, que chamamos de tumor.

Figura 1 - Processo de angiogênese.



No caso do câncer de mama, um indicador dos estágios iniciais é a presença de microcalcificações, que são depósitos de cálcio que se mobilizam do sangue para os tecidos, sofrendo, então, alterações do pH e fixando-se sob a forma de sais de cálcio.

Tumores podem ser malignos ou benignos. De forma geral, células de tumores benignos não se espalham para outras partes do corpo e, caso necessário, podem ser removidas e geralmente não crescem novamente. Entretanto, tumores malignos podem invadir tecidos vizinhos e se espalhar para diversas partes do corpo, processo conhecido como metástase, e, por isso, são mais perigosos.

A idade é o principal fator de risco para o câncer de mama. As taxas de incidência aumentam rapidamente até os 50 anos e, posteriormente, esse aumento ocorre de forma mais lenta (INCA, 2014). Contudo, outros fatores de risco já estão bem estabelecidos, como, por exemplo, aqueles relacionados à vida reprodutiva da mulher (menarca precoce, idade da primeira gestação, anticoncepcionais orais, menopausa tardia e terapia de reposição hormonal), história familiar de câncer da mama e alta densidade do tecido mamário (razão entre o tecido glandular e o tecido adiposo da mama). Além desses, a exposição à radiação ionizante, mesmo em baixas doses, também é considerada um fator de risco, especialmente durante a puberdade (INCA, 2014).

O câncer de mama é uma preocupação mundial e sua detecção em estágios iniciais pode aumentar as chances de cura e sobrevivência da paciente. Neste contexto, existem diversas técnicas de imageamento médico que auxiliam no diagnóstico desta doença, entre elas temos a mamografia, considerada técnica padrão ouro na detecção precoce do câncer de mama, a ultrassonografia da mama, o exame de ressonância magnética e a termografia de mama, esta última ainda pouco utilizada, mas que se mostra como uma técnica bastante promissora.

É recomendado o autoexame e o exame clínico da mama (ECM), visto que o câncer de mama em seus estágios iniciais é geralmente indolor, o que acaba, muitas vezes, sendo negligenciado. Desta forma, a realização do autoexame através da análise, pela própria paciente, do aspecto visual e pela palpação da mama e do ECM, que deve ser realizado por um médico ou enfermeiro treinados, podem auxiliar na detecção precoce de tumores de até 1 centímetro, tendo uma sensibilidade variando de 57% a 83% em mulheres com idade entre 50 e 59 anos, e em torno de 71% nas que estão entre 40 e 49 anos (INCA, 2013). A especificidade varia de 88% a 96% em mulheres com idade entre 50 e 59 anos e entre 71% a 84% nas que estão entre 40 e 49 anos (INCA, 2013).

A sensibilidade do diagnóstico está relacionada à capacidade do método de identificar os verdadeiros positivos, sendo definida como a razão entre o número de verdadeiros positivos e a soma de verdadeiros positivos e falsos negativos, de forma que a sensibilidade é máxima quando o número de falsos negativos é mínimo. Outra variável que pode ser calculada para

analisar a qualidade de um método de classificação é a especificidade, que está relacionada à capacidade do exame de rejeitar, os falsos positivos sendo definida como a razão entre o número de verdadeiros negativos e a soma de verdadeiros negativos e falsos positivos, de forma que a especificidade é máxima quando o número de falsos positivos é mínimo.

No Brasil, a mamografia e o exame clínico das mamas (ECM) são os métodos preconizados para o rastreamento na rotina da atenção integral à saúde da mulher (INCA,2013). A mamografia é a radiografia da mama que permite a detecção precoce do câncer, por ser capaz de mostrar lesões em fases iniciais. É realizada em um aparelho de raios X apropriado, chamado mamógrafo. Nele, a mama é comprimida de forma a fornecer melhores imagens, e, portanto, melhor capacidade de diagnóstico.

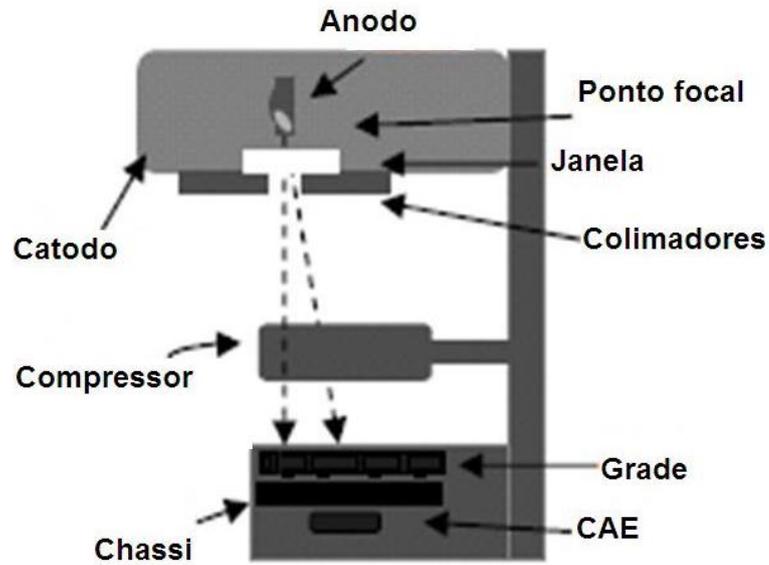
2.2 Mamografia

A mamografia pode ser utilizada como técnica de triagem para detectar alterações da mama em mulheres que não apresentam nenhum sinal nem anomalias observáveis da mama. Trata-se de uma radiografia detalhada do tórax, capaz de detectar anomalias e pode mostrar alterações na mama dois anos antes da paciente ou do médico poder senti-la através do exame clínico.

De forma geral, o exame de mamografia consiste na compressão do seio entre duas placas, cujo objetivo é homogeneizar a espessura da mama, bem como mantê-la imóvel durante o exame, e incidir sobre a mesma radiação X.

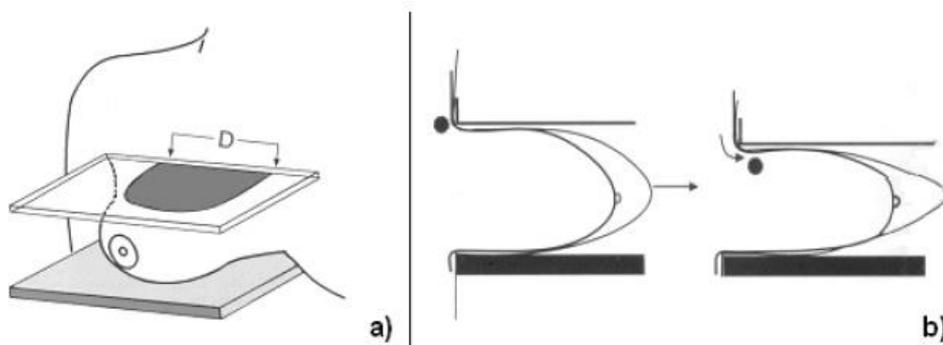
O mamógrafo utiliza um tubo gerador de raios X que é formado basicamente por dois eletrodos que recebem uma tensão da ordem de milhares de volts, de forma a gerar um campo elétrico. Quando aquecido pela passagem da corrente, o catodo (eletrodo negativo) libera elétrons que são acelerados pelo campo elétrico até colidirem com o anodo (eletrodo positivo), gerando um feixe de elétrons. O anodo é revestido por um metal pesado, como o molibdênio, por exemplo, e a região onde o feixe de elétrons atinge esse alvo é chamada de ponto focal. A resolução do equipamento está relacionada com o ponto focal, de forma que, quanto menor o tamanho do ponto focal, melhor a resolução espacial do sistema de imagem.

Figura 2 - Componentes do mamógrafo.



A mamografia é um exame que utiliza baixo kV e alto mAs para gerar melhor contraste, necessário na identificação das estruturas que compõem a mama, que possuem densidades semelhantes. Na realização da mamografia, deve-se utilizar compressão eficiente, entre 13 e 15 kgf, para obtenção de um bom exame. A compressão eficiente da mama contribui para a qualidade da imagem (Figura 3), mantendo a mama imóvel durante o exame, auxiliando na homogeneização e redução da espessura da mama, diminuindo a dose absorvida e o espalhamento dos fótons de raios X, além de aumentar o contraste da imagem ao reduzir a dispersão da radiação e diminuir as distorções ao aproximar a mama do sistema de captação. Isso reduz o ruído da imagem e diminui a sobreposição das estruturas internas da mama na formação da imagem, permitindo que lesões suspeitas possam ser detectadas com maior facilidade.

Figura 3 - a) Compressão da mama. b) Modificações da conformação da mama durante compressão.



Para aumentar a segurança do exame, o colimador é posicionado entre o gerador de raios X e a paciente. Este dispositivo evita a incidência de radiação ionizante em outras partes do corpo da paciente, colaborando também no aumento da qualidade da imagem.

O contraste é a diferença de sinal entre duas zonas adjacentes de uma imagem radiográfica. Este parâmetro depende do objeto, das características do feixe de raios X, da transmissão e do detector. O ruído é definido pela incerteza com a qual o sinal é adquirido. Existem diversas causas para o ruído, estas podem ser ligadas ao feixe de raios X ou ao sistema de aquisição e registro do sinal. O ruído depende também da sensibilidade e transmissão do receptor. De forma geral, podemos dizer que quanto melhor o receptor de imagem, mais perceptível será o ruído captado. Os melhores resultados são obtidos quando o ruído diminui e o sinal de interesse aumenta (STINES et al., 2010). No entanto, existem limites para a diminuição do ruído ligados a características técnicas da cadeia de aquisição de imagem e a necessidade de manter a dose de irradiação suficientemente baixa, conforme o princípio ALARA (As Low As Reasonably Achievable).

Segundo o Ministério da Saúde (INCA, 2007) e pelo sistema de padronização BI-RADS (ACR, 2003), recomenda-se que seja utilizado na descrição da mama o grau de substituição adiposa. A substituição do parênquima mamário pelo tecido adiposo é um processo dinâmico que ocorre na mama da maioria das mulheres, de acordo com a faixa etária. Na mama sem substituição (geralmente de mulheres mais jovens), o parênquima mamário ocupa a maior parte da mama e tem a forma de um triângulo cujo vértice está ligado ao mamilo. Mamas com essa configuração são consideradas densas (Figura 4). Mamas que possuam substituição adiposa menor que 50% são ditas predominantemente densas (Figura 5). Quando a substituição é maior que 50%, a mama é considerada como predominantemente adiposa (Figura 6) e, no caso em que a substituição adiposa é total, a mama é classificada como adiposa (Figura 7).

Os raios X são absorvidos por diferentes partes do corpo em diferentes graus. Assim como em radiografias comuns, tecidos densos, tais como massas, tecido fibroglandular e tecido vascular aparecem mais brilhantes nas imagens radiográficas, enquanto tecidos mais moles, como músculo e gordura, permitem que a radiação os atravesse de maneira mais fácil, impressionando o filme de forma a aparecerem mais escuros. Ou seja, conforme a descrição acima, mamas mais jovens tendem a ser reproduzidas como imagens mais claras na mamografia e estas imagens tendem a escurecer conforme o aumento da constituição adiposa da mama, relacionada ao aumento da idade da paciente, facilitando a visualização de lesões e anomalias.

Visto que a mamografia, principal exame utilizado para o diagnóstico de câncer de mama, possui baixo contraste, inerente à técnica e prejudicado ainda pelo tipo de tecido e constituição mamária da paciente, há uma grande dificuldade de leitura e interpretação do exame. O julgamento do radiologista durante a leitura de uma mamografia depende de seu treinamento, experiência e de critérios subjetivos. De acordo com a literatura, a mamografia tem sensibilidade variável, entre 88% e 93,1%, e especificidade entre 85% e 94,2% (INCA,2007). A sensibilidade depende de fatores como tamanho e localização da lesão, densidade do tecido mamário, qualidade dos recursos técnicos e habilidade de interpretação do radiologista, sendo a especificidade igualmente dependente da qualidade da imagem (INCA,2013). No entanto, a utilização desse exame como método de rastreamento reduz a mortalidade em 25% (INCA,2007).

Figura 4 - Exemplos de mamografias de mama densa.

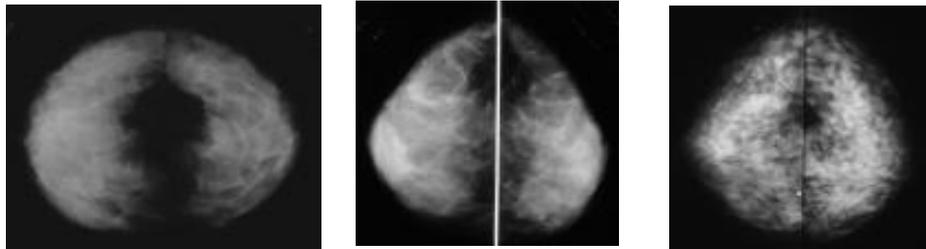


Figura 5 - Mama predominantemente densa.

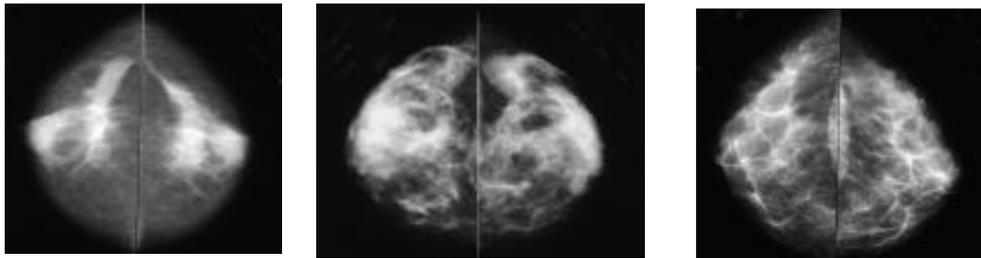


Figura 6 - Mama predominantemente adiposa.

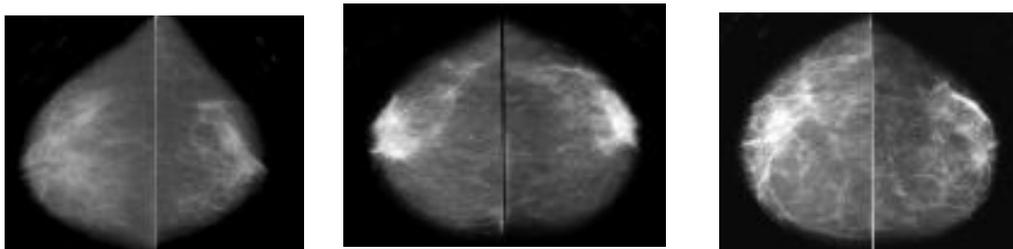
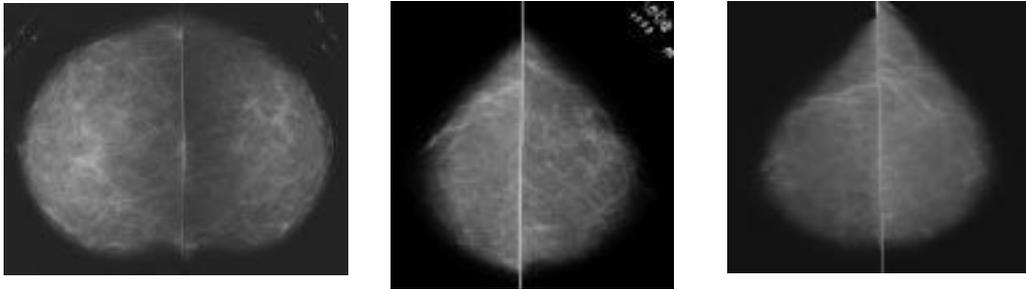


Figura 7 - Mama adiposa.



Dessa forma, sistemas Computacionais de Apoio ao Diagnóstico (Computer Aided Diagnosis, CAD) podem auxiliar os radiologistas na difícil tarefa de leitura e interpretação de imagens de mamografia, de forma a facilitar o processo de interpretação de imagens médicas. O desenvolvimento de sistemas CAD é uma tarefa multidisciplinar, que necessita da utilização de ferramentas de visão computacional, processamento digital de imagens, mineração de dados, aprendizado de máquina e inteligência artificial.

2.2.1 Tipos de lesões e classificação BI-RADS

Para classificação da mama, foi visto que a mesma tem sua constituição ponderada em relação à presença de tecido adiposo e fibroglandular, e esta característica está ligada à idade da paciente. Mulheres jovens tendem a possuir mamas mais densas, havendo a substituição do parênquima por tecido adiposo de forma progressiva com a idade.

A caracterização da lesão é feita, inicialmente, analisando a magnitude da mesma, isto é, se é composta com pequenos pontos de calcificações ou se se apresenta em forma de nódulos e massas. Microcalcificações lobulares geralmente possuem distribuição difusa, representando quase sempre um achado benigno. Já as microcalcificações intraductais apresentam contornos irregulares, podendo variar em tamanho, densidade e morfologia.

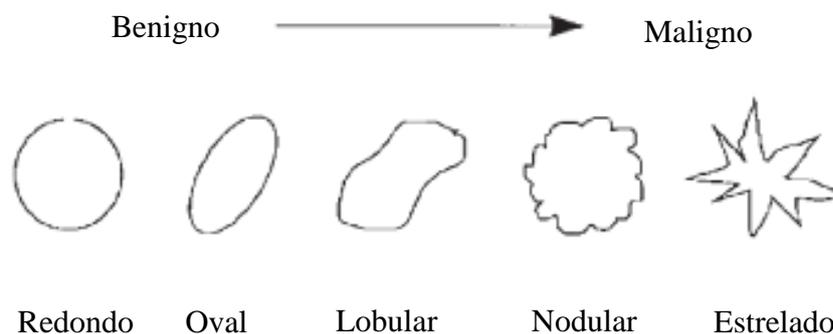
A abordagem para o diagnóstico de microcalcificações mamárias deve incluir tamanho, número, forma, densidade e distribuição. Através da análise da morfologia, é possível classificar as microcalcificações como sendo tipicamente benigna, suspeita ou de alta probabilidade de malignidade. Calcificações vasculares, em formato de pipoca, suturas e tecido gorduroso necrosado são geralmente indícios de achados benignos. Calcificações amorfas ou heterogêneas são casos suspeitos e necessitam de maior análise. Já o caso de calcificações em formato

linear ou ramificado e calcificações pleomórficas (com tamanho e forma irregulares) são indicações de malignidade.

Os nódulos e massas mamários são geralmente benignos e podem ser encontrados em cerca de 39% dos casos de câncer não palpáveis (INCA, 2000, 2007). A caracterização das massas é realizada através da análise das propriedades de forma e da margem da lesão, podendo se apresentar de forma redonda, oval, lobular ou irregular; as margens podem ser circunscritas, microlobulares, indistintas ou espiculadas (BOVIK, 2000). Massas espiculadas e com margens indistintas possuem maior probabilidade de malignidade que massas circunscritas, conforme ilustrado na Figura 8.

A classificação BI-RADS (Breast Image Reporting and Data System), criada pelo trabalho conjunto entre o Colégio Americano de Radiologia (American College of Radiology, ACR), o Instituto Nacional do Câncer (National Cancer Institute, NCI), o Centro de Controle e Prevenção de Doenças (Centers for Disease Control and Prevention, CDC) e a Agência de Administração de comidas e fármacos (Food and Drug Administration, FDA), é o nome do sistema padronizado utilizado para uniformizar os relatos de radiologia quando se analisam as imagens de mamografia (ACR, 2003). Este sistema de classificação é também uma ferramenta de controle de qualidade desenvolvido para normalizar os laudos de exames de mamografia, reduzindo a possibilidade de confusão ou má interpretação de imagens da mama, facilitando o monitoramento dos resultados.

Figura 8 - Aspectos morfológicos de massas mamárias.



A classificação BI-RADS é dividida em categorias, sendo a categoria zero definida como incompleta e as demais (1, 2, 3, 4, 5 e 6) como avaliações completas. Segundo a ACR, a categoria zero é utilizada na mamografia de rastreamento para descrever achados que necessitam

de maior investigação, como avaliação adicional e/ou comparação com mamografias anteriores ou com outro tipo de exames, como ultrassonografia ou ressonância magnética.

A categoria 1 representa um exame com resultado negativo para o câncer, no qual a mama é simétrica e não há presença de massas, distorções de arquitetura ou suspeitas de calcificação. Na categoria 2, os achados são benignos, no entanto, pode-se achar necessário descrever o tipo de achado benigno, como fibroadenomas calcificados, múltiplas calcificações secretórias, lesões, como cistos ou lipomas. A categoria 3 é utilizada para descrever um achado provavelmente benigno, com menos de 2% de risco de malignidade. Na categoria 4, enquadram-se as anormalidades suspeitas e que necessitam de avaliação histológica ou citológica adicional. Esta categoria é aplicada para achados que não possuem a aparência clássica de malignidade, mas que apresenta uma probabilidade de malignidade superior a categoria 3. A categoria 5 representa uma alta propensão a malignidade (probabilidade do câncer superior a 95%). Neste caso, o tratamento cirúrgico pode ser indicado sem a necessidade de confirmação por biópsia. A categoria 6 é restrita a caso em que a lesão identificada através do estudo por imagem é confirmada por biópsia(ACR, 2003).

2.3 Base de dados

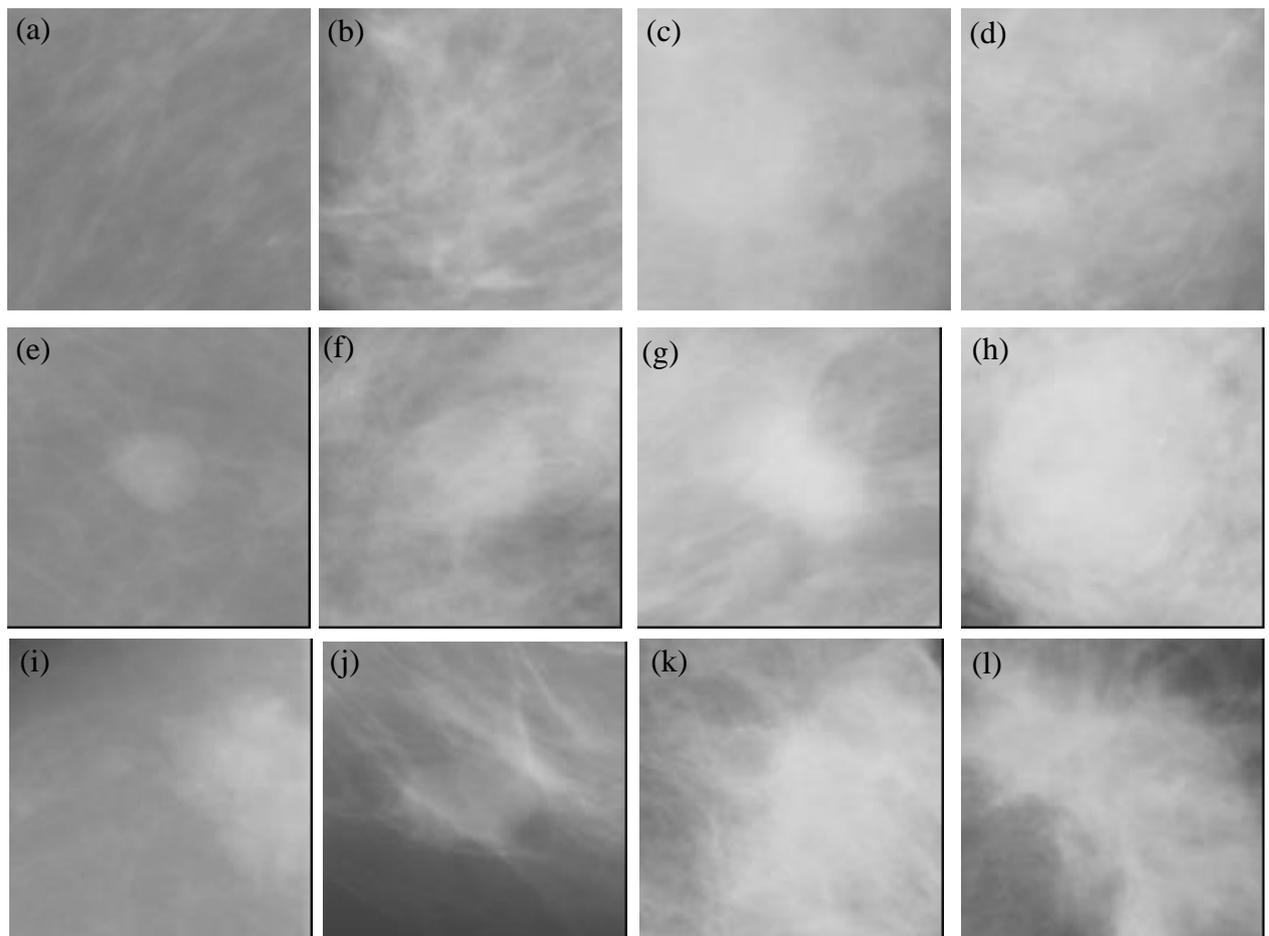
A base de dados IRMA (Image Retrieval in Medical Applications), criada pela Universidade de Tecnologia de Aachen (RWTH Aachen), é a unificação de outras bases de dados online publicamente disponibilizadas para pesquisa, como DDSM (Digital Database for Screening Mammography), MIAS (Mammographic Image Analysis Society), LLNL (Lawrence Livermore National Laboratory) e RWTH (Rheinisch-Westfälische Technische Hochschule, Aachen University, Aachen, Germany Department of Radiology).

A base de dados contém imagens de radiografia digitalizadas, que foram classificadas por radiologistas (OLIVEIRA et al., 2010, DESERNO et al., 2011, 2012). As imagens contidas nessa base são regiões de interesses, extraídas das imagens originais, redimensionadas para imagens de tamanho 128 x 128 pixels. Ao todo, são 2.796 imagens, sendo 2.576 provenientes da base DDSM, 1 da base LLNL, 69 da base RWTH e 150 proveniente da base MIAS.

A base IRMA está dividida em 12 subclasses, cada uma com 233 imagens. As classes são definidas baseadas no tipo de tecido da mama e no tipo do tumor, que pode ser benigno, maligno ou sem lesão. A mama pode ser descrita pela sua densidade em 4 classes, conforme a classificação BI-RADS para densidade da mama, em extremamente adiposa (I), fibroglandular (II), densa (III) e extremamente densa (IV).

Há também uma subdivisão dessas classes, baseadas no formato da lesão, que podem ser não especificadas, para o caso de mamas sem lesão, calcificações, microcalcificações, macrocalcificações, massa circunscrita, massa espiculada, outro tipo de massa, distorção arquitetural e lesão assimétrica.

Figura 9 - Amostra de imagens das 12 classes da base IRMA. Colunas, da esquerda para direita, temos os tipos de mama adiposa, fibroglandular, densa e extremamente densa, respectivamente. Linhas de cima para baixo representam o estágio da lesão iniciando por representação de imagem sem tumor, com tumor benigno e com tumor maligno, respectivamente.



Visto que massas circunscritas são predominantemente benignas e massas espiculadas tendem a ser malignas, utilizaremos amostras de imagens que possuam estes tipos de lesão para a realização dos testes do sistema CAD a ser desenvolvido. Desta forma, trabalharemos com um total de 1.416 amostras da base IRMA, das quais 19% são representativas de lesões circunscritas e destas, 82% são benignas e 19% de amostras representativas de lesões espiculadas, das quais 90% é maligna.

Tabela 1: Divisão da base de dados IRMA utilizada no trabalho estudado conforme tipo de tecido mamário e lesão.

Tipo de tecido mamário	Benigno Circunscrito	Maligno Espiculado	Sem Lesão	Total
Adiposa	66	56	233	355
Fibroglandular	86	95	233	414
Pouco denso	45	63	233	341
Muito densa	32	41	233	306
Total	229	255	932	1416

2.4 Sistemas CAD

As imagens de mamografia são de difícil leitura e interpretação. A mamografia convencional possui sensibilidade de aproximadamente 85% na detecção de tumores, sendo esses resultados dependentes do operador e da sua experiência, bem como de critérios subjetivos do especialista (BLUE CROSS AND BLUE SHIELD ASSOCIATION, 2002). O desenvolvimento de técnicas capazes de aumentar a sensibilidade e reduzir a variabilidade entre especialistas poderão reduzir resultados falsos negativos (BLUE CROSS AND BLUE SHIELD ASSOCIATION, 2002).

Desta forma, sistemas computacionais de apoio ao diagnóstico (Computer Aided Diagnosis, CAD) podem auxiliar o profissional em sua rotina, colaborando para a interpretação de exames de mamografia, detecção e classificação de tumores. Sistemas CAD possuem uma precisão de aproximadamente 80 a 90% em detecção de lesões, sendo melhores na identificação de microcalcificações (precisão entre 91 e 100%) do que de massas (precisão entre 67 e 89%) (BLUE CROSS AND BLUE SHIELD ASSOCIATION, 2002).

Os esquemas CAD podem ser de auxílio à detecção (CADe), colaborando para a localização de estruturas de interesse, ou podem ser sistemas de auxílio ao diagnóstico (CADx), realizando análise das estruturas e classificação das mesmas, colaborando na interpretação dos eventuais achados. Sistemas CAD combinam elementos de inteligência artificial e processamento digital de imagens para aplicações em reconhecimentos de padrões, como detecção de tumores em imagens radiográficas. A combinação de sistemas CAD e do conhecimento do radiologista aumentam a sensibilidade e acurácia de detecção de tumores e lesões.

Todas as etapas envolvidas num sistema de apoio ao diagnóstico são importantes e interdependentes, podendo os métodos utilizados variarem conforme o tipo de imagem trabalhada e do objetivo da proposta. Alguns autores acreditam que o desempenho de um sistema CAD depende principalmente da seleção de atributos e do treinamento do conjunto de dados do que da etapa de classificação propriamente dita (ISLAM et al., 2010).

Sistemas CAD aplicados ao apoio do radiologista no diagnóstico do câncer de mama através da mamografia já são utilizados na prática, principalmente nos Estados Unidos. Sua principal função é servir como uma segunda opinião, alertando o radiologista de regiões suspeitas. O protocolo recomendado para implementação de um sistema CAD é que a consulta à classificação realizada pelo computador ocorra apenas após a análise do exame e do parecer pelo radiologista. Após a primeira análise realizada pelo médico especialista, o sistema CAD é consultado e, então, o médico deve re-analisar as imagens para dar o diagnóstico definitivo. Dessa forma, é preservada e valorizada principalmente a avaliação realizada pelo radiologista.

O primeiro sistema CAD aprovado pelo FDA (Food and Drug Administration) foi o R2 Image Checker, desenvolvido pela empresa Hologic. A precisão de detecção de calcificação pelo sistema foi reportada em 98,5%, com 0,74 falsos positivos por caso (conjunto de quatro imagens). Para detecção de massas, foi reportado uma precisão de 85,7% com 1,32 falsos positivos por caso (BOVIK, 2010).

Outro sistema aprovado pelo FDA em 2002 foi o MammoReader, da empresa Intelligent System Software Inc. (ISSI). A sensibilidade geral reportada pelo fornecedor foi de 89,3%, com 2,32 marcadores falso positivos em caso de presença de câncer e 3,32 marcadores falso positivos nos casos sem câncer de mama (BOVIK, 2010). A terceira empresa a conseguir aprovação para comercialização de um sistema CAD para diagnóstico de câncer de mama foi o Second-Look, fornecido pela empresa CADx Medical Systems. Segundo os dados do fabricante, o sistema possui sensibilidade de 85% (BOVIK, 2010).

A dupla leitura da mamografia é o método mais preciso para o diagnóstico do câncer de mama. No entanto, para países que não dispõem de recursos financeiros nem mão de obra especializada necessários à aplicação desse sistema de rastreamento do câncer de mama, a utilização de sistemas CAD pode ser a solução para redução do número de falso positivos e ao aumento da sensibilidade e precisão do diagnóstico, visto que os sistemas computacionais aplicados à detecção e classificação de câncer de mama utilizado em conjunto com a opinião do médico aumentam o desempenho do sistema de diagnóstico por um único especialista (GILBERT et al., GROMET, 2008).

2.4.1 Análise digital de imagens

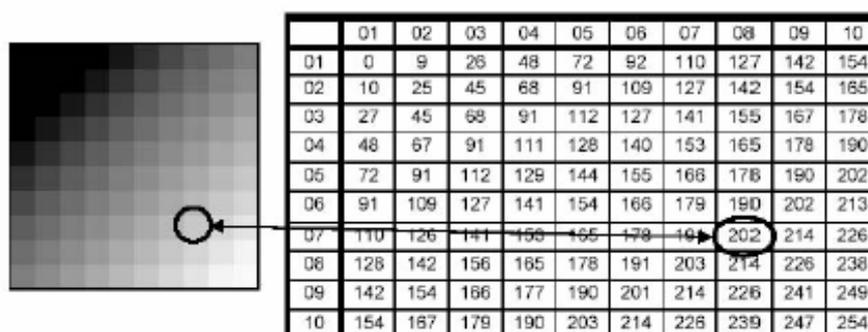
Uma imagem pode ser definida como sendo a representação visual de um objeto. Do ponto de vista matemático, uma imagem é considerada uma função bidimensional $f(x,y)$ onde x e y são coordenadas planas, e a amplitude de f em qualquer par de coordenadas (x,y) é chamada de intensidade ou nível de cinza da imagem no referido ponto. Quando (x,y) e a amplitude de f fazem parte de um conjunto de valores finitos, ou discretos, a imagem é chamada de imagem digital (COSIO et al., 2011).

Uma imagem digital pode ser representada através de uma matriz $N \times M$ onde cada elemento, chamado pixel (Picture Elements), corresponde a um nível de cinza $f(x,y)$ em um determinado ponto da imagem. Na prática, a intensidade do pixel representa uma cor. Em imagens binárias, os valores dos pixels podem assumir os valores 0 ou 1. Para imagens em tons de cinza, os valores de intensidade podem variar de 0 a 255 (OSHIRO e GOLDSCHMIDT, 2008).

As imagens digitais coloridas são, em geral, funções quantificadas e apresentadas como um vetor de três componentes de acordo com o modelo escolhido, de forma que cada componente varia de 0 a 255 de acordo com a cor (Figura 10). Pode-se considerar como exemplo os seguintes modelos (OSHIRO e GOLDSCHMIDT, 2008):

- Modelo RGB (Red, Blue e Green): formado pelas cores primárias;
- Modelo CMY (Cyan, Magenta e Yellow): modelo RGB com cores secundárias;
- Modelo HSI (Hue, Saturation e Intensity): formado por uma cor dominante (matiz), saturação (diluição da matriz na luz branca) e intensidade (quantidade de luz refletida).

Figura 10 - Representação de uma imagem ampliada de 10x10 com 256 tons de cinza.



A imagem pode ser manipulada de diversas formas e para diversos fins. A visão computacional é uma área da computação que busca extrair e inferir informações através de ima-

gens, enquanto o processamento de imagens baseia-se na manipulação e tratamentos de imagens de forma a torná-las mais adequadas para sua utilização na aprendizagem de máquinas e reconhecimento de padrões.

Visão Computacional é a área da ciência que se dedica a desenvolver teorias e métodos voltados à extração automática de informações úteis contidas em imagens. Ou seja, pode ser entendido como um conjunto de métodos e técnicas através dos quais sistemas computacionais podem ser capazes de interpretar imagens. Desta forma, aplicações na área de visão computacional apenas foram possíveis com a evolução tecnológica de computadores e dispositivos de captura de imagens, de forma que grande parte de suas soluções e abordagens são ainda objetos de pesquisa.

O processamento médico de imagens, caracterizado pela extração de informação de imagens para realizar diagnósticos sobre os pacientes é uma área muito promissora para a utilização de ferramentas computacionais, utilizando ferramentas de visão computacional e reconhecimento de padrões. Reconhecimento de padrões é uma área da ciência cujo objetivo é a classificação de objetos dentro de um número de categorias ou classes. Esses objetos de estudo variam de acordo com cada aplicação. Particularmente, trabalharemos com dados contidos em imagens de mamografia. A área abrange tanto a detecção de padrões e escolha entre dois objetos, por exemplo, como a complexa realização da aprendizagem.

A aplicação de reconhecimento de padrões para o aprendizado pode ser supervisionado ou não supervisionado. No primeiro caso, um padrão é considerado como um par de variáveis $\{x, w\}$, onde x é um conjunto de características e w é a classe a que pertence, sendo dito “supervisionado” por ser dado ao sistema o vetor de características e seu rótulo. No caso de sistema de aprendizado não supervisionado, o algoritmo assume que a classe a qual pertence o conjunto de dados não é conhecida. O reconhecimento supervisionado utiliza o conjunto de treinamento para classificar os dados obtidos de acordo com as categorias já existentes e nelas organizá-los. Por sua vez, o reconhecimento não supervisionado utiliza o conjunto de treinamento para criar novas categorias, ao invés de simplesmente separar os dados de acordo com as categorias já existentes.

Um sistema completo de reconhecimento de padrões consiste de um sensor que obtém observações a serem classificadas ou descritas, cujo objetivo principal é eliminar dados desnecessários ou distorcidos fazendo com que sejam apresentados apenas objetos relevantes para o reconhecimento do que está sendo analisado; um mecanismo de extração de características que computa informações numéricas ou simbólicas das observações; e um esquema de

classificação das observações, cujos resultados serão categorizados conforme seus padrões, ou seja, objeto de análise é declarado como pertencente à determinada classe.

Visto que este trabalho propõe o estudo de um algoritmo que auxilie na detecção automática de lesões em imagens de mamografia utilizando ferramentas de visão computacional, processamento de imagens e reconhecimento de padrões, podemos estabelecer algumas etapas fundamentais para a detecção, reconhecimento e classificação de dados mamográficos:

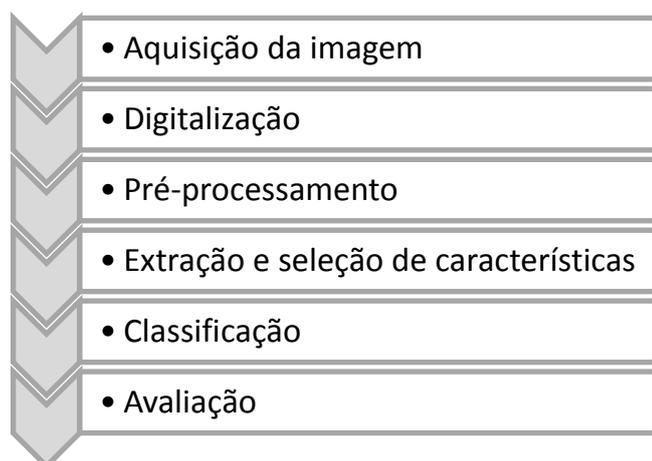
- Aquisição da imagem: trata-se do processo de aquisição da imagem através de sensores e digitalizadores, o primeiro converte a informação ótica em sinal elétrico e o segundo transforma a imagem analógica em digital, imagem esta que pode ser bidimensional, tridimensional ou uma sequência de imagens, colorida ou em tons de cinza, dependendo dos valores dos pixels, que indicam a intensidade da luz.
- Pré-processamento: antes de um método de visão computacional ser aplicado em uma imagem para extrair informação, é geralmente necessário processar a imagem para assegurar-se que ela satisfaz as condições do método, aprimorando a qualidade da imagem, necessitando, para isto, a aplicação de métodos específicos que facilitem a identificação de objetos, redução do ruído da imagem, aumento do contraste, destaque de bordas, contornos, figuras geométricas, etc.
- Detecção e segmentação: processo realizado com a finalidade de destacar regiões relevantes da imagem, possibilitando a seleção de interesse de regiões específicas da imagem e segmentação da mesma para processamento posterior.
- Extração de características: etapa que permite a extração de características matemáticas que compõe a imagem e permite identificá-la, como textura, bordas, formato, movimento, etc.
- Classificação: utilização das informações extraídas da imagem para classificação dos objetos em diferentes categorias, conforme proposta inicial do problema.

A Figura 11 esquematiza as etapas citadas acima, que descrevem o funcionamento da maioria dos detectores de tumores utilizando sistemas CAD (RAHMATI et al., 2012). Visto que sistemas CAD são usados em colaboração com o diagnóstico do radiologista, depois de concluídas as etapas descritas anteriormente, é necessária a intervenção do médico responsável para a interpretação dos resultados (GANESA et al., 2013).

Mamografias não proporcionam um bom contraste entre tecido glandular normal e tecidos malignos. Isto se deve à atenuação dos raios-x, que nesses dois tecidos, não variam muito, especialmente em mulheres jovens ou em mulheres com mamas mais densas. Desta forma, técnicas de filtragem e aprimoramento são importantes no processamento de imagens de mamografia.

A maioria dos algoritmos de detecção consiste em dois estágios: detecção de regiões suspeitas na imagem de mamografia e classificação dessas regiões como tecido normal ou lesão (BOVIK, 2000). O primeiro estágio geralmente possui uma alta sensibilidade, sendo aceitável um grande número de falsos positivos, isto é, regiões classificadas como suspeitas mesmo sendo normal, visto que o segundo estágio pode eliminar a suspeita de regiões normais (BOVIK, 2000).

Figura 11 - Esquema de um sistema CAD.



Em métodos de detecção baseados em região, inicialmente extraímos a região de interesse (ROI) utilizando filtros ou segmentação de imagens. Em seguida, algumas características das imagens são extraídas para cada região e, então, cada região é classificada como suspeita ou normal. Essas características são selecionadas de forma a descreverem informações diagnósticas importantes, como forma e textura da região extraída (BOVIK, 2000).

Resumidamente, características são extraídas das ROIs e, em seguida, cada ROI é classificada como contendo lesão ou tecido normal. Geralmente, utiliza-se para classificação redes neurais artificiais com algoritmos de classificação supervisionados, isto é, primeiramente a rede é treinada com o conjunto de amostras de treinamento e, posteriormente, o desempenho do classificador é avaliado em um conjunto de amostras de teste, geralmente diferente do conjunto de treino.

2.4.2 Segmentação de imagens

Ao analisar as imagens de mamografia é importante distinguir entre regiões suspeitas e seus arredores. O objetivo da etapa de segmentação é simplificar e modificar a representação da imagem em algo que seja mais significativo e mais fácil de analisar. A segmentação é aplicada para localizar objetos e contornos nas imagens. De forma geral, podemos dizer que o objetivo da segmentação é extrair a Região de Interesse (ROI) que contenha a lesão que desejamos classificar.

Utilizamos neste projeto a segmentação através dos Mapas de Kohonen, ou Mapas Auto Organizáveis (Self-Organizing Maps, SOM), algoritmo baseado em aprendizado competitivo. São exemplos de redes neurais artificiais de aprendizado não supervisionado, diferindo destas por utilizar funções de vizinhança para preservar propriedades topológicas do espaço de entrada. As redes SOM são eficientes na visualização de dados de alta dimensionalidade através da redução das dimensões na saída, ou seja, criando um mapa dos dados de entrada.

A rede SOM é composta por neurônios que são associados a um vetor de peso de mesma dimensão que o vetor de entrada. Os neurônios estão dispostos num espaço bidimensional, numa grade retangular. O procedimento para alocar um vetor de entrada no mapa é encontrar o neurônio que apresenta vetor de peso mais próximo (menor distância) do vetor de dados do espaço. Os valores dos pesos nos neurônios são inicializados com valores baixos e aleatórios.

Para que o vetor de peso seja adaptado aos dados de entrada, a rede precisa ser treinada, e para isto, utilizamos aprendizado competitivo. Desta forma, quando uma amostra de treinamento alimenta a rede, a distância Euclidiana de todos os vetores de peso é computada e a rede busca a saída gerada que mais se assemelha aos dados de entrada, obtendo assim o neurônio cujo vetor de peso é semelhante à entrada. Este neurônio e sua vizinhança têm seus pesos modificados com o intuito de aumentar ainda mais a semelhança da saída com o padrão de entrada.

Uma vez apresentado o dado de entrada, cada neurônio computa o seu nível de ativação multiplicando o seu vetor de pesos pelo vetor de entrada, da maneira usual. O neurônio que possuir maior nível de ativação é chamado de vencedor e apenas ele terá atividade diferente de zero na saída da rede, ou seja, o padrão de entrada que estiver sendo apresentado à rede provocará a ativação de apenas um neurônio da rede neural. Pode-se ter mais de um vetor sendo representado pelo mesmo neurônio vencedor, de forma que se diz que esta rede implementa um mecanismo de agrupamento (ou *clustering*) de padrões.

O objetivo do aprendizado em mapas SOM é permitir que diferentes partes da rede respondam de forma semelhante a certos padrões de entrada. Assim sendo, é necessário alimentar a rede com uma grande quantidade de amostras de treinamento que representem de forma satisfatória os tipos de vetores esperados durante o mapeamento (JIANG et al., 2010).

A entrada do mapa será o nível de cinza de cada pixel da imagem. O algoritmo percorrerá toda a imagem fazendo comparações entre a entrada e a saída do mapa e, de forma iterativa, atualizará o peso do neurônio que gerar a saída mais próxima do valor de entrada e seus vizinhos, de forma que o raio da vizinhança abrangida pela atualização irá se reduzindo conforme o andamento do treinamento. O intuito da segmentação utilizando redes SOM é agrupar as regiões da imagem com relação à distribuição de seus níveis de cinza, de forma que tecidos diferentes sejam segmentados, deixando visível a região de interesse, neste caso, as lesões e calcificações mamárias.

As transformações matemáticas são empregadas no mapeamento de funções de um domínio para outro e essas transformações são particularmente importantes em processamento e análise de sinais, pois, no domínio transformado, algumas propriedades relevantes do sinal ficam mais evidentes.

A Transformada Wavelet decompõe um sinal a partir de uma função principal em diferentes componentes de frequência, para que cada componente possa ser analisado em diferentes escalas. As wavelets não se limitam a fornecer a representação do domínio frequência de um sinal. Essas funções também representam o domínio do espaço, propiciando a análise do sinal em qualquer instante, sendo este capaz de revelar aspectos importantes como limites, pontos de inflexão, descontinuidades e similaridade.

Wavelets são conjuntos de bases não lineares, de forma que ao projetar uma função em termos de wavelets as funções de bases são escolhidas de acordo com a função a ser aproximada. Dessa forma, diferentemente das famílias de bases lineares em que as funções de bases são estáticas e utilizadas para todo tipo de função de entrada, as wavelets utilizam um conjunto de bases dinâmico, permitindo que a funções de entrada sejam representadas da melhor forma possível.

As funções de wavelets são obtidas a partir da wavelets mãe, através de dilatações e translações, de forma que uma família de funções base possam descrever outras funções pertencentes ao domínio. Para aplicação em processamento de sinais, algumas propriedades das wavelets são desejáveis, como serem nulas no infinito, possuírem energia finita e ter certo grau de regularidade e que possuam certo número de momentos nulos.

As wavelets de Daubechies são famílias de wavelets ortogonais, talvez a mais conhecida, sendo numerada em função da quantidade de momentos nulos que possui, de forma que o índice em si representa o número de coeficientes que os filtros associados possuem. A wavelet de Haar é bastante simples, semelhante a wavelet de Daubechies D2, com um momento nulo, descrita como um pulso quadrado.

2.4.3 Extração de características

A descrição de imagens de forma sucinta é baseada principalmente em distribuição de intensidades, textura e forma. Apesar de a cor ser um atributo confiável, situações em que esta não fornece muitas informações requerem a descrição através de outros atributos. Uma propriedade desejável do conjunto de características para descrição de imagens é que ele se comporte de maneira semelhante para objetos de mesma natureza que estejam em posição, rotação e escala diferentes. Isto é, que sejam invariantes às transformações geométricas (DA SILVA, 2007).

Nódulos e calcificações aparecem na imagem de mamografia com diferentes características, de forma que lesões malignas geralmente apresentam contornos espiculados e irregulares, enquanto lesões predominantemente benignas são arredondadas ou ovais, com contornos suaves (RAHMATI et al., 2012). Medidas que possibilitem representar formas e contornos auxiliam na classificação de tumores. Podemos distinguir duas técnicas de análise: a análise do formato, que incluem características baseadas na morfologia das lesões e a análise de textura, que nos dá informações sobre a organização espacial de cor ou intensidade de uma imagem ou região de interesse.

A textura pode ser entendida como mudanças na intensidade da imagem, formando padrões repetitivos, podendo ser reconhecida pelo olhar através de características como suavidade, rugosidade e regularidade. A descrição da textura pode ser realizada através de uma abordagem estatística, estimando as propriedades dos valores individuais dos pixels e/ou da interação espacial entre os pixels da imagem; através de uma abordagem estrutural, a qual estabelece o relacionamento de vizinhança, densidade (número de elementos por unidade espacial) e homogeneidade; ou através de uma abordagem espectral, baseados em propriedades de espectros de frequência obtidos por meio de transformadas, como a Transformada de Wavelets.

Para obtenção das características de forma de uma imagem é necessário segmentá-la, de forma que o vetor de características que descreve a imagem passa a ser formado por atributos numéricos, como o formato do contorno, a excentricidade, curvatura e posição do objeto na

imagem. Para melhores resultados na formulação do vetor de atributos é necessário o pré-processamento da imagem, podendo ser seguido por uma etapa de detecção de bordas. Uma das maneiras de descrever a forma de objetos numa imagem é através de momentos invariantes.

A matriz de co-ocorrência de níveis de cinza (Gray Level Co-Occurrence Matrix, GLCM) é uma abordagem estatística amplamente conhecida e utilizada como extrator de padrões de textura, baseada nos descritores de características de textura de Haralick (HARALICK et al., 1973). O método GLCM pode ser entendido como uma matriz de frequência no qual dois pixels, separados por certo vetor, ocorrem na imagem. A distribuição da matriz depende do relacionamento angular e da distância entre pixels.

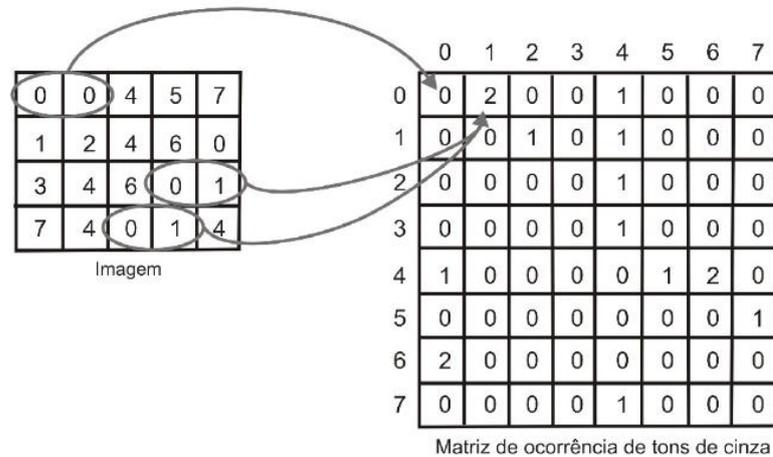
Uma matriz de co-ocorrência $P_d(i,j)$ contabiliza a co-ocorrência de pixels com valores de cinza i e j em uma dada distância d , sendo esta última definida em coordenadas polares (d,α) , onde α pode assumir valores de 0° , 45° , 90° e 135° e d é escolhida conforme a granularidade da imagem ($d=1$ para pixels vizinhos).

Os 14 padrões extraídos por Haralick utilizando a matriz de co-ocorrência são: segundo momento angular, contraste, entropia, momento da diferença inversa, correlação, variância, dissimilaridade, média, soma das variâncias, soma da entropia, diferença de variância, diferença de entropia, medidas de correlação (dois padrões).

A análise da forma pode ser realizada através do método dos momentos, que proporciona uma técnica para decomposição de uma forma arbitrária em um conjunto de características. Essa abordagem se baseia em transformações lineares diretas, sem necessidade de determinação de parâmetros específicos. As técnicas baseadas em momento descrevem as propriedades globais de uma imagem, isto é, a maneira com que as características dos pixels brancos e pretos representam o padrão.

Os momentos de Zernike pertencem à classe dos momentos ortogonais, invariantes à rotação. São valores complexos que podem ser expressos como a soma da parte real e imaginária e que possuem correlação com padrões binários. Os momentos de Zernike podem ser definidos como projeções da função $f(x,y)$ de intensidade da imagem sobre funções de bases ortogonais, no caso, os polinômios de Zernike, que formam um conjunto ortogonal complexo no interior de um círculo unitários, sendo, portanto, limitado pela curva $x^2+y^2 = 1$.

Figura 12 - Matriz de co-ocorrência de níveis de cinza.



Para calcular os momentos de Zernike de uma imagem, ou ROI, o centro dela é considerado como o centro do disco unitário e as coordenadas do pixel são mapeadas para este disco, de forma que os pixels fora dele não são usados no cálculo. As coordenadas polares são, então, descritas num vetor de distâncias ρ , da origem até o ponto da coordenada, formando um ângulo θ entre o eixo x e o vetor ρ . Os momentos de Zernike são invariantes à transformação de rotação e também podem ser invariantes em relação às transformações de escala e translação, se a imagem de entrada passar por uma manipulação prévia baseada na teoria dos momentos geométricos.

Para obter invariância à escala e à translação, a imagem deve ser normalizada utilizando os momentos cartesianos antes do cálculo dos momentos de Zernike. A invariância à translação resulta da modificação da origem para o centro da imagem, utilizando momentos centralizados. Seguindo para a invariância à escala, cada objeto é alterado de forma que sua área ou contagem de pixels seja um valor predeterminado.

2.4.4 Classificação dos atributos

A classificação é uma tarefa de mineração de dados que ocorre quando um objeto precisa ser designado a um grupo ou uma classe predeterminada, baseado nos seus atributos. A tarefa de classificação pode ser dividida num processo de duas etapas. Na primeira etapa, a fase de treinamento, um classificador é construído para descrever um conjunto predeterminado de classes. Para isto, utiliza-se um algoritmo que "aprende" a partir do conjunto de treinamento (BONIFACIO, 2010). A segunda etapa é a avaliação do classificador, na qual a precisão da predição é avaliada através de um conjunto de teste. O conjunto de teste é formado a partir de registros selecionados randomicamente, do conjunto de dados gerais e independentes dos re-

gistros de treinamento, ou seja, não são usadas para construir o classificador (BONIFACIO, 2010).

Uma vez que os atributos relacionados às lesões são extraídos, eles são utilizados como entrada para um classificador que deverá realizar a tarefa de classificação dos dados em três tipos de classes: normal, benigno e maligno. As redes neurais artificiais (Artificial Neural Networks, ANNs), como redes de função de base radial (Radial Basis Function, RBF) e perceptrons de múltiplas camadas (Multi Layer Perceptron, MLP), possuem boa performance em classificação de massas (BOUJELBEN et al., 2009).

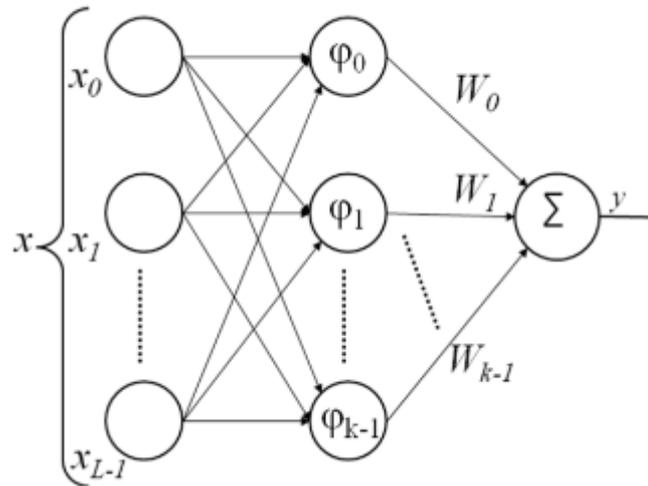
As redes MLP são bastante comuns nas áreas de medicina e engenharia biomédica, utilizadas para solução de problemas complexos sem a necessidade de modelos matemáticos e conhecimentos aprofundados sobre o mecanismo envolvido. Apesar da popularidade das redes MLP, esta requer uma grande quantidade de tempo para seu treinamento e para tarefa de determinação dos pesos dos neurônios. Assim como para redes MLP, as redes RBF necessitam atribuir pesos aos seus neurônios, mas estas possuem menos pesos para treinar e são atribuídos aos neurônios uma distribuição, reduzindo a quantidade de tempo e memória necessário ao treinamento de redes RBF (AZAR e EL-SAID, 2013).

As redes RBF são do tipo *feed-forward* não-linear, e sua estrutura básica é formada por três camadas com papéis diferentes. A camada de entrada, que contém nós que conectam a rede a seu ambiente, a camada escondida, que forma um conjunto de funções que constituem uma base arbitrária para os vetores de entrada, aplicando uma transformação não-linear do espaço de entrada para o espaço escondido e a camada de saída, que realiza uma transformação linear dos dados provenientes da camada oculta, fornecendo a resposta da rede ao padrão de ativação aplicado na entrada.

O funcionamento da camada escondida está baseado no teorema de Cover (HAYKIN, 1999) sobre a separabilidade dos padrões, onde o mesmo afirma que um problema complexo de classificação de padrões dispostos não-linearmente em um espaço de alta dimensionalidade tem maior probabilidade de ser linearmente separável do que em um espaço de baixa dimensionalidade (BONIFACIO, 2010).

Nas redes RBF, a função de ativação de cada neurônio da camada escondida é função da distância entre seus vetores de peso e de entrada. O treinamento dessas redes é realizado em dois estágios: o primeiro estágio utiliza funções de ativação não lineares, isto é, funções de base radial não supervisionada. No segundo estágio, são determinados os pesos da camada de saída, que é uma etapa linear.

Figura 13 - Arquitetura de uma rede RBF.



Os perceptrons de múltiplas camadas (Multilayer Perceptrons, MLP) são constituídos por nós fontes, que formam a camada de entrada da rede, por uma ou mais camadas escondidas e uma camada de saída. São redes do tipo *feedforwards*, isto é, nenhum neurônio de uma dada camada agirá como entrada para um neurônio de uma camada anterior, ou seja, são redes que não apresentam laços de realimentação.

O número de nós fonte na camada de entrada da rede é determinado pela dimensionalidade do espaço de observação, que é responsável pela geração dos sinais de entrada. O número de neurônios na camada de saída é determinado pela dimensionalidade requerida da resposta desejada.

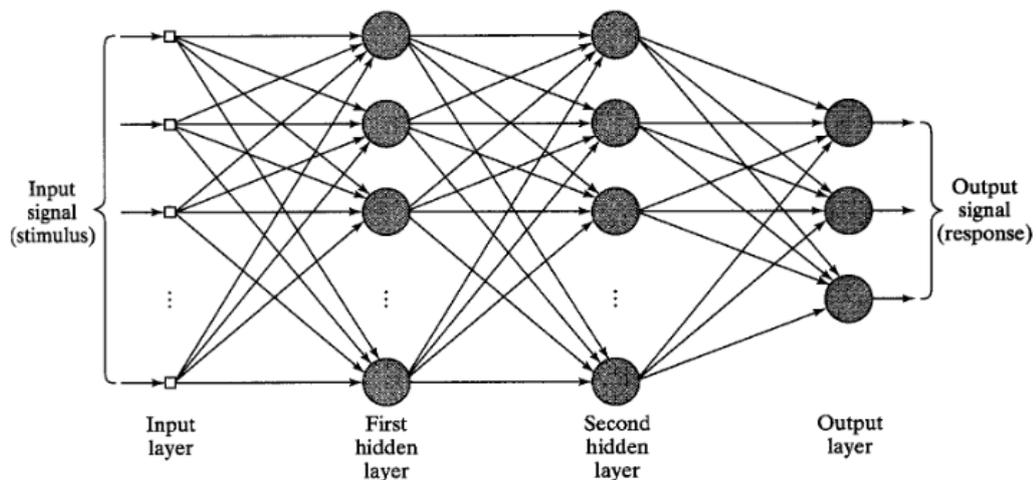
Numa rede MLP típica, os nós de entrada possuem uma função de ativação linear e não possuem limitadores (*thresholds*). Cada nó das unidades oculta e de saída possuem um limiar associado em adição aos pesos atribuídos. Os nós das unidades ocultas possuem função de ativação não linear enquanto as saídas têm função de ativação linear. Desta forma, cada sinal aplicado aos nós em camadas subsequentes tem a entrada original multiplicada por um peso e associado a um limiar. A função de ativação sigmoideal é uma escolha comum para a função de ativação não linear das camadas escondidas de uma rede MLP. O princípio geral da rede é que os dados presentes na entrada sofram modificações ao passar pelas camadas escondidas de forma que o sinal de saída seja apto a indicar a classe apropriada para cada sinal de entrada (AZAR e EL-SAID, 2013).

O algoritmo mais conhecido e utilizado para o treinamento de redes MLP é o algoritmo *backpropagation*, com paradigma supervisionado. A popularidade deste algoritmo resulta de sua relativa simplicidade de implementação. Na medida em que o conjunto de dados usado

para treinar uma rede MLP seja grande o suficiente para ser representativo do ambiente no qual a rede está inserida, a rede treinada através do algoritmo *backpropagation* desenvolverá a capacidade de generalizar. Esta capacidade permite à rede MLP apresentar um desempenho satisfatório quando é alimentada com dados de teste retirados do mesmo espaço de entrada que os dados de treino, mas não previamente apresentados à rede. Segundo (AZAR e EL-SAID, 2013), as redes MLP apresentam melhor desempenho do que as redes RBF na fase de validação para a classificação de mamografias.

Redes de otimização sequencial mínima (Sequential minimal optimization, SMO) são algoritmos utilizados para resolução de problemas de programação quadrática, desenvolvido por John Platt em 1998 (PLATT, 1998), e é utilizado para o treinamento de Máquinas de Vetores de Suporte (Support Vector Machine, SVM).

Figura 14 - Arquitetura de redes MLP.



Máquinas de vetores de suporte constroem um hiperplano (ou conjunto de hiperplanos em um espaço de alta dimensionalidade) que são utilizados para classificação, separando padrões. Uma boa separação significa um hiperplano que possui a maior distância até o ponto de teste mais próximo de cada classe. De forma geral, quanto maior a margem, menor o erro de generalização do classificador.

O SVM padrão é um classificador linear binário não probabilístico e toma como entrada um conjunto de dados e prediz, para cada entrada dada, qual de duas possíveis classes a entrada faz parte. Dados um conjunto de exemplos de treinamento, cada um marcado como pertencendo a uma das duas classes.

cente a uma de duas categorias, um algoritmo de treinamento do SVM constrói um modelo que atribui novos exemplos a uma categoria ou outra.

SVM é utilizado na determinação do hiperplano para separação dos padrões. Visto que uma grande parte dos problemas não são linearmente separáveis no espaço (Figura 16), utiliza-se o mapeamento do espaço do problema em espaços de maiores dimensões, presumindo-se que a separação dos padrões em espaços de maior dimensionalidade seja uma tarefa possível e mais fácil. O mapeamento realizado pelo SVM é definido pela função kernel $k(x,y)$, de forma que torne o mapeamento adequado para determinado espaço-problema. Os hiperplanos num espaço de alta dimensionalidade são definidos como um conjunto de pontos cujo produto escalar com um vetor nesse espaço seja uma constante.

Figura 15 - Padrões linearmente separáveis.

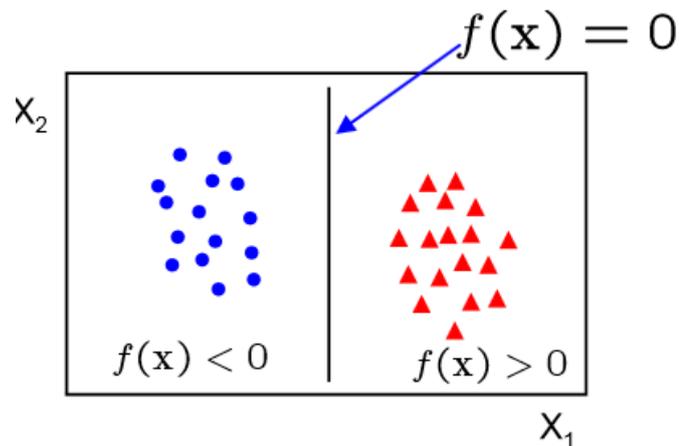
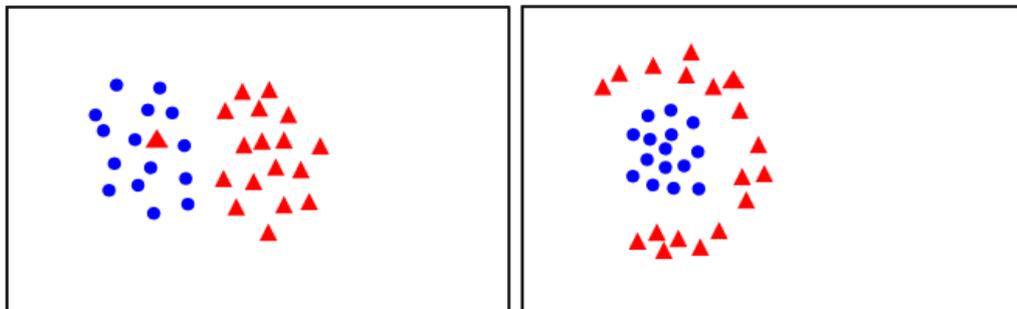


Figura 16 - Dados não linearmente separáveis.



Apesar do SVM padrão ser um classificador linear, é possível sua aplicação para classificação de padrões não lineares utilizando o truque kernel. O algoritmo é semelhante ao padrão,

exceto que nesse caso substituiremos o produto escalar por uma função kernel não linear. Os kernels não lineares mais conhecidos são o polinomial e o gaussiano (RBF).

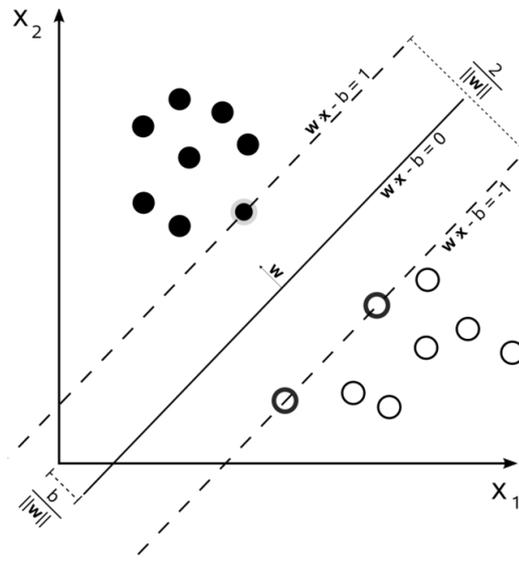
$$\Phi: \mathbb{R}^d \rightarrow \mathbb{R}^n \quad (1)$$

$$K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \quad (2)$$

$$f(x) = W^T \Phi(x) + b \quad (3)$$

Onde Φ é a função de mapeamento do espaço-problema (\mathbb{R}^d) em um espaço de maior dimensionalidade (\mathbb{R}^n), K é a função kernel, w é o peso do vetor e b é o viés determinado no processo de treinamento.

Figura 17 - Margem do hiperplano que separa dois padrões. Amostras localizadas na margem são chamadas de vetores de suporte.



O treinamento de SVM necessita da solução da otimização de complexos problemas de programação quadrática (PQ). O SMO possibilita a partição de grandes problemas PQ em uma série de PQ menores, que são resolvidos analiticamente. Desta forma, a quantidade de memória necessária para o SMO é linear em relação ao conjunto de treinamento, permitindo ao SMO resolver problemas PQ grandes e complexos com a utilização de pouca memória.

2.5 Métricas de desempenho em classificação

A validade de um método de diagnóstico refere-se a quanto, em termos quantitativos, um determinado método é útil para diagnosticar um evento ou predizê-lo. Para isto, comparam-se os resultados do método com os de um padrão. Neste trabalho, estaremos comparando os re-

sultados do sistema CAD com a leitura de exames de mamografia realizadas por um único especialista, que é o método de diagnóstico mais comum e aceito.

O software de mineração de dados utilizado para classificação, Weka 3.6, (HALL et al., 2009) apresenta, ao fim da tarefa de classificação, um relatório de desempenho no qual informa algumas medidas para análise do classificador, como a taxa de verdadeiros positivos (VP), taxa de falso positivo (FP), precisão, revocação (recall), medida F_1 e área sob a curva ROC para cada classe analisada. Além dessas informações, também nos é mostrado a matriz confusão dos resultados da classificação. Dessa forma, possuímos diversas métricas para avaliação dos classificadores.

A taxa de verdadeiros positivos indica a quantidade de amostras que possuem a característica em análise, no nosso caso refere-se às imagens com lesão, que foram corretamente classificadas como “doente”. Já a taxa de falso positivo, indica a quantidade de amostras que possuem a característica em análise, mas que foram classificadas como normais ou não possuidoras de tal característica. A quantidade de instâncias corretamente classificadas é a medida da média da taxa de verdadeiros positivos para as três classes.

A medida de precisão (PREC) combina as taxas de VP e FP, conforme a Equação 4, e representa a acurácia com a qual uma determinada classe foi prevista pelo classificador. Já a taxa de *recall* (REC) indica, do total de amostras com lesões (VP+FN), quantas foram corretamente selecionadas (VP). A medida F_1 , conhecida como medida harmônica, combina a precisão e o recall

$$PREC = \frac{VP}{VP + FP} \quad (4)$$

$$REC = \frac{VP}{VP + FN} \quad (5)$$

$$F_1 = 2 \times \left(\frac{PREC \times REC}{PREC + REC} \right) \quad (6)$$

Uma forma de expressar graficamente a relação entre a sensibilidade e a especificidade é através da construção da curva ROC (Receiver Operating Characteristic curve). Sua construção é feita colocando-se os valores de sensibilidade no eixo das ordenadas e o complemento da especificidade (1-ES), ou seja, a proporção de falsos positivos, no eixo das abscissas. Quanto mais próxima a curva estiver do canto superior esquerdo do gráfico, melhor será o poder discriminatório do teste e quanto mais distante, até o limite da diagonal do gráfico, pior será o seu poder de discriminar doentes e não doentes.

As curvas ROC são muito utilizadas para comparar dois (ou mais) testes diagnósticos para a mesma doença. Nesse caso, o poder discriminatório do teste, ou seja, a acurácia global, pode ser mensurado através do cálculo da área sob a curva ROC. Quanto maior for a área, melhor será o teste diagnóstico (OBUCHWSK, 2003).

Em um teste diagnóstico ideal, com 100% de sensibilidade e 100% de especificidade, a área total seria igual a 1 (um), ou seja, a curva seria coincidente com o lado esquerdo e o topo do gráfico. Neste caso, o teste não possuiria erro, identificando perfeitamente todos os indivíduos sadios e doentes. Para um teste sem nenhuma utilidade, a curva seria uma linha reta em diagonal, partindo do canto inferior esquerdo até o canto superior direito. Neste caso, o teste diagnóstico não possuiria nenhum poder para discriminar doentes de não-doentes.

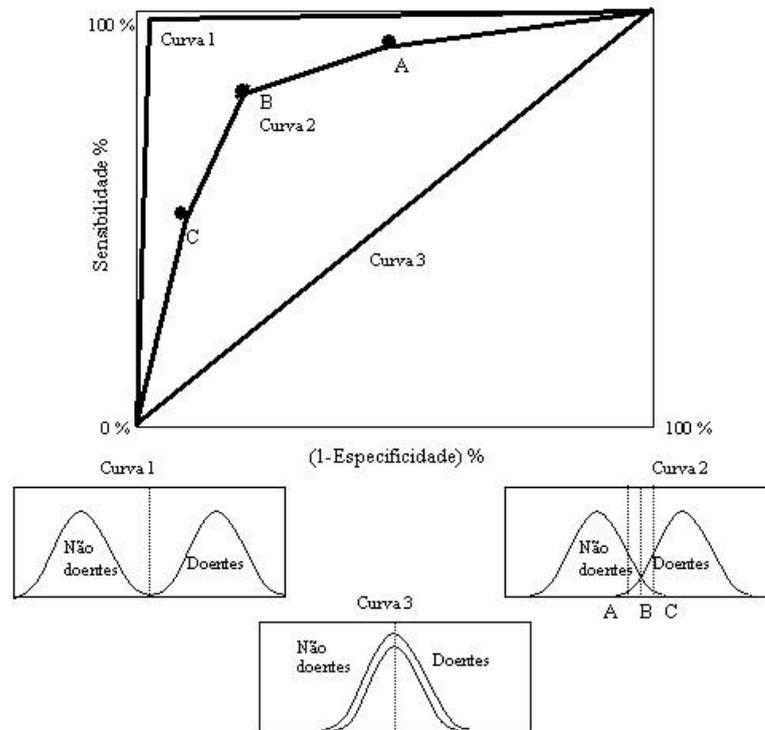
Outra ferramenta utilizada para analisar a qualidade de um método de diagnóstico é o índice *kappa* (κ), que é uma medida de concordância que nos fornece uma ideia do quanto os métodos de classificação se afastam daqueles esperados, fruto do acaso, indicando-nos assim o quão legítimas são as interpretações. O índice de Cohen é uma medida utilizada estritamente para comparações entre duas leituras, enquanto o índice de Fleiss pode ser aplicado para mais de dois observadores (ANTONIO e CRESPI, 2010). O valor de *kappa* pode ser calculado através da matriz confusão.

A matriz confusão ideal é aquela em que apenas a diagonal principal está preenchida, indicando que não houve erro de classificação. No nosso estudo, trabalhamos com três classes (sem lesão, com lesão benigna e com lesão maligna), de forma que temos uma matriz confusão conforme mostrado abaixo.

Tabela 2: Matriz confusão genérica.

Classes	a	b	c	Total
A	aa	ab	ac	A1
B	ba	bb	bc	B1
C	ca	cb	cc	C1
Total	A2	B2	C2	T

Figura 18 - Representação de curvas ROC.



O índice *kappa* relaciona a proporção de vezes em que a classificação dos dados é realizada de forma correta ($P(C)$) com a proporção de vezes em que se espera que os classificadores sejam coerentes entre si ($P(E)$). A equação para o cálculo do índice *kappa* através da matriz confusão é vista abaixo (FLEISS, 1973 apud ANTONIO e CRESPI, 2010):

$$\kappa = \frac{P(C) - P(E)}{1 - P(E)} \quad (7)$$

$$P(C) = \frac{aa + bb + cc}{T} \quad (8)$$

$$P(E) = \frac{(A1 \times A2) + (B1 \times B2) + (C1 \times C2)}{T^2} \quad (9)$$

Os valores de κ variam de “-1” a “1”, sendo “-1” indicativo de total discordância dos dados, “0” representa não haver concordância além do puro acaso, e “1” representa correlação perfeita. Um eventual valor de *kappa* negativo sugere que a concordância encontrada foi menor do aquela esperada por acaso. Em alguns estudos foram possíveis mensurar o valor de *kappa* para mamografia. Lazarus et al. (2006) obtiveram um índice médio de 0,34 (LAZARUS et al., 2006), enquanto Ooms et al. (2007) alcançaram valores médios de *kappa* de 0,74 (OOMS et al., 2007).

Landis e Koch (1977) formularam uma tabela atribuindo uma descrição qualitativa a valores de *kappa*, cuja descrição referente à uma concordância entre observadores foi considerada como substancial para *kappa* entre 0,60 e 0,79, e considerada quase perfeita para valores de *kappa* entre 0,80 e 1,00 (LANDIS e KOCH, 1977).

Capítulo 3

Estudo proposto

Este estudo propõe a utilização de ferramentas que envolvem segmentação, extração de características e classificação de imagens de mamografia a fim de prever a ausência de câncer de mama ou sua presença, distinguindo entre lesões benignas e malignas. Trabalhou-se com três abordagens para segmentação da imagem, uma utilizando mapas SOM e outra utilizando transformada de wavelets com um e dois níveis. Para cada sistema de segmentação foram testados três abordagens para extração de características. Os atributos de textura extraídos pela matriz de co-ocorrência foram utilizados para descrição da imagem, outra proposta utilizou a descrição da imagem pelas suas características de forma, extraídas pelos momentos de Zernike e também foi elaborado um descritor misto, que gera o vetor de características através dos atributos de textura e de forma.

Cada uma dessas nove bases de vetores de atributos geradas foi utilizada como entrada para diferentes redes neurais artificiais (ANN) para fins de classificação. Foram realizados testes com três modelos de ANN diferentes, as redes RBF, SMO com kernel polinomial e RBF e redes MLP, com 1 e 2 camadas ocultas com diferentes quantidades de neurônios nessas camadas. Dessa forma, foi possível testar a eficiência de nove classificadores para cada um dos nove vetores de características gerados para um conjunto de imagens.

Este estudo consiste ainda em verificar a melhora na classificação do conjunto de imagens utilizado quando este é dividido em quatro subconjuntos mediante a característica tecidual e densidade da mama. Ao ser utilizada a base completa, não é levado em consideração o tipo da mama em análise, informação que influencia na extração de atributos e criação do vetor de atributos. Dessa forma, espera-se obter melhores resultados de classificação quando a base de dados está separada conforme o tipo da mama. Mais ainda, será realizada uma análise visando verificar se realmente há uma redução da eficiência na classificação de mamas extremamente densas, devido ao baixo contraste entre tecido patológico e tecido normal.

Esta seção será iniciada pela descrição dos softwares utilizados para geração dos vetores de atributos e classificação. Em seguida serão descritos os experimentos realizados e serão apresentados os principais resultados das classificações, verificando e analisando qual combinação para segmentação, extração de atributos e classificadores são mais apropriados para determinado tipo de mama e qual, de forma geral, apresentou os melhores resultados.

3.1 Softwares utilizados

3.1.1 CONVIRMA

Este aplicativo foi desenvolvido pelo grupo de pesquisa, encabeçado pelo professor Wellington Pinheiro dos Santos e possibilita o processamento dos dados obtidos da base IRMA, que é composto, além das imagens utilizadas, por um arquivo de dados contendo as informações sobre cada imagem, como o tipo de mama, o tipo da lesão e o grau de severidade e a classificação da mesma.

Na tela inicial é possível carregar um arquivo de dados no formato *Comma separated values* (CSV), conforme Figura 19. Em seguida, é aberta a tela que permite a seleção de dados que permitem elaborar a base de dados que iremos trabalhar (Figura 20). Foram geradas bases com os quatro tipos de mama e uma base contendo todas as imagens, nos três estágios de lesão (sem lesão, com lesão maligna e benigna) do tipo massa circunscrita e espiculada.

Figura 19 - Tela inicial do aplicativo ConvIRMA.



Uma vez geradas as bases de dados que serão utilizadas, passamos para etapa de parametrização da configuração de processamento dos dados (Figura 21). Apesar das diversas operações disponíveis no aplicativo, trabalhamos apenas utilizando pré-processamento por mapas SOM e transformada de wavelets e atributos gerados pelos descritores de Haralick e pelo momento de Zernike.

Ao selecionar as configurações desejadas o aplicativo realiza a segmentação pelo modelo selecionado e gera os atributos conforme o descritor. Após o processamento e balanceamento dos dados o arquivo pode ser salvo no formato *Attribute-Relation File Format* (ARFF), que é o tipo de arquivo lido pelo Weka.

Figura 20 - Seleção de informação para geração da base de dados.

Seleção do tipo de mama

Densidade:

- I - Adiposa
- II - Fibroglandular
- III - Densa
- IV - Extremamente densa

Estágio do lesão:

- BIRADS 1 - Inexistente
- BIRADS 2 - Benigna
- BIRADS 5 - Maligna

Tipo de lesão:

- Calcificação
- Massa circunscrita
- Massa espiculada
- Outra massa

OK Cancelar

Figura 21 - Seleção das configurações de processamento da base de dados.

Seleção de atributos

Pré-processamento:

- Nenhuma operação
- Equalização de histograma
- Transformação quadrática
- Transformação de raiz quadrada
- Método de Otsu
- Mapa SOM de Kohonen
- Transformada de Wavelets
- Transformada de Wavelets morfológica

Pós-processamento:

- Balanceamento de classes

Atributos:

- Histograma de Gradientes Orientados (HOG)
- Padrões Locais Binários (LBP)
- Matriz de Co-Ocorrência (GLCM)
- Momentos de Hu da GLCM
- Momentos de Hu
- Momentos de Haralick
- Histograma
- Espectro morfológico de padrões
- Momentos de Zernike

OK Salvar Salvar ARFF Cancelar

3.1.2 WEKA

O pacote de software Weka (Waikato Environment for Knowledge Analysis) é uma coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados, desenvol-

vido pela Universidade de Waikato, Nova Zelândia. O software contém ferramentas para pré-processamento, classificação, regressão visualização de dados. Weka é um software aberto licenciado pela GNU General Public Lincese (HALLet al., 2009).

Figura 22 - Tela inicial do Weka 3.6.11.



O software é escrito em linguagem Java™ e lê dados no formato Attribute-Relation File Format (ARFF). A interface permite interação com dados produzindo resultados visuais, como tabelas e curvas, de forma que a tela inicial permite a seleção de quatro modos diferentes de utilização do Weka. Para a tarefa de classificação, utilizamos a opção “Explorer”.

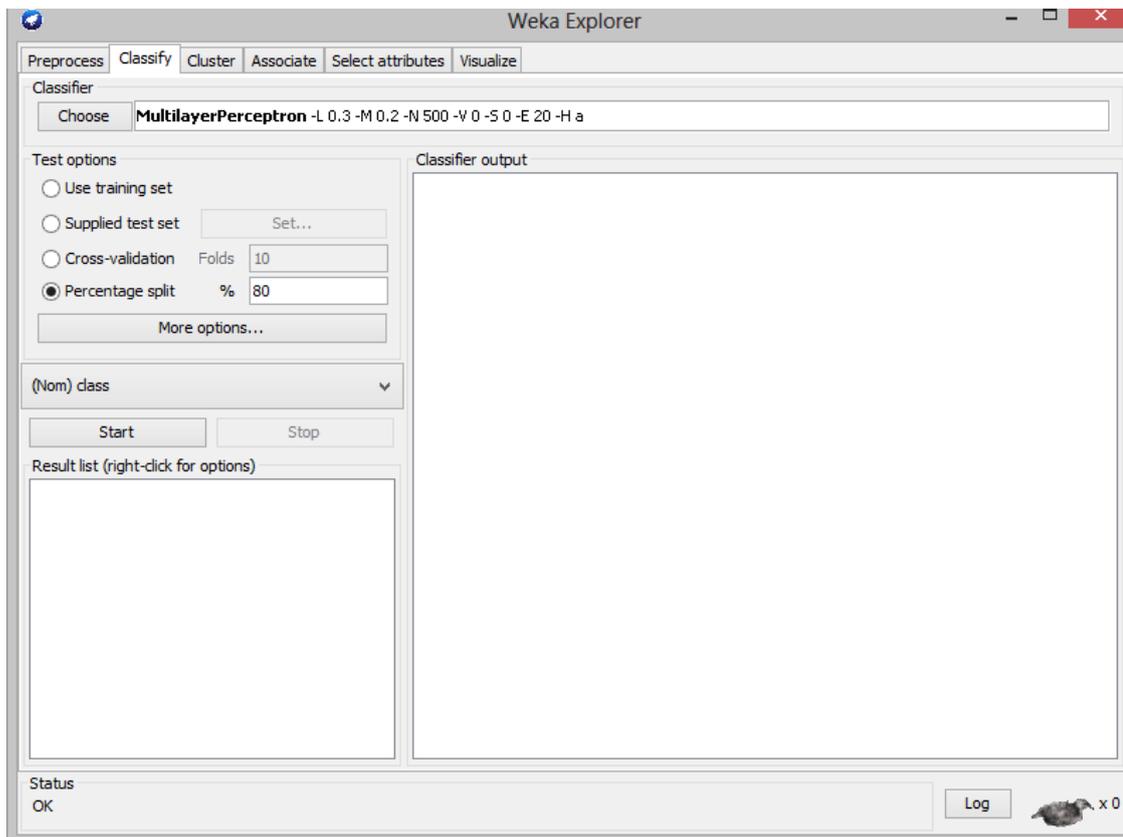
Para iniciar a análise dos dados, é necessário inicializar o arquivo “.arff” que será utilizado. Na opção de trabalho “Explorer”, é possível realizar um pré-processamento dos dados, selecionando filtros e verificando a relação dos atributos e das instâncias contidas na base de dados. Além disso, o Weka contém opções para classificação dos dados, clusters e visualização dos dados e resultados.

Após a seleção do arquivo para análise e feito o pré-processamento, é realizada a parametrização do classificador, determinando a forma de teste, os atributos que desejamos prever e o classificador que será utilizado. Para avaliação dos resultados da classificação é necessário observar se os dados de treinamento são diferentes dos dados utilizados no teste do classificador.

Visto que o conjunto de dados é único, foi realizada a partição do conjunto em dados de treinamento e dados de teste. Desta forma, foi montado um conjunto de treinamento do classificador com 80% dos dados, de forma que a avaliação do poder de predição do classificador foi testada em 20% dos dados. A seleção de ambos os conjuntos é realizada de forma randômica, com valor da semente igual a 1.

Uma vez selecionado os parâmetros para classificação, o processo de aprendizado é iniciado. O Weka elabora o modelo do classificador, treinando-o e posteriormente testa-o, fornecendo como resultado um sumário estatístico da classificação.

Figura 23 - Aba de classificação do Weka.



Foram testados três classificadores da biblioteca do Weka, as redes RBF, SMO e MLP. O classificador RBF implementa o algoritmo de classificação por função de base radial, treinada de forma completamente supervisionada utilizando a classe de otimização do Weka por mínimo erro quadrado com o método BFGS (Broyden–Fletcher–Goldfarb–Shanno).

O classificador SMO implementa o algoritmo de otimização sequencial mínima de John Platt para o treinamento do classificador de vetor de suporte. De forma geral, o algoritmo substitui todos os valores ausentes e transforma atributos nominais em binários. Também é realizada a normalização de todos os atributos.

O classificador MLP utiliza *backpropagation* para classificação das instâncias, com todos os nós sendo sigmóides. Dentre os diversos parâmetros ajustáveis, temos o número de camadas escondidas. No entanto, maior quantidade de camadas implica em mais tempo de proces-

samento. O algoritmo possui meios de redução do tempo de processamento, como a utilização de atualização utilizando gradiente conjugado descendente ao invés de BFGS. Além disso, para valores de delta dentro do intervalo de tolerância especificado pelo usuário para o passo do *backpropagation*, o gradiente não é atualizado para aquela instância, o que auxilia na redução do tempo de processamento.

3.2 Experimentos

Conforme descrito na seção 2.3, utilizamos apenas 1.416 amostras da base de dados IRMA, compostas por imagens de mamas sem lesão, com lesões benignas e malignas do tipo massa espiculada, circunscrita. Separamos a base IRMA em 4 subgrupos conforme a composição tecidual da mama: Adiposa (I), Fibroglandular (II), Densa (III) e Predominantemente densas (IV).

Cada classificador foi aplicado a um subgrupo da base de IRMA, de forma que a classificação é feita levando em consideração o tipo de tecido que constitui a mama. Dessa forma, para comparar os classificadores, foi analisado como os mesmos se comportam classificando cada grupo.

Visto que o quantitativo de imagens em cada subgrupo é diferente, foi realizado um balanceamento das classes para que todos os grupos de teste possuíssem 233 amostras. O balanceamento da rede foi realizado a partir da criação de novos vetores de atributos, aplicando operadores de cruzamento e mutação a um vetor existente utilizando o método de otimização por evolução diferencial em sua formulação clássica. O algoritmo é iniciado com uma população inicial e a geração de novas soluções é feita através de um operador de mutação que realiza combinação linear entre três soluções da população e um operador de cruzamento, que mistura as coordenadas do vetor gerado pelo operador de mutação e uma quarta solução da população, resultando em um novo vetor de atributos. Dessa forma, criando novos vetores de atributos baseado nos vetores diretamente extraídos das imagens, foi possível igualar a quantidade de amostras para cada classe.

Após a organização da base de dados, seguiu-se as demais etapas de um sistema CAD. Realizou-se a segmentação das imagens utilizando dois métodos, os Mapas de Kohonen (SOM) e transformada de Wavelets (WA) com um e dois níveis de decomposição. As redes SOM utilizadas permitem a segmentação da imagem em 15 classes, de forma que o aprendizado ocorreu fixando-se o número de iterações para 50 e a taxa de aprendizagem para 0,1.

A aproximação da decomposição de imagem em série de wavelets foi realizada usando filtros lineares passa baixas e passa altas de tamanho 3×3 . Os termos da série foram então armazenados como bandas em uma imagem multiespectral, de forma que a última banda corresponde à última aproximação. As demais correspondem às imagens de detalhes. Para transformação com um nível, obtemos quatro imagens referentes aos detalhes verticais, detalhes horizontais, detalhes diagonais e ao coeficiente de aproximação. No caso da decomposição em dois níveis, são geradas sete imagens. No fim da etapa de segmentação a imagem é redimensionada, ajustando a borda para minimizar a quantidade de pixels escuros, de forma que a imagem contenha a maior quantidade de informações relevantes para classificação.

Após a segmentação das imagens, aplicou-se a etapa de extração de atributos em três combinações: extração dos atributos de textura das imagens utilizando o descritor de Haralick, utilização dos momentos de Zernike (ZE) para extração de atributos de forma e a terceira abordagem utiliza uma combinação desses dois extratores, elaborando uma base que utilizará as características de forma e de textura (HA-ZE) em conjunto para formação do vetor de características.

O teste utilizando diferentes características para descrever a imagem possibilita a comparação entre as abordagens de utilização de características de forma e de textura para descrição e classificação da imagem, permitindo analisar igualmente a proposta de que um vetor com mais informações (vetor único formado por atributos de textura e de forma) descreve melhor a imagem, melhorando o desempenho do classificador.

Abordou-se a descrição das características de textura extraíndo 13 atributos utilizando o descritor de Haralick, enquanto a abordagem da análise da forma e contorno foi obtida extraíndo 32 atributos de forma através da aplicação dos momentos de Zernike, de forma que são calculados os termos complexos do polinômio para, em seguida, serem calculados os atributos da imagem. Cada abordagem, individualmente e a combinação desses atributos, foi utilizada para alimentar os classificadores baseados em redes neurais artificiais.

O descritor de Haralick calcula os atributos de textura para uma determinada inclinação da matriz de co-ocorrência de níveis de cinza (GLCM), que é uma matriz simétrica montada considerando uma vizinhança quadrada de dimensão 3×3 . Os atributos de Haralick foram calculados para quatro inclinações (0° , 45° , 90° e 135°), desta forma, o vetor de atributos é composto por 52 instâncias. Os 13 atributos extraídos utilizando o descritor de Haralick foram: segundo momento angular, correlação, variância, momento diferencial inverso, entropia, con-

traste, variância da diferença, entropia da diferença, informação da correlação 1, informação da correlação 2, media da soma, entropia da soma, variância da soma.

Cada abordagem para segmentação e extração de características influencia na quantidade de instâncias geradas para classificação, pois a utilização de mapas SOM gera uma matriz simétrica, que implica na duplicação na quantidade de atributos extraídos, enquanto a utilização da transformada de wavelets com um e dois níveis modifica a quantidade de imagens de onde são extraídas as características, sendo quatro e sete imagens, respectivamente. Dessa forma, elaborou-se a Tabela 4 com a quantidade de instâncias geradas para cada abordagem, com “SOM” representando a utilização de mapas auto-organizáveis, “WA”, a transformada de wavelets com o índice informando a quantidade de níveis, “ZE” fazendo referência aos momentos de Zernike e “HÁ” ao descritor de Haralick.

Tabela 3: Quantidade de instâncias por abordagem.

	HA	ZE	HA-ZE
SOM	104	64	168
WA-1	208	128	336
WA-2	364	224	588

A etapa de segmentação e extração de atributos foi realizada utilizando o aplicativo Con-vIRMA. Em seguida, para etapa de classificação, foi utilizado o software de mineração de dados, Weka. Realizamos teste com três tipos de classificadores, as redes RBF, MLP e SMO. Utilizamos os parâmetros das redes conforme disponibilizado na biblioteca do software Weka 3.6. Para todos os classificadores, montamos o grupo de treinamento aleatório com 80% das amostras, deixando os demais 20% para o teste do classificador.

Realizaram-se testes para redes MLP com uma única camada e com duas camadas escondidas. Além disso, testamos diferentes quantidades de neurônios para cada camada escondida: “a” neurônios, correspondentes à metade da soma da quantidade de atributos com a quantidade de classes, “i” neurônios referentes à quantidade de atributos e “o” neurônios relativo à quantidade de classes. Para a rede SMO, utilizamos dois kernels diferentes, o polinomial e o kernel RBF.

Para cada teste realizado, foram armazenados os valores das métricas de desempenho disponibilizadas pelo Weka, como o tempo para construção do modelo do classificador, taxa de verdadeiros positivos para cada classe e seu valor médio (porcentagem das instâncias cor-

retamente classificadas), precisão, índice *kappa* e matriz confusão. Essas métricas foram utilizadas para a comparação do desempenho entre os classificadores a fim de verificar qual rede neural artificial melhor se aplica à classificação de determinado tipo de mama.

3.2.1 Experimento 1: Análise da textura

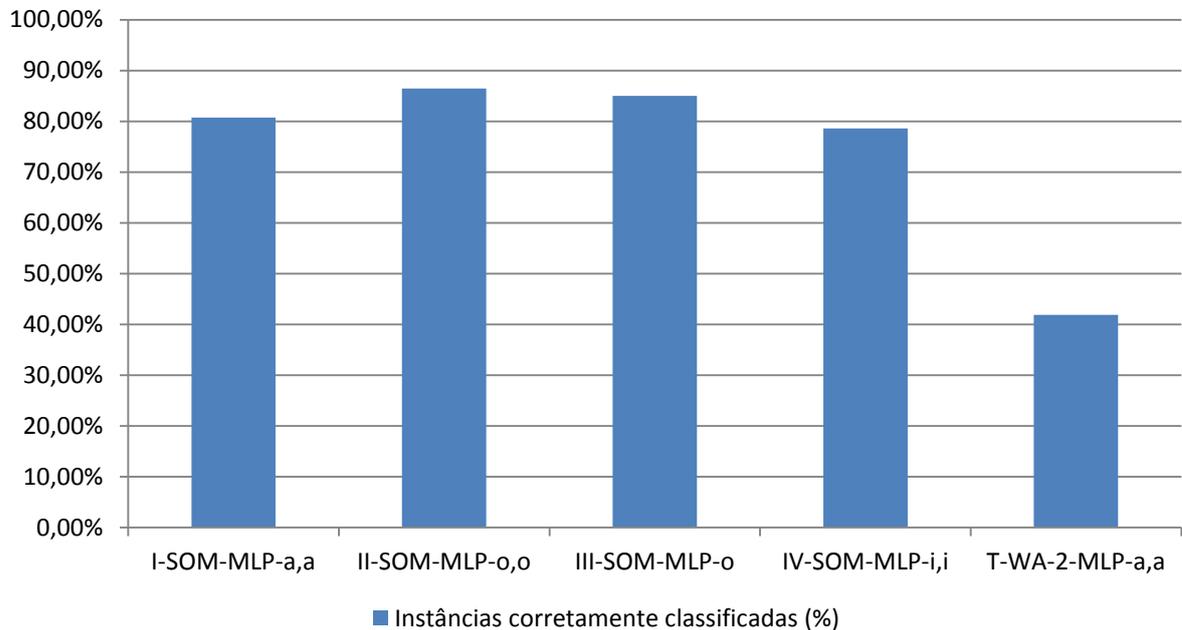
O primeiro experimento analisou a classificação das imagens de mamografia no caso em que as mesmas são descritas por suas características de textura. Foram realizados 27 experimentos para cada tipo de mama, que variaram em relação à ferramenta utilizada para segmentação e aos classificadores aplicados. As bases para cada tipo de mama possuem 699 instâncias a serem classificadas (a base de dados que contém todos os tipos de mama possui 4 vezes o número de instâncias que as bases individualmente, ou seja, 2.796 instâncias).

Conforme o Figura 24, para mamas adiposas (I) o classificador por rede neural artificial do tipo perceptron de múltiplas camadas com duas camadas escondidas e 54 neurônios em cada uma dessas camadas (MLP-a,a) classificou corretamente 80,71% das instâncias. Para mamas fibroglandulares (II), o melhor classificador também foi o MLP, nesse caso com 3 neurônios em cada uma das duas camadas ocultas (MLP-o,o), resultando em 86,43% de instâncias corretamente classificadas.

No caso de mamas pouco densas (III), o classificador com melhor taxa de classificação foram as redes MLP com apenas uma camada oculta com 3 neurônios (MLP-o), acertando 85% da base. As redes MLP também demonstraram maior satisfação na classificação da base de mamas extremamente densas (IV), com 78,57% de acertos para rede com duas camadas ocultas com 105 neurônios em cada camada (MLP-i,i).

De forma geral, foi possível verificar que a segmentação utilizando Mapas Auto Organizáveis (SOM) possibilitou que os atributos extraídos utilizando o descritor de Haralick obtivessem melhor desempenho na classificação. Também foi possível verificar que, de forma geral para cada tipo de mama, os classificadores por redes neurais artificiais do tipo perceptron de múltiplas camadas (MLP) mostram-se como as melhores redes para classificação de vetores de características de textura.

Figura 24 - Melhores configurações do sistema CAD para classificação de atributos de textura.



3.2.2 Experimento 2: Análise da forma

Neste segundo experimento foram gerados vetores de atributos de forma através dos momentos de Zernike, compostos por 32 características. Semelhante ao experimento anterior, foram testadas 27 combinações de configuração do sistema CAD. Visto que a quantidade de instâncias geradas para classificação dos atributos de forma são menores que no experimento anterior, o tempo de processamento para extração das características utilizando o ConvIRMA e o tempo para construção do modelo e teste do classificador pelo Weka foi menor.

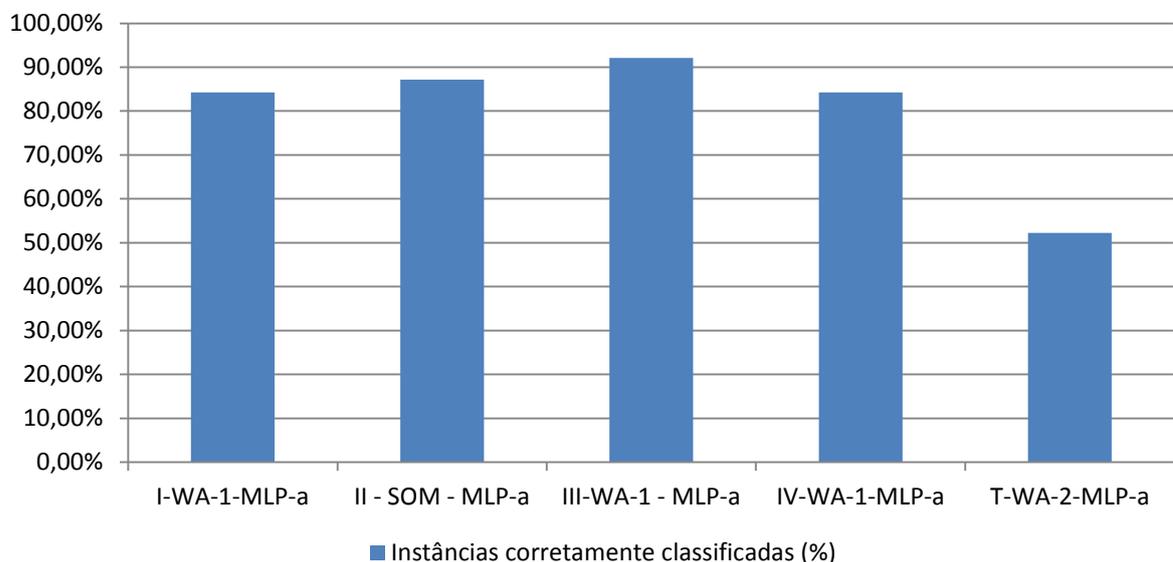
Analisando o Figura 25 é possível verificar que para a classificação de mamas adiposas a configuração que obteve maior percentual de lesões corretamente classificadas (84,29%) foi a partir da utilização de transformada de wavelets com 1 nível e classificador por redes MLP com uma camada oculta contendo 66 neurônios. Esta configuração se mostrou a mais eficiente também para a classificação de mamas pouco densas, com 92,14% de acertos e para mamas muito densas, alcançando para este último um percentual de dados corretamente classificados igual a 84,29%.

Para classificação de mamas fibroglandulares a melhor configuração dentre as testadas é composta por segmentação utilizando mapas SOM e classificação utilizando redes MLP com uma camada escondidas contendo 64 neurônios, possibilitando um percentual de 87,14% instâncias corretamente classificadas. Ao serem utilizadas as 1.416 imagens, desconsiderando o

tecido predominante e densidade de mama, foi obtido uma taxa de 52,24% das imagens corretamente classificadas utilizando a transformada de wavelets com dois níveis e redes neurais MLP com uma camada oculta composta por 244 neurônios.

Em comparação com o percentual de instâncias corretamente classificadas utilizando descritores de textura, obteve-se uma melhora na classificação das cinco bases testadas. Dessa forma, é possível concluir que, para a base de dados IRMA, a descrição da imagem pelas suas características de forma se mostraram mais úteis para alimentação do classificador, possibilitando uma maior quantidade de dados corretamente classificados.

Figura 25 - Percentual de classificações corretas de lesões utilizando descritores de forma.



3.2.3 Experimento 3: Análise da forma e da textura

Neste experimento foi proposta a classificação das imagens de mamografia através da análise de forma e de textura. Foram gerados vetores de características com 45 componentes, os 13 de textura e os 32 de forma, conforme visto anteriormente. Da mesma forma como os dois experimentos anteriores, foram utilizadas segmentações por mapas SOM, gerando 168 instâncias para classificação, e por transformada de wavelets, com um e dois níveis de decomposição, com 336 e 588 instâncias, respectivamente.

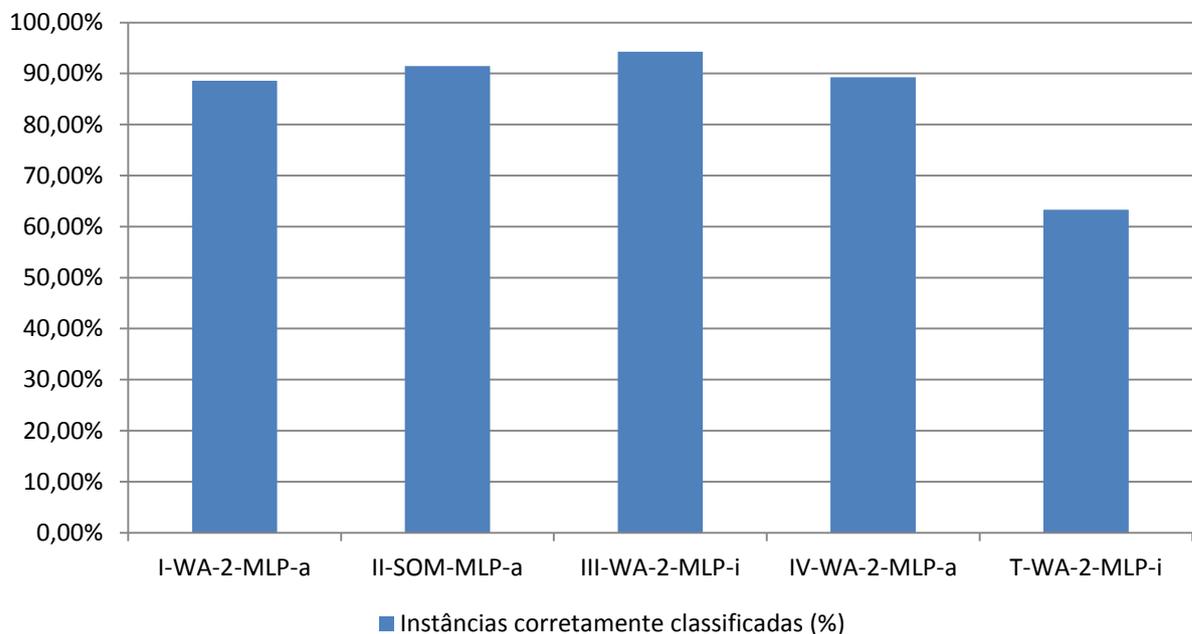
Semelhante ao experimento 2, melhores resultados de classificação foram obtidos para mapas adiposas, pouco e muito densas e para a base de dados completa, utilizando segmentação com transformada de wavelets com dois níveis de decomposição. Para mamas adiposas obteve-se 88,57% de instâncias corretamente classificadas utilizando redes MLP com uma

camada oculta com 296 neurônios como classificador (I-WA-2-HA-ZE-MLP-a). Para esta mesma configuração, obteve-se 89,29% de acertos do classificador para mamas muito densas (IV-WA-2-HA-ZE-MLP-a).

Para mamas pouco densas, a rede MLP também foi a que apresentou melhor resultado de classificação. Neste caso, com 588 neurônios na camada oculta foi possível atingir 94,29% de dados corretamente classificados (III-WA-2-HA-ZE-MLP-i). Para mamas fibroglandulares o melhor resultado da classificação das lesões foi realizado pela segmentação com mapas auto organizáveis de Kohonen e redes MLP com 292 neurônios na camada oculta, atingindo 91,43% de dados corretamente classificados (II-SOM-HA-ZE-MLP-a).

Mesmo durante os testes em que foi utilizada toda a base de dados, desconsiderando o tipo de mama e sua composição tecidual e densidade, a rede MLP com 588 neurônios na camada escondida resultou numa taxa razoável de acertos, com 63,33% de instâncias corretamente classificadas.

Figura 26 - Percentual de classificações corretas de lesões utilizando descritores de forma e de textura.



É possível verificar que houve uma melhora no percentual de classificações corretas utilizando vetores de características contendo atributos de textura e de forma para todas as 5 bases testadas. Percebe-se, então, que a combinação de informações sobre a textura com a forma das regiões de interesse segmentadas possibilitou a geração de vetores de características mais

completos, que aumentaram o poder discriminatório das redes neurais utilizadas como classificadores.

3.3 Resultados

A composição tecidual da mama influencia no contraste, na leitura e interpretação da imagem de mamografia, o que pode vir, conseqüentemente, a dificultar o diagnóstico. Vimos que mamas de mulheres jovens possuem em sua composição mais tecido do tipo glandular, o que as caracteriza como mamas do tipo muito densas. Como as lesões e calcificações tem densidade semelhante ao tecido glandular, há pouca diferença nos padrões de absorção de raios X, de forma que a imagem produzida através da mamografia em mamas muito densas mostram pouco contraste com as lesões.

Em contrapartida, mulheres mais velhas, em que a substituição do parênquima mamário é superior a 50%, possuem mamas predominantemente adiposas. Dessa forma, mamografias de mulheres com mamas predominante adiposas apresentam melhor contraste com as lesões e calcificações.

Visto que a composição tecidual da mama afeta a imagem a ser diagnosticada, foi sugerido um sistema de classificação que considere a composição da mama em análise, informação que pode ser indiretamente obtida através da associação com a idade da paciente, dado este que é clinicamente acessível. Desta forma, abordou-se o problema de classificação de imagens de mamografia por tipo de mama.

Foi realizado um total de 405 experimentos, variando entre três abordagens de segmentação de imagem (mapas auto-organizáveis, transformadas de wavelets com um e dois níveis), três tipos de vetores de características e nove configurações de classificadores, aplicados a cinco bases (uma para cada um dos quatro tipos de mama e uma base contendo as imagens de todos os tipos de mama). Pode-se, então, realizar uma análise da melhor configuração testada para classificação de determinado tipo de mama, analisar qual tipo de segmentação melhor combina com determinado vetor de característica, bem como verificar o classificador com melhores resultados gerais.

As análises realizadas foram feitas baseadas na estatística disponibilizada pelo Weka quanto à quantidade de instâncias corretamente classificadas, índice kappa, área sob a curva ROC e demais métricas geradas pelo software. O índice *kappa* permite avaliar a concordância dos dados e o resultado de forma geral, de forma que *kappa* superior a 0,60 demonstra boa concordância entre os classificadores (LANDIS e KOCH, 1977). No caso analisado, compara-

se a classificação pelo sistema CAD implementado com a classificação do especialista disponibilizada juntamente com a imagem na base IRMA.

Foram também utilizadas a porcentagem de instâncias corretamente classificadas para verificar o desempenho do classificador e a área sob a curva ROC para analisar o poder discriminatório e a acurácia global dos testes, de forma que maior a área, melhor o teste de diagnóstico.

A Tabela 4 relaciona as três configurações do sistema CAD que obtiveram melhor percentual de instâncias corretamente classificadas para cada tipo de mama, apresentando também os valores dos respectivos índices *kappa* e do tempo para construção do modelo do classificador. Os demais resultados para todos os testes realizados podem ser encontrados no Apêndice A.

Tabela 4: Configuração dos melhores classificadores.

Tipo de mama	Configuração do Sistema	Corretamente classificado (%)	Índice kappa	Tempo (segundos)
Adiposa	WA-2-HA-ZE-MLP-a	88,57%	0,8283	2.217
	SOM-HA-ZE-MLP-a	87,86%	0,8166	130
	WA-2-HA-ZE-MLP-o	87,14%	0,8064	23
Fibroglandular	SOM-HA-ZE-MLP-a	91,43%	0,8703	135
	SOM-HA-ZE-MLP-i	90,71%	0,8593	251
	WA-2-HA-ZE-MLP-a	89,29%	0,8387	2.291
Densa	WA-2-HA-ZE-MLP-i	94,29%	0,9139	6.163
	WA-2-HA-ZE-MLP-a	93,57%	0,9031	1.182
	WA-2-HA-ZE-MLP-o	93,57%	0,9032	25
Extremamente densa	WA-2-HA-ZE-MLP-a	89,29%	0,8387	1.886
	WA-1-HA-ZE-MLP-a	87,86%	0,8169	497
	WA-1-HA-ZE-MLP-i	87,86%	0,8169	1.307

É possível perceber que os maiores valores para o índice kappa correspondem igualmente ao maior percentual de instâncias corretamente classificadas, de forma que para todos os casos obtivemos um valor de kappa acima de 0,8, relevando a coerência do classificador. O tempo da construção do modelo está associado ao tipo de segmentação e ao classificador, de

forma que quanto mais instâncias alimentam o classificador, maior o tempo de processamento. Da mesma forma que maior a quantidade de camadas ocultas e neurônios nessas camadas, no caso de redes MLP, maior também o tempo de processamento.

As Figuras 24, 25 e 26, conjuntamente com a Tabela 4 nos permite observar que, de forma geral para todos os tipos de mama, a configuração que obteve melhores resultados foram aquelas que utilizaram a transformada de wavelets, com maior quantidade de atributos na entrada da rede neural e classificador com redes MLP. Além disso, verificamos que as mamas com densidades intermediárias, ou seja, do tipo fibroglandulares e pouco densas, obtiveram resultados na classificação das lesões bastante satisfatórias, acima de 90%.

No entanto, o percentual de dados corretamente classificados apenas nos fornece uma informação geral sobre a classificação, visto que trabalhamos tentando prever quais imagens continham mamas sem lesão e quais possuíam lesão, devendo esta ser classificada como benigna ou maligna. Para verificar o desempenho da classificação podemos analisar a matriz confusão e a área sob a curva ROC.

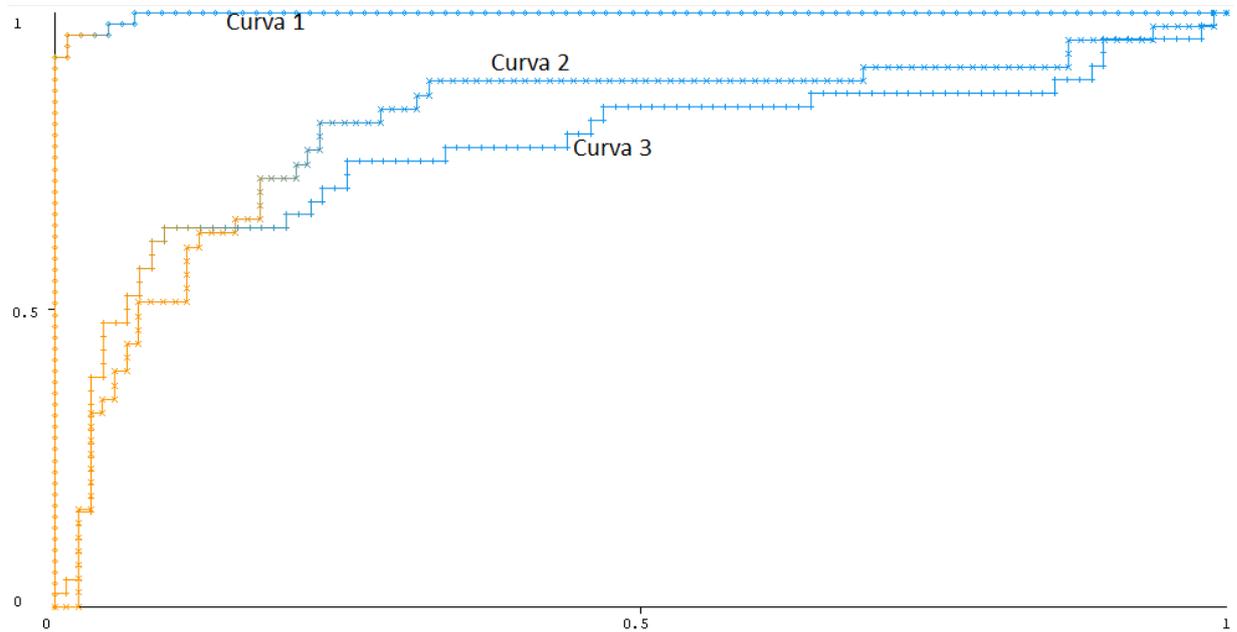
Para mamas adiposas o sistema com melhor quantidade de dados corretamente classificados utilizou transformada de wavelets com 2 níveis de decomposição, vetor de características composto por 588 instâncias, classificados por redes MLP com 296 neurônios na camada escondida (WA-2-HA-ZE-MLP-a). Pela matriz confusão, verificamos que a rede neural classificou incorretamente 9,4% das mamas normais como possuindo lesões benignas, 14% das imagens com lesão benigna ele atribuiu incorretamente malignidade e 9% das mamas com lesão maligna foram classificadas como tumores benignos.

Ao serem analisadas a curva ROC e a área sob a curva para cada uma das três classes possíveis, é verificado que a configuração montada para classificação de mamas adiposas possui uma área sob a curva ROC bem próximo da área máxima que pode ser alcançada, ou seja, 1. Ainda, é possível perceber que a classificação das mamas normais foi mais bem sucedida que as outras duas classes possíveis.

Tabela 5: Matriz confusão e área sob a curva ROC para classificação I-WA-2-HA-ZE-MLP-a.

Normal	Benigno	Maligno		Área sob curva ROC
48	5	0	Normal	0,992
1	36	6	Benigno	0,947
0	4	40	Maligno	0,973

Figura 27 - Curva ROC para mamas adiposas. Curva 1: mamas sem lesão. Curva 2: Lesões benignas. Curva 3: Lesões malignas.

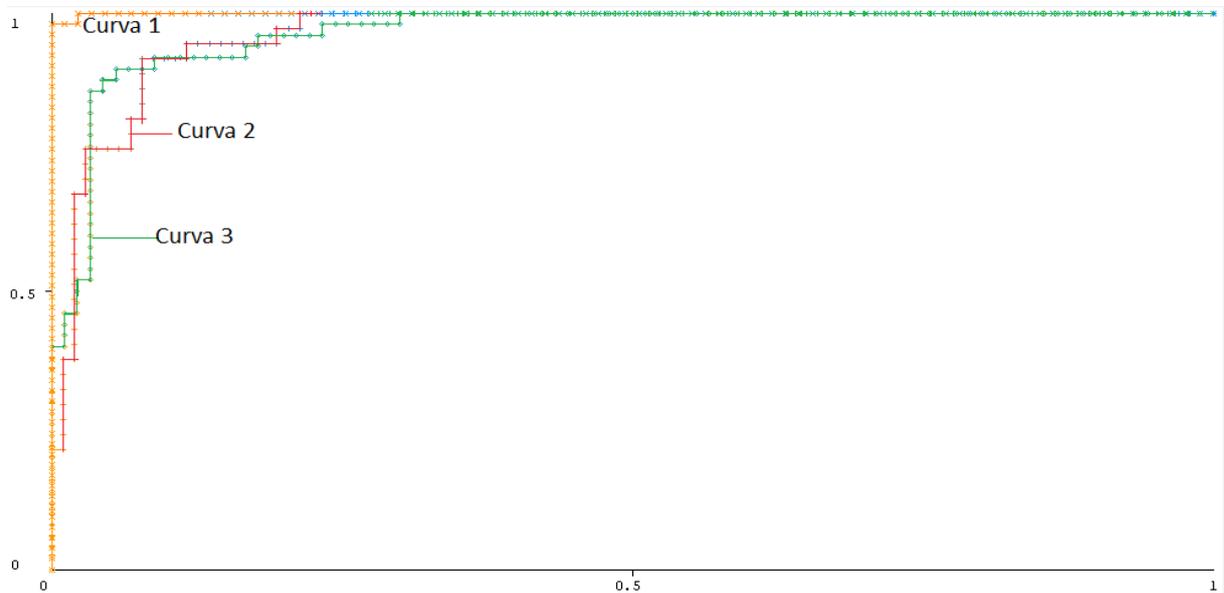


Pela Tabela 5, nota-se que a melhor classificação de tumores para mamas fibroglandulares ocorreu para a configuração com segmentação utilizando mapas auto-organizáveis de Kohonen, classificação com redes MLP, com 296 neurônios na camada escondida, do vetor de características composto por 588 atributos de forma e textura (II-SOM-HA-ZE-MLP-a). Pela Figura 28 e pela Tabela 6 vê-se uma classificação quase perfeita para as imagens sem lesão. Na diferenciação entre lesão maligna e benigna, no entanto, percebe-se que houve uma pequena confusão da rede neural, classificando incorretamente como lesão maligna 5 imagens que continham tumores benignos. Da mesma forma, houve a falha na classificação de 6, das 50 imagens malignas, que foram incorretamente classificadas como lesões benignas.

Tabela 6: Matriz confusão e área sob curva ROC para classificação II-SOM-HA-ZE-MLP-a.

Normal	Benigno	Maligno		Área sob curva ROC
52	1	0	Normal	1,00
0	32	5	Benigno	0,975
0	6	44	Maligno	0,981

Figura 28 - Curva ROC para mamas fibroglandulares. Curva 1: mamas sem lesão. Curva 2: Lesões benignas. Curva 3: Lesões malignas.

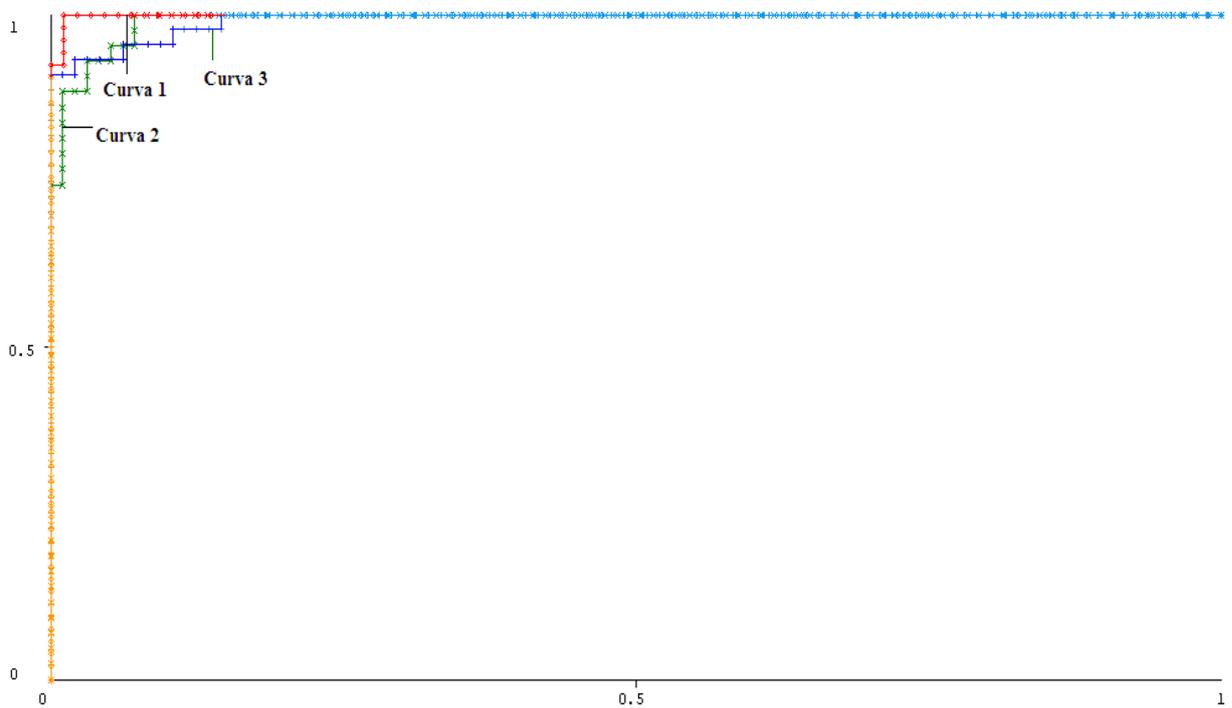


Para mamas pouco densas o percentual de classificações corretas alcançado pela melhor configuração (III-WA-2-HA-ZE-MLP-i) foi de 94,29%, com índice *kappa* de 0,9139, o que demonstra boa concordância do classificador com o padrão. Observando a Figura 29 e a Tabela 7, referente à curva ROC para classificação de mamas pouco densas, verifica-se, novamente houve maior precisão para a classificação de imagens de mamas normais, apesar de não ter sido de 100% como ocorreu para mamas fibroglandulares (Tabela 6).

Tabela 7: Matriz confusão e área sob curva ROC para classificação III-WA-2-HA-ZE-MLP-i.

Normal	Benigno	Maligno		Área sob curva ROC
51	1	1	Normal	0,999
0	40	3	Benigno	0,994
1	2	41	Maligno	0,992

Figura 29 - Curva ROC para mamas pouco densas. Curva 1: mamas sem lesão. Curva 2: Lesões benignas. Curva 3: Lesões malignas.

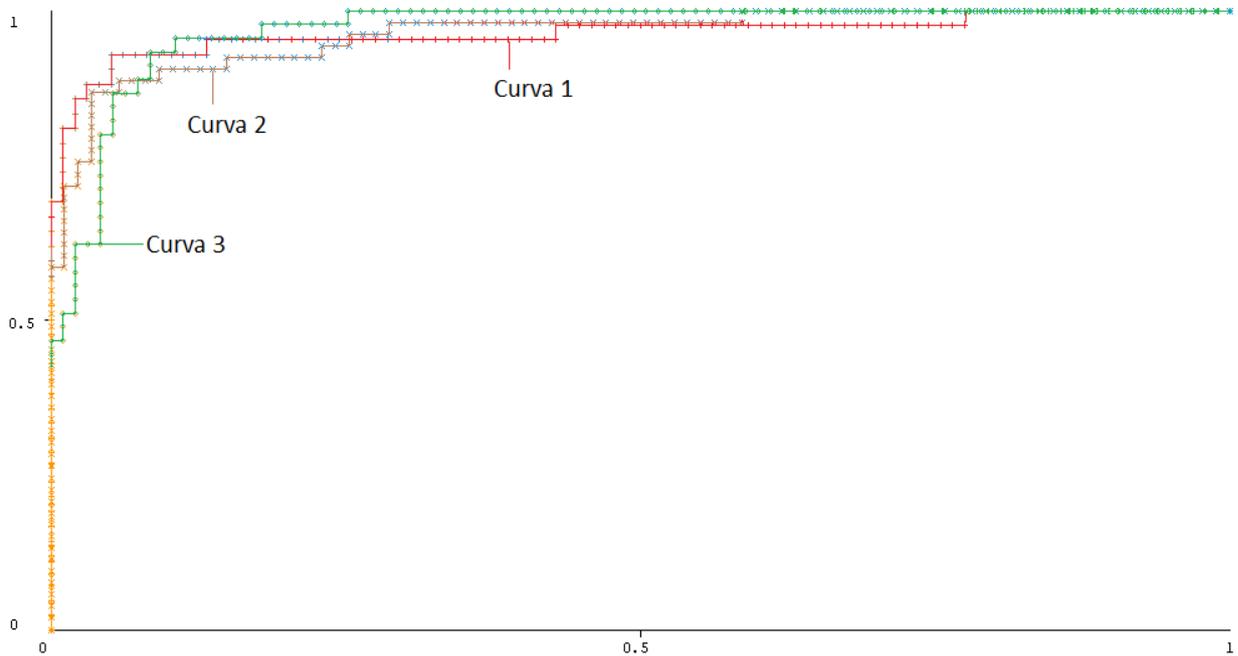


Conforme visto, a configuração que resultou na melhor classificação de mamas muito densas foi utilizando a transformada de wavelets com 2níveis e redes MLP com uma camada oculta (IV-WA-2-HA-ZE-MLP-a), retornando 89,29% corretamente classificados. Para este tipo de mama percebeu-se maior dificuldade da rede neural em classificar mamas sem lesão, atribuindo erroneamente 6 imagens de mamas normais à classe de lesões malignas. Notou-se também um número maior do que para os demais tipos de mama de lesões benignas classificadas incorretamente como normais.

Tabela 8: Matriz confusão e área sob curva ROC para classificação IV-WA-2-HA-ZE-MLP-a.

Normal	Benigno	Maligno		Área sob curva ROC
46	0	7	Normal	0,962
4	37	1	Benigno	0,962
0	3	42	Maligno	0,967

Figura 30 - Curva ROC para mamas muito densas. Curva 1: mamas sem lesão. Curva 2: Lesões benignas. Curva 3: Lesões malignas.



As curvas ROC de cada tipo de mama possibilitam visualizar e diferenciar o poder de classificação da rede neural conforme o tipo de lesão. Para os quatro tipos de mama analisados, verificamos que todos os classificadores possuem maior facilidade na distinção entre as imagens sem lesão das imagens com lesão. No entanto, para diferenciar lesões benignas de malignas há uma menor precisão. Estes resultados estavam dentro do esperado, uma vez que um *threshold* maior pode ser utilizado para separar imagens normais, sem objetos, de imagens com tumores. No entanto, o ajuste fino na diferenciação da forma e da textura da lesão para classificação de malignidade requer um limiar para diferenciação menor, implicando numa maior taxa de erro do classificador.

Capítulo 4

Conclusões e Discussões

Cânceres de mama são tradicionalmente conhecidos por possuírem uma alta taxa de mortalidade entre as mulheres. Alguns dos sinais característicos do câncer de mama que os radiologistas podem buscar são a presença de agrupamentos de microcalcificações, massas e distorções arquiteturais. As massas podem ser descritas pelas características da sua forma e margem e de textura. Lesões com margens espiculadas possuem maior tendência à malignidade, enquanto lesões circunscritas e com bordar arredondadas e características homogêneas são predominantemente benignas.

A mamografia é um exame de baixo contraste, variando para diferentes densidades da mama. Mulheres jovens tendem a possuir mamas densas, que tem densidade semelhante às lesões, dificultando ainda mais a identificação, localização e classificação da lesão. Por isto, acreditamos que o estudo da classificação separadamente por tipo de mama pode aumentar a taxa de detecção e classificação de tumores em mamografia.

Conforme foi visto anteriormente, se for realizada a classificação de tumores desconsiderando a constituição tecidual da mama, os classificadores não apresentam um bom desempenho, obtendo uma taxa de acerto máxima em torno dos 60%, com índice *kappa* próximo à 0,60, o que indica uma concordância razoável com o teste padrão utilizado como referência, neste caso, a opinião de especialistas.

Por outro lado, ao se levar em consideração a pré-classificação da mama por densidade, associado à idade da paciente, obtemos resultados mais satisfatórios. Para mamas extremamente densas, que possuem maior dificuldade na análise da mamografia, obtivemos valor de *kappa* de 0,84 e precisão média de 89% para amostras descritas através de suas características de forma e textura e classificadas por redes MLP, sendo segmentadas utilizando transformada de wavelets com 2 níveis de decomposição. A mesma configuração também se mostrou a mais eficiente para classificação de mamas adiposas, resultando numa precisão média de 90% e mamas pouco densas, com precisão média de 94%. Apenas para mamas fibroglandulares foi obtida uma melhor classificação das lesões utilizando segmentação por mapa auto-organizáveis de Kohonen, com precisão média de 91%.

A partir dos resultados é possível verificar que a classificação com maior índice de acerto deu-se para mamas fibroglandulares e densas. Tais resultados se mostraram conforme com a

teoria apresentada, uma vez que em mamas muito densas o contraste entre tecido normal e tecido lesionado é muito baixo, dificultando a diferenciação dos mesmos, conforme visto na Tabela 8.

No entanto, a classificação de mamas adiposas ficou aquém do esperado. Apesar de ter alcançado uma precisão geral alta na classificação, era esperado a obtenção de uma precisão de 100% na diferenciação entre lesão e tecido sadio, o que não foi, no entanto, o obtido, conforme Tabela 5.

Os experimentos permitiram corroborar o pressuposto no início do trabalho em relação aos atributos utilizados na descrição da imagem. A utilização das características de forma e de textura para a descrição da região de interesse forneceu à rede neural maiores informações, permitindo alcançar melhores resultados de classificação das imagens e das lesões. Ainda foi possível verificar que a associação da transformada de wavelets com 2 níveis de decomposição (que gera mais atributos que são utilizados para descrição das imagens) com redes neurais artificiais do tipo perceptron de múltiplas camadas possibilitou, de forma geral, a melhor configuração do sistema para obtenção de melhores resultados de classificação de tumores.

No entanto, foi visto que a utilização deste sistema CAD apenas foi efetiva ao aplicarmos em bases de dados previamente organizadas em relação ao tipo de constituição tecidual e densidade da mama. De forma geral, os classificadores obtiveram uma precisão média de classificação das imagens de mamografia analisadas acima de 90% e mostram desempenho ainda melhor (próximo à 100%) na distinção entre amostras sem lesão das amostras contendo lesão, independente da severidade da mesma.

Visto que sistemas CAD tendem a ser utilizados como apoio ao diagnóstico, muitas vezes servindo como segunda opinião para o médico especialista e considerando os resultados obtidos pela classificação de imagens de mamografia utilizando redes neurais artificiais em conjunto de imagens com mamas de densidades semelhantes, acreditamos que a associação desse sistema com o diagnóstico do especialista pode resultar em um número ainda menor de falsos positivos e numa classificação mais precisa para o rastreamento e diagnóstico de câncer de mama.

4.1 Trabalhos futuros

Os resultados apresentados nesse trabalho foram bastante significativos e permitiram analisar diferentes instâncias que fazem parte do processo de classificação de imagens. Acreditamos ser possível melhorá-los aplicando uma etapa de otimização das redes neurais utilizadas

na classificação. Verificamos que é possível alcançar uma boa taxa de classificação mesmo em imagens com pouco ou nenhum pré-processamento, associando descritores de características com redes neurais artificiais.

Desta forma, desejamos, em trabalhos futuros, aplicar métodos de computação evolucionária, como algoritmos genéticos e otimização por enxames, para melhorar o desempenho das redes neurais artificiais utilizadas na classificação através da otimização dos paradigmas de treinamento da rede.

Além disso, uma contribuição para alcançar resultados mais fidedignos pode surgir a partir de pesquisas que visem à classificação, de forma automática, do tipo de mama a partir de informações como idade da paciente e análise da imagem da composição tecidual e densidade da mama. Essa etapa pode ser aplicada antes da classificação das lesões, visto que os resultados obtidos neste trabalho mostram que os processos de segmentação, extração de características e classificação possuem melhor desempenho na classificação de lesões para cada tipo pré-definido de mama. Dessa forma, teríamos um sistema CAD mais completo, capaz de identificar o tipo de mama e a composição tecidual da mesma, classificando-a segundo sua densidade, e utilizar essa informação para aplicar a melhor combinação de descritores de características e redes neurais artificiais para a classificação de lesões em imagens de mamografia.

Referências

- ACR - American College of Radiology. **Breast Imaging Reporting and Data System Atlas (BI-RADS® Atlas)**. Mammography, 4^o Edição, 2003.
- ANTONIO, A. L.M e CRESPI, C.M. **Predictors of interobserver agreement in breast imaging using the Breast Imaging Reporting and Data System (BI-RADS)**. Breast Cancer Res Treat páginas 539–546. Abr. 2010.
- AZAR, A. T. e EL-SAID, S.A. **Probabilistic neural network for breast cancer classification**. NeuralComput&Applic, páginas1737–1751, 2003.
- BLUE CROSS AND BLUE SHIELD ASSOCIATION. **Computer-Aided Detection (CAD) in Mammography**. Technology Evaluation Center Assessment Program, Volume 17, N^o 17, Dez. 2002.
- BONIFÁCIO, F.N. **Comparação entre as Redes Neurais Artificiais MLP, RBF e LVQ na classificação de dados**. Dissertação em Ciência da Computação, Universidade Estadual do Oeste do Paraná (UNIOESTE), 2010.
- BOUJELBEN, A., CHAABANI, A. C., TMAR, H. **Feature Extraction from contours shape for tumor analyzing in Mammographic images**. Digital Image Computing: Techniques and Applications, 2009.
- BOVIK, A. **Handbook of Image and Video Processing**. Academic Press, 1^o Edição, 2000.
- COSIO, F. A., HEVIA, N., LIRA, E., GARCIA, C., et al., **Mammographic image analysis and computer assisted biopsy of breast tumors**.4th International Conference on Biomedical Engineering and Informatics, 2011.
- DA SILVA, C.Y.V.W. **Extração de características de imagens médicas utilizando wavelets para mineração de imagens e auxílio ao diagnóstico**. Dissertação de mestrado, Instituto de Ciências Matemáticas e de Computação (ICMC-USP), USP, São Carlos, 2007.

DESERMO, T.M., SORION, M., DE OLIVEIRA, J.E.E., DE ARAÚJO, A. **Towards computer-aided diagnostics of screening mammography using content-based image retrieval.** Proc IEEE SIBGRAPHI 2011; 211-9.

DESERNO, T.M., SORION, M., DE OLIVEIRA, J.E.E., DE ARAÚJO. **Computer-aided diagnostics of screening mammography using content-based image retrieval.** Proc SPIE 2012; 8315:271-9.

EDDAOUDI, F., REGRAGUI, F., MAHMOUDI, A., LAMOURI, N. **Masses Detection Using SVM Classifier Based on Textures Analysis.** Applied Mathematical Sciences, Vol. 5, 2011, no. 8, 367 – 379

FLEISS, J.L. **Statistical methods for rates and proportions.** New York: John Wiley & Sons; 1973. pp. 598–626.

GANESAN, K., ACHARYA, U. R. A., CHUA, C. K., MIN, L. C., ABRAHAM, K.T. e NG, KH. **Computer-Aided Breast Cancer Detection Using Mammograms: A Review.** IEEE Reviews in Biomedical Engineering, vol. 6, 2013.

GILBERT F. J. et al. **Single Reading with Computer-Aided Detection for Screening Mammography.** The New England Journal of Medicine, 2008.

GROMET, M. **Comparison of Computer-Aided Detection to Double Reading of Screening Mammograms: Review of 231,221 Mammograms.** American Journal of Radiology, 2008.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEM, I. H. **The WEKA Data Mining Software: An Update;** SIGKDD Explorations, Volume 11, Issue 1, 2009.

HARALICK, R. M., SHANMUGAN, K., DISTEIN, IH. **Textural Features for Image Classification.** IEEE Transactions on Systems, Man and Cybernetics. Vol. SMC-3, N° 6, Novembro 1973, pp. 610-621

S. HAYKIN. **Neural networks.** 2nd.edition. Prentice Hall, 1999

INCA - Instituto Nacional de Câncer, Ministério da Saúde, Coordenação de Prevenção e Vigilância (Conprev). **Falando sobre Mamografia**. Viva Mulher – Programa Nacional de Controle do Câncer do Colo do Útero e de Mama. Rio de Janeiro, 2000.

INCA - Instituto Nacional de Câncer, Ministério da Saúde. **Mamografia: da prática ao controle**. Rio de Janeiro, 2007.

INCA - Instituto Nacional do Câncer, Ministério da Saúde, Disponível em: **<http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama>**. Acesso em: 02 set. 2013.

INCA - Instituto Nacional de Câncer, Ministério da Saúde, Coordenação de Prevenção e Vigilância. **Estimativa 2014: Incidência de Câncer no Brasil**. Rio de Janeiro, 2014.

ISLAM, M.J., AHMADI, M., SID-AHMED, M. A. **An efficient automatic mass classification method in digitized mammograms using artificial neural network**. International Journal of Artificial Intelligence & Applications (IJAIA), Vol.1, No.3, July 2010. DOI : 10.5121/ijaia.2010.1301 1

JIANG, J., Trundle, P., REN, J. **Medical image analysis with artificial neural networks**. Computerized Medical Imaging and Graphics, Vol 34, Issue 8, páginas617–631, 2010.

KILIC, N., GORGEL, P., UCAN, O. N., SERTBAS, A. **Mammographic Mass Detection using Wavelets as Input to Neural Networks**. Springer Science, J Med Syst, Vol. 34, páginas1083–1088, 2010.

LANDIS, J.R. e KOCH,G.G. **The measurement of observer agreement for categorical data**.Biometrics.Vol.33, páginas 159-74, Mar 1977.

LAZARUS, E, MAINIERO, M.B, SCHEPPS, B., KOELLIKER, S.L., LIVINGSTON, L.S. **BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value**. Radiology, Vol. 239, Maio 2006.

OBUCHOWSK, N. A. **Receiver Operating Characteristic Curves and Their Use in Radiology**. Radiology, Vol. 229, páginas 3-8, 2003.

OLIVEIRA, J.E.E., MACHADO, A.M.C., CHAVEZ, G.C, LOPES, A.P.B., DESERNO, T.M., DE ARAUJO, A. **MammoSys: a content-based image retrieval system using breast density patterns**. *Comput Methods Programs Biomed*, Vol 99, Issue 3, páginas 289-297, 2010.

OOMS, E.A, ZONDERLAND, H.M, EIJKEMANS, M.J, KRIEGE, M., MAHDAVIAN, D. B, BURGER, C.W., ANSIK, A.C. **Mammography: interobserver variability in breast density assessment**. *Breast*, Vol 16, Issue6, Dec. 2007.

OSHIRO, S. H. e GOLDSCHMIDT, R. R. **Processamento de Imagens**. Monografias em Sistemas e Computação n°05/2008. Instituto Militar de Engenharia, Rio de Janeiro, 2008.

Platt, J. C. **Fast Training of Support Vector Machines using Sequential Minimal Optimization**. *Advances in Kernel Methods - Support Vector Learning*, MIT Press, 1998.

RAHMATI, P., ADLER, A., HAMARNEH, G. **Mammography segmentation with maximum likelihood active contours**. *Medical Image Analysis*, Vol. 16, páginas 1167–1186, 2012.

RAPELYEA J. A. e BREM R.F. **Clinical Utility of CAD Systems for Breast Cancer**. In: LI Q, NISHIKAWA R M. *Computer-Aided Detection and Diagnosis in Medical Imaging*. CRC Press, 2015. P. 345 - 356.

STINES, J., NOEL, A., GALUS, X. **Technologie en mammographie**. EMC - Radiologie et Imagerie Médicale: Génito-urinaire - Gynéco-obstétricale – Mammaire, 2010.

TAHMASBI, A., SAKI, F., SHOKOUHI, S.B. **Classification of benign and malignant masses based on Zernike moments**. Elsevier. *Computers in Biology and Medicine*, 2011.

TZIKOPOULOSA, S.D. et al. **A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry**. Elsevier. *Computer Methods and Programs in Biomedicine*, 2011.P. 47 – 63.

APÊNDICE A – RESULTADO DAS SIMULAÇÕES DOS SISTEMAS CAD ESTUDADOS

Tipo de mama	Segmentação e atributos	Classificador	Tempo para construção do modelo	Instâncias corretamente classificadas	Kappa
Densa	WA-2-HA-ZE	MLP-i	6163,51	94,29%	0,9139
Densa	WA-2-HA-ZE	MLP-a	1881,94	93,57%	0,9031
Densa	WA-2-HA-ZE	MLP-o	25,41	93,57%	0,9032
Densa	WA-1-HA-ZE	MLP-o,o	22,73	92,14%	0,8815
Densa	WA - 1 - ZE	MLP-a	178,87	92,14%	0,8812
Densa	WA - 1 - ZE	MLP-i	230,8	92,14%	0,8813
Densa	WA-2-HA-ZE	MLP-a,a	3996,39	92,14%	0,8817
Densa	WA-2-HA-ZE	MLP-o,o	25,89	92,14%	0,8815
Fibroglandular	SOM - HA-ZE	MLP-a	135,08	91,43%	0,8703
Densa	WA - 1 - ZE	MLP-a,a	150,77	90,71%	0,8596
Densa	WA-2-HA-ZE	MLP-i,i	11571,67	90,71%	0,8602
Fibroglandular	SOM - HA-ZE	MLP-i	251,31	90,71%	0,8593
Densa	WA-2-ZE	MLP-a	326,3	90,00%	0,8492
Densa	SOM - HA-ZE	MLP-a	117,48	89,29%	0,8385
Densa	WA-1-HA-ZE	MLP-a	490,56	89,29%	0,8383
Densa	WA-1-HA-ZE	MLP-i,i	4180,86	89,29%	0,8381
Densa	WA-1-HA-ZE	MLP-o	14	89,29%	0,8384
Fibroglandular	WA-1-HA-ZE	MLP-i	5438,96	89,29%	0,8384
Fibroglandular	WA-2-HA-ZE	MLP-a	2291,35	89,29%	0,8387
Fibroglandular	WA-2-HA-ZE	MLP-i	5728,33	89,29%	0,8384
Extremamente Densa	WA-2-HA-ZE	MLP-a	1886,45	89,29%	0,8387
Densa	WA-2-HA-ZE	SMOpoly	0,89	88,57%	0,828
Adiposa	WA-2-HA-ZE	MLP-a	2217,03	88,57%	0,8283
Densa	WA-1-HA-ZE	MLP-a,a	1467,39	88,57%	0,8276
Densa	WA-1-HA-ZE	MLP-i	1383,64	88,57%	0,8275
Densa	WA - 1 - ZE	MLP-o	13,39	88,57%	0,8269
Fibroglandular	SOM - HA-ZE	MLP-a,a	189,82	88,57%	0,8267
Fibroglandular	SOM - HA-ZE	MLP-i,i	600,84	88,57%	0,8267
Fibroglandular	SOM - HA-ZE	SMOpoly	0,22	88,57%	0,8273
Adiposa	SOM - HA-ZE	MLP-a	130,07	87,86%	0,8166
Densa	SOM - HA-ZE	MLP-a,a	190,25	87,86%	0,8171
Densa	SOM - HA-ZE	MLP-i,i	593,46	87,86%	0,8170
Densa	WA-2-ZE	MLP-i	469,63	87,86%	0,8168
Extremamente Densa	WA-1-HA-ZE	MLP-a	497,1	87,86%	0,8169
Extremamente Densa	WA-1-HA-ZE	MLP-i	1307,11	87,86%	0,8169
Adiposa	WA-2-HA-ZE	MLP-i	4145,95	87,14%	0,8068
Adiposa	WA-2-HA-ZE	MLP-o	22,72	87,14%	0,8064
Densa	SOM - HA-ZE	MLP-i	220,24	87,14%	0,8062
Densa	WA - 1 - ZE	SMOpoly	0,41	87,14%	0,8053

Densa	WA-2-ZE	MLP-o,o	9,01	87,14%	0,8061
Extremamente Densa	WA-2-HA-ZE	MLP-i	6554,09	87,14%	0,8067
Fibroglandular	SOM - ZE	MLP-a	24,69	87,14%	0,8053
Densa	WA-1-HA-ZE	SMOpoly	0,64	86,43%	0,7953
Densa	WA-2-ZE	MLP-o	10,45	86,43%	0,7953
Densa	WA-2-ZE	SMOpoly	0,91	86,43%	0,7951
Fibroglandular	SOM- HA	MLP-o,o	8,58	86,43%	0,7935
Fibroglandular	WA-1-HA-ZE	MLP-a,a	4024,71	86,43%	0,795
Fibroglandular	WA-2-HA-ZE	MLP-a,a	4532,31	86,43%	0,7950
Adiposa	WA-2-HA-ZE	MLP-o,o	20,92	85,71%	0,7842
Densa	WA - 1 - ZE	MLP-i,i	1040,24	85,71%	0,7838
Densa	WA-2-ZE	MLP-a,a	700,78	85,71%	0,7849
Extremamente Densa	WA-2-HA-ZE	SMOpoly	1,19	85,71%	0,7843
Fibroglandular	SOM- HA	MLP-i	219,48	85,71%	0,7859
Fibroglandular	SOM - ZE	MLP-i	42,04	85,71%	0,7829
Fibroglandular	WA-1-HA-ZE	MLP-i,i	16271,81	85,71%	0,785
Fibroglandular	WA-1-HA-ZE	MLP-o,o	26,89	85,71%	0,7862
Fibroglandular	WA-2-HA-ZE	MLP-o,o	28,64	85,71%	0,7862
Fibroglandular	SOM- HA	MLP-o	7,82	85,57%	0,7525
Adiposa	SOM - HA-ZE	MLP-i,i	830,14	85,00%	0,7733
Adiposa	SOM - HA-ZE	SMOpoly	0,32	85,00%	0,7723
Densa	SOM- HA	MLP-i,i	606,28	85,00%	0,7737
Densa	SOM- HA	MLP-o	8,77	85,00%	0,7739
Densa	SOM - ZE	MLP-i	35,56	85,00%	0,7742
Fibroglandular	SOM- HA	MLP-a,a	186,75	85,00%	0,775
Fibroglandular	SOM- HA	SMOpoly	0,23	85,00%	0,7727
Fibroglandular	SOM - HA-ZE	MLP-o,o	6,53	85,00%	0,7746
Fibroglandular	WA-2-HA-ZE	SMOpoly	0,99	85,00%	0,7738
Adiposa	SOM - HA-ZE	MLP-i	254,78	84,29%	0,7623
Adiposa	SOM - ZE	MLP-a,a	36,7	84,29%	0,7637
Adiposa	WA-1-ZE	MLP-a	176,25	84,29%	0,7635
Adiposa	WA-2-HA-ZE	SMOpoly	0,94	84,29%	0,7632
Densa	SOM- HA	MLP-a	118,83	84,29%	0,7631
Densa	SOM- HA	MLP-i	232,97	84,29%	0,7632
Densa	SOM - HA-ZE	MLP-o	6	84,29%	0,7632
Densa	SOM - ZE	MLP-a	18,38	84,29%	0,7631
Densa	WA-2-ZE	MLP-i,i	1595,87	84,29%	0,7631
Extremamente Densa	WA-1-ZE	MLP-a	131,33	84,29%	0,7638
Extremamente Densa	WA-2-HA-ZE	MLP-o	27,38	84,29%	0,7636
Extremamente Densa	WA-2-HA-ZE	MLP-o,o	25,54	84,29%	0,7633
Extremamente Densa	WA-2-ZE	MLP-a	228,29	84,29%	0,7633
Fibroglandular	SOM- HA	MLP-a	110,96	84,29%	0,7641
Fibroglandular	SOM - HA-ZE	MLP-o	6,19	84,29%	0,7636
Adiposa	WA-2-ZE	SMOpoly	0,91	83,57%	0,7524
Densa	SOM - HA-ZE	SMOpoly	0,28	83,57%	0,7523

Densa	WA - 1 - ZE	MLP-o,o	12,04	83,57%	0,7515
Extremamente Densa	SOM - ZE	MLP-i	38,06	83,57%	0,7529
Extremamente Densa	WA-2-HA-ZE	MLP-i,i	10771,21	83,57%	0,752
Fibroglandular	SOM- HA	MLP-i,i	576,13	83,57%	0,754
Fibroglandular	SOM - ZE	MLP-a,a	38,25	83,57%	0,7539
Fibroglandular	SOM - ZE	RBF	0,31	83,57%	0,7514
Densa	SOM - ZE	MLP-a,a	29,48	82,96%	0,742
Adiposa	SOM - HA-ZE	MLP-a,a	248,96	82,86%	0,7405
Adiposa	WA-1-HA-ZE	MLP-a	586,28	82,86%	0,7415
Adiposa	WA-1-ZE	RBF	1,03	82,86%	0,7419
Densa	SOM - ZE	MLP-i,i	77,79	82,86%	0,7409
Densa	SOM - ZE	MLP-o	2,22	82,86%	0,7413
Extremamente Densa	SOM - ZE	MLP-i,i	80,18	82,86%	0,7413
Extremamente Densa	WA-1-HA-ZE	MLP-a,a	1052,26	82,86%	0,741
Extremamente Densa	WA-2-HA-ZE	MLP-a,a	5029,44	82,86%	0,7421
Extremamente Densa	WA-2-ZE	SMOpoly	0,61	82,86%	0,7434
Fibroglandular	SOM - ZE	MLP-i,i	97,35	82,86%	0,7442
Fibroglandular	SOM - ZE	SMOpoly	0,15	82,86%	0,7406
Fibroglandular	WA-1-HA-ZE	MLP-o	93	82,86%	0,7415
Fibroglandular	WA-2-HA-ZE	MLP-o	28,1	82,86%	0,7415
Adiposa	WA-1-HA-ZE	MLP-a,a	959,51	82,14%	0,7307
Adiposa	WA-2-ZE	MLP-a	337,66	82,14%	0,7308
Adiposa	WA-2-ZE	MLP-o,o	13,76	82,14%	0,7315
Densa	SOM - ZE	RBF	0,86	82,14%	0,7311
Extremamente Densa	WA-1-ZE	MLP-i	312,95	82,14%	0,7315
Extremamente Densa	WA-2-ZE	MLP-i	476,68	82,14%	0,7304
Fibroglandular	SOM - HA-ZE	RBF	0,55	82,14%	0,7339
Fibroglandular	WA-2-ZE	MLP-i	610,61	82,14%	0,7298
Fibroglandular	WA-2-ZE	RBF	2,26	82,14%	0,7303
Adiposa	WA-1-HA-ZE	MLP-i	1276,23	81,43%	0,72
Adiposa	WA-2-HA-ZE	MLP-a,a	4258,9	81,43%	0,7205
Densa	SOM- HA	SMOpoly	0,16	81,43%	0,72
Densa	SOM - ZE	SMOpoly	0,24	81,43%	0,7197
Densa	WA-1-HA	MLP-i,i	1355,95	81,43%	0,7194
Extremamente Densa	WA-2-ZE	MLP-o,o	8,83	81,43%	0,7205
Fibroglandular	WA-1-HA-ZE	MLP-a	498,52	81,43%	0,7226
Fibroglandular	WA-2-ZE	MLP-a,a	582,39	81,43%	0,7185
Fibroglandular	WA-2-ZE	MLP-i,i	2062,4	81,43%	0,7194
Fibroglandular	WA-2-ZE	MLP-o	17,15	81,43%	0,7189
Adiposa	WA-1-ZE	MLP-a,a	271,27	81,33%	0,7206
Adiposa	SOM- HA	MLP-a,a	114,41	80,71%	0,7094
Adiposa	SOM - HA-ZE	MLP-o,o	7,54	80,71%	0,7089
Adiposa	SOM - ZE	MLP-a	22,13	80,71%	0,7105
Adiposa	SOM - ZE	MLP-i	51,86	80,71%	0,7104
Adiposa	WA-1-HA-ZE	SMOpoly	0,66	80,71%	0,708

Adiposa	WA-2-HA-ZE	MLP-i,i	15442,89	80,71%	0,71
Adiposa	WA-2-ZE	RBF	2,03	80,71%	0,7095
Densa	WA-1-HA	MLP-a,a	382,61	80,71%	0,7084
Densa	WA-2-ZE	RBF	2,48	80,71%	0,7094
Fibroglandular	SOM - ZE	MLP-o,o	2,89	80,71%	0,7112
Fibroglandular	WA-1-ZE	MLP-a	174,88	80,71%	0,7112
Fibroglandular	WA-2-ZE	MLP-a	326,68	80,71%	0,7075
Fibroglandular	WA-2-ZE	MLP-o,o	16,75	80,71%	0,7086
Fibroglandular	WA-2-ZE	SMOpoly	0,72	80,71%	0,7984
Fibroglandular	WA-1-HA	MLP-o	8,79	80,71%	0,7094
Adiposa	SOM - ZE	MLP-i,i	98,33	80,00%	0,6989
Adiposa	WA-2-ZE	MLP-i	696,19	80,00%	0,6987
Densa	SOM- HA	RBF	0,62	80,00%	0,6983
Densa	WA - 1 - ZE	RBF	0,62	80,00%	0,6983
Densa	WA-2-ZE	SMOrbf	1,28	80,00%	0,6976
Extremamente Densa	WA-1-HA-ZE	MLP-i,i	4210,96	80,00%	0,6981
Extremamente Densa	WA-1-HA-ZE	SMOpoly	0,57	80,00%	0,6972
Extremamente Densa	WA-2-ZE	MLP-a,a	422,29	80,00%	0,6993
Fibroglandular	SOM - ZE	MLP-o	2,87	80,00%	0,696
Fibroglandular	WA-1-ZE	RBF	1,2	80,00%	0,6981
Adiposa	WA-1-HA	MLP-a,a	361,36	79,29%	0,6859
Adiposa	WA-1-HA	SMOpoly	0,94	79,29%	0,686
Adiposa	WA-1-ZE	MLP-o	12,71	79,29%	0,6881
Adiposa	WA-1-ZE	MLP-o,o	13,65	79,29%	0,6879
Adiposa	WA-1-ZE	SMOpoly	0,66	79,29%	0,6875
Adiposa	WA-2-ZE	MLP-a,a	408,66	79,29%	0,6868
Densa	WA-1-HA	MLP-i	424,3	79,29%	0,6865
Densa	WA - 2- HA	MLP-a	690,1	79,29%	0,6869
Extremamente Densa	SOM - HA-ZE	MLP-a	109,85	79,29%	0,6867
Extremamente Densa	SOM - HA-ZE	MLP-i	246,89	79,29%	0,6868
Fibroglandular	SOM - HA-ZE	SMOrbf	0,55	79,29%	0,6927
Fibroglandular	WA-1-HA-ZE	SMOpoly	0,55	79,29%	0,6904
Fibroglandular	WA-1-ZE	MLP-a,a	313,39	79,29%	0,6901
Fibroglandular	WA-1-ZE	MLP-i	331,89	79,29%	0,6888
Adiposa	WA-1-HA-ZE	MLP-i,i	3992,44	78,57%	0,6777
Extremamente Densa	SOM- HA	MLP-i,i	629,76	78,57%	0,676
Extremamente Densa	WA-1-ZE	RBF	1,28	78,57%	0,6773
Fibroglandular	WA-1-HA	MLP-a,a	336,57	78,57%	0,6771
Adiposa	SOM- HA	MLP-i,i	400,1	77,86%	0,6657
Adiposa	WA-1-HA	MLP-a	197,94	77,86%	0,6642
Adiposa	WA-1-HA	MLP-i	519,58	77,86%	0,6646
Adiposa	WA-2-HA-ZE	SMOrbf	0,86	77,86%	0,6654
Adiposa	WA-2-ZE	MLP-i,i	2201,8	77,86%	0,6658
Densa	WA - 2- HA	MLP-i	1487	77,86%	0,6651
Extremamente Densa	WA-1-HA-ZE	MLP-o	15,8	77,86%	0,6672

Extremamente Densa	WA-2-ZE	MLP-i,i	1564,82	77,86%	0,6672
Extremamente Densa	WA-2-ZE	RBF	1,06	77,86%	0,6638
Densa	WA-1-HA	MLP-a	208,39	77,14%	0,6539
Fibroglandular	WA-1-HA	SMOpoly	0,42	77,14%	0,6541
Fibroglandular	WA-1-ZE	MLP-i,i	738,12	77,14%	0,6595
Fibroglandular	WA-2-HA	MLP-i	1459,46	77,14%	0,6527
Fibroglandular	WA-2-HA-ZE	RBF	1,59	77,14%	0,6493
Adiposa	SOM –HA	MLP-i	246,31	76,43%	0,6445
Adiposa	SOM - HA-ZE	MLP-o	7,23	76,43%	0,6429
Adiposa	WA-1-HÁ	MLP-i,i	2661,23	76,43%	0,644
Adiposa	WA-1-HÁ	MLP-o,o	19,37	76,43%	0,6426
Adiposa	WA-1-ZE	MLP-i	352,54	76,43%	0,6447
Adiposa	WA-1-ZE	MLP-i,i	924,34	76,43%	0,6448
Densa	SOM –HA	SMOrbf	1,25	76,43%	0,6429
Densa	SOM - HA-ZE	RBF	1,34	76,43%	0,6442
Densa	WA-1-HA	MLP-o	9,12	76,43%	0,6434
Densa	WA - 2– HA	MLP-a,a	1562,58	76,43%	0,6423
Densa	WA - 2– HA	MLP-i,i	5359,62	76,43%	0,6431
Extremamente Densa	SOM – HA	MLP-a	134,54	76,43%	0,6438
Extremamente Densa	SOM – ZE	MLP-a	20,29	76,43%	0,6455
Extremamente Densa	SOM - ZE	RBF	0,59	76,43%	0,6465
Adiposa	WA-1-HA-ZE	MLP-o	11,87	75,71%	0,6343
Adiposa	WA-1-HA-ZE	MLP-o,o	11,99	75,71%	0,6346
Adiposa	WA-1-HA-ZE	RBF	1,47	75,71%	0,6328
Densa	SOM - HA-ZE	SMOrbf	0,91	75,71%	0,6327
Densa	WA - 2– HA	SMOpoly	0,64	75,71%	0,6312
Extremamente Densa	SOM - HA-ZE	MLP-i,i	655,74	75,71%	0,6316
Fibroglandular	WA-1-HA	MLP-a	204,8	75,71%	0,6335
Fibroglandular	WA-2-HA	MLP-a	645,19	75,71%	0,6337
Fibroglandular	WA-2-HA	MLP-i,i	5706,35	75,71%	0,6321
Adiposa	SOM– HA	MLP-a	114,65	75,00%	0,623
Adiposa	SOM - ZE	RBF	0,61	75,00%	0,6243
Adiposa	WA-1-HA	MLP-o	12,27	75,00%	0,6205
Adiposa	WA-2-ZE	MLP-o	15,93	75,00%	0,6223
Densa	WA-2-HA-ZE	SMOrbf	0,85	75,00%	0,6194
Extremamente Densa	SOM– HA	MLP-i	302,88	75,00%	0,6212
Extremamente Densa	SOM - HA-ZE	MLP-a,a	190,28	75,00%	0,6214
Extremamente Densa	SOM - ZE	MLP-a,a	29,79	75,00%	0,6243
Extremamente Densa	WA-2-ZE	MLP-o	9,35	75,00%	0,6236
Fibroglandular	WA-1-HA	MLP-i,i	1218,63	75,00%	0,6249
Adiposa	SOM - HA-ZE	RBF	0,82	74,29%	0,6144
Densa	WA-1-HA	MLP-o,o	8,25	74,29%	0,6101
Extremamente Densa	SOM– HA	MLP-a,a	207,93	74,29%	0,6095
Extremamente Densa	SOM - ZE	MLP-o	2,47	74,29%	0,6124
Extremamente Densa	WA-1-HA-ZE	MLP-o,o	13,46	74,29%	0,6099

Extremamente Densa	WA-1-ZE	MLP-a,a	279,15	74,29%	0,613
Fibroglandular	SOM– HA	RBF	0,66	74,29%	0,6164
Fibroglandular	WA-1-HA	MLP-i	382,66	74,29%	0,6125
Fibroglandular	WA-1-ZE	MLP-o	11,87	74,29%	0,6149
Fibroglandular	WA-1-ZE	SMOpoly	0,7	74,29%	0,6161
Fibroglandular	WA-2-HA	MLP-a,a	1550,4	74,29%	0,6118
Fibroglandular	WA-2-HA-ZE	SMOrbf	1	74,29%	0,6124
Adiposa	SOM– HA	MLP-o	9,35	73,57%	0,599
Adiposa	WA-2-HA	MLP-i	2082,59	73,57%	0,5976
Adiposa	WA-2-ZE	SMOrbf	1,1	73,57%	0,5967
Densa	WA - 1 - ZE	SMOrbf	0,87	73,57%	0,5988
Densa	WA - 2– HA	MLP-o	17,72	73,57%	0,5984
Extremamente Densa	WA-1-HA	MLP-a,a	365,78	73,57%	0,6006
Fibroglandular	WA-1-ZE	MLP-o,o	13,85	73,57%	0,6054
Fibroglandular	WA-2-HA	MLP-o	16,18	73,57%	0,6021
Adiposa	WA-1-HA-ZE	SMOrbf	0,66	72,86%	0,5879
Adiposa	WA-2-HA	MLP-a	682,23	72,86%	0,5883
Densa	SOM - ZE	MLP-o,o	2,42	72,86%	0,5906
Densa	WA-1-HA-ZE	RBF	1,25	72,86%	0,5908
Extremamente Densa	WA-1-ZE	MLP-i,i	580,91	72,86%	0,5922
Fibroglandular	WA-1-HA	MLP-o,o	7,97	72,86%	0,5917
Adiposa	SOM - ZE	MLP-o	2,89	72,14%	0,5809
Adiposa	SOM - ZE	SMOpoly	0,24	72,14%	0,5755
Adiposa	WA-2-HA-ZE	RBF	1,65	72,14%	0,5784
Extremamente Densa	WA-1-ZE	SMOpoly	0,59	72,14%	0,5805
Fibroglandular	WA-1-HA-ZE	RBF	0,74	72,14%	0,5763
Fibroglandular	WA-1-HA-ZE	SMOrbf	0,56	72,14%	0,5833
Fibroglandular	WA-2-HA	SMOpoly	1,29	72,14%	0,5706
Extremamente Densa	WA-1-HA	MLP-i	392,49	71,44%	0,5694
Adiposa	SOM - HA-ZE	SMOrbf	0,67	71,43%	0,5668
Densa	SOM - HA-ZE	MLP-o,o	5,42	71,43%	0,5689
Densa	WA-1-HA	SMOpoly	0,55	71,43%	0,5664
Extremamente Densa	SOM– HA	MLP-o	6,25	71,43%	0,5663
Extremamente Densa	WA-2-HA	MLP-i,i	5232,91	71,43%	0,5676
Adiposa	SOM– HA	SMOpoly	0,56	70,71%	0,5545
Adiposa	WA-2-HA	MLP-i,i	5827,98	70,71%	0,555
Adiposa	WA-2-HA	MLP-o,o	16,01	70,71%	0,5546
Extremamente Densa	SOM– HA	RBF	0,47	70,71%	0,5608
Extremamente Densa	WA-2-HA	MLP-a	697,15	70,71%	0,5564
Adiposa	SOM– HA	RBF	1,09	70,00%	0,5509
Densa	WA-1-HA-ZE	SMOrbf	1,04	70,00%	0,5442
Densa	WA - 2– HA	MLP-o,o	16,17	70,00%	0,5432
Extremamente Densa	WA-1-HA	MLP-i,i	1290,97	70,00%	0,548
Extremamente Densa	WA-2-HA	MLP-i	2073,73	70,00%	0,5459
Extremamente Densa	WA-2-HA	RBF	0,75	70,00%	0,5471

Fibroglandular	WA-2-ZE	SMOrbf	1,06	70,00%	0,5503
Adiposa	SOM - ZE	MLP-o,o	3	69,29%	0,5362
Adiposa	WA-2-HA	MLP-a,a	1155,72	69,29%	0,533
Extremamente Densa	SOM- HA	MLP-o,o	8,27	69,29%	0,5373
Fibroglandular	WA-1-HA	RBF	0,96	69,29%	0,535
Adiposa	SOM- HA	MLP-o,o	6,52	68,57%	0,5251
Adiposa	WA-1-HA	RBF	1,95	68,57%	0,5219
Adiposa	WA-1-ZE	SMOrbf	0,87	68,57%	0,5198
Densa	SOM- HA	MLP-a,a	217,86	68,57%	0,528
Densa	SOM- HA	MLP-o,o	9,45	68,57%	0,5251
Extremamente Densa	SOM - HA-ZE	RBF	0,72	68,57%	0,5278
Extremamente Densa	WA-1-ZE	MLP-o,o	6,48	68,57%	0,5281
Extremamente Densa	WA-2-HA	MLP-a,a	1340,38	68,57%	0,5243
Adiposa	SOM - ZE	SMOrbf	0,58	67,86%	0,5084
Extremamente Densa	SOM- HA	SMOpoly	0,31	67,86%	0,3063
Adiposa	WA-2-HA	SMOpoly	0,69	67,14%	0,4965
Densa	WA-1-HA	RBF	0,77	67,14%	0,5022
Extremamente Densa	WA-1-HA-ZE	RBF	0,75	67,14%	0,5058
Extremamente Densa	WA-2-HA-ZE	RBF	1,51	67,14%	0,502
Fibroglandular	WA-2-HA	MLP-o,o	18,46	67,14%	0,5151
Extremamente Densa	WA-2-HA-ZE	SMOrbf	0,98	66,43%	0,4849
Fibroglandular	SOM- HA	SMOrbf	0,84	66,43%	0,5138
Densa	WA-2-HA-ZE	RBF	1,63	65,71%	0,4789
Extremamente Densa	WA-1-HA	MLP-o,o	8,31	65,71%	0,4806
Extremamente Densa	WA-1-ZE	MLP-o	6,73	65,71%	0,487
Fibroglandular	WA-2-HA	RBF	1,59	65,71%	0,4714
Adiposa	WA-2-HA	MLP-o	18,94	65,00%	0,463
Extremamente Densa	WA-1-HA	MLP-a	206,54	65,00%	0,4727
Extremamente Densa	SOM - HA-ZE	SMOpoly	0,36	64,29%	0,4574
Adiposa	SOM- HA	SMOrbf	1,09	63,57%	0,4471
Adiposa	WA-2-HA	RBF	1,1	63,57%	0,4456
Extremamente Densa	WA-2-ZE	SMOrbf	0,69	63,57%	0,4486
Base completa	WA-2-HA-ZE	MLP-i	7446,92	63,33%	0,5999
Extremamente Densa	SOM - HA-ZE	MLP-o	5,62	62,14%	0,4324
Extremamente Densa	WA-1-HA	RBF	0,75	62,14%	0,4281
Extremamente Densa	SOM - HA-ZE	MLP-o,o	5,77	61,43%	0,4245
Extremamente Densa	WA-1-HA	MLP-o	8,6	61,43%	0,4175
Extremamente Densa	WA-2-HA	MLP-o	17,54	61,43%	0,4154
Extremamente Densa	WA-2-HA	SMOpoly	0,7	61,43%	0,4142
Base completa	WA-2-HA-ZE	SMOpoly	14,66	61,00%	0,5744
Extremamente Densa	SOM - ZE	SMOpoly	0,34	60,71%	0,4031
Fibroglandular	WA-1-ZE	SMOrbf	0,78	60,71%	0,4245
Base completa	WA-2-HA-ZE	MLP-a	3484,94	60,29%	0,5668
Densa	WA - 2- HA	RBF	0,83	60,00%	0,4029
Base completa	WA-1-HA-ZE	MLP-a	2240,2	59,03%	0,5529

Extremamente Densa	WA-1-HA	SMOpoly	0,63	58,57%	0,3683
Adiposa	WA-1-HA	SMOrbf	0,89	57,86%	0,3544
Extremamente Densa	SOM - HA-ZE	SMOrbf	0,66	57,86%	0,3629
Extremamente Densa	WA-1-HA-ZE	SMOrbf	0,69	57,86%	0,3574
Extremamente Densa	WA-2-HA	MLP-o,o	16,74	57,86%	0,3594
Fibroglandular	SOM - ZE	SMOrbf	0,7	57,86%	0,3982
Densa	WA - 2- HA	SMOrbf	0,79	57,14%	0,3501
Densa	WA-1-HA	SMOrbf	0,62	56,43%	0,3402
Base completa	WA-1-HA-ZE	MLP-i	5005,54	56,35%	0,5238
Extremamente Densa	WA-2-HA	SMOrbf	0,68	55,71%	0,3251
Adiposa	WA-2-HA	SMOrbf	0,94	54,29%	0,2974
Base completa	WA-1-HA-ZE	MLP-a,a	4062,98	54,03%	0,4981
Base completa	WA-2-HA-ZE	SMOrbf	4062,98	54,03%	0,4981
Base completa	SOM - HA-ZE	MLP-a,a	8854,04	53,67%	0,4943
Base completa	WA-2-HA-ZE	MLP-a,a	8854,04	53,67%	0,4943
Base completa	WA-2-HA-ZE	MLP-i,i	24231,34	53,67%	0,4945
Extremamente Densa	WA-1-HA	SMOrbf	0,74	52,86%	0,2805
Base completa	WA-1-HA-ZE	MLP-i,i	16209,25	52,77%	0,4846
Base completa	WA-2-ZE	MLP-a	903,86	52,24%	0,4789
Base completa	WA-2-ZE	MLP-i	1746,84	52,06%	0,4768
Base completa	WA-1-HA-ZE	SMOpoly	9,41	51,88%	0,4752
Base completa	SOM - HA-ZE	MLP-i	954,7	50,98%	0,465
Base completa	WA-1-ZE	MLP-i	591,42	50,63%	0,4611
Base completa	WA-1-ZE	MLP-a	380,92	49,74%	0,4514
Extremamente Densa	SOM- HA	SMOrbf	1,08	49,29%	0,2498
Fibroglandular	WA-1-HA	SMOrbf	0,74	49,29%	0,2589
Fibroglandular	WA-2-HA	SMOrbf	1,32	49,29%	0,249
Base completa	SOM - HA-ZE	MLP-i,i	2544,9	49,02%	0,4436
Base completa	WA-1-ZE	MLP-i,i	1405,89	47,76%	0,4298
Base completa	SOM - HA-ZE	MLP-a	6,56	45,98%	0,4109
Base completa	WA-2-ZE	MLP-a,a	1575,23	45,80%	0,4089
Base completa	WA-2-ZE	SMOpoly	7,23	45,62%	0,407
Base completa	SOM - HA-ZE	RBF	15,84	44,72%	0,3958
Base completa	SOM -ZE	MLP-i,i	340,96	44,36%	0,3931
Base completa	WA-2-ZE	MLP-i,i	5658,37	43,65%	0,385
Base completa	SOM - HA-ZE	SMOpoly	2,72	43,29%	0,3809
Base completa	WA-2-ZE	RBF	4918,81	43,11%	0,3795
Densa	SOM - ZE	SMOrbf	0,79	42,86%	0,1639
Extremamente Densa	WA-1-ZE	SMOrbf	1,17	42,86%	0,172
Base completa	SOM -ZE	MLP-i	175,05	41,86%	0,3662
Base completa	WA-2-HA	MLP-a,a	5341,17	41,86%	0,3659
Base completa	WA-1-ZE	MLP-a,a	601,94	41,50%	0,3618
Base completa	SOM -ZE	MLP-a	100,12	41,14%	0,3584
Base completa	SOM -ZE	MLP-a,a	204,36	40,97%	0,3566
Base completa	WA-2-HA	MLP-i,i	17682,72	40,97%	0,3558

Base completa	WA-1-ZE	RBF	12057,21	40,61%	0,3522
Base completa	SOM- HA	MLP-i	378,71	40,25%	0,3487
Base completa	WA-2-HA-ZE	MLP-o	154,83	40,07%	0,3464
Base completa	SOM -ZE	RBF	22104,4	39,53%	0,34
Base completa	WA-1-HA	MLP-i,i	5069,08	39,53%	0,3402
Base completa	SOM- HA	MLP-a	575,47	39,36%	0,3389
Base completa	WA-1-HA	MLP-a	904,33	39,36%	0,3382
Base completa	SOM - HA-ZE	MLP-o	76,84	39,00%	0,3343
Base completa	SOM - HA-ZE	MLP-o,o	88,53	39,00%	0,335
Base completa	WA-1-ZE	SMOpoly	6,37	39,00%	0,3346
Base completa	SOM- HA	MLP-a,a	35,69	38,99%	0,3342
Base completa	WA-1-HA	MLP-i	1522,54	38,82%	0,3325
Base completa	WA-2-HA	MLP-i	5989,6	38,46%	0,3282
Base completa	SOM- HA	SMOpoly	3,67	38,10%	0,3239
Base completa	WA-2-HA-ZE	MLP-o,o	141,57	37,92%	0,3228
Base completa	WA-1-HA	MLP-a,a	1589,05	37,75%	0,3205
Base completa	SOM- HA	MLP-i,i	838,84	37,39%	0,318
Base completa	WA-1-HA-ZE	RBF	46,04	37,21%	0,3153
Base completa	WA-1-HA-ZE	MLP-o,o	144,85	36,67%	0,3093
Base completa	SOM- HA	RBF	1796,09	36,49%	0,3061
Base completa	SOM -ZE	SMO poly	1,69	36,49%	0,3078
Extremamente Densa	SOM - ZE	MLP-o,o	2,4	36,43%	0,0852
Base completa	WA-2-HA	MLP-a	2940,26	36,14%	0,3031
Base completa	SOM- HA	MLP-o	49,65	35,96%	0,3021
Base completa	SOM -ZE	MLP-o,o	41,27	35,24%	0,2937
Base completa	SOM -ZE	MLP-o	37,97	33,99%	0,2798
Base completa	WA-1-HA-ZE	MLP-o	175,92	33,99%	0,2799
Base completa	SOM- HA	MLP-o,o	56,8	33,81%	0,279
Base completa	WA-2-ZE	MLP-o,o	117,75	33,27%	0,2721
Base completa	WA-2-HA	SMOpoly	9,95	32,92%	0,2684
Base completa	WA-2-HA-ZE	RBF	616,99	32,38%	0,263
Base completa	WA-2-ZE	MLP-o	103,96	32,38%	0,2621
Base completa	WA-1-HA	MLP-o	88,29	32,02%	0,2588
Base completa	WA-1-ZE	MLP-o	65,44	32,02%	0,2581
Base completa	WA-2-HA	MLP-o	183,52	32,02%	0,2589
Base completa	WA-2-HA	MLP-o,o	185,57	32,02%	0,2589
Base completa	WA-1-HA	SMOpoly	5,24	30,77%	0,2454
Base completa	WA-1-HA	MLP-o,o	96,35	30,23%	0,2397
Extremamente Densa	SOM - ZE	SMO rbf	0,72	30,00%	0,4603
Base completa	WA-1-ZE	MLP-o,o	81,16	29,34%	0,2299
Base completa	WA-2-HA	RBF	39,51	28,98%	0,2261
Base completa	SOM - HA-ZE	SMOrbf	11,97	28,26%	0,2215
Base completa	WA-1-HA	RBF	426,74	28,09%	0,2165
Base completa	SOM- HA	SMOrbf	25,82	22,54%	0,1603
Base completa	WA-2-ZE	SMOrbf	11,37	22,54%	0,1593

Base completa	WA-2-HA	SMOrbf	14,13	21,29%	0,1406
Base completa	WA-1-ZE	SMOrbf	19,07	19,32%	0,125
Base completa	WA-1-HA	SMOrbf	14,32	18,25%	0,1112
Base completa	SOM -ZE	SMO rbf	13,55	8,94%	0,0153

Legenda:

WA: Segmentação com transformada de Wavelets (índice indica quantidade de decomposições).

HA: Extração de características de textura com o descritor de Haralick.

ZE: Extração de características de forma da imagem pelo momento de Zernike.

SOM: Segmentação com mapas auto-organizáveis.

SMOpoly: Classificador SMO com kernel polinomial.

SMOrbf: Classificador SMO com kernel RBF.

APÊNDICE B – ARTIGO PUBLICADO NO XXIV CONGRESSO BRASILEIRO DE ENGENHARIA BIOMÉDICA

CLASSIFICAÇÃO DE MAMOGRAFIAS UTILIZANDO EXTRAÇÃO DE ATRIBUTOS DE TEXTURA E REDES NEURAIS ARTIFICIAIS

I. M. M. Fernandes e W. P. dos Santos

Departamento de Engenharia Biomédica, Universidade Federal de Pernambuco, Brasil
e-mail: bella.mfernandes@gmail.com

Resumo: O câncer de mama continua sendo um problema significativo de saúde pública. O exame de mamografia tem aumentado o prognóstico desta neoplasia. No entanto, as imagens de mamografia são de difícil interpretação, de forma que os sistemas CAD podem auxiliar o radiologista na sua leitura, aumentando a sensibilidade e a precisão do diagnóstico. Visamos desenvolver um sistema CAD para a detecção e classificação de lesões em mamografias. Utilizamos a base de dados IRMA, separada em quatro grupos organizados segundo a constituição do tecido mamário, obtendo-se classificadores específicos por tipo de tecido. Essa abordagem se justifica, pois os quatro tipos de mama estão comumente relacionados à faixa etária da paciente, uma informação facilmente disponível na prática clínica. Analisamos o poder de classificação de diversos classificadores baseados em redes neurais artificiais a fim de identificar a estrutura que melhor detecta e classifica lesões suspeitas em imagens de mamografia por tipo de mama, obtendo mais de 80% de instâncias corretamente classificadas.

Palavras-chave: Mamografia, sistema CAD, redes neurais artificiais.

Abstract: *The breast cancer continues to be a significant public health problem. Mammography has increased the prognosis of this neoplasia. However, the mammographic images are difficult to be read, so the CAD systems would help the radiologist to interpret those images, increasing the sensibility and accuracy of the diagnosis. The aim of this article is to develop a CAD system to detect and classify lesions in mammograms. We used the IRMA database, separated in four groups organized according to the mammary tissue composition, drawing specific classifiers by tissue type. This approach is justified because those four groups of breasts are usually attached to the age of the patient, an easily obtained information. We analyzed the classification strength of different*

classifiers based on artificial neural networks to help identifying the best structure to detect and classify suspicious lesion on mammograms, achieving more than 80% of instances well classified.

Keywords: *Mammogram, CAD system, artificial neural networks.*

Introdução

De acordo com o Instituto Nacional do Câncer (INCA), o câncer de mama é o segundo mais frequente no mundo e o mais comum entre as mulheres [1]. Apesar de não haver formas preconizadas de prevenção do câncer de mama, quanto mais cedo detectado e diagnosticado, maiores as chances de cura do paciente. No Brasil, a mamografia e o exame clínico das mamas (ECM) são os métodos indicados para o rastreamento na rotina da atenção integral à saúde da mulher.

De acordo com a literatura, a mamografia tem sensibilidade e especificidade variáveis entre 88% e 93,1%, e entre 85% e 94,2%, respectivamente [1]. A sensibilidade depende de fatores como tamanho e localização da lesão, densidade do tecido mamário, qualidade dos recursos técnicos e habilidade de interpretação do radiologista, sendo a especificidade igualmente dependente da qualidade da imagem [2]. De forma geral, a leitura e interpretação de imagens de mamografia são tarefas que exigem experiência do radiologista, e seu julgamento depende de treinamento, experiência e critérios subjetivos. Mesmo experts bem treinados apresentam uma taxa acertos de classificação de 65 à 75% entre observadores [3]. No entanto, a utilização desse exame como método de rastreamento reduz a mortalidade em 25% [1].

Dessa forma, sistemas Computacionais de Apoio ao Diagnóstico (*Computer Aided Diagnosis* – CAD) podem auxiliar os radiologistas na difícil tarefa de leitura e interpretação de imagens de mamografia, de forma a automatizar o processo de

interpretação de imagens médicas. De 65 a 90% das biopsias de tecidos com suspeita de câncer mostram-se benignas. Desta forma, é importante o desenvolvimento de sistemas CADs que permitam a distinção entre lesões malignas e benignas [2, 4].

Propomos neste artigo um sistema automatizado para a classificação de lesões a partir de imagens de mamografia, aplicando mapas de Kohonen para a segmentação da imagem, extração de atributos a partir dos momentos de Haralick e de Zernike e classificação utilizando redes neurais artificiais. A base de dados foi separada em quatro grupos de imagens, organizadas segundo o tipo de tecido mamário, visto que mamas de composições diferentes não são comparáveis e que a divisão por tipo de mama reduz a carga de processamento no classificador. O artigo é organizado de forma que é apresentada, na seção seguinte, uma breve descrição das etapas de um sistema CAD. Em seguida, exporemos os resultados encontrados e a conclusão.

Materiais e métodos

A combinação sistemas CAD com o conhecimento e expertise do radiologista pode melhorar significativamente a acurácia da detecção e classificação de lesões suspeitas em imagens de mamografia. De forma geral, a elaboração de sistemas CAD envolvem quatro grandes etapas: pré-processamento da imagem digitalizada, segmentação, extração e seleção de atributos, e classificação do tipo de lesão [4].

Utilizamos a base de dados IRMA (*Image Retrieval in Medical Applications*) [5, 6], elaborada a partir de um projeto desenvolvido pela Universidade de Aachen (RWTH Aachen), na Alemanha, que reúne imagens de diversas outras bases de dados, disponíveis para consulta na internet. Esta base de dados é formada pelas Regiões de Interesse (*Region of Interest*, ROI) de imagens de radiografias digitalizadas da mama, que foram classificadas por radiologistas e redimensionadas para imagens de tamanho 128 x 128 pixel. Ao todo, são 2.796 imagens, organizadas em 12 classes, sendo 2.576 provenientes da base DDSM (*Digital Database for Screening Mammography*), 1 da base LLNL (*Lawrence Livermore National Laboratory*), 69 da base RWTH e 150 proveniente da base MIAS (*Mammographic Image Analysis Society*).

Pré-processamento – Sabemos que as imagens de mamografia possuem baixo contraste, devido principalmente à composição do tecido mamário, isto é, mamografias não possuem bom contraste entre tecido glandular normal e tecidos com características de malignidade, devido à similaridade da atenuação dos raios-x desses dois tipos de tecido, principalmente em mulheres jovens com mamas densas. Do ponto de vista clínico, a

constituição do tecido mamário está relacionado à idade do paciente, uma vez que a substituição do tecido glandular por tecido adiposo ocorre de forma gradativa com a idade, sendo esta informação acessível ao especialista. Desta forma, a base de dados IRMA foi separada em quatro subgrupos, levando em consideração a constituição do tecido mamário: adiposo (I), fibroglandular (II), denso (III) e extremamente denso (IV) [7], obtendo-se assim classificadores específicos por tipo de mama.

Das 2.796 imagens disponíveis na base IRMA, utilizamos 1.416 amostras, separadas em imagens sem lesão e com lesão do tipo massa espiculada e massa circunscrita, por serem mais representativas em caráter de benignidade (massas circunscritas) e malignidade (massas espiculadas) [8]. Das imagens utilizadas, 355 são amostras de mamas adiposas, 414 são de mamas com tecido fibroglandular, 341 de mamas densas e 306 de mamas extremamente densas.

Segmentação– A segunda etapa da detecção de lesões utilizando sistemas CAD é separar as regiões suspeitas do tecido mamário, músculo e demais estruturas presentes na imagem mamográfica, para então extrair as regiões de interesse (ROI) e localizar as lesões. As regiões de suspeita de lesão são aquelas que são mais claras e brilhantes que seus arredores, com densidade uniforme, forma regular (variando em tamanho) e bordas difusas [3].

A segmentação de imagem pode ser realizada por diferentes técnicas, em sua maioria baseadas em descontinuidades e similaridades dos níveis de cinza da imagem.

Nesse artigo propomos a abordagem de segmentação utilizando mapas auto organizáveis de Kohonen (*Self-Organizing Maps*, SOM), que são redes de treinamento não supervisionadas e que, por isto, utilizam apenas os dados de entrada, sem necessidade de conhecimento prévio sobre os padrões, natureza e quantidades das regiões a serem segmentadas. O mapa SOM projeta espaços de alta dimensionalidade em espaços de baixa dimensionalidade, preservando a relação de vizinhança entre padrões. Após a segmentação usando SOM, redimensionamos a imagem para que a mesma enquadre a região lesionada, de forma que as bordas da lesão venham a se tornar a própria borda da imagem.

Extração de atributos– A terceira etapa da detecção de lesões é a extração e seleção de atributos, que podem ser calculados a partir de algumas características da ROI, como tamanho, forma, densidade e suavidade das bordas. Visto que o espaço de atributos é grande e complexo, devido à diversidade dos tecidos normais e das anormalidades, é necessário selecionar quais dos atributos extraídos são significantes para a detecção de lesões. Os atributos redundantes precisam ser eliminados para melhorar o desempenho do classificador, uma vez que o uso excessivo de

atributos pode aumentar a complexidade do classificador e reduzir sua performance [3].

A extração e seleção de atributos que serão utilizados para a classificação é um passo fundamental, visto que o bom desempenho do sistema CAD depende mais da otimização desta etapa do que do método de classificação [3]. De acordo com as características escolhidas, o espaço de atributos pode ser baseado em intensidade, na geometria ou na textura. Atributos extraídos das características dos níveis de cinza, forma e textura da lesão e de seus arredores podem ser expressos como descritores matemáticos e utilizados pelo classificador para distinção de lesões benignas e malignas [4, 3].

Os extratores de atributos baseados em textura são amplamente utilizados, permitindo o cálculo de entropia, correlação, variância, energia e outros [3]. Utilizamos a extração de características de textura de Haralick (HA), baseada na matriz de Co-Ocorrência de níveis de cinza (*GreyLevelCo-Occurrence Matrix*, GLCM), e os momentos de Zernike para extração de 32 atributos. Este último forma um mapeamento em um conjunto complexo de polinômios de Zernike, e estes são ortogonais entre si, de forma que os momentos de Zernike (ZE) possibilitam representar as propriedades da imagem sem redundância ou sobreposição de informação entre os momentos [8].

Ao final da etapa de extração dos atributos, realizamos um balanceamento linear das imagens, uma vez que a base IRMA possui quantidades diferentes de imagens sem lesão, com lesão benigna e lesão maligna. Desta forma, equilibramos o conjunto das imagens de forma que todos contenham 233 amostras para cada tipo de classe.

Classificação e avaliação dos resultados— Uma vez que os atributos relacionados às lesões são extraídos, eles são utilizados como entrada para um classificador que deverá realizar a tarefa de classificação dos dados em três tipos de classes (normal, benigno e maligno). Classificadores como discriminantes lineares e redes neurais artificiais (*Artificial Neural Networks*, ANN) possuem boa performance em classificação de massas [3]. As ANNs utilizadas possuem uma camada de entrada com 32 nodos, que são referentes à quantidade de atributos extraídos da imagem, a camada de saída com os 3 nodos possíveis de classificação e uma ou mais camadas escondidas entre elas. Utilizamos amostras da base de dados para realizar o treinamento da rede, com informações de entrada e saída conhecidas, necessitando determinar os pesos dos nodos da camada escondida. Utilizamos para o treinamento o método de validação cruzada por *k-folds*, com *k* igual a 10. Uma vez determinado os pesos dos neurônios e treinada a rede, a mesma está pronta para realizar a classificação. De acordo com a literatura, os três tipos de classificadores ANN mais utilizados em aplicações de processamento de

imagens de mamografia são redes neurais de três camadas com *backpropagation*, redes de função de base radial (*Radial BasisFunction*, RBF) e perceptrons de múltiplas camadas (*MultiLayerPerceptron*, MLP) [3, 9].

Realizamos neste artigo um comparativo entre três classificadores a fim de verificar o poder de classificação dos mesmos. Além dos classificadores RBF e MLP citados anteriormente, ensaiamos também com o classificador de Otimização Sequencial Mínima (*SequentialMinimalOptimization*, SMO), utilizado principalmente em treinamentos de Máquinas de Vetores de Suporte (*Support Vector Machine*, SVM). A etapa de classificação foi realizada com a ferramenta Weka 3.6, disponível gratuitamente. Utilizamos a configuração da rede RBF conforme disponível na biblioteca do Weka 3.6. Para o classificador SMO, testamos duas variações do kernel, o polinomial e o RBF (SMO-KPM e SMO-KRBF). Para o classificador MLP, testamos seu poder de classificação com 1, 2 e 3 camadas escondidas, cada uma com a quantidade de neurônios definida como a metade da soma do número de classes e atributos, neste caso 18 neurônios em cada camada escondida.

Resultados e Discussão

A partir dos gráficos de 1 a 4, verificamos que a classificação, para os quatro subgrupos da base IRMA, baseados na composição do tecido mamário, apresentou melhores resultados com o classificador MLP. Uma vez que o aumento do número de camadas aumenta a complexidade da rede, o tempo de processamento e custo computacional, podemos considerar que o classificador com melhor custo é a rede MLP com apenas uma camada escondida (MLP-1), que obteve uma média de acertos de 84,62% para a segmentação da base utilizando os extratores de atributos de Haralick e Zernike e 81,37% para segmentação realizada unicamente com Zernike. Notamos também que a divisão da base IRMA em 4 subgrupos, organizados segundo o tipo de mama, reduziu a carga sobre o classificador, o que mostrou-se favorável à classificação.

No entanto, é importante levar em consideração o custo computacional envolvido para obtenção dos resultados. Conforme visto na tabela 1, a classificação do grupo no qual foi utilizado conjuntamente os momentos de Haralick e de Zernike para extração de atributos resultou num tempo de construção do classificador maior que no caso em que usamos apenas os atributos extraídos pelos momentos de Zernike. Vimos ainda que ambos os métodos realizam uma boa segmentação, permitindo à rede MLP classificar bem as imagens de entrada. Desta forma, notamos que os momentos de Zernike são bastante promissores para a extração

de características de imagens de mamografia, reduzindo o tempo de processamento e custo computacional necessário para sua implementação.

Tabela 1: Tempo médio, em segundos, para a construção do classificador pelo Weka 3.6.

Classificador	HA-ZE	ZE
RBF	1,11	0,81
SMO-KPM	0,49	0,37
SMO-KRBF	1,07	1,01
MLP-1	266,52	36,71
MLP-2	406,58	59,26
MLP-3	547,38	82,38

Gráfico 1: Instâncias corretamente classificadas para subgrupo de mamas adiposas (I).

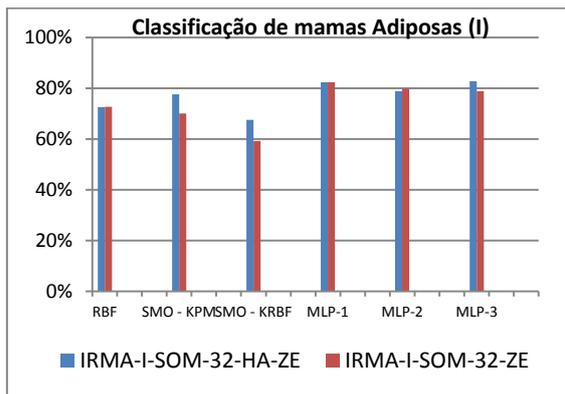


Gráfico 2: Instâncias corretamente classificadas para subgrupo de mamas fibroglandulares (II).

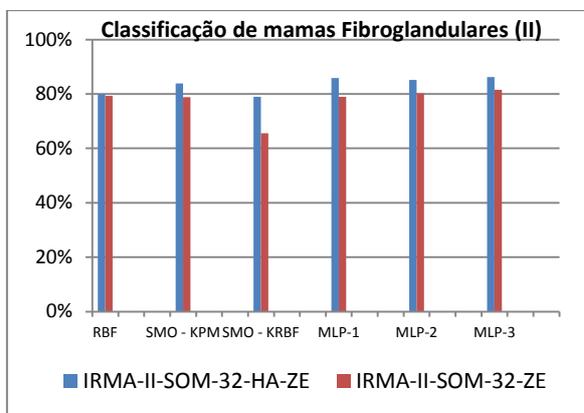


Gráfico 3: Instâncias corretamente classificadas para subgrupo de mamas densas (III).

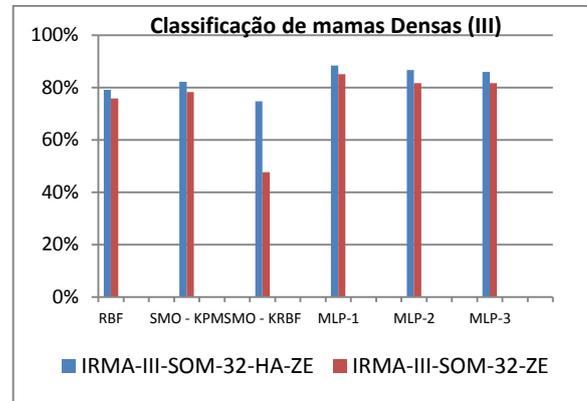
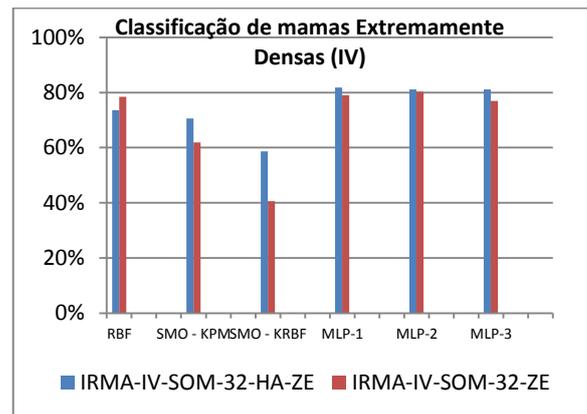


Gráfico 4: Instâncias corretamente classificadas para subgrupo de mamas extremamente densas.



Conclusão

Verificamos que, para aplicações em detecção de lesões em imagens de mamografia, realizada de forma independente para cada subgrupo da base de dados, o método apresentado, com determinação de ROIs e segmentação por mapas SOM e extração de atributos pelos momentos de Zernike e classificação por redes MLP, aplicado a cada tipo de mama separadamente resultou numa taxa de classificação superior à 80%. Esse estudo permitiu verificar o potencial de redes MLP para classificação de massas em mamografia e a vantagem na aplicação no classificador de forma específica por tipo de mama. Acreditamos que essas redes podem ser otimizadas para melhorar seu poder de classificação e acurácia. Pretendemos alcançar, em estudos futuros, uma taxa de classificação de 90%.

Referências

- [1] Ministério da Saúde. Instituto Nacional de Câncer. “Mamografia: da prática ao controle”. Rio de Janeiro: INCA, 2007.
- [2] Ganesan K, Acharya UR, Chua CK, Min LC, Abraham KT, e Kwan-HoongNg. Computer-Aided Breast Cancer Detection Using Mammograms: A Review. IEEE Reviews in Biomedical Engineering, vol. 6, 2013
- [3] Cheng HD, Shi XJ, Min R, Hu LM, Cai XP e Du HN. Approaches for automated detection and classification of masses in mammograms. Pattern Recognition, vol. 39, 2006.
- [4] Shanthu S, Bhaskaran AV. A Novel Approach for detecting and Classifying Breast Cancer. In: International Journal of Intelligent Information Technologies (IJIT) 2013; 9(1), 21-39.
- [5] de Oliveira JE, Machado AM, Chavez GC, Lopes AP, Deserno TM, Araújo A de A. MammoSys: A content-based image retrieval system using breast density patterns. Comput Methods Programs Biomed 2010; 99(3): 289-97.
- [6] de Oliveira JEE, Soiron M, DesernoTM. Texture patterns extracted from digitizes mammograms of different BI-RADS classes. ProcSpie 2012; 8315: 217-9.
- [7] Vállez N, BuenoG, Déniz O, Esteve P, Rienda MA, Pastor C. Automatic Breast Tissue Classification Based on BIRADS Categories. Digital Mammography/ IWDM, volume 6136 of Lecture Notes in Computer Science, page 259-266. Springer, (2010).
- [8] Tahmasbi A, Saki F, Shokouhi SB. Classification of benign and malignant masses based on Zernike moments. Computer in Biology and Medicine, vol 41, 2011.
- [9] Mustra M., Grgic M., Delac, K. Breast Density Classification Using Multiple Feature Selection. Automatika, vol. 53, N°4, 2012.