JAMILSON BISPO DOS SANTOS

Pesquisa de Similaridades em Imagens Mamográficas com Base na Extração de
Características

São Paulo
2013

JAMILSON BISPO DOS SANTOS

Pesquisa de Similaridades em Imagens Mamográficas com Base na Extração de
Características

Tese apresentada à Escola Politécnica
da Universidade de São Paulo para
obtenção do título de Doutor em
Ciências.

Área de Concentração:
Sistemas Digitais

Orientador:
Prof. Dr. Jorge Rady de Almeida Junior

São Paulo
2013

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, de junho de 2013.

Assinatura do autor _____

Assinatura do orientador _____

FICHA CATALOGRÁFICA

Santos, Jamilson Bispo dos

Pesquisa de similaridade em imagens mamográficas com base na extração de características / J.B. dos Santos. -- versão corr. -- São Paulo, 2013.

105 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Computação e Sistemas Digitais.

1.Imageamento (Bioengenharia) 2.Mamografia I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Computação e Sistemas Digitais II.t.

AGRADECIMENTO

Ao Professor Dr. Jorge Rady de Almeida Jr. pelo crédito, paciência e empenho em mim depositados.

Ao pessoal do GAS que sempre ajudou com dicas e sugestões.

A professora Beatriz pelas informações e dicas importantes na conclusão do trabalho

A minha Esposa Luciana e aos meus filhos Mariana e Joao Vitor pela compreensão.

A minha mãe Maria e meu pai José que sempre me apoiaram.

A todos que direta ou indiretamente me apoiaram nesta caminhada e aos que colaboraram com dicas, conselhos e ideias para a execução deste trabalho.

RESUMO

Este trabalho apresenta uma estratégia computacional para a consolidação do treinamento dos radiologistas residentes por meio da classificação de imagens mamográficas pela similaridade, analisando informações dos laudos realizados por médicos experientes, obtendo os atributos extraídos das imagens médicas. Para a descoberta de padrões que caracterizam a similaridade aplicam-se técnicas de processamento digital de imagens e de mineração de dados nas imagens mamográficas. O reconhecimento de padrões tem como objetivo realizar a classificação de determinados conjuntos de imagens em classes. A classificação dos achados mamográficos é realizada utilizando Redes Neurais Artificiais, por meio do classificador *Self-Organizing Map* (SOM). O presente trabalho utiliza a recuperação de imagens por conteúdo (CBIR- *Content-Based Image Retrieval*), considerando a similaridade em relação a uma imagem previamente selecionada para o treinamento. As imagens são classificadas de acordo com a similaridade, analisando-se informações dos atributos extraídos das imagens e dos laudos. A identificação da similaridade é obtida pela extração de características, com a utilização da transformada de *wavelets*.

Palavras-chave: *Data Mining*. Processamento Digital de Imagens. Reconhecimento de padrões. CBIR.

ABSTRACT

This work presents a computational strategy to consolidate the training of residents radiologists through the classification of mammographic images by similarity, analyzing information from reports made by experienced physicians, obtaining the attributes extracted from medical images. For the discovery of patterns that characterize the similarity apply techniques of digital image processing and data mining in mammographic images. Pattern recognition aims to achieve the classification of certain sets of images in classes. The classification of mammographic is performed using Artificial Neural Networks, through the classifier Self-Organizing Map (SOM). This work uses the image retrieval (CBIR-Content-Based Image Retrieval), considering the similarity in relation to an image already selected for training. The images are classified according to similarity, analyzing attribute information extracted from the images and reports. The identification of similarity was obtained by feature extraction, using the technique of wavelet transform.

Keywords: Data Mining. Digital Image Processing. Pattern Recognition. CBIR.

LISTA DE ILUSTRAÇÕES

Figura 1 - Estrutura da mama.....	22
Figura 2 - Matrizes de vizinhança	29
Figura 3 - Matriz de definição de verdadeiro e falso, positivo e negativo (SANTOS, 2006). .	31
Figura 4 - Características extraídas por sistema de recuperação por conteúdo.....	39
Figura 5 - Processo de Extração de Características de uma imagem.....	41
Figura 6 - Rede SOM com 16 neurônios na camada de saída (mapa 4 X 4).....	44
Figura 7 - Fundamentações Conceituais.....	45
Figura 8 - Premissas e proposições associadas deste trabalho.....	46
Figura 9 - Sistema Mammo.....	50
Figura 10 - Arquitetura da realização de laudo pelo médico experiente com o uso do MammoAdviser	52
Figura 11 - Arquitetura da realização de laudo pelo residente com o uso do MammoAdviser	53
Figura 12 - Modelo Entidade Relacionamento	55
Figura 13 – Módulo de segmentação da imagens MammoSegm	56
Figura 14 - Módulo de extração e classificação de imagens, MammoSort.....	57
Figura 15 - Esquema geral do método desenvolvido	58
Figura 16 - Esquema de funcionamento da etapa de pré-processamento	58
Figura 17 - Esquema de decomposição por análise em multiresolução	59
Figura 18 - Esquema de avaliação do conjunto de coeficientes (SILVA, 2009)	60
Figura 19 - Modulo de visualização de imagens segmentadas, MammoViewer	61
Figura 20 - Ilustração do vetor de características e informações do processo SOM.....	64

LISTA DE TABELAS

Tabela 1 - Classe BI-RADS 0.....	65
Tabela 2 - Classe BI-RADS 1.....	65
Tabela 3 - Classe BI-RADS 2.....	65
Tabela 4 - Classe BI-RADS 3.....	66
Tabela 5 - Classe BI-RADS 4.....	66
Tabela 6 - Classe BI-RADS 5.....	66
Tabela 7 - Classe BI-RADS 6.....	66

LISTA DE SIGLAS

AMR	- Análise de Multiresolução
BI-RADS™	- Breast Imaging Reporting and Data System
CAD	- <i>Computer-Aided Diagnosis</i>
CBIR	- <i>Content Based Image Retrieval</i>
CCD	- Crânio Caudal Esquerda
CCE	- Crânio Caudal Esquerda
DICOM	- <i>Digital Imaging and Communication in Medicine</i>
DTF	- <i>Discrete Fourier Transform</i>
DW	- Data Warehouse
DM	- Data Mining
DMa	- Data Mart
EM	- Eletromagnético
EMD	- <i>Empirical Mode Decomposition</i>
FN	- Falso Positivo
FP	- Falso Negativo
KDD	- <i>Knowledge Discovery in Databases</i>
HIS	- <i>Hospital Information System</i>
HHT	- Transformada Hilbert-Huang
HSA	- Hilbert Spectral Analysis
IM	- Imagens Médicas
IA	- Inteligência Artificial
INCA	- Instituto Nacional do Câncer
MI	- Momentos Invariantes
MLD	- Médio Lateral Direita
MLE	- Médio Lateral Esquerda
MLP	- <i>Multilayer Perceptron's</i>
OLAP	- On Line Application Process
OLTP	- On Line Transaction Process
QualIM	-Laboratório de Qualificação de Diagnóstico por Imagem Médicas
RB	- Rede Bayesiana

RI	- Regiões de Interesse
RIS	- <i>Radiology Information System</i>
RM	- Ressonância Magnética
RNA	- Redes Neurais Artificiais
ROC	- <i>Receiver Operating Characteristics</i>
SBD	- Sistema de Banco de Dados
SGBD	- Sistema de Gerenciador de Banco de Dados
SISMAMA	- Sistema de Informação do Câncer de Mama
SOM	- <i>Self-Organizing Map</i>
UNIFESP	- Universidade Federal de São Paulo
US	- Ultrassonografia
VN	- Verdadeiro Negativo
VP	- Verdadeiro Positivo
TC	- Tomografia Computadorizada
TW	- Transformada <i>Wavelet</i>
WFT	- <i>Windowed Fourier Transform</i>
WWW	- <i>World Wide Web</i>

LISTA DE EQUAÇÕES

Equação 1.....	32
Equação 2.....	33
Equação 3.....	33
Equação 4.....	33
Equação 5.....	33

SUMÁRIO

1.	Introdução	14
1.1.	Motivação e Justificativa	16
1.2.	Objetivos	17
1.3.	Estrutura do trabalho	18
2.	Câncer, Imagem Mamográfica, e a Interpretação Diagnóstica.	19
2.1.	Câncer	19
2.2.	A Estrutura da Mama	21
2.3.	Mamografia	22
2.4.	Padronização BI-RADS®	23
2.4.1.	Nódulo	24
2.4.2.	Calcificações	26
3.	Imagem Digital	28
3.1.	Imagens Médicas na Medicina Moderna	30
3.2.	Terminologia na Avaliação de Imagens Médicas	30
3.3.	Análise de Imagens Médicas	31
3.4.	Crítérios de Avaliação de Imagens Médicas	32
3.5.	Análise ROC	34
3.6.	Séries Temporais	35
3.7.	Análise de Fourier	35
3.8.	Transformada de Fourier	36
3.9.	Transformada <i>Wavelet</i>	37
3.10.	Análise dos momentos	37
3.11.	Extração e Análise de Atributos	38
3.12.	Extratores de atributos	41
3.13.	Mapa Auto-Organizável	42

4.	Proposta para o Reconhecimento de Imagens Mamográficas	45
4.1.	Premissas Adotadas	45
4.1.1.	Premissa 1	47
4.1.2.	Premissa 2	48
4.1.3.	Premissa 3	48
4.1.4.	Premissa 4	49
4.2.	Arquitetura Geral da Proposta	49
4.3.	Detalhamento da Proposta	51
4.3.1.	Modulo MamoSegm	55
4.3.2.	Modulo MammoSort	56
4.3.2.1.	Etapa de Pré-Processamento	58
4.3.2.2.	Etapa Wavelets	59
4.3.2.3.	Etapa SOM	60
4.3.2.4.	Pós-Processamento	61
4.3.3.	Modulo MammoViewer para visualização das imagens e dos laudos	61
4.4.	Formato e utilização do vetor de características	62
4.4.1.	Base de dados para os experimentos de categorização e classificação	62
4.4.2.	Categorização da imagem mamográfica com o Mapa SOM	64
4.5.	Ambiente	64
4.6.	Avaliação do processo de classificação das imagens mamográficas	65
5.	Considerações Finais	67
5.1.	Trabalhos Futuros	68
	Referencias Bibliográficas	69
	Apêndice A – Especificação Funcional	76

1. Introdução

A partir da descoberta dos Raios-X, pelo físico alemão Wilhelm Conrad Röntgen em 1895, foi possível visualizar as estruturas internas do corpo humano (GLASSER, 1993). O uso dos Raios-X marcou o início da Radiologia Médica (CHO; JONES; SINGH, 1993 apud Santos, 2006 p. 1).

Segundo Lehmann (LEHMANN; MEINZER; TOLDOXDORF, 2004) a evolução dos sistemas geradores de imagens e das tecnologias de computação muito contribuíram para o desenvolvimento na área da radiologia. As modalidades Tomografia Computadorizada (TC) e ressonância magnética (RM) geram imagens no formato digital, as quais podem ser tratadas como qualquer arquivo digital (SIEGEL; KOLODNER, 2003). As imagens podem ser vistas em monitores de computador com a possibilidade da utilização de ferramentas computacionais que realizam atividades de pós-processamento.

O processamento da imagem no formato digital (GONZALES; WOODS, 2010) é um tema abordado em trabalhos em diversas áreas, como por exemplo: Informática Médica e Engenharia Biomédica. Um dos problemas identificados no processamento digital dessas imagens consiste na dificuldade de comparar o desempenho das diferentes metodologias (SANTOS; FURUIE, 2005).

As imagens médicas geradas no formato digital podem ser armazenadas nos diversos tipos de mídias disponíveis ou ainda serem impressas para a execução do diagnóstico. O aumento no uso de imagens pelos radiologistas contribuiu para a criação do conceito de Sistema de Arquivamento e Comunicação Imagem (*Picture Archival and Communication System* - PACS) (SIEGEL; KOLODNER, 2003).

O sistema PACS realiza o armazenamento e visualização de imagens médicas em um ambiente integrado, permitindo a distribuição das imagens em uma rede digital (MÜLLER *et al.*, 2004; FURUIE *et al.*, 2003).

O treinamento na leitura e interpretação de uma imagem médica, pelos residentes, é uma atividade que está sendo pesquisada no laboratório QuaLIM da UNIFESP e tem como um dos objetivos preparar os residentes na realização de diagnósticos confiáveis. Um dos problemas identificados no processo de treinamento

é a dificuldade dos residentes em detectar e diferenciar o seu objeto de diagnóstico com outros que apresentam patologias similares.

A recuperação de imagens por conteúdo (do inglês *Content-Based Image Retrieval – CBIR*) é uma técnica que tem sido bastante pesquisada nos últimos anos (MÜLLER *et al.*, 2004). A utilização dessas técnicas tem como um dos objetivos permitir a mineração de dados (*Data Mining*) para classificar as imagens recuperadas pela similaridade.

Os sistemas de recuperação de imagens por conteúdo utilizam indexação a partir da avaliação do conteúdo visual das imagens (MÜLLER *et al.*, 2004).

Data Mining é uma tecnologia usada para revelar informações estratégicas em grandes quantidades de dados. É usada em diversas áreas, tais como: análise de riscos, marketing direcionado, controle de qualidade, análise de dados científicos, diagnósticos médicos, etc. *Data Mining* compreende um processo automatizado de captura e análise de enormes conjuntos de dados, para então extrair um significado. Esta técnica está sendo utilizada para descrever características do passado, assim como prever tendências para o futuro (FIGUEIRA, 1998).

Na área médica, diversos estudos estão sendo realizados com a utilização de *Data Mining* como ferramenta no auxílio ao diagnóstico. Em Antoine (ANTOINE; ZAIANE; COMAN, 2001) foram utilizadas técnicas de *Data Mining* para a detecção de tumores em imagens médicas.

Para a elaboração dos laudos, a partir das imagens, destacam-se os seguintes aspectos:

- Dificuldade para elaborar comparações com outras imagens médicas;
- Qualidade e/ou validade dos conjuntos de dados que tiveram dois ou mais laudos, visando a geração de um laudo considerado padrão (caso em que os médicos convergiram para um mesmo diagnóstico). Caso exista divergência, se possível, novo laudo deverá ser realizado por um terceiro médico. Caso a divergência persista, as imagens em questão devem ser retiradas do conjunto de imagens a ser utilizado para treinamento;
- Possibilidade de utilizar métodos de processamento automáticos;
- Dificuldade no acompanhamento da evolução do residente no treinamento;

- Dificuldade na utilização de imagens previamente selecionadas e de acordo com a complexidade necessária para o estágio de aprendizado do residente.

1.1. Motivação e Justificativa

O desenvolvimento de sistemas computacionais com o objetivo de buscar indícios que podem estar relacionados à presença de tumores está sendo realizado em diversos centros de pesquisas. Entre 10% e 30% de mulheres que tiveram câncer de mama e foram submetidas à mamografia tiveram laudos negativos em seus mamogramas (GIGER, 2000).

Para os especialistas, diagnosticar lesões em uma mamografia é uma atividade de difícil execução, uma vez que as imagens mamográficas possuem baixo contraste, além de possuir outros fatores que influenciam em sua qualidade. A mamografia é considerada o método mais eficaz para detecção do câncer de mama. A dificuldade na classificação de lesões como malignas ou benignas, a partir das imagens mamográficas, é relativamente baixa (HADJISKI, et al., 1999).

A classificação das imagens digitais pode ser realizada por meio do reconhecimento e descoberta de padrões. A área de reconhecimento de padrões visa a descoberta automática de regularidade nos dados por meio de algoritmos computacionais e, a partir dessa regularidade, realizar ações para classificar os dados em diferentes categorias (BISHOP, 2006).

Com o objetivo de aprimorar os resultados obtidos na classificação, utiliza-se a abordagem de aprendizagem de máquina, na qual um conjunto de dados de treinamento é utilizado para elaborar critérios para a tomada de decisão. O aprendizado pode ser supervisionado ou não. No aprendizado supervisionado, os dados de treinamento e as classes são conhecidos, enquanto que no aprendizado não supervisionado o objetivo é descrever as associações e a quantidade de classes resultantes não são conhecidas (DUDA; HART; STORK, 2001) (THEODORIDIS; KOUTROUMBAS, 2009).

Outro fator de motivação foi o de dar continuidade ao trabalho, com a mesma finalidade, que estava sendo desenvolvido por Lasca, Patrocínio, Rady e Bitelli, (LASCA; PATROCÍNIO; RADY; BITELLI, 2008).

1.2. Objetivos

O objetivo deste trabalho consiste na proposição de uma metodologia que permita efetuar o reconhecimento de padrões em imagens mamográficas. Tal objetivo é cumprido por meios da identificação e classificação de imagens mamográficas similares, utilizando técnicas de processamento de imagens, envolvendo a Transformada *Wavelet*. Na descoberta de padrões são ainda aplicadas técnicas de mineração de dados, especificamente o SOM (*Self-Organizing Map*).

Como objetivo complementar está o desenvolvimento de um sistema computacional de reconhecimento dos padrões, cuja principal finalidade é servir como plataforma de treinamento dos residentes em radiologia na interpretação e comparação de imagens mamográficas.

O reconhecimento de padrões visa realizar a classificação de determinados conjuntos de dados em classes, de acordo com similaridade. A classificação é realizada utilizando Redes Neurais Artificiais, por meio do SOM (*Self-Organizing Map*) (SILVA, 2009). A precisão da classificação é verificada utilizando o cálculo da área sob a curva *Receiver Operating Characteristics* (ROC). O sistema permite a recuperação de imagens por conteúdo (CBIR- *Content-Based Image Retrieval*), considerando a similaridade em relação a uma imagem previamente selecionada para o treinamento do residente.

A identificação da similaridade é obtida com a extração de características de textura e forma, por meio da transformada *wavelet*. Uma análise de diversas transformada *wavelets* é feita, com o objetivo de verificar qual a melhor transformada *wavelet* para essa tarefa, e a integração do método proposto com trabalho realizado por Silva (SILVA, 2009).

O laudo efetuado pelo residente e a leitura de imagens similares deverá ser incorporada ao treinamento dos residentes em radiologia, sendo que a seleção dos casos clínicos poderá ser feita de acordo com o grau de complexidade. O treinamento de cada residente deverá acompanhado, visando identificar graus de dificuldade à sua avaliação, de acordo com a evolução no treinamento.

1.3. Estrutura do trabalho

A partir das considerações expostas, este trabalho foi dividido em duas partes: a primeira consta de uma revisão bibliográfica a respeito da base teórica, enquanto que a segunda parte contém detalhamento sobre a técnica proposta, assim como os resultados obtidos e discussões sobre os mesmos. Segue um breve resumo de cada capítulo, com o objetivo de facilitar a localização dos assuntos abordados:

No Capítulo 2 são introduzidas informações relativas à descrição do exame mamográfico, o que se busca ao analisar em uma mamografia, a padronização de laudos mamográficos, além de conceitos básicos sobre imagem digital.

No Capítulo 3 são introduzidas as técnicas de extração de características existentes nas imagens médicas digitais com o uso da transformada *wavelet* e suas aplicações na caracterização de imagens.

No Capítulo 4 apresenta-se a proposta central desta tese, composta pela proposição de uma técnica que permita efetuar o reconhecimento de padrões em imagens mamográficas, bem como pelo sistema computacional desenvolvido..

No Capítulo 5 são apresentados os principais resultados obtidos e as considerações finais, assim como os trabalhos futuros.

2. Câncer, Imagem Mamográfica, e a Interpretação Diagnóstica.

Neste capítulo são apresentados os conceitos básicos ligados ao câncer, notadamente ao câncer de mama e suas imagens mamográficas.

2.1. Câncer

O câncer de mama representa uma das principais causas de morte em mulheres nos países ocidentais, segundo o Instituto Nacional de Câncer (INCA, 2010).

Em alguns países estão sendo desenvolvidos programas de rastreamento periódicos com o objetivo de detectar precocemente a doença. A chance de cura do câncer aumenta em até 40% quando o mesmo é descoberto precocemente (DENGLER; BEHRENS; DESAGA, 1993).

Segundo dados do INCA, (INCA, 2011c) em 2008 ocorreram mais de 11.860 óbitos no Brasil, causados pelo câncer de mama, sendo 11.735 mulheres e 125 homens. O número de casos de câncer de mama no Brasil em 2010 foi de 49.240.

Os tumores de mama geralmente são descobertos pela própria mulher, apalpando suas mamas de forma acidental. Nesses casos os tumores já estão em estágio avançado, dificultando assim o seu tratamento segundo o INCA (INCA, 2000). O tamanho do tumor no diagnóstico inicial influenciará diretamente na sobrevida do paciente. A identificação do tumor no estágio inicial pode evitar cirurgias mais agressivas (BAUER; IGOT; LE, 1980).

Os procedimentos básicos no rastreamento e no diagnóstico do câncer de mama são (NUNES, 2001):

- Autoexame das mamas;
- Exame clínico das mamas;
- Mamografia.

Com o objetivo de eliminar dúvidas em diagnósticos, além dos procedimentos citados, podem ser utilizados os seguintes exames:

- Ultrassonografia;
- Exame citológico;
- Exame histopatológico (biopsia).

O uso da mamografia de rotina auxiliou na redução de cerca de 20% na taxa de mortalidade pelo câncer, nos Estados Unidos na década de 1990 (AGUILLAR; BAUAD; MARANHÃO, 2009).

Segundo Aguillar, Buad e Maranhão (AGUILLAR; BAUAD; MARANHÃO, 2009) a mamografia é um dos exames de mais difícil interpretação, pois o câncer de mama é uma doença heterogênea, apresentando diversas formas de visualização da glândula mamária, em cada mulher. A detecção do câncer de mama em um estágio precoce reduz, comprovadamente, a mortalidade para este tipo de câncer.

A busca de mecanismos automatizados que possam auxiliar no diagnóstico precoce do câncer incentivou pesquisadores a atuarem nessa área. Surgem, nos anos 1980, os conceitos de *Computer Diagnosis and Automated Diagnoses* (DOI, 2004).

O desenvolvimento de esquemas computadorizados de auxílio ao diagnóstico (*CAD – Computer-Aided Diagnoses*) começou duas décadas atrás, sendo utilizado na tentativa de detectar e classificar lesões em imagens radiográficas (LI et al., 2005).

A utilização do esquema CAD pelos radiologistas visa obter uma segunda opinião na tentativa de detecção do câncer (DOI, 2004). Estudos realizados constataram que, com a utilização do esquema CAD houve um aumento na taxa de reavaliação de 6,5% a 7% e um aumento de 19,5% na taxa de detecção de câncer (FREER; ULISSEY, 2001).

Segundo Patrocínio (PATROCÍNIO, 2004), para que se possa classificar um achado mamográfico com a utilização de um esquema CAD é importante considerar: a qualidade e as técnicas de aquisição de imagem, o pré-processamento, o

processamento das imagens, a identificação das regiões de interesse, a extração das características e atributos de imagem, a seleção das características que representam cada padrão e a utilização de técnicas de classificação que apresentem resultados satisfatórios.

2.2. A Estrutura da Mama

Para discorrer sobre imagens mamográficas é importante entender a estrutura e composição da mama.

Conforme ilustrado na Figura 1, a mama contém glândulas secretoras que se estendem da 2ª à 6ª costela e do osso esterno à linha axilar média, sendo formada por lóbulos, ductos e estroma. A glândula mamária, gordura, vasos sanguíneos, nervos e linfáticos compõem o volume da mama. O parênquima, tecido mamário, é composto por aproximadamente 20 a 30 glândulas alveolares compostas ou lobos, tendo como principal função a secreção do leite (INCA 2011). Os lobos são constituídos por lóbulos, menores, onde é produzido o leite. O leite flui dos lóbulos até o mamilo através de dutos denominado de galactóforos (POP, 2011).

Separando a glândula da pele da parede torácica está o estroma, o qual é constituído por tecido adiposo (gordura) e tecido conjuntivo que o rodeia, tendo por objetivo suportar os ductos, lóbulos, vasos sanguíneos e linfáticos (POP, 2011).

As mulheres mais jovens apresentam mama com maior quantidade de tecido glandular, tornando esse órgão mais denso e firme. A substituição progressiva do tecido mamário por tecido gorduroso acontece ao se aproximar a menopausa (INCA, 2000).

Figura 1 - Estrutura da mama

2.3. Mamografia

Para que seja formada a imagem mamográfica é necessária a interação dos fótons de raios-X com a matéria. No caso do corpo humano, os fótons são absorvidos, de forma diferente pelos ossos, músculos e gordura, levando a diferentes atenuações e diferentes intensidades de radiação no plano da imagem. A formação de contrastes é obtida pelas diferentes intensidades, as quais permitem a identificação da estrutura anatômica (NUNES, 2001).

Conforme Patrocínio (PATROCÍNIO, 2004) a identificação de estruturas de interesse depende de fatores físicos e técnicos, como a iluminação, que deve ser adequada, além do que o aparelho de imagem mamográfica que deve estar devidamente calibrado.

Na mamografia, apesar de a sensibilidade ser alta, ocorre a perda de 10% a 15% dos casos de câncer que são identificáveis pelo exame clínico (KOPANS, 2000).

A mamografia inclui a projeção craniocaudal e a projeção oblíqua. A projeção craniocaudal permite observar a maior parte do tecido mamário com exceção da porção alta dos quadrantes superiores e o prolongamento axilar. Já a projeção oblíqua externa inclui toda a mama, o prolongamento axilar e a parte da região axilar (PINOTTI; RICCI; PINOTTI, 2006).

Segundo Aguillar, Buad e Maranhão (AGUILLAR; BAUAD; MARANHÃO, 2009) a mamografia é considerada um dos exames de mais difícil interpretação pelas diferentes formas de apresentação da glândula mamária, em cada mulher. Para permitir um diagnóstico adequado é necessário garantir a qualidade da imagem, a qual pode ser definida como a habilidade da mesma em indicar mudanças muito pequenas de contraste no tecido leve e revelar a existência de calcificações associadas à patologia (KEMP et al. 2003).

O baixo contraste e o tamanho reduzido das características de malignidade de achados, como microcalcificações, são considerados fatores que dificultam a interpretação das imagens mamográficas (SHEN, 1993).

2.4. Padronização BI-RADS®

Com o objetivo de padronizar o laudo mamográfico, em 1995, o modelo de normalização de laudos mamográficos denominado BI-RADS™ (*Breast Imaging Reporting and Data System*) foi adotado pelo Colégio Americano de Radiologia (ACR, 2003).

Por meio de estudo realizado no período de agosto de 1999 a novembro de 2001 foi possível identificar e subdividir os achados mamográficos em (VIEIRA; TOIGO, 2002):

- Classe 0: avaliação incompleta; uma nova avaliação é necessária. Geralmente utilizado em exames de rastreamento;
- Classe 1: Achados mamográficos negativos; mamografia normal;
- Classe 2: Achados mamográficos benignos; calcificações vasculares, calcificações cutâneas, calcificações com centro lucente, fibroadenoma calcificado, cisto oleoso (esteatonecrose), calcificações de doença secretória

(“plasma cell mastitis”), calcificações redondas (acima 1 mm), calcificações tipo “Milk of calcium”, fios de sutura calcificados, linfonodo intramamário;

- Classe 3: Achados mamográficos provavelmente benignos; nódulo de densidade baixa, contorno regular, limites definidos e dimensões não muito grandes, calcificações monofórmicas e isodensas sem configurar grupamento com características de malignidade;
- Classe 4: Achados mamográficos suspeitos; nódulo de contorno bocelado ou irregular e limites pouco definidos, micro calcificações com pelomorfismo incipiente, densidade assimétrica, algumas lesões espiculadas. É uma classe subdivida em:
 - Classe 4A: Baixa suspeição, porém indicada a intervenção;
 - Classe 4B: Suspeição intermediária necessária correlação com achados radiológicos;
 - Classe 4C: Suspeição moderada nódulos irregulares;
- Classe 5: Achados mamográficos altamente suspeitos; nódulo denso e espiculado, micro calcificações pleomórficas agrupadas, micro calcificações pleomórficas seguindo trajeto ductal, ramificadas, tipo letra chinesa;
- Classe 6: Malignidade confirmada: lesões com biopsia realizada confirmando a malignidade.

Os fatores que podem dificultar a interpretação da mamografia são a composição tecidual da mama, o conhecimento da anatomia da mama e da patologia da glândula mamária (AGUILLAR; BAUAD; MARANHÃO, 2009).

Com o objetivo de definir o que são achados mamográficos expõe-se, a seguir, o conceito de nódulos e microcalcificações.

2.4.1. Nódulo

Nódulo é uma lesão que pode ser visualizada de duas formas diferentes. Se o mesmo aparecer em apenas uma incidência é chamado de densidade, enquanto não for confirmada, por outras imagens (KEMP; BARACAT; ROSTAGNO, 2003).

Se o nódulo se encontrar com suas margens obscurecidas pelo tecido fibroglandular em uma ou nas duas incidências pode ser difícil sua detecção. Nesta situação haverá necessidade de incidências mamográficas adicionais.

A caracterização de um nódulo inclui as seguintes características: forma, margem, densidade, tamanho, localização e achados associados (AGUILLAR; BAUAD; MARANHÃO, 2009).

Com relação à forma, a classificação adotada para os nódulos é a seguinte (KEMP; BARACAT; ROSTAGNO, 2003):

- Redonda: esférico, circular, glandular;
- Oval: elíptica;
- Lobular: contorno com ondulações;
- Irregular: forma não pode ser caracterizada;
- Distorção arquitetural: a arquitetura normal está distorcida, sem lesão definida visível.

No que se refere à margem, a classificação adotada é (KEMP; BARACAT; ROSTAGNO, 2003):

- Definida: circunscritas, margem bem definidas;
- Micro lobulada: semicírculo, produzindo pequenas ondulações;
- Encoberta: obscura, escondida, por sobreposição de tecido fibroglandular adjacente;
- Indefinida: a pouca definição das margens sugere infiltração pela lesão;
- Espiculada: linhas radiadas partindo das margens da lesão.

Finalmente, segundo a densidade, a classificação seguida é (KEMP; BARACAT; ROSTAGNO, 2003):

- Alta densidade;
- Baixa densidade;
- Conteúdo gorduroso.

2.4.2. Calcificações

As calcificações, como achados mamográficos, começaram a ser identificadas em 1951, quando Raul Leborgne descreveu que as calcificações estavam frequentemente associadas ao câncer de mama (AGUILLAR; BAUAD; MARANHÃO, 2009).

A mamografia é a única modalidade segura para a identificação de calcificações. Apesar das microcalcificações representarem 25% das lesões malignas, aproximadamente 70% a 80% dos procedimentos realizados em função de sua detecção revelam processos benignos (AGUILLAR; BAUAD; MARANHÃO, 2009)

As três etapas utilizadas na validação radiológica das calcificações são (AGUILLAR; BAUAD; MARANHÃO, 2009):

- Detecção: capacidade de encontrar as calcificações na mamografia;
- Diagnóstico: capacidade de classificar essas calcificações como benignas ou malignas;
- Decisão da conduta a ser sugerida no laudo: baseada nas etapas anteriores.

Segundo Aguillar, Buad e Maranhão (AGUILLAR; BAUAD; MARANHÃO, 2009) pode-se definir calcificações como: a calcificação mamária é composta por depósitos de cálcio que se mobilizam do sangue para os tecidos, sofrendo alterações do pH e fixando-se sob a forma de sais de cálcio. Do ponto de vista bioquímico, podem ser divididos em dois grupos:

- Tipo 1: cristais de oxalato de cálcio (10%);
- Tipo 2: cristais de fosfato de cálcio (90%).

As calcificações, constituídas por cristais oxalato de cálcio, dificilmente são identificadas nos estudos de rotina, em parte porque estes cristais podem desintegrar-se durante o processamento histológico. Porém são prontamente vistas sob luz polarizada, como estruturas birrefringentes. Frequentemente estão associadas a alterações benignas e seu achado é importante, pois assegura que a área de interesse foi adequadamente avaliada. São intraluminais, relacionadas com cistos ou microcistos.

As calcificações, constituídas por cristais de fosfato de cálcio são facilmente identificadas no estudo microscópico de rotina, por meio da coloração com hemotoxilina-eosina. Aparecem como estruturas basófilas, pois quando eosinófilos associam-se a material proteináceo, ainda não calcificado. Relacionam-se com processos benignos e malignos. Parecem desenvolver-se como resultado de degeneração celular e necrose. Contribui com mais de 70% dos achados mamográficos (AGUILLAR; BAUAD; MARANHÃO, 2009).

As calcificações associadas com malignidade geralmente são pequenas e necessitam de lentes de aumento (KEMP; BARACAT; ROSTAGNO, 2003).

3. Imagem Digital

Uma imagem digital é uma matriz de duas dimensões de inteiros não negativos $f(x, y)$, $1 \leq x \leq M$ e $1 \leq y \leq N$, onde M e N são valores positivos inteiros representando os números das colunas e linhas, respectivamente. O *pixel* é a menor representação da imagem, também chamado de elemento da imagem, o qual corresponde ao valor de $f(x, y)$, dado um valor de x e de y. Se a imagem for de três dimensões a matriz será representada pela função $f(x, y, z)$ e o elemento da imagem é chamado de *voxel* (HUANG, 2004).

Se a imagem for gerada por um processo físico, os valores de intensidade são proporcionais à energia irradiada por uma fonte real, ondas eletromagnéticas, por exemplo (GONZALES; WOODS, 2010).

Para converter uma imagem contínua em relação às coordenadas x e y, além de sua amplitude, é necessária a digitalização dos valores de coordenadas, processo chamado de amostragem e da digitalização dos valores de amplitude, processo denominado quantização (GONZALES; WOODS, 2010). O processo de amostragem consiste em discretizar as coordenadas espaciais da cena em pontos para a formação da imagem digital. Cada um desses pontos é chamado de *pixel*. É necessário conhecer o tamanho do *pixel* para a realização da amostragem.

A operação de quantização, também utilizada na transformação de uma cena em imagem digital, consiste em discretizar o brilho da cena para um valor de escala de cinza utilizada.

O resultado obtido com a aplicação da operação de amostragem e da quantização pode ser armazenado na forma de uma matriz. As posições dos elementos da matriz representam os *pixels* da imagem digital e os seus valores representam a tonalidade de cinza da imagem digital.

Conforme Gonzáles (GONZALES; WOODS, 2010) existem, basicamente, três formas de representar $f(x, y)$:

- Dois eixos determinando a localização espacial e um terceiro representando os valores de intensidade de f ;
- Níveis de cinza de cada ponto proporcionais ao valor da intensidade f ;
- Na forma de matriz.

A matriz gerada a partir da imagem pode ser utilizada para processamento e desenvolvimento de algoritmos, por meio dos relacionamentos entre seus *pixels*.

Uma informação importante que pode ser extraída da matriz é a vizinhança de *pixel*, pois algumas operações são baseadas na suposição de que determinado objeto é formado por *pixels* vizinhos, uns aos outros. A representação das três vizinhanças de um determinado *pixel* p de uma imagem localizado no ponto (x, y) de N (GONZALES; WOODS, 2010) pode ser representada por:

- Vizinhança de 4: formada por todos os *pixels* que fazem fronteira de borda com p na horizontal e na vertical.
- Vizinhança diagonal: formada por todos os *pixels* que fazem fronteira de borda com p através de suas diagonais;
- Vizinhança 8: formada pela união dos *pixels* vizinhos de 4 e vizinhos diagonais a p através de suas diagonais.

Pode-se representar as três vizinhanças de p , conforme a Figura 2. Os vizinhos são destacados em tons de cinza, em (a) para os vizinhos de 4; em (b) para os vizinhos diagonais; e em (c) para os vizinhos de 8.

Figura 2 - Matrizes de vizinhança

Um relacionamento importante entre os pixels é o da conectividade. A conectividade é importante na localização de regiões com semelhança na tonalidade de níveis de cinza, pois permite identificar quais pixels estão conectados uns aos outros (GONZALEZ; WOODS, 2010).

Diversas operações utilizam o valor do nível de cinza dos pixels – $f(x, y)$, portanto elas trabalham no domínio espacial da imagem.

A utilização de imagens médicas combinada com laudos ou exames laboratoriais forma um acervo rico em informações que podem ser utilizadas tanto para fins didáticos, como para pesquisa (FURUIE et al. 2003).

A criação e manutenção de uma base de dados de imagens e laudos diversificados são importantes para a realização do treinamento dos médicos residentes. Desta forma, este capítulo tem o objetivo de apresentar o processo de aquisição e de visualização das imagens médicas.

3.1. Imagens Médicas na Medicina Moderna

A utilização de imagens médicas para realização de diagnósticos de doenças tornou-se um instrumento de apoio e suporte aos médicos e radiologistas. Imagens digitais são produzidas cada vez em maior quantidade e utilizadas para o diagnóstico e terapia. Apenas o Departamento de Radiologia do Hospital Universitário de Genebra produziu mais de 12.000 imagens em um dia, em 2002 (MÜLLER, et al., 2004). Portanto é necessário um repositório devidamente classificado e associado aos laudos realizados, para ser utilizado como apoio à decisão (FURUIE et al., 2003). No presente trabalho o repositório de dados é utilizado como base para o aprendizado dos radiologistas em fase de treinamento.

No início dos anos 1980 surgiu o conceito de armazenamento de imagens em conjunto com um sistema de comunicação denominado PACS. O sistema PACS consiste na aquisição, armazenamento e exibição de imagens integradas por uma rede computadores e aplicações. O sistema permite o acesso às imagens por vários usuários (HUANG, 2004).

3.2. Terminologia na Avaliação de Imagens Médicas

Em um processo de avaliação de imagens médicas pode-se identificar duas tarefas: detecção e estimativa. A detecção tem como objetivo avaliar a presença ou não de algum tipo de anomalia. A segunda tarefa tem relação com o peso e a comparação com determinado parâmetro. O parâmetro utilizado na comparação é definido como *Gold Standard* ou *Ground Truth* (SANTOS, 2006). O termo *Gold* também poderá ser utilizado no caso do diagnóstico em uma imagem médica ter sido realizado por mais um avaliador especializado.

Neste trabalho o termo *Gold Standard* está sendo substituído pelo termo Padrão.

As denominações utilizadas na avaliação em imagens médicas estão relacionadas com a suspeita da presença de uma anomalia, definida como positivo, e sua ausência, definida como negativo. Caso a anomalia se confirme, ter-se-á a condição de verdadeiro, caso contrario ter-se-á a condição de falso. As combinações possíveis estão apresentadas na Figura 3 (SANTOS, 2006).

		Condição Observada	
		Presença de anomalia	Ausência de anomalia
Condição Real	Presença de anomalia	Verdadeiro Positivo	Falso Negativo
	Ausência de anomalia	Falso Positivo	Verdadeiro Negativo

Figura 3 - Matriz de definição de verdadeiro e falso, positivo e negativo (SANTOS, 2006).

3.3. Análise de Imagens Médicas

A utilização de padrões e sistematização na descrição dos tecidos mamários tem como objetivo padronizar a nomenclatura dos laudos e propor condutas. O sistema BI-RADS (ACR, 2002) é um exemplo dessa padronização.

Mesmo com a utilização de uma padronização na avaliação de imagens médicas pelos radiologistas, ainda existe uma grande variação nos diagnósticos (SCHIMIDT et al., 2001). Portanto uma mesma imagem médica pode ser classificada de forma diferente entre dois ou mais radiologistas.

A classificação ou categorização de imagens médicas em grupos pré-definidos é dividida em manual ou automática (LEHMANN et al., 2005).

A categorização manual de uma imagem médica é feita por médicos ou radiologistas com a utilização de sistemas de padronização do tipo BI-RADS. A utilização de sistemas de arquivamento e sistemas auxilia no processo de categorização, por meio da inclusão de informações pelo médico, diretamente nas imagens.

A categorização automática de imagens médicas pode ser definida como um mapeamento de imagens em suas classes sem a intervenção de um médico ou um radiologista (SILVA, 2009).

A interpretação e análise de imagens médicas por médicos e radiologistas está intimamente associada com a sua experiência. Os principais parâmetros considerados na interpretação de imagens médicas são (SANTOS, 2006):

- Simetria (deslocamento de estruturas a partir de uma linha média);
- Diferenças em estruturas normais (estruturas anatomicamente correta, mas com alterações topológicas);
- Conhecimento dos padrões típicos associados a uma patologia particular;
- Parâmetros anormais que podem não pertencer a uma categoria conhecida.

O processo de automação na análise de imagens médicas está relacionado com problemas de visão computacional e de processamento de imagens médicas.

3.4. Critérios de Avaliação de Imagens Médicas

Um novo diagnóstico médico deve ter o seu poder de discriminação comprovado em relação à doença que se deseja detectar. A acurácia obtida pode não ser suficiente para comprovar como o método utilizado se comporta com relação aos falsos positivos (FP) ou aos falsos negativos (FN). Os resultados previstos pelo diagnóstico são comparados com os resultados reais, obtidos por outro método mais preciso, por exemplo, um radiologista com mais experiência. Para se avaliar, de forma objetiva, a precisão do diagnóstico em uma imagem médica é necessário medir a sensibilidade, a especificidade e a acurácia, as quais podem ser obtidas da seguinte forma (SABATINI, 1995):

- Acurácia: a proporção de predições corretas (soma de verdadeiros positivos - VP e verdadeiros negativos - VN):

$$\frac{VP+VN}{VP+VN+FP+FN}$$

Equação 1

- Sensibilidade: a proporção de verdadeiros positivos (VP), ou seja, a medida da capacidade do método de decisão de prever a condição patológica para aqueles casos que realmente a apresentam (VP + FN):

$$\frac{VP}{VP + FN} \quad \text{Equação 2}$$

- Especificidade: a proporção de verdadeiros negativos (VN), ou seja, a medida da capacidade do método de decisão de apontar ausência da condição para aqueles casos (FP + VN) que realmente não a têm:

$$\frac{VN}{FP + VN} \quad \text{Equação 3}$$

- Preditividade positiva: a proporção de verdadeiros positivos (VP) em relação a todas as predições positivas (VP + FP):

$$\frac{VP}{VP + FP} \quad \text{Equação 4}$$

- Preditividade negativa: a proporção de verdadeiros negativos (VN) em relação a todas as predições negativas (VN + FN):

$$\frac{VN}{VN + FN} \quad \text{Equação 5}$$

Geralmente, para qualquer tarefa de decisão médica, a sensibilidade e a especificidade variam em direções opostas, ou seja, quando o teste é muito sensível a positivos, tende a gerar muitos falsos positivos.

Assim como é preciso avaliar o diagnóstico médico, faz-se necessário avaliar os classificadores utilizados para a descoberta de padrões, objeto dessa tese. Para avaliar os classificadores uma das técnicas utilizadas é a análise ROC.

3.5. Análise ROC

Para a avaliação dos resultados do processo de classificação na área médica é utilizada uma técnica de análise de resultados, desenvolvido por Metz (METZ, 1986).

A análise ROC (*Receiver Operation Characteristics*) é uma técnica para visualizar, organizar e selecionar o desempenho dos classificadores (FAWCETT, 2005). A análise ROC tem sido utilizada para visualizar e analisar o comportamento de sistema de diagnósticos médicos como em Patrocínio (PATROCINIO, 2004).

A análise ROC avalia a qualidade das imagens médicas em relação à habilidade de classificar como positivo ou negativo a presença de uma anomalia ou doença (METZ, 1986).

Segundo Evans (EVANS, 1981) existem duas possibilidades em se tratando de avaliação da imagem: a imagem contém informação que indica uma doença ou anormalidade ou a imagem não contém essa informação. Portanto o diagnóstico de verdadeiro positivo (VP) indica que o classificador identifica um problema que realmente existe e o verdadeiro negativo (VN), quando o classificador não identifica sinal de problema na imagem e realmente não existe.

Se o classificador identificar uma anomalia que não existe o diagnóstico se torna falso-positivo (FP). Ou se o classificador identificar a ausência de anomalia e elas existem, o diagnóstico se torna falso negativo (FN).

Uma possibilidade é a utilização de gráficos. Os gráficos possibilitam uma melhor visualização das várias dimensões do problema. Apesar dos gráficos da análise ROC serem simples podem ocorrer interpretações incorretas a seu respeito.

Para se construir um gráfico é necessário plotar a taxa de verdadeiro positivo no eixo das ordenadas – eixo y - e a taxa de falsos positivos no eixo das abscissas – eixo x.

No presente trabalho a análise ROC será utilizada para avaliar o classificador.

3.6. Séries Temporais

Uma série temporal pode ser definida como uma coleção de observações de uma variável, feitas sequencialmente ao longo do tempo. Uma das características mais importantes deste tipo de dado é que as observações são dependentes, sendo necessário analisar e modelar esta dependência. Informações de séries temporais estão sendo utilizadas em diversos campos do conhecimento, como economia (preços diários de ações), epidemiologia (números mensal de casos de meningite), medicina (eletrocardiograma, eletroencefalograma), entre outros.

Características particulares a este tipo de dado são (EHLERS, 2011):

- Observações correlacionadas são mais difíceis de analisar e requerem técnicas específicas;
- É necessário levar em conta a ordem temporal das observações;
- Fatores complicadores como presença de tendências e variação sazonal ou cíclica pode ser difícil de estimar ou remover;
- A seleção de modelos pode ser bastante complicada, e as ferramentas podem ser de difícil interpretação;

Uma série temporal pode ser contínua, quando as observações são feitas continuamente no tempo e pode ser discreta, quando as observações são feitas em tempos específicos (EHLERS, 2011).

As séries temporais podem ser analisadas a partir de movimentos descritos como: variações aleatórias, sazonalidade, ciclo e tendência.

Na análise de séries temporais, o principal objetivo é o de aproximar uma função do tempo por uma combinação linear de harmônicos. A análise de Fourier é utilizada para estudar funções ou sinais determinísticos (MORETTI, 1999).

3.7. Análise de Fourier

O objetivo da análise de sinais é o de extrair informações relevantes de um sinal. A extração de informações é feita através de uma transformação.

A análise de Fourier pode ser analisada a partir de movimentos descritos como: variações aleatórias, sazonalidade, ciclo e tendência. A análise de Fourier decompõe o sinal na base ortogonal formada pelas funções $e^{-j2\pi ft}$.

A análise de Fourier tradicionalmente tem sido utilizada para resolver equações diferenciais. O objetivo é o de aproximar uma função $f(t)$ por uma combinação linear de componentes senoidais, cada um com sua frequência. A análise de Fourier é utilizada em processos estocásticos, que é uma família de variáveis aleatórias. Para processos não estacionários outros sistemas ortogonais podem ser úteis, como as transformadas de *wavelets* (MORETTI, 1999). O termo *wavelets* apareceu no trabalho de A. Haar (1909). Haar construiu a primeira família de transformadas *wavelets* ortogonal de suporte compacto. Na sequência são discutidos dois tipos de transformada..

3.8. Transformada de Fourier

A transformada de Fourier tem como objetivo transformar um sinal (função) do domínio de espaço para o domínio da frequência (GONZALES e WOODS, 2010);

A Transformada de Fourier de um sinal contínuo apresenta algumas deficiências, uma delas é o fato de não estar localizado no domínio do espaço. Desta forma ela não pode representar as mudanças que ocorrem no sinal, no espaço. A transformada de Fourier está baseada na integração de toda a função para o cálculo de cada frequência. Se o sinal (sinais estacionários) não mudasse durante o tempo, não haveria problema, mas alguns sinais com características importantes são não estacionários ou transitórios (CASTAÑON, 2003).

A solução desse problema foi proposta por Dennis Gabor (GABOR, 1946), adaptando a transformada de Fourier para analisar uma pequena porção do sinal em um tempo. A transformada de Gabor utiliza a janela gaussiana. Essa solução foi denominada de transformada por janela de Fourier (*Windowed Fourier Transform – WFT*).

Um problema identificado é que alguns sinais precisam que o sinal da janela possa ser alterado com o objetivo de identificar o conteúdo local da frequência (DAUBECHIES, 1992).

3.9. Transformada *Wavelet*

As transformadas *wavelets* têm como base a utilização de janelas com regiões de dimensão variável e a transformada de Fourier tem como base uma função de duração limitada para uma janela de tempo constante para todas as frequências.

A Transformada *Wavelet* está sendo muito utilizada como alternativa para o processamento de sinais. Uma das razões para sua utilização é que com uma quantidade mínima de valores é possível representar toda a imagem. Os valores podem ser utilizados na montagem do vetor de características para a representação de uma imagem (FARIAS, 1998).

Transformadas wavelets, sob a ótica da análise de sinais são pequenas ondas. A utilização de imagens médicas combinadas com laudos ou exames laboratoriais forma um acervo rico em informações que podem ser utilizadas tanto para fins didáticos como para pesquisa (FURUIE et al. 2003).

Historicamente a análise de transformadas *wavelets* se apresenta em origens diferentes (MEYER, 1993). Em 1909 A. Haar faz menção às transformadas *wavelets*, que ficaram no anonimato por diversos anos.

As transformadas *wavelets* são funções matemáticas que decompõem o dado em diferentes componentes de frequência, permitindo analisar cada componente com uma resolução combinada a sua escala. A transformada discreta de *wavelet* utiliza duas abordagens (DAUBACHIES, 1992): sistemas redundantes discretos e ortogonais. No sistema ortogonal é considerada a estratégia de multiresolução, que é abordada neste trabalho.

3.10. Análise dos momentos

Uma técnica alternativa utilizada para o processo de extração de características que não necessita que a imagem seja previamente segmentada é a Análise de Momentos. Os momentos e suas funções são uteis na análise da forma e são caracterizados como valores numéricos e utilizados para fornecer descrições da distribuição dos pontos contidos em uma imagem (ALVES, 2004).

Além dos momentos regulares existem os momentos centrais utilizados para imagens digitais (KHOTANZAD; HONG, 1990).

Um dos problemas para o reconhecimento de padrões é a falta de invariância das características. Um objeto que tenha sido rotacionado ou tenha tido sua escala alterada precisa ser reconhecido da mesma forma.

Os momentos são invariantes quanto à rotação, translação e escala, sendo utilizados em várias aplicações da área de visão computacional para localização de objetos, reconhecimento de formato, orientação e tamanho. As técnicas que utilizam a análise de momentos descrevem as propriedades globais de uma imagem.

Hu (HU, 1961) apresentou um trabalho onde empregava a teoria dos momentos no processamento de imagens baseado na teoria das invariâncias algébricas.

Um momento pode ser definido com uma soma ponderada de um padrão $f(x)$ aplicado sobre todo o campo de entrada, com os pesos iguais a algum polinômio x . Hu (HU, 1961) desenvolveu a teoria dos Momentos Invariante (MI).

Zernike (KHOTANZAD; HONG, 1999) e Hu (HU, 1961) são as principais técnicas utilizadas.

Algumas formas apresentam simetrias similares em diversas posições rotacionais o que é um problema no reconhecimento de objetos invariante à posição. HU (HU, 1962) propôs calcular momentos invariantes à rotação.

Para torná-los invariantes à translação é necessário calculá-los em relação ao centro da massa da figura. Para tornar os momentos invariantes às mudanças de translação e escala é necessário normalizar a equação e normalizá-los em relação à área da figura (PATROCINIO, 2004). Hu derivou combinações de momentos que são invariantes à orientação, à escala e à posição. Os sete momentos de Hu são invariantes à escala, rotação e translação.

3.11. Extração e Análise de Atributos

As informações em uma imagem médica podem ser avaliadas por meio da extração de características da imagem. Estas características podem ser agrupadas em quatro categorias:

- Espectrais;

- Espaciais;
- De contexto;
- Temporais.

As características temporais estão relacionadas com a radiância de um determinado alvo, que varia em função do comprimento da onda as características de um pixel ou de uma região da imagem ao longo do tempo. As características espaciais estão relacionadas com textura, forma e dimensões. As características de contexto utilizam informações obtidas de regiões vizinhas à área que está sendo analisada.

O processo de extração de características implica na criação de um vetor que represente os principais componentes de uma imagem. Técnicas de análise de imagens são utilizadas para extração dos valores das características (SILVA, 2009).

A Figura 4 apresenta algumas das características extraídas por sistemas de recuperação de imagem por similaridade.

Característica	Objetivo	Referência
Densidade	Medida de níveis de cinza referentes aos pixels pertencentes à identificação da mama	GATO (2004).
Desvio Padrão	Média do Contraste calculado sobre uma área com k-vizinho	GONZALES (2010).
Entropia	Quantidade média de informações obtida por saída de uma única fonte.	GONZALES (2010).
Histograma	Conjunto de números indicando o percentual de pixels de uma determinada intensidade naquela imagem	GONZALES (2010) MARQUES FILHO (1999).
Medida de Área	Quantidade de pixels de um objeto em uma área.	GATO (2004).
Textura	Frequência de ocorrência de um determinado par de níveis de cinza.	MARQUES (2002)
Média dos valores dos pixels	Valor médio dos níveis de cinza os pixels pertencentes à mama	KINOSHITA (2004)

Figura 4 - Características extraídas por sistema de recuperação por conteúdo

A extração das características da imagem é uma etapa importante na recuperação de imagens por conteúdo, pois fornece subsídios para a consulta por similaridade (GATO; NUNES; SCHIEBEL, 2004) (MARQUES; TRAINA, 2006).

A utilização do reconhecimento de padrões, na área médica, é muito utilizada para a identificação de tumores, visualização de contornos e para auxílio ao diagnóstico médico. (PATROCINIO, 2004).

Para que seja possível identificar padrões e classificar automaticamente as imagens, é comum extrair diversas medidas de cada padrão e representá-las na forma de vetor. Os atributos ou características que farão parte do vetor são extraídos com a utilização de técnicas de processamento de sinais e armazenados em um banco de dados. O número de atributos dependerá das propriedades que se deseja caracterizar. Segundo Loew (LOEW, 2000), apud (CASTAÑON, 2003) a representação da imagem por meio de um vetor de característica deve atender às seguintes considerações:

- Reduzir a dimensionalidade dos dados;
- Ressaltar aspectos da imagem;
- Ser invariante às transformações

Para a descrição de uma imagem as características mais utilizadas em sistemas CBIR são baseadas na cor, textura e forma, sendo que a recuperação de imagens está baseada nessas características e em características derivadas (Kinoshita, 2004). O resultado da extração é representado por um vetor, denominado de vetor de características. Para a comparação entre os vetores de características é necessária a utilização de uma função de distância que calcule a similaridade. Funções de distância são usadas em diversas áreas, incluindo reconhecimento de padrões, redes neurais e estatísticas (WILSON; MARTINEZ, 1997)

Em sistemas de recuperação de imagem por conteúdo o processo de extração de características é importante, uma vez que é desse processo que depende a representação correta da imagem no ambiente computacional. As características extraídas das imagens devem ser equivalentes às características a serem analisadas pelo especialista (SILVA, 2009).

3.12. Extratores de atributos

A partir da identificação das características obtidas da imagem é necessária a execução de pré-processamento e processamento para a verificação da descoberta, ou não, de áreas de interesse. A Figura 5 ilustra o processo de extração de características.

Figura 5 - Processo de Extração de Características de uma imagem

O processo de seleção de atributos deve extrair as características mais significativas. Dois itens devem ser destacados nesse processo (GONZALES; WOODS, 2010):

- Transformar o espaço de atributos, o que reduz a dimensão dos vetores de características com o objetivo de torná-los mais representativos em uma classe;
- Calcular a distância entre as classes e medidas de informação

A técnica de transformação é discutida em (FUKUNAGA, 1990) e a técnica de cálculo da distância é discutida em (MASCARENHAS; VELASCO, 1989).

Em uma classificação, os atributos devem representar todas as granularidades desejadas não se restringindo à cor, forma ou à textura. Portanto, é necessário utilizar uma técnica que possa extrair características de diversas categorias e que sejam invariantes com relação a contraste, escala, rotação e translação. O problema da invariância a contraste pode ser resolvido aplicando normalização (SILVA, 2009).

Um dos atributos que pode ser analisado é o da textura. Haralick (HARALICK, 1979), em seu experimento, utilizou imagens que cobriam regiões do ártico. As imagens foram analisadas utilizando o atributo de textura com funções de auto correlação. Os resultados obtidos (coeficiente de correlação de 0,99) demonstraram

a capacidade da utilização de funções matemáticas de reproduzir a percepção da visão humana.

Em se tratando de processamento de sinais existe uma série de ferramentas, podendo-se citar a análise de Fourier e a Transformada *Wavelet* (TW).

Uma das dificuldades no processo de extração e seleção de características é identificar um conjunto reduzido de características que descrevam a imagem com precisão.

O processo de extração e seleção das características é um dos passos mais críticos no processo de identificação de objetos na imagem. Um grande problema é identificar um conjunto reduzido de características para descrever a imagem com precisão e extrair padrões.

Este trabalho propõe métodos de extração de características de imagens médicas utilizando Transformadas *Wavelets*. No capítulo seguinte é feita uma breve explicação do conceito de SOM e o seu uso para a categorização.

Geralmente as imagens são recuperadas com base em informações textuais. A extração de características da imagem pode ser utilizada para a recuperação de imagens. Com base em suas características, a imagem será indexada e recuperada quando necessário. Recuperação por similaridade é aquela que se considera quando dois elementos são semelhantes entre si. A função de distância visa definir a similaridade entre dois elementos (BUENO, 2001).

3.13. Mapa Auto-Organizável

O Mapa Auto-Organizável (Self-Organizing Map – SOM) é uma arquitetura de rede neural que utiliza o aprendizado não supervisionado (HAYKIN, 1999), (KOHONEN, 2001). O SOM está sendo utilizado em tarefas de agrupamento (BOSCARIOLLI; SILVA, DEL-MORAL-HERNANDES, 2006) e de classificação (SILVA, 2009) (SILVA et al., 2006).

O Mapa Auto-Organizável foi desenvolvido por Tuevon Kohonen em 1987 (KOHONEN, 2001). Em um Mapa Auto Organizável, os neurônios são colocados nos nós de uma rede de uma ou duas dimensões. Os neurônios tornam-se

seletivamente ajustados para vários padrões de entrada ou classes de padrões de entrada no curso de um processo competitivo de aprendizagem (HAYKIN, 2001)

SOM é uma ferramenta de software eficaz para a visualização de dados de alta dimensão. Na sua forma básica, produz um gráfico de similaridade dos dados de entrada. Ele converte as relações não lineares estatísticos de dados de alta-dimensão em simples relações geométricas dos seus pontos de imagem em baixa dimensão, usualmente de duas dimensões (KOHONEN, 2001). Padrões de entrada similares serão mapeados em regiões especialmente próximas do mapa de saída. SOM é uma ferramenta apropriada para visualização e exploração em alta dimensão (SILVA, 2009).

Som é uma rede em duas camadas onde na entrada das informações é utilizado um vetor de características e tendo como saída um mapa de neurônios.

A Figura 6 representa um pequena rede SOM de 4 X 4 nós conectados a uma camada de entrada representando um vetor de duas dimensões.

Cada nó da camada de saída conterá um vetor de pesos de mesma dimensão do vetor de entrada. Os dados de teste consistem de vetor, V , de n elementos:

$$V_1, V_2, V_3 \dots V_n$$

Os vetores de pesos podem ser inicializados aleatoriamente ou através de outros métodos discutidos em (KOHONEN, 2001).

O processo de treinamento pode ser caracterizado da seguinte forma, (KOHONEN, 2001):

1. Inicialização dos pesos de cada nó da camada de saída;
2. Seleção aleatória de um vetor de características retirado do conjunto de dados para treinamento;
3. O nó que apresenta a menor distância no espaço euclidiano entre seu vetor de pesos e o vetor de características é declarado como nó vencedor e será conhecido como a unidade com melhor casamento (Best-Matching Unit – BMU);
4. Para garantir as relações de similaridade entre os vetores de características e os nós do mapa os pesos das unidades vizinhas ao neurônio vencedor também serão atualizados;

5. O passo 2 sofrerá n interações.

Conforme SILVA (SILVA, 2009) após a realização do treinamento do mapa SOM as imagens categorizadas serão associadas aos respectivos neurônios vencedores a partir da utilização da distância Euclidiana entre os vetores de pesos do mapa e os vetores de características das imagens utilizadas no treinamento. Desta maneira obtém-se uma melhor associação entre os neurônios vencedores para cada vetor de características.

Figura 6 - Rede SOM com 16 neurônios na camada de saída (mapa 4 X 4).

4. Proposta para o Reconhecimento de Imagens Mamográficas

Neste capítulo apresenta-se a técnica proposta para efetuar o reconhecimento de padrões em imagens mamográficas. Utilizam-se para a identificação e classificação de imagens mamográficas similares, técnicas de processamento de imagens que envolvem a Transformada *Wavelet* e *mapas* auto organizáveis (SOM) para classificação das imagens mamográficas.

4.1. Premissas Adotadas

Após a revisão bibliográfica deste trabalho foram formuladas premissas e proposições com o objetivo de elaborar uma metodologia acerca de como resolver a questão principal do trabalho e de suas questões recorrentes. A Figura 7 representa o modelo teórico sobre a questão principal deste trabalho. Nessa figura são identificadas as áreas que envolveram a construção da técnica proposta.

Figura 7 - Fundamentações Conceituais

A Figura 8 tem como objetivo demonstrar que a problemática da pesquisa está relacionada com o conjunto de fundamentações conceituais capazes de demonstrar os aspectos críticos e necessários para responder à questão central da pesquisa que é a recuperação de imagens pela similaridade.

Figura 8 - Premissas e proposições associadas deste trabalho

Com base na questão principal foram formuladas quatro premissas básicas, que têm como objetivo demonstrar o relacionamento entre as áreas de processamento de imagem mamográficas, mineração de dados e classificação das informações para a construção da metodologia proposta.

As premissas da pesquisa são formuladas com base na revisão bibliográfica desenvolvida com o propósito de efetuar a ligação entre a questão principal da pesquisa e as proposições do trabalho.

4.1.1. Premissa 1

O treinamento dos médicos residentes na Universidade Federal de São Paulo utilizava imagens mamográficas, inicialmente no formato TIFF, as quais eram acessadas utilizando o software QuaLIM® construído por Pires (PIRES, 2003).

O software QuaLIM utiliza um banco de dados de imagens composto de casos reais e caso simulados. O termo caso está associado a um grupo de imagens de um paciente realizado em um determinado período com todas as imagens mamográficas necessárias à realização do laudo.

O software QuaLIM® está estruturado em duas fases: uma fase de detecção de estruturas que simulam microcalcificações e fibras dispostas aleatoriamente, utilizando as imagens geradas em filmes e digitalizadas; e a segunda fase de interpretação das imagens mamográficas digitais, na identificação de estruturas de interesse.

As imagens utilizadas pelo software QuaLIM® foram convertidas para o formato DICOM utilizando-se o software EZDICOM. Cada grupo de imagens foi associado a um diretório correspondente ao número do caso executado previamente.

O trabalho realizado por Munerati (MUNERATI, 2011) teve como objetivo avaliar a situação atual do sistema QuaLIM® e criar uma solução que viesse atender às necessidades do momento e sugerir e criar novas funcionalidades. Com base nesse estudo foi construído um aplicativo denominado de MammoAdviser.

O processo de aquisição automático das imagens mamográficas que são utilizadas nesse trabalho está em fase desenvolvimento e não é objeto de discussão

neste trabalho. As imagens mamográficas são armazenadas de acordo com o caso (termo acima descrito) a que pertencem. Com as imagens armazenadas, organizadas e identificadas pelo número do caso correspondente, é realizada a avaliação e a geração de dois laudos por dois médicos experientes. Ao final, caso haja concordância desses laudos, é armazenado um único laudo que é utilizado para comparações e identificado como padrão, sendo que a atualização na base de dados é feita pelo módulo MammoAdviser. O laudo padrão é utilizado pelos médicos residentes para tomarem ciência do laudo realizado pelos médicos com maior experiência. Caso haja discordância nos dois laudos acima descritos, é solicitado um novo laudo a um terceiro médico e caso esse novo laudo discorde de ambos os laudos anteriores, as mamografias desse laudo são descartadas.

Para responder à questão central desta Tese, a proposta é a utilização de um processo de transformação da imagem, na qual são utilizadas técnicas de processamento de imagens com o uso das Transformadas *Wavelet*. A técnica utilizada tem como base o trabalho desenvolvido por Silva (SILVA, 2009).

4.1.2. Premissa 2

Cada imagem mamográfica processada está diretamente associada a um determinado caso. Após o processo de transformação das imagens faz-se necessária a categorização das imagens, de acordo com a classe BI-RADS. Em cada imagem processada é identificada a qual classe BI-RADS a mesma pertence. Para atender a essa premissa foi criado o módulo Mammosort que executa a etapa de categorização e classificação das imagens. A classificação é feita por meio dos mapas SOM, que geram como resultado a distância Euclidiana entre as imagens mamográficas. Para cada imagem mamográfica são selecionadas, no universo de todas as demais mamografias, as quatro imagens, pertencentes a casos diferentes, com a menor distância entre elas.

4.1.3. Premissa 3

Após o processo de classificação e categorização das imagens mamográficas, conforme descrito na premissa 2, deve ser atualizada a base de

dados do sistema, contendo informações sobre o número do caso a que pertence a imagem sob análise, bem como os números dos casos a que pertencem as imagens identificadas na premissa 2. Dessa forma, é importante destacar que apesar de o processo de identificação da similaridade estar associado a cada imagem mamográfica, a exibição para o residente leva em conta o caso a que pertence cada imagem.

4.1.4. Premissa 4

Para que o médico residente aprimore suas habilidades, foi desenvolvido o módulo MammoViewer que exibe quais são as imagens mais similares a uma determinada imagem mamográfica previamente selecionada conforme descrito na premissa 3. O médico residente deve informar o número do caso que ele está avaliando, sendo que o módulo MammoViewer deve apresentar as imagens que apresentam maior grau de similaridade dentre todas as pesquisadas. Juntamente com as imagens de cada caso, também é disponibilizado o laudo textual redigido para cada um deles. Em associação ao módulo MammoViewer está previsto o módulo de nome MammoSegm que tem como objetivo permitir que o médico experiente identifique as áreas de interesse para a realização do laudo. A segmentação, nesse caso, não é um processo automatizado, pois depende da atuação do médico experiente. Caso o médico residente necessite visualizar informações sobre a imagem, as mesmas são obtidas por meio do módulo MammoSegm.

4.2. Arquitetura Geral da Proposta

Assim, ao módulo MammoAdviser, desenvolvido em outro trabalho, (MUNERATI, 2011), foram agregados os módulos MammoSegm, MammoSort e MammoViewer, implementando a proposta central objeto desta tese, conforme ilustra a Figura 9.

Figura 9 - Sistema Mammo

Para que seja possível a busca por similaridade, as imagens originais dos laudos passam por um processamento e classificação. Com o laudo modelo e as imagens processadas, o médico residente pode realizar o seu treinamento. Após a execução do laudo, com sua inclusão na base de dados, o residente pode solicitar ao sistema a exibição de imagens com o mesmo grau de complexidade e similaridade (caso exista na base de dados alguma imagem segmentada que possua uma região de interesse para o laudo, a imagem poderá ser exibida).

Todos os laudos são armazenados na base de dados do sistema para a realização de avaliações estatísticas do aprendizado dos residentes, individual ou coletivamente.

A solução da aplicação foi dividida em dois ambientes de desenvolvimento. Para a comunicação com o usuário final, a aplicação foi desenvolvida utilizando a linguagem de programação Java® (SUN, 2006) com a utilização do Netbeans como ambiente de desenvolvimento. Juntamente com o Java foi utilizada a API (*Application Program Interface*) JAI (Java Advanced Image) (SUN, 2006), permitindo a

implementação do processo de classificação com o uso de Redes Neurais Artificiais, por meio de um módulo desenvolvido utilizando a linguagem MATLAB®.

A base de dados utilizada está sendo gerenciada pelo sistema gerenciador de bancos de dados MS SQL\SERVER 2008 R2 da Microsoft®.

4.3. Detalhamento da Proposta

Para toda mamografia presente na base de dados é preparado um laudo modelo por dois médicos experientes, chamados somente médicos, a partir deste ponto, conforme a

Figura 10. Caso exista alguma divergência entre os médicos, se possível, será solicitada uma terceira avaliação. Se isso não for possível a mamografia não será desconsiderada para efeito de treinamento.

Após a criação do laudo modelo, pelos médicos, as imagens estarão disponíveis aos médicos residentes, denominados como residentes a partir de agora, executarem o seu treinamento onde lhes serão apresentadas imagens médicas de acordo com o seu nível de especialização.

O módulo MammoAdviser (MUNERATTI, 2011), representado na

Figura 11 tem como principais objetivos:

- Permitir a comparação de resultados entre os profissionais;
- Criar etapas de treinamento sensíveis as dificuldade de cada profissional;
- Gerar resultados automatizados de desempenho do profissional;

Figura 10 - Arquitetura da realização de laudo pelo médico experiente com o uso do MammoAdviser

A solução proposta adiciona ao módulo MammoAdviser três módulos independentes: um para segmentação das imagens pela região de interesse, denominado de MammoSegm, o módulo MammoViewer para exibição das imagens e módulo MammoSort, responsável na extração de características das imagens e de classificação das mesmas.

Figura 11 - Arquitetura da realização de laudo pelo residente com o uso do MammoAdviser

O módulo MammoAdviser (MUNERTI, 2001) está sendo utilizado pelos residentes para a geração do laudo de treinamento e pelos médicos, para a geração do laudo considerado padrão. O laudo de treinamento é comparado com o laudo padrão com objetivo de avaliar o desempenho do médico residente no treinamento. Em virtude da identificação da necessidade de manutenção desse módulo, optou-se em construir os outros módulos alterar o Módulo MammoAdviser.

Visando a segmentação das imagens construiu-se o módulo MammoSegm, cujo objetivo é permitir ao médico marcar, nas imagens, regiões de interesse que apresentem alguma anomalia. Essas informações são utilizadas para comparações

e avaliação do laudo executado pelo médico residente. As informações das regiões das imagens de interesse são armazenadas na base de dados e utilizadas para a exibição das partes que tenham a mesma similaridade e também para auxiliar o médico residente na visualização de áreas que tenham interesse no laudo final.

O módulo MammoViewer tem como finalidade exibir as imagens médicas com a mesma similaridade e as imagens que foram segmentadas, por região de interesse, no ato da execução do laudo padrão pelos médicos. Com as informações disponíveis no laudo realizado por médicos, as imagens são recuperadas pela similaridade. Juntamente com a imagem segmentada pelos médicos com a região que possa ter alguma anomalia, o residente terá condições de utilizar essas informações para o desenvolvimento de seus laudos.

O módulo MammoSort é o responsável pelo tratamento das imagens. Nessa etapa executa-se o processamento das de imagens e a comparação por conteúdo (CBIR), considerando a similaridade em relação a uma imagem previamente selecionada para o treinamento. A seleção da imagem para a realização do treinamento está relacionada como o estágio de aprendizado do residente. Caso as imagens recuperadas pelo CBIR para comparação não estejam de acordo com as imagens esperados pelo médico residente, tem-se o que se define como gap semântico. Diversos trabalhos foram pesquisados para identificar qual o extrator de característica que apresentasse melhor desempenho de classificação e de qualidade na recuperação de imagens. Para extração das características das imagens médicas optou-se por utilizar a Transformada *Wavelet*. As imagens são agrupadas de acordo com a classificação BI-RADS® utilizando Redes Neurais Artificial do tipo *Self-Organizing Maps* (SOM) (SILVA et al., 2006).

Para a classificação das imagens de mesma similaridade existe uma série de classificadores automáticos, que podem ser estatísticos como o *Knn* (DUDA; HART; STORK, 2001) ou que envolvam algum tipo de aprendizado (HAYKIN, 2001).

No presente trabalho é utilizada a rede neural SOM, com base no trabalho de SILVA (SILVA, 2009) que apresenta como vantagens a capacidade de aprendizado não supervisionado e agrupamento.

A seguir é apresentado um o detalhamento da proposta desta tese, que compreende uma visão geral das técnicas utilizadas, bem como o processamento

das imagens que envolvem a classificação e a segmentação da imagem médica. O apêndice A contém a especificação funcional utilizada para a implementação da solução.

4.3.1. Módulo MammoSegm

A realização do laudo pelos médicos, resulta na atualização da base de dados, o que é feito por meio do módulo MammoAdviser. Com o intuito de auxiliar o treinamento do residente, o médico, por meio do módulo MammoSegm poderá selecionar e criar imagens apenas com áreas de interesse para cada uma das imagens da mamografia.

Para a implementação do módulo MammoSegm foi necessária a criação de três tabelas: IMAGEM_SEGMENTADA, que armazena o código e a localização da imagem segmentada; IMAGEM_OBSERVACAO, que armazena um código e informações relevantes sobre a imagem e a tabela TIPO_ANOMALIA que armazena informações do tipo de anomalia, conforme a Figura 12.

Figura 12 - Modelo Entidade Relacionamento

Para cada imagem, o médico pode identificar diversas áreas de interesse. Inicialmente o módulo contém as informações sobre a presença de nódulos e microcalcificações. O sistema atualiza as informações nas tabelas o que implica em não alterar as características originais da imagem, pois as informações das áreas de interesse são atualizadas na base de dados do sistema, conforme a Figura 13.

Na tabela Imagem_Segmentada informa-se a localização, na imagem, da observação do médico. Para cada observação o médico pode incluir um pequeno texto com as observações que achar necessárias.

Figura 13 – Módulo de segmentação da imagens MammoSegm

4.3.2. Modulo MammoSort

O módulo MammoSort disponibiliza uma interface para que sejam informados quais são os parâmetros a serem utilizados no processo de extração de características e de classificação. A Figura 14 exemplifica a execução do módulo.

Para a realização do processo de extração e de classificação a implementação do módulo foi dividida em diversas etapas: pré processamento, etapa wavelet e etapa SOM.

Figura 14 - Módulo de extração e classificação de imagens, MammoSort

A seguir, discute-se, em detalhes, a metodologia envolvida no trabalho. A discussão envolve todas as etapas de preparação, extração de características das imagens médicas e a fase de classificação, bem como a atualização na base de dados para a realização das pesquisas pelos médicos residentes. Um esquema geral do método desenvolvido pode ser visualizado na Figura 15.

Figura 15 - Esquema geral do método desenvolvido

Nas subseções seguintes são detalhadas cada uma das etapas que formam o método, iniciando-se com a etapa de pré-processamento.

4.3.2.1. Etapa de Pré-Processamento

Com o objetivo de padronizar as imagens aplica-se uma operação de normalização de níveis de cinza nas mesmas, conforme ilustra Figura 16.

Figura 16 - Esquema de funcionamento da etapa de pré-processamento

Ao final de da etapa de pré-processamento tem inicio a etapa de wavelet. Nessa etapa, o conjunto de imagens é submetido ao extrator de atributos que representam cada classe de imagens, utilizando Transformada *Wavelets*.

Essa metodologia foi utilizando com sucesso em Silva (SILVA, 2009) com imagens médicas. Para cada imagem será executado um pré-processamento para realçar a região patológica.

4.3.2.2. Etapa Wavelets

A proposta do trabalho é categorizar as imagens médicas. Utiliza-se um extrator de características que tem como base a transformada *wavelet* e um classificador, utilizando mapa SOM (mapas auto-organizáveis).

A extração de características é realizada por meio da transformada *wavelets*, utilizando análise em multiresolução, na qual uma determinada imagem é decomposta em quatro diferentes conjuntos de coeficientes, conforme representado na Figura 17.

Os quatro coeficientes representam:

- cA_n representa as baixas frequências da imagem;
- cA_n^h representa os coeficientes horizontais;
- cA_n^v representa os coeficientes verticais;
- cA_n^d representa os coeficientes diagonais.

Figura 17 - Esquema de decomposição por análise em multiresolução

O critério do número de decomposições (parâmetro n , na figura 17) utilizado é ilustrado no esquema da Figura 18 (SILVA, 2009). Para cada nível de decomposição, calculam-se os coeficientes de aproximação, cA_n , das imagens para treinamento do mapa SOM e das imagens para testes do mapa SOM. A partir desses parâmetros são extraídas e calculadas as medidas MQE (erro médio de quantização), TE (erro topográfico), IDB (índice de Davies-Boudin) e taxa de acerto.

Figura 18 - Esquema de avaliação do conjunto de coeficientes (SILVA, 2009)

Inicialmente é utilizada a *wavelet* Haar para a definição da decomposição. Após a definição do tipo de *wavelet* mãe e dos coeficientes, as características das imagens de treinamento são extraídas.

4.3.2.3. Etapa SOM

A definição do tamanho do mapa do SOM é feito pelas medidas utilizadas na definição das características MQE, TE, IDB e taxa de acerto para os mapas 4 X 4 , 6 X 6, 8 X 8, 10 X 10, 12 X 12. (SILVA, 2009).

A medida de qualidade empregada no agrupamento é o Índice de Davies-Boudin (IDB) (DAVID; BOULDIN, 1979). Neste caso, o mapa é segmentado em grupos utilizando a metodologia de Vesanto e Alhoniemi (VESANTO; ALHOHIEMI, 2000), que utiliza o classificador *K-médias* para agrupar os vetores de pesos e o IDB para avaliar a qualidade do agrupamento.

As características das imagens das imagens são representadas pelos coeficientes que resultaram nas melhores medidas e com a melhor família *wavelet* (SILVA, 2009).

4.3.2.4. Pós-Processamento

Nessa etapa as informações geradas nos processos anteriores são atualizadas nas bases de dados.

4.3.3. Modulo MammoViewer para visualização das imagens e dos laudos

O módulo MammoViewer disponibiliza uma interface para que sejam recuperadas as imagens da base de dados, de acordo com a necessidade do residente ou do médico que executará o processo de segmentação e validação do processo de similaridade conforme ilustra Figura 19.

Figura 19 - Modulo de visualização de imagens segmentadas, MammoViewer

As imagens de cada caso estão localizadas em um determinado diretório. O processo de extração das características implica, primeiramente, na criação de um diretório, no qual o nome de cada imagem é associado ao caso que pertence. Nesse momento as imagens são processadas para a geração do vetor de características.

No vetor de características associa-se o número da classificação BI-RADS ao qual pertence o caso. Desta forma, antes de entrar no processo de classificação, a imagem na passa por um processo de identificação a qual categoria ela pertence. Nos itens abaixo será feita uma descrição dos procedimentos adotados para a geração das imagens de mesma similaridade.

4.4. Formato e utilização do vetor de características

Todas as imagens envolvidas no processo de categorização passam por um extrator, resultando na geração de um vetor de características. Neste trabalho, como informado anteriormente, utiliza-se a análise de Multirresolução (AMR) por meio da transformada *wavelet* de Haar para a geração desse vetor.

A AMR decompõe a imagem em 'n' níveis, em conjuntos de coeficientes, conforme já descrito no item 4.3.2.2.

O tamanho padronizado para o vetor de características, em seu nível 1 ($n=1$) é de 1024 posições, no nível 2 em 512 posições, nível 2 em 256 posições, e assim por diante.

Como neste trabalho o objetivo foi o de criar uma técnica e não o de comparar as diversas etapas na geração dos coeficientes C_n , não foi comparado os vetores gerados em cada um dos níveis de geração dos mesmos. O vetor de características utilizou o coeficiente n de valor 5, com base no trabalho de Silva (SILVA, 2009)

4.4.1. Base de dados para os experimentos de categorização e classificação

Para a validação da metodologia proposta foi utilizada uma base de dados real disponível no laboratório da UNIFESP (QualIM). A base de dados é composta de

994 casos reais. Cada caso é composto por quatro imagens mamográficas padrão DICOM. Para efeito de padronização, nos testes foram utilizadas as imagens mamográficas com as seguintes características: Crânio Caudal Esquerda (CCE), Crânio Caudal Direita (CCD), Médio Lateral Esquerdo (MLE) e Médio Lateral Direito (MLD). As outras imagens, nessa fase de validação, foram descartadas.

Nos testes realizados no escopo desta Tese, os vetores das imagens foram gerados independente da característica da imagem, conforme especificado no parágrafo anterior.

Todas as imagens mamográficas utilizadas têm tamanho de 10.757 bytes, independente do tipo de imagem mamográfica.

Para o processo de classificação de imagens, as mesmas foram classificadas por radiologistas experientes, sendo que essa informação está disponível na base de dados dos laudos, dessa maneira cada imagem está associada a uma classe BI-RADS específica.

Para a realização da classificação com o uso do SOM as imagens foram separadas em 20% para treinamento e 80 % para teste.

A distribuição de imagem pela classe não é uniforme. A inclusão de novas imagens no processo de classificação e categorização pode ser realizada a qualquer momento. A inclusão das informações na base de dados está associada à distância Euclidiana, isto é, as informações podem ser incluídas na base de dados e a extração das informações estará associada ao valor da distância na qual é armazenada. Caso o processo seja repetido, as informações serão adicionadas na base de dados de casos similares.

Com as informações dos casos similares o módulo MammoViewer exibe todas as imagens mamográficas associada a cada caso.

4.4.2. Categorização da imagem mamográfica com o Mapa SOM

No pré-processamento foi executado um processo de mapeamento, no qual cada imagem foi rotulada de acordo com a classe BI-RADS. Essa informação foi retirada do laudo textual e associada a cada vetor de características de cada imagem mamográfica. Essa associação é feita com base no laudo executado pelo medico experiente e está armazenada na base de dados de laudos textuais.

Na Figura 20 está ilustrado o formado do resultado da seleção dos melhores neurônios, obtidos por meio do processo de classificação com o uso do mapa SOM. A coluna BMU (Best Match Unit) representa os neurônios vencedores obtidos a partir da distância Euclidiana entre os vetores de características.

Imagem	Característica 1	Característica 2	...	Característica n	BMU 1	BMU 2	...	BMU Ultimo
1								
2								
3								
4								
...								
x								

Figura 20 - Ilustração do vetor de características e informações do processo SOM

4.5. Ambiente

O ambiente de desenvolvimento utilizado foi o Java® (J2SE, Java 2 Standard Edition) com o aplicativo NETBEANS® (www.netbeans.org). A versão utilizada do NETBEANS® é a 7.0.1.

Os dados são armazenados em uma Base de Dados que utiliza o sistema gerenciador de banco de dados da Microsoft®, SQL SERVER 2008 R2.

Para a utilização dos métodos de treinamento do classificador será utilizado o software MATLAB® (www.mathworks.com) com o pacote de funções somtoolbox (SOMTOOLBOX, 2007) para implementar os mapas auto-organizáveis (SOM).

4.6. Avaliação do processo de classificação das imagens mamográficas

Os resultados a seguir foram utilizados para caracterizar o processo de classificação. A base de dados atual é não é tão grande e os laudos realizados pelos médicos experientes apresentam divergências, o que pode acarretar em divergência no processo de classificação.

Os resultados obtidos estão sintetizados nas Tabelas 1 a 6.

Classe BI-RADS	0
Amostras	53
Verdadeiros Positivos	115
Falsos Positivos	97

Tabela 1 - Classe BI-RADS 0

Classe BI-RADS	1
Amostras	9
Verdadeiros Positivos	6
Falsos Positivos	30

Tabela 2 - Classe BI-RADS 1

Classe BI-RADS	2
Amostras	36
Verdadeiros Positivos	36
Falsos Positivos	108

Tabela 3 - Classe BI-RADS 2

Classe BI-RADS	3
Amostras	0
Verdadeiros Positivos	0
Falsos Positivos	0

Tabela 4 - Classe BI-RADS 3

Classe BI-RADS	4
Amostras	5
Verdadeiros Positivos	0
Falsos Positivos	20

Tabela 5 - Classe BI-RADS 4

Classe BI-RADS	5
Amostras	8
Verdadeiros Positivos	0
Falsos Positivos	32

Tabela 6 - Classe BI-RADS 5

Classe BI-RADS	6
Amostras	0
Verdadeiros Positivos	0
Falsos Positivos	0

Tabela 7 - Classe BI-RADS 6

A técnica proposta esta preparada para a alteração dos coeficientes no processo de extração de características, assim como no tamanho do mapa SOM utilizado para a classificação.

Conforme se pode notar pelas tabelas acima, houve divergência significativa entre a classificação do sistema aqui proposto, com a verdadeira classe BI-RADS de cada caso, conforme já descrito no início desta seção.

5. Considerações Finais

Conseguiu-se cumprir os objetivos estabelecidos para este trabalho, ou seja, criar uma técnica que permita o reconhecimento de similaridade em imagens mamográficas.

De forma a cumprir esse objetivo, foram utilizadas diversas técnicas, dentre elas as de processamento de imagens, envolvendo a Transformada *Wavelet*, bem como as técnicas de mineração de dados, especificamente o SOM (*Self-Organizing Map*) com base nas imagens mamográficas.

Além disso, como implementação prática, houve o desenvolvimento de um sistema computacional de reconhecimento de padrões em imagens mamográficas, possibilitando oferecer uma plataforma de treinamento dos residentes em radiologia na interpretação de imagens mamográficas.

Além do aprendizado proporcionado, podem ser citados como contribuições deste trabalho:

- Servir como referência para estudos futuros na utilização de outros extratores de características e classificadores; e
- Disponibilização de um sistema capaz de auxiliar o médico e o médico residente na elaboração de laudos médicos.

Apesar de implementar o uso da arquitetura SOM, este trabalho apresenta a possibilidade da utilização de qualquer mecanismo de classificação de agrupamento da forma em que esta concebido, onde o processo de classificação resulta na criação de uma base de dados atualizada da imagens similares.

O sistema possui potencial como ferramenta de auxílio ao trabalho de radiologistas, especialmente por permitir a comparação com imagens de mesmas características. Além disso, a estrutura implementada é flexível à incorporação e definição de novas técnicas para a descrição das imagens, permitindo a inclusão de outros classificadores.

5.1. Trabalhos Futuros

O sistema foi concebido de tal forma que a substituição do processo de classificação e extração de características pode ser modificada a qualquer momento, portanto outros métodos de classificação que por ventura seja mais poderão ser realizados.

O sistema permite sua internacionalização o que implica na adaptação para outras línguas.

Outro processo de validação poderá ser incorporado através da seleção de informações textuais uma vez que deverá ser incorporado o módulo MammoSegm, que permitirá a inclusão de informações úteis nas imagens, com o objetivo de aumentar a taxa de acerto no processo de classificação.

Podem ainda, os resultados serem utilizados na comparação para a elaboração de relatórios estatísticos que possam ajudar a identificar onde foram encontradas as maiores dificuldade na elaboração do laudo pelos residentes.

O esquema genérico implementado pode ser aplicado em outros tipos de imagens, com o objetivo de avaliar o desempenho dos algoritmos de classificação e de agrupamento.

Além disso, podem ser implementadas novas funcionalidades ao sistema Mammo. Uma delas está no fato de que a aplicação foi concebida para utilização local, mas pode vir a ser disponibilizada através da Internet.

Referencias Bibliográficas

- AGUILLAR, V.; BAUAD, S.; MARANHÃO, N. **Mama: diagnóstico por imagem: mamografia, ultrassonografia, ressonância magnética**. Rio de Janeiro: Revinter, 2009.
- ACR, AMERICAN COLLEGE OF RADIOLOGY. ***Illustrate breast imaging report and data system (BI-RADS™)***, 4ª Ed. Reston VA:American College of Radiology, 2003.
- ALBUZ, E.; KOCALAR, E.; KHOKHAR, A. A. Scalable color image indexing and retrieval using vetor wavelets. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, v. 13, n. 5, p. 851-860, 2001.
- ANTOINE, M.; ZAIANE, R. M.; COMAN, A.; ***Application of Data Mining Techniques for Médical Image. Second International Workshop on Multimedia Data Mining***, San Francisco, USA, Ago. 2001
- BAUER, W.; IGOT, J.P; LE, G.Y. ***Chronologic du cancer mammaire utilisant un Modele de Croissance de Gompertz***. *Ann Anat Pathol* vol , p. 39-56, 1980. (resgatar todo artigo)
- BISHOP, C. M. ***Pattern Recognition and Machine Learning***. Springer Science+Business Media, LCC, 2006.
- BUENO, J. M.; LAND, J. M.; ALTMAN, D. G. ***Statistical Methods for assessing agreement between two methods of clinical measurement***. Tese (Doutorado em Ciencia da Computação) – Escola de Engenharia de São Carlos. Universidade de São Paulo, São Carlos, 2001.
- CABENA, P.; JADJINIAM, P; STADLER, R; VERTHEES, J; ZANASI, A. ***Discovering Data Mining – From Concepts to Implementation***. Upper Sadle: Prentice Hall, 1998.
- CASTAÑON, C. A. ***Recuperação de imagens por conteúdo através de análise multiresolução por Wavelets***. Dissertação (Mestrado em Ciências de Computação e matemática Computacional) – Instituto de Ciências de Matemáticas e de Computação, São Carlos, 2003.
- CHO, Z.H.; JONES,J.P.;SINGH, M. ***Foundations of Médical Imaging***. New York-USA:Jophn Wiley ans Sons, 1993, 586 p.
- DAUBECHIEIES, I. ***Ten Lectures on Waveltes***, volume 61, CBNS-NSF Regional Convergence Series in Applied Mathematics, Rutger University and AT&T Bell Laboratories, 1992;

DENGLER, J.; BEHRENS, S.; DESAGA, J.F. **Segmentation of microcalcifications in mammograms**. IEEE Transactions on Medical Imaging v.12, n.4 p.632-634, 1993.

DINIZ, C.A .R; NETO, F. L. **Data Mining: Uma Introdução**. São Paulo: ABE – Associação Brasileira de Estatística, 2000

DOI, K. “**Overview on Research and Development Computer-Aided Diagnostic Schemes**” Semin Ultrasound CT MRI 25:404-410, 2004.

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification**, 2nd ed. New York: John Wiley & Sons, 2001.

DUNCAN, J. S.; AYACHE N. **Médical image analysis: Progress over two decades and the challenges ahead**. IEEE Transactions on Pattern Analysis and Machine intelligence 2000; vol. 22 n.1, pp 85-106, 2000.

EHLERS, R. S. **Análise de Séries Temporais**. Disponível em <<http://www.icmc.usp.br/~ehlers/stemp/stemp.pdf>> Acesso em 13 jul. 2011.

EVANS, A.L.; **The Evaluation of medical Images**, Adma Hilger Ltd, Bristol, Grato Britain, 1981.

FARIAS, M. C. Q.; **Aplicação da Transformada Wavelet na Compressão de Imagens**. Dissertação de Mestrado apresentado a Faculdade de Engenharia Elétrica e da Computação, UNICAMP, 1998.

FAWCETT, T. **An Introduction to ROC analysis**. Pattern Recognition Letters, pp. 861-874, 2005.

FIGUEIRA, R.M.A. **Miner: Um software de Inferência de Dependências Funcionais**. Monografia apresenta na Universidade Federal do Rio de Janeiro, 1998.

FREER, T.W.; ULISSEY, M. J.; “**Screening Mammography with Computer-aided Detection: Prospective Study of 12.860 Patients in a Community Breast Center**” Radiology 220:781-786, Setembro, 2001

FUKUNAGA, K.; **Introduction to Statistical Pattern Recognition**, Academic Press, Second Edition, 1990.

FURUIE, S.; REBELLO;M.S.; GUTIERREZ, M.A.; MORENO, R.A.; MOTTA, G.H.M.B.; BERTOZZO, N.; NARDON, F.B.; FIGUEIREDO, J.C.B.; OLIVEIRA, P.P.M.; TACHINARDI, U. **Prontuário eletrônico de pacientes: integrando informações clínicas e imagens médicas**. Revista Brasileira de Engenharia Médica, v.19 p.103-115, Dez. 2003.

GABOR, D. ; ***Theory of Communication***. J.IEE., vol 93 (Part III), p 439-457, 1946.

GATO, H. E. R.; NUNES, F. L. S.; SCHIEBEL H.; **Uma proposta de Recuperação de Imagens Mamográficas Baseada em Conteúdo**. Anais do IX Congresso Brasileiro de Informática em Saúde, 2004.

GIGER, M.L.; ***Computer-aided diagnosis of breast lesions in medical images***. Computing in Science & Engineering, v.2,n.5,p.39-45, 2000.

GLASSER, O. ***Wilhelm Conrad Röntgen and the Early History of Röntgen Rays***. São Francisco-USA: Norman Publising, 1993.

GONZALES, R.C.; WOODS, R. E.; **Processamento Digital de Imagem**. Ed. Pearson Printece Hall, 2010

HAN, J.; KAMBER, M. ***Data Mining Concepts and Techniques***, Morgan Kaufmann São Francisco, CA, 2001.

HARALICK, R. M.; ***Statistical and Structural Approaches to Texture – Proceeding of the IEEE***, v.67, n. 5, May, pp 786-805, 1979.

HAYKIN, S. **Redes Neurais Princípios e Prática**, 2.ed. Bookman, 2001.

HU, M. K. **Pattern recognition by moment invariants**. IRE Transactions on Information Theory, v 49, pp. 1428, 1961.

HU, M. K. **Visual Pattern recognition by moment invariants**. IRE Transactions on Information Theory, v. 8, pp. 179-187, 1962.

HUANG, H.K.; ***Pacs and Imagins Informatics***, Ed. Wiley-Liss, 2004.

INCA, **INSTITUTO NACIONAL DO CANCER**. Ministério da Saúde. Falando sobre doenças de mama, 2000.

INCA, **INSTITUTO NACIONAL DO CANCER**. Ministério da Saúde. Mamografia: da pratica ao controle, 2007.

INCA, **INSTITUTO NACIONAL DO CANCER**. Ministério da Saúde. Disponível em <http://www1.inca.gov.br/conteudo_view.asp?id=322>. Acesso em 17 mar. 2010.

INCA, **INSTITUTO NACIONAL DO CANCER**. Ministério da Saúde. Disponível em <http://www.inca.gov.br/pqrt/download/tec_int/cap2_p2.pdf>. Acesso em 05 abr. 2011a.

INCA, **INSTITUTO NACIONAL DO CANCER**. Ministério da Saúde. Disponível em <<http://www.inca.gov.br/estimativa/2010>>. Acesso em 21 fev. 2011b.

INCA, **INSTITUTO NACIONAL DO CANCER**. Ministério da Saúde. Disponível em <http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama/cancer_mama++>. Acesso em 25 out. 2011c.

KEMP, C.; BARACAT, F.F.; ROSTAGNO, R. **Lesões Não Palpáveis da Mama Diagnóstico e Tratamento**. 1ª. Ed., Rio de Janeiro, Revinter , 2003

KIM, C.; CHUNG, C. W. **A multi-step approach for partial similarity search inlarge image data using histogram intersection**. In: Information and software technology, v. 45, 4 ed. Elsevier, p.203-215(13), 2003;

KINOSHITA, S.K; PEREIRA JR., R.R.; RODRIGUES, J.A.H.; AZEVEDO-MARQUES, P.M. **Recuperação Baseada em Conteúdo de Imagens Mamográficas: Atributos Visuais de Forma, Espectrais no domínio de Randon e Granulometria**. IX Congresso Brasileiro de Informática em Saúde, 2004, Ribeirão Preto, CBIS2004, 2004.

KOHONEN, T. **Self-Organizing Maps**. 3 Ed. Springer, 2001.

KOPANS, D. B.; **Imagem de Mama** 2ª edição Revinter , 2000. Confirmar

LASCA, V.; PATROCÍNIO, A. C.; RADY, J. ; BITELLI, R. **Busca de Imagens por grau de complexidade em banco de imagens mamográficas**, 2008.

LEHMANN, T. M. **From plastic to gold: a unified classification scheme for reference standards in medical image processing**. In: *Proceedings of SPIE Image Processing*. San Jose, CA, USA:SPIE, p. 1819-1827, 2002.

LEHMANN, T.M.;MEINZER, H.P.; TOLDOXDORF,T. **Advances in biomédical image analysis: Part, Present nad future challenges**. *Methods of Information in Medicine*, n.43, p.308-314, 2004.

LEHMANN, T.M.;GULDA, M. O.; DESELAERSB, T.; KEYSERSB, D. SCHUBETC, H.; SPITZERA, K.; NEYB, H.; WEINC, B. B. **Automatic categorization of medical images for content-based retrieval and data mining**. *Computerized Médical Imaging and Graphics*, v.29, p.143-155, 2005.

LI Q, LI F, SUZUKI K, SHIRAISHI J, ABE, H.; ENGELMANN, R, et al. **Computer-aided diagnosis in thoracic CT**. *Semin Ultrasound CT MR*. 26(5):357-63, 2005.

MARQUES FILHO, O.; VIERA NETO, H. **Processamento Digital de Imagens** . Ed. Brasport, 1999.

MARQUES. J.; TRAINA. A. J. M. **Realimentação de Relevância: Integração do Conhecimento do especialista com a Recuperação de Imagens por Conteúdo**. In: VI Workshop de Informática Médica – WIM’2006 (junto ao VI Simpósio Brasileiro

de Qualidade de Software – SBQS), 2006, Vitoria Anais do VI Workshop de Informática Médica WIM'2006, Vitoria, v.1 p. 83-93, 2006.

MARQUES, P. M. A.; HONDA, M. H. RODRIGUES, J. A. H. SANTOS, R. R. TRAINA, A. J. M. TRAINA JR. C.; BUENO, J. M. **Recuperação de Imagem Baseado em Conteúdo: Uso de atributos de Textura para Caracterização de Microcalcificações Mamográficas**. Revista Brasileira de Radiologia, vol. 35, pp 93-98, 2002.

MASCARENHAS, N. D. A.; VELASCO, F. R. D. ; **Processamento Digital de Imagens**, Editora Kapelusz S.A., Buenos Aires, Argentina, 1989

METZ, C. **ROC Methodology in Radiologic Imaging**. Investigate Radiology, v. 21, n.9, p.720-733, Sep. 1986.

METZ, C. **ROC Analysis in medical imaging: a tutorial review of the literature**. Radiological Physics and Technology, Out, 2007.

MÜLLER, H.;MICHOUX, N. BANDON, D.; GEISSBUHLER, A. **A review of content-based image retrieval system in medical applications-clinical benefits and future directions**. International Journal of Medical Informatic, v. 73 p. 1-23, 2004.

MUNERATI, C. **Remodelagem de um Banco de Dados e Imagens Digitais utilizados para o Treinamento de Residentes em Mamografia**. Dissertação de Mestrado apresentado à Escola Paulista de Medicina - UNIFESP, 2011.

NUNES, F. L. **Investigações em Processamento de Imagens Mamográficas para Auxílio ao Diagnóstico de Mamas Densa**. Tese de Doutorado apresentado ao Instituto de Física de São Carlos, 2001

PATROCINIO, ANA CLAUDIA. **Classificador automáticos de achados mamográficos em imagens digitais de mamas densas utilizando técnicas híbridas**. Tese (Doutorado em Engenharia Elétrica) – Escola de Engenharia de São Carlos. Universidade de São Paulo, São Carlos, 2004.

PINOTTI, J.A.; RICCI, M.D.; PINOTTI, M. **Diagnóstico das lesões Mamaria**. Ed. Revinter, 2006.

PIRES, S. R. **Software Gerenciador de uma Base de Dados e de Imagens Mamográficas Classificadas Segundo um Índice de Qualidade**. Dissertação (Mestrado em Ciências) - Universidade Federal de São Paulo, 2003.

POP, **Portal de Oncologia Portugues**, Disponível em <<http://www.pop.eu.com/portal/publico-geral/tipos-de-cancro/cancro-da-mama/qual-a-estrutura-normal-da-mama.html>>. Acesso em 05 abr. 2011.

SANTOS, M. **Ambiente para avaliação de algoritmos de processamento de imagens médicas** Tese (Doutorado em Engenharia Elétrica) – Escola Politécnica da Universidade de São Paulo, 2006.

SANTOS, M. ; FURUIE, S. S. **Medical image database for software and algorithm evaluation**. In *SPIE Proceeding: Medical Image 2005*. San Diego, CA: SPIE, v. 5748, 2005.

SCHIMIDT, R. A.; JIANG, Y. ;NISHIKAWA, R.M.; ALICIA Y. TOLEDANO, A. Y.; Doi, K. **Potential of Computer-aided Diagnosis to Reduce Variability in Radiologists' Interpretations of Mammograms Depicting Microcalcifications**. Radiology v.220, pp. 787–794, 2001.

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

SIEGEL , E. L.; KOLONDNER,R.M. **Filmless Radiology – A future Perspective**. MJAFI, vol. 59, n. 3 p 187-188, 2003

SILVA, C. Y. **Extração de características de imagens médicas utilizando wavelets para mineração de imagens e auxílio ao diagnóstico**. Dissertação (Mestrado em Ciências) – Instituto de Ciências de Matemáticas e de Computação, São Carlos, 2007.

SILVA, L. A. da; **Categorização de imagens médicas baseada em transformação wavelet e mapas auto-organizáveis**. Tese (Doutorado em Engenharia Elétrica) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Sistemas Eletrônicos, São Paulo, 2009.

SILVA, L. A. da; MORENO, R.; DEL-MORAL-HERNANDEZ, E.; FURUIE, S. S. **Mapa de Características Auto-Organizável para Sistemas de recuperação de Imagens por Conteúdo**. Proceeding of CBEB 2006 – XX Congresso Brasileiro de Engenharia Biomédica, p. 254-257, 2006.

SILVA, L. A. da; MORENO, R.; DEL-MORAL-HERNANDEZ, E.; FURUIE, S. S. **Medical Image Categorization Based on Wavelet Transform and Self-Organizing Map**. Seventh International Conference on Intelligent System Design and Applications, ISDA2007, Rio de Janeiro, RJ, p. 353-356, 2007.

SILVA, L. A. da; SANDMANN, H.; DEL-MORAL-HERNANDEZ, E. **A Self-Organizing Architecture of Recursive Elements for Continuous Learning**. IEEE World Congress on Computational Intelligence. International Joint Conference on Neural Network. Hong Kong, China, p. 2784-2791, 2008.

SOMTOOLBOX. **SOM Toolbox a function package for Matlab5 implementing the Self-Organizing Map (SOM)**, Out, 2007. Http <http://www.cis.hut.fi/somtoolbox/>

TAN, P.; STEINBACH, M. KUMAR, V. **Introdução ao Data Mining Mineração de Dados**. Rio de Janeiro: Editora Ciência Moderna Ltda. 2009.

THEODORIDIS, S.; KOUTROUMBAS, K. ***Pattern Recognition***. Ed. Academic Press, 2009.

VIEIRA, A. V. TOIGO, F. T. **Classificação BI-RADS™: Categorização de 4968 Mamografias**. RadiolBras vol. 35 n.4 São Paulo, 2002.

Apêndice A – Especificação Funcional

Especificação Funcional

Mammo

Versão 1.0

Jamilson Bispo dos
Santos

Índice

1.	Histórico de revisão	79
2.	Introdução	80
2.1.	Descrição geral da solução	80
2.2.	Referências	80
3.	Requisitos funcionais.....	81
3.1.	Atores	81
3.1.1.	Administrador.....	81
3.1.2.	Médico.....	81
3.1.3.	Pacs.....	81
3.1.4.	Residente.....	81
3.2.	Casos de Uso.....	82
3.2.1.	Visão geral	82
	Descrição	82
	Diagrama	83
3.2.2.	Manter do cadastro de usuários (Manter Usuários).....	84
	Descrição	84
	Diagrama	85
	Modelo(s)	85
3.2.3.	Autenticação do Usuário (Autenticar Usuários).....	88
	Descrição	88
	Diagrama	88
	Modelo(s)	89
3.2.4.	Carga de Imagens Médicas (Carregar Imagem)	91
	Descrição	91
	Diagrama	91
	Modelo(s)	91
3.2.5.	Classificação de imagens de acordo com a similaridade (Realizar análise médica) 92	
	Descrição	92
	Diagrama	93
	Modelo(s)	93
3.2.6.	Realizar treinamento	95
	Descrição	95
	Diagrama	95
	Modelo(s)	96

4. Telas do sistema	99
4.1. Mapa de navegação do sistema Mammo Mapa com os nomes das telas.....	99
4.2. Tela Acesso ao sistema Mammo	100
4.3. Tela para seleção dos laudos (Caso).....	101
4.4. Tela para visualização das imagens similares.....	102
4.5. Tela para visualização do laudo executado.....	103
4.6. Tela para visualização das imagens segmentadas	104
4.7. Tela para visualização do laudo.....	105

1. Histórico de revisão

Versão	Data	Autor	Descrição
1.0	01/02/2013	Jamilson Bispo dos Santos	Criação do documento

2. Introdução

2.1. Descrição geral da solução

O objetivo deste documento é descrever o sistema Mammo.

O sistema estará acessível por computadores que possuam o aplicativo instalado. É oferecida uma interface intuitiva e prática. A ferramenta ficará disponível a todos os médicos e residentes envolvidos no treinamento de residentes em radiologia.

O sistema será alimentado com imagens médicas digitalizadas ou obtidas através do sistema PACS, as imagens digitalizadas serão utilizadas para a realização de laudo pelos médicos residentes em radiologia.

Para toda imagem médica serão realizados dois laudos, por médicos diferentes, resultando em um único laudo que será utilizado como base de comparação com os laudos gerados pelo residente.

Para os casos em que se identificou alguma anomalia o médico poderá gerar uma nova imagem contendo o segmento da imagem com a anomalia.

As imagens serão classificadas e agrupadas de acordo com a sua similaridade com o objetivo de serem utilizadas pelo residente para efeitos comparativos. Haverá a disponibilidade de extração de relatório de evolução do treinamento.

2.2. Referências

3. Requisitos funcionais

3.1. Atores

3.1.1. Administrador

O administrador é o usuário com acesso irrestrito às imagens e aos laudos previamente realizados. Ele também é o responsável pela gestão dos usuários autorizados a acessar o sistema.

3.1.2. Médico

O médico é o usuário responsável pela criação do laudo prévio das imagens. A partir das suas informações é gerado um laudo padrão utilizado para comparação com o laudo realizado pelo residente. Ele também é o responsável pela segmentação das imagens.

3.1.3. Pacs

O Pacs é o sistema utilizado pelas entidades de saúde para armazenar as imagens médicas. As imagens por ele armazenadas são utilizadas para criação da base de dados do sistema MammoAdviser.

3.1.4. Residente

O ator residente representa todos os médicos em fase de treinamento que podem utilizar o sistema para comparação de seus laudos com os laudos gerados por médicos experientes, visando assim o aprimoramento de seus diagnósticos.

3.2. Casos de Uso

3.2.1. Visão geral

Descrição

O sistema deve ser utilizado para a realização do treinamento de médicos residentes em radiologia.

O sistema Mammo contempla as seguintes funcionalidades:

- Manutenção do cadastro de usuários (Manter Usuários);
- Autenticação do usuário (Autenticar Usuários);
- Carga de imagens médicas (Carregar Imagem);
- Classificação de imagens de acordo com a similaridade (Realizar Análise Médica);
- Geração de imagens segmentadas (Selecionar Caso);
- Pesquisa de imagens (Visualizar Imagens Similares);
- Geração do laudo template (padrão) (Realizar Laudo Template);
- Geração do laudo dos médicos residentes (Realizar Laudo Treinamento);
- Relatórios de acompanhamento do treinamento (Gerar Relatórios de Avaliação do Residente).

Diagrama

Diagrama de caso de uso para realização de treinamento para médicos residentes

3.2.2. Manter do cadastro de usuários (Manter Usuários)

Descrição

Esta funcionalidade permite que o administrador do sistema cadastre os usuários, atribuindo à eles o perfil adequado, de acordo com a função por ele exercida na instituição. Além da inclusão do usuário, é permitida a consulta e a alteração de seus dados.

Segue a distribuição das funcionalidades, de acordo com os perfis de usuários:

Residente: para os médicos em treinamento.

Os usuários cadastrados com este perfil têm acesso às funcionalidades específicas para realização de seu treinamento, são elas:

- Autenticação do usuário;
- Pesquisa de imagens;
- Geração do laudo dos médicos residentes;
- Relatórios de acompanhamento do treinamento.

Médico: para os médicos experientes.

Estes usuários têm acesso às seguintes funcionalidades:

- Geração de imagens segmentadas;
- Geração do laudo template (padrão).

Administrador: responsável pela execução de atividades administrativas.

Tem acesso às seguintes funcionalidades:

- Manutenção do cadastro de usuários;
- Carga de imagens médicas;
- Classificação de imagens de acordo com a similaridade.

Diagrama

Diagrama de caso de uso de Manter Usuários

Modelo(s)

Nome	Manter Usuários	
Atores	Administrador.	
Pré-condições	Usuário com perfil de administrador logado no sistema.	
Pós-condições	Manutenção do cadastro de usuários realizada.	
Fluxo Principal	1. Administrador	Seleciona o botão para acesso à funcionalidade de manutenção do cadastro de usuários.
	2. Sistema	Exibe o Menu para seleção da opção desejada: - Inclusão; - Alteração; - Consulta.
	3. Administrador	Seleciona a opção para inclusão de novo usuário.
	4. Sistema	Verifica que foi selecionada a opção de inclusão.
	5. Sistema	Chama o caso de uso "Incluir Usuário"
Fluxo Alternativo A	3. Administrador	Seleciona a opção para alteração de usuário.
	4. Sistema	Verifica que foi selecionada a opção de alteração.
	5. Sistema	Chama o caso de uso "Alterar Usuário"
Fluxo Alternativo B	3. Administrador	Seleciona a opção para consulta de usuário.
	4. Sistema	Verifica que foi selecionada a opção de consulta.
	5. Sistema	Chama o caso de uso "Consultar Usuário"

Nome	Incluir Usuário	
Atores	Administrador.	
Pré-condições	O Administrador selecionou a opção de inclusão de usuários.	
Pós-condições	Inclusão de novo usuário realizada.	
Fluxo Principal	1. Sistema	Exibe a tela para preenchimento dos dados do usuário.
	2. Administrador	Preenche as informações do usuário.
	3. Sistema	Verifica que as informações fornecidas atendem aos requisitos.
	4. Sistema	Solicita a confirmação da inclusão.
	5. Administrador	Confirma a inclusão.
	6. Sistema	Efetiva as informações na base de dados.
Fluxo Alternativo A	3. Sistema	Verifica que as informações fornecidas não atendem aos requisitos.
	4. Sistema	Solicita a correção da informação de acordo com as regras de negócio.
	5. Sistema	Volta ao passo 1 do Fluxo Principal.

Nome	Alterar Usuário	
Atores	Administrador.	
Pré-condições	O Administrador selecionou a opção de alteração de usuários.	
Pós-condições	Alteração dos dados do usuário realizada.	
Fluxo Principal	1. Sistema	Exibe a tela para alteração dos dados do usuário.
	2. Sistema	Solicita a identificação do usuário que será alterado.
	3. Administrador	Informa o código do usuário.
	4. Sistema	Localiza os dados do usuário na base de dados.
	5. Sistema	Exibe a tela preenchida com os dados do usuário recuperados da base de dados.
	6. Administrador	Altera a informação desejada, de acordo com as regras de negócio.
	7. Sistema	Verifica que as informações fornecidas atendem aos requisitos.
	8. Sistema	Solicita a confirmação da alteração.
	9. Administrador	Confirma a alteração.
	10. Sistema	Efetiva as informações na base de dados.
Fluxo Alternativo A	7. Sistema	Verifica que as informações fornecidas não atendem aos requisitos.
	8. Sistema	Solicita a correção da informação de acordo com as regras de negócio.
	9. Sistema	Volta ao passo 6 do Fluxo Principal.
Fluxo Alternativo B	4. Sistema	Não localiza os dados do usuário na base de dados.
	5. Sistema	Informa que o usuário não foi localizado.
	6. Sistema	Volta ao passo 1 do Fluxo Principal.

Nome	Consultar Usuário	
Atores	Administrador.	
Pré-condições	O Administrador selecionou a opção de consulta de usuários.	
Pós-condições	Consulta dos dados do usuário realizada.	
Fluxo Principal	1. Sistema	Exibe a tela para consulta dos dados do usuário.
	2. Sistema	Solicita a identificação do usuário.
	3. Administrador	Informa o código do usuário.
	4. Sistema	Localiza os dados do usuário na base de dados.
	5. Sistema	Exibe a tela preenchida com os dados do usuário recuperados da base de dados.
Fluxo Alternativo A	4. Sistema	Não localiza os dados do usuário na base de dados.
	5. Sistema	Informa que o usuário não foi localizado.
	6. Sistema	Volta ao passo 1 do Fluxo Principal.

3.2.3. Autenticação do Usuário (Autenticar Usuários)

Descrição

O sistema deve disponibilizar uma tela para efetuar o login no Sistema, onde o usuário deve informar o nome de acesso e a sua senha. É disponibilizada a opção, para quando o usuário esquecer a senha, o sistema automaticamente enviar para um e-mail previamente cadastrado, uma nova senha do usuário.

Diagrama

Diagrama de caso de uso para autenticar usuário

Modelo(s)

Nome	Autenticar usuário	
Atores	Usuário.	
Pré-condições	Sistema instalado.	
Pós-condições	Usuário logado no sistema MammoAdviser.	
Fluxo Principal	1. Usuário	Seleciona o Ícone para informação dos dados de acesso ao sistema.
	2. Sistema	Abre a tela de acesso ao Sistema com os campos disponíveis: - nome de acesso - senha de acesso.
	3. Usuário	Informa os dados de acesso.
	4. Sistema	Disponibiliza a opção: - esqueci minha senha - efetuar acesso ao Sistema
	5. Sistema	Verifica que o usuário selecionou a opção efetuar acesso ao sistema.
	6. Sistema	Chama caso de uso "Efetuar acesso ao sistema".
Fluxo Alternativo A	5. Sistema	Verifica que o usuário selecionou a opção esqueci a minha senha.
	6. Sistema	Chama caso de uso "Enviar senha".

Nome	Efetuar Acesso Sistema	
Atores	Usuário.	
Pré-condições	O usuário selecionou a opção para efetuar acesso ao sistema.	
Pós-condições	Usuário autenticado no Sistema.	
Fluxo Principal	1. Sistema	Verifica que o nome de acesso informado existe na tabela de usuários do Sistema e está ativo.
	2. Sistema	Verifica que a senha de acesso informada é a mesma encontrada na tabela de usuários do Sistema, para o nome de acesso acima.
	3. Sistema	Atualiza o campo data/hora do último acesso da tabela usuários do Sistema com a data e hora atual.
	4. Sistema	Libera o acesso às demais funcionalidades do sistema.
	5. Sistema	Finaliza o caso de uso.
Fluxo Alternativo A	1. Sistema	Verifica que o usuário informado não é válido.
	2. Administrador	Ajusta a informação do usuário.
	3. Sistema	Volta ao passo 1 do Fluxo Principal.
Fluxo Alternativo B	2. Sistema	Verifica que a senha informada não é válida.
	3. Administrador	Corrige a senha informada.
	4. Sistema	Volta ao passo 1 do Fluxo Principal.

Nome	Enviar senha	
Atores	Usuário.	
Pré-condições	É necessário ter selecionado a opção esqueci minha senha.	
Pós-condições	Nova senha enviada para o e-mail do usuário.	
Fluxo Principal	1. Sistema	Verifica que o nome de acesso informado existe na tabela de usuários do Sistema e está ativo.
	2. Sistema	Seleciona o e-mail cadastrado para o nome de acesso informado.
	3. Sistema	Gera uma nova senha para o usuário.
	4. Sistema	Atualiza o registro do usuário encontrado com a nova senha gerada.
	5. Sistema	Grava os dados.
	6. Sistema	Envia o e-mail de notificação.
	7. Sistema	Apresenta a mensagem de sucesso na geração da nova senha.
	8. Sistema	Finaliza o caso de uso.
Fluxo Exceção de	5. Sistema	Caso erro na gravação dos dados desfaz a transação do caso de uso, e exibe mensagem de erro.
	6. Sistema	Finaliza o caso de uso.
Fluxo Exceção de	6. Sistema	Caso erro no envio do e-mail apresenta mensagem de erro.
	7. Sistema	Finaliza o caso de uso.

3.2.4. Carga de Imagens Médicas (Carregar Imagem)

Descrição

Para realização do treinamento se faz necessário carregar as imagens a serem avaliadas pelos Médicos e pelos Residentes, no sistema Mammo.

Esta função é realizada pelo Administrador, que deve ter acesso à base de dados do sistema PACS de onde as imagens são extraídas.

Diagrama

Diagrama de caso de uso carregar imagem

Modelo(s)

Nome	Carregar Imagem	
Atores	Administrador	
Pré-condições	Existem imagens disponíveis para carga no Sistema PACS.	
Pós-condições	Imagens médicas carregadas na base de dados do sistema Mammo.	
Fluxo Principal	1. PACS	Disponibiliza base de dados de imagens médicas para carga.
	2. Administrador	Aciona a carga das imagens médicas no sistema Mammo.
	3. Sistema	Efetiva a carga das imagens médicas no sistema Mammo.
	4. Administrador	Verifica que a carga foi realizada com sucesso.
	6. Sistema	Verifica que a opção para classificação das imagens foi selecionada.
	7. Sistema	Chama o caso de uso "Classificar imagens por similaridade".

Nome	Classificar imagens por similaridade	
Atores	Administrador	
Pré-condições	A opção de classificação de imagens por similaridade foi selecionada.	
Pós-condições	A base de dados de imagens classificada por similaridade foi gerada.	
Fluxo Principal	1. Sistema	Associa as imagens ao caso.
	2. Sistema	Extraí características.
	3. Sistema	Associa a imagem à classificação BI-RADS definida pelo laudo template.
	3. Sistema	Classifica as imagens por similaridade.
	4. Administrador	Verifica que a classificação foi realizada com sucesso.
	5. Administrador	Finaliza o caso de uso.
Fluxo de Exceção	4. Sistema	Verifica que houve falha na classificação das imagens.
	5. Sistema	Exibe mensagem de erro.
	6. Sistema	Finaliza o caso de uso.

3.2.5. Classificação de imagens de acordo com a similaridade (Realizar análise médica)

Descrição

Para cada caso dos quais as imagens foram carregadas, os médicos devem realizar análise e cadastrar no sistema Mammo seus respectivos laudos, bem como, gerar a segmentação destas imagens para que sirvam como modelo para comparação com o resultado do laudo elaborado pelo Residente.

Diagrama

Diagrama de caso de uso realizar análise médica

Modelo(s)

Nome	Realizar análise médica	
Atores	Médico	
Pré-condições	Estar logado no sistema com perfil de Médico.	
Pós-condições	O laudo do Médico e a segmentação das imagens foram gerados para as imagens analisadas.	
Fluxo Principal	1. Médico	Seleciona um caso para realização da análise médica.
	2. Sistema	Disponibiliza as seguintes opções de análise para o caso selecionado: - Segmentação da imagem; - Geração do laudo template.
	3. Médico	Seleciona o tipo de análise médica desejada.
	4. Sistema	Verifica que a opção de segmentação de imagem foi selecionada.
	5. Sistema	Chama o caso de uso "Gerar Imagem Segmentada".
Fluxo Alternativo A	4. Sistema	Verifica que a opção para realizar laudo template foi selecionada.
	5. Sistema	Chama o caso de uso "Realizar Laudo Template".

Nome	Gerar Imagem Segmentada	
Atores	Médico	
Pré-condições	Ter sido selecionada a opção para geração de imagens segmentadas.	
Pós-condições	As imagens analisadas foram segmentadas.	
Fluxo Principal	1. Sistema	Disponibiliza as imagens para segmentação.
	2. Médico	Seleciona a imagem para a segmentação.
	3. Médico	Analisa a imagem e gera a segmentação.
	4. Sistema	Armazena a imagem segmentada.
Fluxo de Exceção	4. Sistema	Verifica que houve falha na geração da imagem segmentada.
	5. Sistema	Exibe mensagem de erro.
	6. Sistema	Finaliza o caso de uso.

Nome	Realizar Laudo Template	
Atores	Médico	
Pré-condições	Ter sido selecionada a opção para realização do laudo template.	
Pós-condições	O laudo foi gerado para o caso analisado.	
Fluxo Principal	1. Sistema	Disponibiliza as telas para preenchimento do laudo.
	2. Médico	Preenche as informações do laudo.
	3. Médico	Confirma as informações do laudo.
	4. Sistema	Verifica que as informações atendem aos requisitos.
	5. Sistema	Efetiva o laudo na base de dados.
Fluxo de Exceção	5. Sistema	Verifica que houve falha na efetivação do laudo.
	6. Sistema	Exibe mensagem de erro.
	7. Sistema	Finaliza o caso de uso.

3.2.6. Realizar treinamento

Descrição

Para cada caso dos quais as imagens foram carregadas, os residentes podem realizar análise e cadastrar no sistema Mammo seus respectivos laudos.

Os residentes têm à disposição, uma base de dados com outros casos de imagens similares, que podem ser acessadas para comparação com as imagens do caso cuja análise está em andamento.

O laudo padrão gerado pelo médico também esta acessível, para que o residente possa compará-lo com o laudo por ele realizado.

O residente pode buscar as imagens segmentadas dos casos analisados, onde pode visualizar detalhes destacados pelo médico considerados importantes para o diagnóstico.

O residente tem à sua disposição, relatórios gerados pelo módulo MammoAdviser com avaliações de seu desempenho no treinamento.

Diagrama

Diagrama de caso de uso realizar treinamento

Modelo(s)

Nome	Selecionar Laudo	
Atores	Residente	
Pré-condições	Estar logado no sistema com perfil de Residente.	
Pós-condições	Residente provido de informações complementares para análise de um caso específico, melhorando sua capacitação técnica e a qualidade dos laudos por ele gerados.	
Fluxo Principal	1. Residente	Seleciona um caso para realização da análise.
	2. Sistema	Disponibiliza as seguintes opções de análise para o caso selecionado: - Visualização das imagens do caso executado; - Visualização das imagens similares; - Visualizar das imagens segmentadas.
	3. Residente	Seleciona a opção desejada.
	4. Sistema	Verifica que a opção para visualização da imagem do caso executado foi selecionada.
	5. Sistema	Chama o caso de uso "Visualizar Imagem Residente".
Fluxo Alternativo A	4. Sistema	Verifica que a opção para visualização das imagens similares foi selecionada.
	5. Sistema	Chama o caso de uso "Visualizar Imagens Similares".
Fluxo Alternativo B	4. Sistema	Verifica que a opção para visualização das imagens segmentadas foi selecionada.
	5. Sistema	Chama o caso de uso "Visualizar Imagens Segmentadas".

Nome	Realizar Laudo Treinamento	
Atores	Residente	
Pré-condições	Ter sido acessada a funcionalidade para realização do laudo.	
Pós-condições	O laudo foi gerado para o caso analisado.	
Fluxo Principal	1. Sistema	Disponibiliza as telas para preenchimento do laudo.
	2. Residente	Preenche as informações do laudo.
	3. Residente	Confirma as informações do laudo.
	4. Sistema	Verifica que as informações atendem aos requisitos.
	5. Sistema	Efetiva o laudo na base de dados.
Fluxo Exceção de	5. Sistema	Verifica que houve falha na efetivação do laudo.
	6. Sistema	Exibe mensagem de erro.
	7. Sistema	Finaliza o caso de uso.

Nome	Visualizar Imagem Residente	
Atores	Residente	
Pré-condições	Ter sido selecionada a opção para visualização da imagem do caso executado.	
Pós-condições	As imagens do caso executado serão disponibilizadas para análise.	
Fluxo Principal	1. Sistema	Disponibiliza as imagens do caso executado para análise.
	2. Residente	Seleciona a imagem desejada.
	3. Residente	Analisa a imagem selecionada.
	4. Residente	Finaliza o caso de uso.

Nome	Visualizar Imagens Similares	
Atores	Residente	
Pré-condições	Ter sido selecionada a opção para visualizaçãodas imagens similares.	
Pós-condições	As imagens similares serão disponibilizadas para análise.	
Fluxo Principal	1. Sistema	Disponibiliza as imagens similares para análise.
	2. Residente	Seleciona a imagem desejada.
	3. Residente	Analisa a imagem selecionada.
	4. Residente	Finaliza o caso de uso.

Nome	Visualizar Imagens Segmentadas	
Atores	Residente	
Pré-condições	Ter sido selecionada a opção para visualização das imagens segmentadas.	
Pós-condições	As imagens segmentadas serão disponibilizadas para análise.	
Fluxo Principal	1. Sistema	Disponibiliza as imagens similares para análise.
	2. Residente	Seleciona a imagem desejada.
	3. Residente	Analisa a imagem selecionada.
	4. Residente	Finaliza o caso de uso.

Nome	Consultar Laudo	
Atores	Residente	
Pré-condições	Ter sido selecionada a visualização de imagem de um caso.	
Pós-condições	Os dados do laudo serão disponibilizados para visualização.	
Fluxo Principal	1. Residente	Seleciona a opção de consulta do laudo.
	2. Sistema	Exibe a tela com as informações do laudo.
	3. Sistema	Finaliza o caso de uso.

Nome	Avaliar laudo do residente	
Atores	Residente	
Pré-condições	Ter sido selecionada a opção para avaliação do laudo.	
Pós-condições	O laudo do residente será comparado com o laudo padrão.	
Fluxo Principal	1. Sistema	
	2. Residente	
	3. Residente	
	4. Residente	

Nome	Gerar relatórios de avaliação do residente	
Atores	Residente	
Pré-condições	Ter sido selecionada a opção para geração de relatórios.	
Pós-condições	O relatório de avaliação do residente será gerado.	
Fluxo Principal	1. Sistema	
	2. Residente	
	3. Residente	
	4. Residente	

4. Telas do sistema

4.1. Mapa de navegação do sistema Mammo Mapa com os nomes das telas

Mapa de navegação do sistema Mammo

4.2. Tela Acesso ao sistema Mammo

Esta tela representa o caso de uso “Autenticar Usuário”.

Tela Acesso ao sistema Mammo

4.3. Tela para seleção dos laudos (Caso)

Esta tela representa o caso de uso “Selecionar Laudo”.

Tela de Seleção do Caso a ser avaliado

4.4. Tela para visualização das imagens similares

Esta tela representa o caso de uso “Visualizar Imagens Similares”.

Tela para visualização das imagens Similares

4.5. Tela para visualização do laudo executado

Esta tela representa o caso de uso “Visualizar Imagem Residente”.

Tela para visualização da imagem do laudo executado

4.6. Tela para visualização das imagens segmentadas

Esta tela representa o caso de uso “Visualizar Imagens Segmentadas”.

Tela para visualização das imagens segmentadas

4.7. Tela para visualização do laudo

Esta tela representa o caso de uso “Consultar Laudo”.

Tela para visualização dos dados do laudo.