



Rede Nordeste de Biotecnologia
Programa de Pós-graduação em Biotecnologia

FRANCISCO ADELTON ALVES RIBEIRO

OTIMIZAÇÃO DE IMAGENS MAMOGRÁFICAS PARA AUXÍLIO AO DIAGNÓSTICO
MÉDICO

TERESINA
2017

FRANCISCO ADELTON ALVES RIBEIRO

OTIMIZAÇÃO DE IMAGENS MAMOGRAFICAS PARA AUXILIO AO DIAGNÓSTICO
MÉDICO

Tese de Doutorado apresentada ao curso de Pós-Graduação RENORBIO - Rede Nordeste de Biotecnologia da Universidade Federal do Piauí, como requisito para obtenção do título de Doutor.

Área de Concentração: Biotecnologia em Saúde

Orientação: Dr. Benedito Borges da Silva

TERESINA
2017

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Biblioteca Comunitária Jornalista Carlos Castello Branco
Serviço de Processamento Técnico

R484o Ribeiro, Francisco Adelson Alves.
Otimização de imagens mamográficas para auxílio ao diagnóstico médico / Francisco Adelson Alves Ribeiro. -- 2017.
122 f.

Tese (Doutorado) – Programa de Pós-Graduação em Biotecnologia – RENORBIO, Universidade Federal do Piauí, Teresina, 2017.

“Orientação: Dr. Benedito Borges da Silva.”

1. Mamografia. 2. Câncer de mama – Diagnóstico.
3. Processamento de imagens. 4. Redes neurais artificiais.
I. Título.

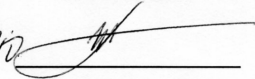

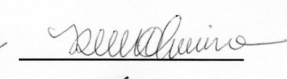


CDD 616.994 49

FOLHA DE APROVAÇÃO – DEFESA DE TESE

ALUNO: FRANCISCO ADELTON ALVES RIBEIRO

TÍTULO DO PROJETO: "Otimização de imagens mamográficas para auxílio ao diagnóstico médico"

PROFESSOR ORIENTADOR: Prof. Dr. Benedito Borges da Silva

BANCA EXAMINADORA:	CONCEITO	ASSINATURA
Prof. Dr. Benedito Borges da Silva - UFPI (Presidente)	<u>SATISFATORIO</u>	
Prof. Dr. Francisco das Chagas Alves Lima - UFPI (Examinador)	<u>Satisfatorio</u>	
Profa. Dra. Rita de Cassia Meneses Oliveira - UFPI (Examinadora)	<u>Satisfatorio</u>	
Prof. Dr. Ailton Mendes Conde Junior – UFPI (Examinador)	<u>SATISFATORIO</u>	
Prof. Dr. Joaquim Evêncio Neto - UFPI (Examinador)	<u>SATISFATORIO</u>	

DATA DA AVALIAÇÃO: 31 de março de 2017.

HORÁRIO: 08:30h

LOCAL: Auditório do Núcleo Integrado de Morfologia e Pesquisas com Células-tronco
(NUPCelt/CCA/UFPI)

Dedico esta conquista a meu Pai, **Francisco Paulo Ribeiro** (*in memoriam*) e minha mãe **Maria Odinéa Alves Ribeiro**, por todo amor, carinho, apoio, ensinamentos e incentivo, minha Doutora da vida.

A minha esposa, Adriana de Oliveira Ribeiro, por sempre me incentivar e apoiar em todas as minhas decisões, pela paciência e compreensão nos momentos que estive ausente. Aos meus filhos, Vanessa de Oliveira Ribeiro e Murilo de Oliveira Ribeiro, pelo carinho, apoio e incentivo dado ao longo de minha jornada.

DEDICATÓRIA ESPECIAL

À meu orientador, prof. **Dr. Benedito Borges da Silva**, mestre na verdadeira acepção da palavra, pela brilhante orientação deste trabalho, pelos ensinamentos de fundamental importância em minha formação ética, profissional e científica e pelo exemplo de dedicação, amizade, competência e liderança.

AGRADECIMENTOS

Agradeço a **Deus**, por permitir que nosso objetivo fosse alcançado.

Ao meus amigos Otilio Paulo da Silva Neto e Elmo de Jesus Nery Junior, pela amizade, palavras de encorajamento, sempre disposto a ajudar no desenvolvimento do software e escutar minhas preocupações, enfim, amigos para todas as horas.

Ao professor Francisco das Chagas Alves Lima, por acreditar em meu projeto e pelas palavras de incentivo durante esta pesquisa.

À coordenação e servidores do curso de Pós-Graduação da RENORBIO – Rede Norte e Nordeste de Biotecnologia da Universidade Federal do Piauí (UFPI), pela oportunidade e apoio recebido durante a realização da pesquisa.

Ao Reitor e vice-Reitora da UFPI pelo incentivo à Pós-Graduação.

Aos Coordenadores e Professores da Renorbio Ponto Focal Piauí pela condução do Curso.

Aos colegas do Curso de Doutorado da Renorbio, Carla Solange, Luana Mota, Camila Revoredo, Fabiane Sampaio, Gilmara Rodrigues, Danylo, Conceição, Cléciton e demais colegas do curso.

Às secretárias do Renorbio, Deusilene e Eliana.

À Edilene, secretária do Mestrado Ciências e Saúde.

Às funcionárias da Ginecologia e Mastologia do HGV, Efigênia, Toinha, Socorro e Eugênia.

Por fim, a todos que de forma direta ou indireta contribuíram à realização deste sonho e por toda ajuda na conclusão desta pesquisa.

O câncer de mama é a neoplasia maligna mais comum em mulheres de países ocidentais após o câncer de pele não melanoma. No mundo, há mais de 1,5 milhão de casos novos diagnosticados anualmente e mais de 500 mil mortes pela doença, sendo que no Brasil, em 2016, foram estimados 57.960 novos casos de câncer de mama e cerca de 14.206 mortes em mulheres pela doença. A mamografia é o exame diagnóstico padrão que permite a descoberta precoce que reduz a mortalidade em cerca de 30%. Todavia, os sistemas computacionais podem melhorar a acurácia da imagem mamográfica na detecção das neoplasias. O presente estudo teve por objetivo desenvolver um software para otimização de imagens mamográficas e auxílio ao diagnóstico médico. Esta tese foi estruturada em quatro capítulos. O capítulo I trata de uma prospecção sobre software para processamento de imagem, buscando indícios do estado atual da técnica via cenário de patentes depositadas relacionadas a softwares de diagnóstico e processamento de imagens mamográficas no Brasil e no mundo. Foram encontrados 107 registros de patentes, sendo 72 resultantes da busca relacionadas ao termo "*Software diagnosis breast cancer*", 32 relacionadas ao termo "*Software mammography*" e 3 ao termo "*Software processing images*", demonstrando a necessidade de desenvolvimento de softwares para processamento de imagens mamográficas. O capítulo II mostrou as técnicas de treinamento e aprendizado de máquina com uso de redes neurais artificiais, para detectar massas em imagens mamográficas, utilizando extração de características de texturas. Os resultados obtidos comprovaram que o algoritmo *Random Forest* obteve o melhor índice de acerto, no que diz respeito à sensibilidade, especificidade e acurácia, atingindo 92.10%, 93.50% e 92.90%, respectivamente. O capítulo III trata de um mapeamento sistemático sobre algoritmos utilizados nas etapas de processamento de imagens mamomográficas. As buscas foram realizadas nas bases de dados *Web of Science*, *Engineering Village* e *Scopus*. Os resultados obtidos foram 199 artigos dos quais, após uma seleção rigorosa realizada por três especialistas em processamento de imagens, apenas 14 artigos foram indicados como estudos relevantes para o mapeamento. Os algoritmos Otsu e o PSO destacaram-se na etapa de segmentação, os quais foram utilizados em 9 dos 14 artigos aceitos, já na etapa de extração das características obteve destaque o algoritmo de *Haralick*, utilizado em 7 das 14 publicações e por fim na etapa de classificação por aprendizagem supervisionada, destacou-se como melhor classificador o algoritmo de *Support Vector Machines*, utilizado em 5 dos 14 artigos aceitos. O capítulo IV trata do Certificado de Registro de Programa de Computador.

Palavras-chave: processamento de imagens, mamografia, câncer de mama, redes neurais artificiais.

ABSTRACT

Breast cancer is the most common malignancy among women in Western countries, following non-melanoma skin cancer. There are currently more than 1.5 million new cases diagnosed worldwide and more than 500 thousand deaths are due to the disease. In Brazil, in 2016, it was estimated that 57 960 new cases of breast cancer occurred in women and about 14 206 deaths were due to disease. Mammography is the standard diagnostic exam which permits early tumor detection and reduces mortality in about 30% of cases. Nevertheless, computer systems may improve the degree of accuracy of mammographic images in tumor detection. The aim of the current study was to develop software to optimize mammographic images and assist in medical diagnosis. This thesis was structured into four chapters. Chapter I was dedicated to the search for software used in image processing, seeking indications of the current status of technique through patent deposits which were related to software used in diagnosis and mammographic image processing in Brazil and worldwide. One-hundred and seven patent registrations were found, 72 of which resulted from the search related to the term "*Software diagnosis breast cancer*", 32 were related to the term "*Software mammography*" and 3 were related to the term "*Software processing images*", demonstrating the need to develop software for image processing in mammography. Chapter II covered training and machine learning techniques with the use of artificial neural networks, for mass detection in mammographic images, using texture feature extraction. The results obtained confirmed that Random Forest algorithm obtained the best index of success, regarding sensitivity, specificity and accuracy, achieving 92.10%, 93.50% and 92.90%, respectively. Chapter III was dedicated to systematic mapping of algorithms used in stages of mammographic image processing. The search was performed in the Web of Science, Engineering Village and Scopus databases, resulting in 199 articles. After a rigorous selection performed by three specialists in image processing, only 14 articles were indicated and regarded as studies relevant to mapping. Otsu and PSO algorithms were highlighted in the segmentation stage, used in 9 of the 14 articles accepted. In contrast, in the feature extraction stage, Haralick algorithm was highlighted and used in 7 of the 14 publications. Finally, in the classification phase of supervised learning, the best classifier was the Support Vector Machines algorithm, used in 5 out of the 14 articles accepted. Chapter IV focused on the Certificate of Registration of Computer Programs.

Key words: images processing. mammographic. breast cancer. artificial neural networks.

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

ACR	<i>American College of Radiology</i>
BI-RADS	<i>Breast Imaging Reporting and Data System</i>
CAD	Diagnóstico auxiliado por computador, do inglês: <i>Computer Aided Diagnosis</i>
CADe	Detecção auxiliada por computador, do inglês: <i>Computer Aided Detector</i>
CADx	Diagnóstico auxiliado por computador, do inglês: <i>Computer Aided Diagnosis</i>
CC	Crânio-caudal
CLAHE	<i>Contrast Limited Adaptive Histogram Equalization</i>
DDSM	<i>Digital Database for Screening Mammography</i>
FDC	Função de Distribuição Cumulativa
IPC	Classificação Internacional de Patentes
MLO	Médio-Lateral-Oblíqua
MM	Morfologia Matemática
PSO	<i>Particle Swarm Optimization</i>
RFP	Redução de Falsos Positivos
ROIs	Regiões de Interesse
SMOTE	<i>Synthetic Minority Oversampling Technique</i>
SVM	Maquina de Vetores de Suporte, do inglês: <i>Support Vector Machines</i>

LISTA DE FIGURAS E QUADROS

2. REVISÃO DA LITERATURA

Figura 1. Ilustração do processo de metástases das células cancerígenas.	28
Figura 2. Ilustração dos tipos de anomalias em uma imagem mamográfica.	29
Figura 3. Estrutura anatômica da mama.	30
Figura 4. Incidência das mamas.	32
Figura 5. Ilustração de anomalias mamárias.	34
Figura 6. Mamógrafo: (a) analógico e (b) digital.	35
Figura 7. Ilustração da redistribuição do histograma de CLAHE.	39
Figura 8. Figura ilustrativa do <i>Graph Clustering</i> com 3 subgrafos, representando cada um os clusters.	44
Figura 12. Formas de elementos estruturantes.	45
Quadro 1: Avaliação BI-RADS das lesões mamárias (JATOI, 2010).	34

CAPÍTULO 1

Figura 1. Número de patentes por país.	62
Figura 2. Quantidade de patentes depositadas por país de origem.	62
Figura 3. Classificação internacional de patentes por CIP.	63

QUADRO

Quadro 1. Quantidade de registros obtidos com os termos, por resumos e título.	61
---	----

CAPÍTULO 2

Figura 1. Resultado das etapas de melhoramento da imagem mamográfica.....	74
Figura 2. Imagem mamográfica binarizada.	75
Figura 3. Resultado dos clusters gerados.....	76
Figure 4. Resultado do algoritmo Graph Clustering.....	78

QUADRO

Quadro 1. Etapas de classificação.....	80
--	----

CAPÍTULO 3

Figura 1. Artigos obtidos por base de busca.....	90
Figura 2. Publicações por ano.....	91
Figura 3. Algoritmos mais utilizados na segmentação de imagens.	91
Figura 4. Algoritmos utilizados na extração de características das imagens.....	92
Figura 5. Classificadores de redes neurais artificiais	93

SUMÁRIO

1. INTRODUÇÃO GERAL	22
2. REVISÃO DA LITERATURA	27
1. Câncer de Mama	28
1.1. Tipos de Lesões mamárias.....	28
2. Mama	29
3. Mamografia	31
4. Rastreamento Mamográfico	32
5. CAD – <i>Computer Aided Diagnosis</i>	35
6. Processamento Digital de Imagens	36
6.1. Aquisição de Imagens.....	37
6.2. Pré-Processamento das Imagens	37
6.2.1. Filtro de Média	38
6.2.2. Contrast-Limited Adaptive Histogram Equalization	38
6.3. Crescimento de Região.....	39
6.4. Segmentação.....	40
6.4.1. Algoritmo de Otsu	40
6.4.2. Particle Swarm Optimization.....	42
6.4.3. Graph Clustering.....	43
6.5. Pós-Processamento.....	45
6.6 Extração de Atributos.....	46
6.7 Classificação	47
7.0 Justificativa	47
REFERÊNCIAS	48
CAPÍTULO 1.....	55
PROSPECÇÃO DE SOFTWARE PARA PROCESSAMENTO DE IMAGEM MAMÁRIA	56
RESUMO.....	57
ABSTRACT	58
I. INTRODUÇÃO	59
II. METODOLOGIA	60
III. RESULTADOS.....	61

IX. CONCLUSÃO.....	63
REFERÊNCIAS	64
CAPÍTULO 2.....	68
DETECÇÃO DE MASSAS EM IMAGENS MAMOGRÁFICAS UTILIZANDO EXTRAÇÃO DE CARACTERÍSTICAS DE TEXTURAS E REDES NEURAS.....	69
RESUMO.....	70
ABSTRACT	71
I. INTRODUÇÃO	72
II. MATERIAIS E MÉTODOS.....	73
III. EXPERIMENTOS.....	79
IV. RESULTADO E DISCUSSÃO	80
V. CONCLUSÃO	81
REFERÊNCIAS	81
CAPÍTULO 3.....	84
MAPEAMENTO SISTEMÁTICO DE ALGORITMOS UTILIZADOS EM PROCESSAMENTO DE IMAGENS MAMOMOGRÁFICAS	85
RESUMO.....	86
ABSTRACT	87
I. INTRODUÇÃO	88
II. MATERIAIS E MÉTODOS.....	89
III. RESULTADOS.....	90
IV. CONCLUSÃO	93
REFERÊNCIAS	94
CAPÍTULO 4.....	96
REGISTRO DE PROGRAMA DE COMPUTADOR.....	97
CONSIDERAÇÕES FINAIS	98
ANEXOS.....	101
ANEXO A - Aprovação pelo Comitê de Ética e Pesquisa.....	102
ANEXO B - Notificação de Aceite - Software Surver for Breast Image Processing	105
ANEXO C - Certificado de Publicação de Artigo.....	108
ANEXO D - Certificado de Apresentação.....	107
ANEXO E - Notificação de Aceite - Masse detection in mammographic images using texture feature extraction and neural networks	108
ANEXO F - Submissão - Systematic Mapping of Algorithms Used in Mammographic Image Processing	109

ANEXO G - Publicação de Concessão do Registro de Software.....	102
APÊNDICES	112
APÊNDICE A - Ficha de entrevista para coleta de dados com as pacientes.....	113
APÊNDICE B - Termo de Consentimento Livre e Esclarecido	117
APÊNDICE C - Diagrama de Contexto - MedMamo.....	120
APÊNDICE D - Diagrama de casos de uso - MedMamo.....	121
APÊNDICE E - Diagrama de Classes do Projeto - MedMamo.....	122

1. INTRODUÇÃO GERAL

O câncer de mama é a doença maligna que mais comumente acomete as mulheres em todo o mundo, tanto em países desenvolvidos quanto em desenvolvimento. Aproximadamente 1,7 milhões de novos casos foram diagnosticados em 2012, o que representa cerca de 12% de todos os novos casos de câncer e 25% de todo tipo de tumor em mulheres, com cerca de 522.000 mortes (DESANTIS et al., 2014). No Brasil, para o ano de 2016, foram estimados 57.960 casos novos de câncer de mama e cerca de 14.206 mortes em mulheres pela doença, com um risco estimado de 56,20 casos novos para cada 100 mil mulheres (INCA, 2016). Entre 2008 e 2012 houve um crescimento da incidência do câncer de mama de 20%, enquanto a mortalidade teve um aumento de 14%. Mais da metade das mortes ocorreram em regiões menos desenvolvidas e estima-se que essas proporções irão aumentar até 2025 (WHO, 2013).

O declínio da mortalidade por câncer de mama nos países mais ricos deve-se principalmente ao acesso das mulheres a mamografia de rastreamento em programas organizados e ao acesso aos tratamentos adjuvantes após a cirurgia, tais como quimioterapia, radiologia, terapia endócrina e biológica (BERRY et al., 2005). Os tumores de mama geralmente são descobertos pela própria mulher, apalpando suas mamas de forma acidental. Nesses casos os tumores já estão em estágio mais avançado, dificultando assim o tratamento. O diagnóstico precoce não apenas influencia o prognóstico, mas propicia cirurgia menos mutilante e com sobrevida comparáveis a intervenções cirúrgicas mais dramáticas e agressivas (BERRY et al., 2005).

As altas taxas de mortalidade pelo câncer de mama despertam interesse na seleção de pacientes de alto risco, visando o estabelecimento de estratégias redutoras de risco e de diagnóstico precoce da doença, pois a alta mortalidade significa uma sobrevida reduzida, principalmente devido ao diagnóstico da neoplasia em estádios avançados (JUSTO et al., 2013).

Existem dificuldades na prevenção do câncer de mama devido à grande variedade de fatores de risco e das características genéticas dos tumores. Novas metodologias para detecção de câncer de mama vêm sendo estudadas, principalmente abordagens que podem ser realizadas em países onde a grande parte da população sobrevive com baixa renda (BORCHARTT, 2013).

O reconhecimento de fatores de risco tem ajudado a realizar o diagnóstico precoce da doença, principalmente pela mamografia, que associado com as

estratégias terapêuticas têm sido imperativo na redução das taxa de mortalidade por câncer de mama (VEISY et al., 2014). Contudo, somente uma fraca correlação tem sido consistentemente demonstrada entre os fatores de risco estabelecidos para o câncer de mama e a mortalidade pelo rastreamento mamográfico (WALKER et al., 2013).

Segundo Berry (2005), a melhor forma de redução da mortalidade pelo câncer de mama é o diagnóstico precoce, que reduz não só a mortalidade como facilita o tratamento. Dentre os métodos de diagnóstico precoce, o mais utilizado é a mamografia, que é um método simples, barato, e disponível em praticamente todas as regiões do Brasil, principalmente em serviços privados e com uma certa escassez em serviços públicos. Todavia, o grande problema da mamografia é sua sensibilidade, cujo resultado falso negativo pode chegar a 15% e em até 30% em mulheres em idade reprodutiva, resultando em biópsias desnecessárias, com traumatismos físicos e emocionais de difícil reparação para a paciente (JUSTO et al., 2013).

Os principais métodos atualmente utilizados na detecção do câncer de mama são o exame físico e a mamografia. Destes, a mamografia é a mais indicado para detecção precoce da doença. Várias campanhas acontecem em todo o Brasil para lembrar a importância da detecção e do tratamento precoces. Muitas mulheres são motivadas desde cedo para que façam o autoexame e mamografia de rastreamento como exames de rotina a partir dos 40 anos (INCA, 2011).

Com o uso do rastreamento mamográfico foi observada uma redução na taxa de mortalidade ocasionada por essa doença (BERRY et al., 2005). A mamografia, portanto, é o método padrão para o rastreamento e diagnóstico precoce do câncer de mama, consistindo de um exame radiológico que gera uma imagem em tons de coloração cinza da mama, que o especialista analisa e identifica de forma visual onde se encontra a lesão ou as lesões (BERRY et al., 2005).

Todavia, a taxa de falso negativo é de 10% a 15%, podendo chegar até 30% em mulheres com mamas densas (INCA, 2011). De uma forma geral, a leitura e interpretação de imagens da mamografia são tarefas que exigem experiência do radiologista, e seu julgamento depende de treinamento, experiência e critérios subjetivos. Mesmo *experts* bem treinados apresentam uma taxa de acertos de classificação com variação de 65 a 75% entre observadores (JUSTO et al., 2013), no entanto a utilização da mamografia como método de rastreamento para reduzir a

mortalidade em cerca de 25%, mesmo com uma taxa de falso negativo supracitada (INCA, 2008).

Estima-se que o período entre o surgimento do tumor e seu crescimento até ele se tornar palpável, alcançando cerca de 1 cm de diâmetro, é de aproximadamente 8 anos, onde a partir daí ele duplica de tamanho a cada período de 3 a 4 meses, conforme descrito por Abreu e Koifman (2002). Se não for tratado, o tumor gera metástases que são focos de tumor em outros órgãos, mais comumente nos ossos, pulmões e fígado (MEDEIROS et al., 2014). Após três ou quatro anos do descobrimento do tumor pela palpação, o não tratamento do câncer pode ocasionar o óbito do paciente (INCA, 2008). A avaliação correta do tamanho da lesão tem um papel importante no planejamento do tratamento do câncer de mama, podendo evitar cirurgias mutilantes, tais como a mastectomia (BANDEIRA et al., 2011). O tamanho da lesão além de permitir a realização de biópsia está associado ao prognóstico da doença (MATHENGE et al., 2014), podendo permitir a realização de uma cirurgia conservadora sem mutilação física e emocional da mutilante cirurgia radical.

A presença de falso positivos e negativos na mamografia, com repercussões físicas e emocionais supracitadas para as paciente, tem motivado a busca de recursos de imagens complementares à mamografia visando sua otimização. A propósito, técnicas computacionais motivaram o surgimento de diversas pesquisas ao longo das últimas décadas, no sentido de desenvolver sistemas computacionais para auxiliar especialistas a desempenhar seu papel de interpretar as imagens radiológicas, que são conhecidas como sistemas CADe (*Computer Aided Detection*) e CADx (*Computer Aided Diagnosis*), e presentes em diversos centros de diagnóstico por imagem, aumentando as taxas de acerto na identificação precoce de doenças graves, como o câncer de mama (TANG et al., 2009).

A utilização de técnicas de inteligência artificial, como Redes Neurais Artificiais (RNA), mostra grande potencial em aplicações de leitura automática de mamografias. A associação dessas técnicas com reconhecimento de padrões, classificação e algoritmos de detecção pode ser aplicada no desenvolvimento de sistemas CADs (NISHIKAWA, 2010).

Diante do exposto, este estudo teve por objetivo desenvolver um software para otimização de imagens mamográficas e auxílio ao diagnóstico médico, utilizando aprendizado de máquina para classificação de nódulos mamários.

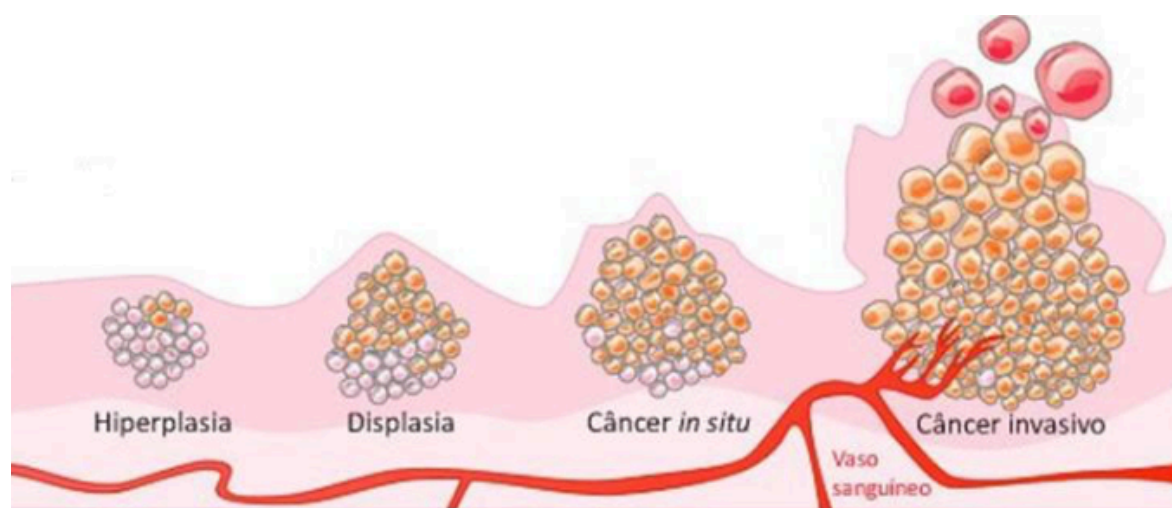
O conteúdo descrito neste estudo está estruturado em forma de capítulos, conforme segue: O capítulo I trata de uma prospecção sobre os softwares para processamento de imagem, buscando os registros de patentes “Registro de Programa de Computador” relacionados a softwares de diagnóstico e processamento de imagens mamográficas. Deste capítulo foi gerado um artigo aceito e publicado intitulado “*Software Survey for Breast Image Processing*”. O capítulo II descreve os algoritmos utilizados para detectar massas em imagens mamográficas, utilizando extração de características de texturas, no qual foi possível identificar o melhor índice de acerto no que diz respeito à sensibilidade, especificidade e acurácia, para classificar nódulos mamários. O artigo deste capítulo, já aceito, é intitulado “*Masses detection in mammographic images using texture feature extraction and neural networks*”. O capítulo III descreve um mapeamento sistemático sobre algoritmos utilizados nas etapas de processamento de imagens mamomográficas, com o artigo submetido intitulado “*Systematic Mapping of Algorithms Used in Mammographic Image Processing*”. O capítulo IV trata do Certificado de Registro de Programa de Computador.

2. REVISÃO DA LITERATURA

1. Câncer de Mama

Câncer é uma palavra isolada e usada para descrever cerca de 200 doenças diferentes que afetam os órgãos ou sistemas pelo corpo (SPEECHLEY, 2000). O câncer de mama tem início quando há um erro no processo de divisão celular, criando-se múltiplas células alteradas de forma desordenada (Figura 1). É possível que a atividade funcional comprometida do gene de supressão tumoral p53 seja um dos mecanismos responsáveis pelo desequilíbrio entre a proliferação celular e a morte celular programada ou apoptose. Estas células alteradas multiplicam-se desordenadamente e circulam pelos tecidos adjacentes, gerando aglomerados de células neoplásicas, que obstruem vasos linfáticos e veias, podendo se disseminar pela corrente sanguínea comprometendo órgãos à distância, num processo conhecido como metástases (KUTANZI et al., 2010).

Figura 1. Ilustração do processo de metástases das células cancerígenas.



Fonte: (HAGEN, 2010)

1.1. Tipos de Lesões mamárias

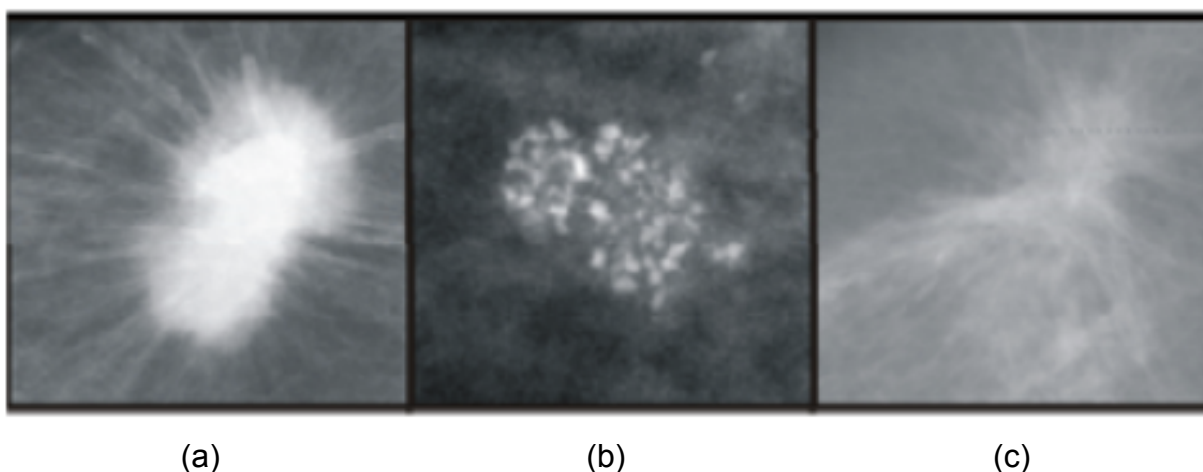
Os tipos de lesões detectadas em uma imagem mamográfica, podem ser observados na Figura 2, e os achados referente à imagem, podem ser descritos como:

- Nódulo: Qualquer opacidade com algum contorno arredondado ou não e caracterizado segundo os cortonos e a densidade (Figura 2a);

- Microcalcificações: pequenos depósitos de cálcio classificados de acordo com sua morfologia e distribuição (Figura 2b);
- Distorção de arquitetura: espiculações em uma região da mama ou uma retração focal do contorno parenquimatoso denso (Figura 2c).

Das lesões supracitadas, para efeito deste estudo, serão levados em consideração apenas os nódulos como escopo do trabalho e não serão levadas em consideração as distorções e microcalcificações.

Figura 2. Ilustração dos tipos de anomalias em uma imagem mamográfica: (a) Nódulo espiculado; (b) Microcalcificações; (c) Distorção de arquitetura.



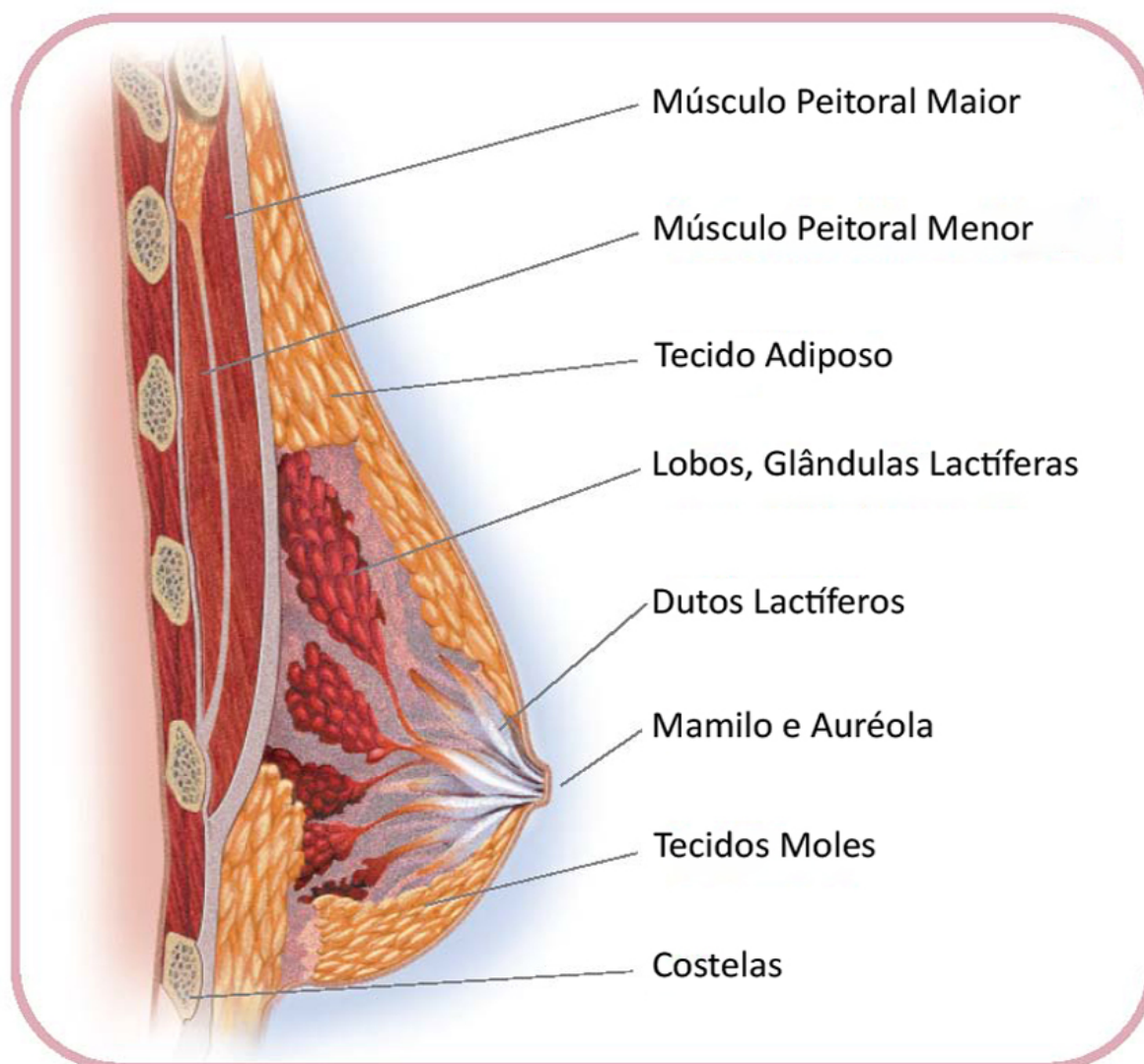
Fonte: (HEATH et al., 2001)

2. Mama

A anatomia da mama, figura 3, é descrita da seguinte forma: mamas são glândulas exócrinas, que fazem parte do sistema reprodutor, lembrando que a mama esta presente nas mulheres e nos homens, porém se desenvolve na mulher. As mamas são estruturas complexas constituídas por tecido glandular onde é produzido o leite, rodeado por tecido gorduroso e por tecido de sustentação (SPENCE, 1991).

De acordo com Ribas e Melo (2006), a mama está localizada entre as camadas superficial e profunda da tela subcutânea, superiormente na parede torácica, na qual está ventralmente situada aos músculos peitoral maior, serrátil anterior e oblíquo externo do abdome. As mamas se estendem verticalmente da segunda até a sexta costela, e transversalmente do esterno à linha axilar média, figura 3.

Figura 3. Estrutura anatômica da mama.



Fonte: (NETTER, 2007)

A mama da mulher sofre alterações na sua constituição à medida que esta vai envelhecendo. Quando adolescente, a mama é constituída majoritariamente por tecido fibroglandular e muito pouca percentagem de gordura. Com o amadurecimento da adolescente, ocorre uma substituição lenta e gradual do tecido fibroglandular por tecido gorduroso, até que, na pós-menopausa, a mama é maioritariamente constituída por tecido gorduroso (NETTER, 2007).

Para a análise e estudo da mama são utilizados alguns métodos de diagnósticos por imagem como, a ultrassonografia, a mamografia e a ressonância magnética. Contudo, neste estudo será dada ênfase à mamografia.

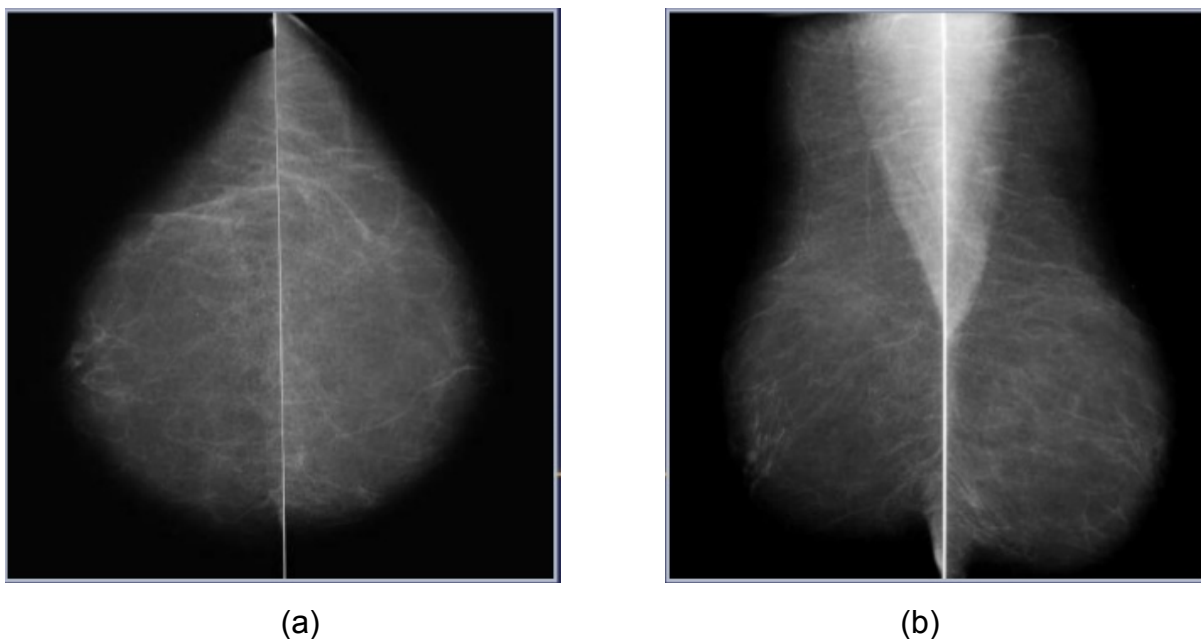
3. Mamografia

Nas últimas décadas, a tecnologia de imagem tem evoluído muito. Atualmente, podemos detectar várias patologias antes da percepção do paciente. A mamografia convencional usa o sistema de écran-filme, associado a um equipamento específico para o exame de mamografia, com gravação da imagem na película através de reações de agentes químicos e é aceita como o método padrão para o rastreamento mamográfico populacional na mulher assintomática sem câncer de mama (PISANO et al., 2005).

A mamografia digital é um método mais recente no qual a imagem é obtida por raios-X em aparelho especialmente desenhado para este fim, com um detector individual para o equipamento, sendo obtida uma imagem digital. O custo do equipamento é cerca de três a quatro vezes maior que equipamento para a mamografia convencional. Pisano et al. (2005) publicaram um estudo com mais de 49.000 mulheres comparando a mamografia digital em relação à mamografia convencional. A acurácia global das técnicas parece ser muito similar como um todo. Todavia, analisando as mulheres mais jovens (< 50 anos) e com mamas mais densas parece haver uma vantagem a favor da mamografia digital. Contudo, a mamografia digital é um método mais caro. Recentemente foi publicado um modelo de custo-efetividade, sob a perspectiva da sociedade americana e do *Medicare*, que mostrou que o uso de mamografia digital custo-efetivo principalmente em mulheres jovens (TOSTESON et al., 2008).

A mamografia constitui uma forma particular de radiografia, que trabalha com níveis de tensões e correntes em intervalos específicos, destinada a registrar imagens da mama a fim de diagnosticar a presença ou ausência de estruturas que possam indicar patologias. Segundo o INCA (2005), em uma mamografia, duas incidências de cada mama são indispensáveis: uma visão crânio-caudal (figura 4a), e uma médio-lateral-oblíqua (figura 4b).

Figura 4. Incidências das mamas: (a) crânio-caudal; (b) médio-lateral.



Fonte: (SANTOS, 2002)

No entanto, a incidência médio-lateral-oblíqua (MLO) é a mais eficaz, pois mostra uma quantidade maior de tecido mamário e inclui estruturas mais profundas do quadrante súpero-externo e do prolongamento axilar (CAMPUS, 2006).

A incidência crânio-caudal (CC) tem como objetivo incluir todo o material pósteromedial, complementando a médio-lateral-oblíqua. Permite também mais compressão da mama, uma vez que não inclui a axila, resultando em uma definição superior da arquitetura mamária e de lesões. Os radiologistas estudam as incidências crânio-caudais e as médio-laterais aos pares de modo a permitir a comparação de regiões simétricas, pois qualquer assimetria pode ser indício de patologia (CAMPUS, 2006).

4. Rastreamento Mamográfico

O câncer de mama é um problema de saúde pública. Neste sentido, estratégias de rastreamento mamográfico são aceitas como procedimento de escolha na maioria dos países. Desta forma, novas abordagens que agreguem maior acurácia ao diagnóstico dentro de um impacto econômico e social aceitável devem ser investigadas.

No ano de 2012, foi estimado cerca de 1,7 milhões de novos casos e 522 mil mortes pela doença, representando 25% de todos os casos de câncer e 15% de

todas as mortes por mulheres. No Brasil para 2016 foi estimado 57.960 casos novos e 14.388 óbitos pela doença (INCA, 2016). Nesta perspectiva, segundo o Ministério da Saúde, no Brasil o rastreamento mamográfico é oficialmente recomendado em mulheres entre 50 e 69 anos no máximo de dois em dois anos e em pacientes de alto risco anualmente a partir dos 40 anos (INCA, 2016).

O objetivo principal de qualquer programa de rastreamento mamográfico populacional de câncer é muito claro e direto – detectar casos de câncer em estágios iniciais quando são tratáveis e com chances de cura. Todavia, alcançar esse objetivo no sistema de saúde é bem complexo. Para um teste de triagem ser útil, ele deve detectar o caso de neoplasia no estágio inicial e deve melhorar o desfecho desses pacientes tratados precocemente (CAMPUS, 2006).

Com o objetivo de melhorar a eficácia dos programas de rastreamento mamário, implementar a qualidade dos laudos emitidos pelos médicos radiologistas e reconhecendo a necessidade de fornecer relatórios claros que permitissem a aquisição de dados confiáveis, a *American College of Radiology (ACR)* com a colaboração dos comitês *American Medical Association, American College of Surgeons, Centers for Disease Control and Prevention, College of American Pathologists, College of Surgeons, Food and Drug Administration e National Cancer Institute*, desenvolveram, no ano de 1993, o sistema BI-RADS (*Breast Imaging Reporting and Data System*) para uniformização e a padronização da linguagem usada nas imagens de mama e a avaliação das categorias para os achados mamários (mamografia, ultrassonografia e ressonância magnética). Esse sistema é amplamente utilizado no mundo todo como uma linguagem padrão entre mastologistas, radiologistas e ginecologistas para evitar ambiguidades ou distorções de interpretação de laudos mamográficos (CAMPUS, 2006).

O BI-RADS foi concebido inicialmente como uma ferramenta mamográfica, mas em 2003 sofreu uma atualização e foi disponibilizado também para a ultrassonografia e ressonância magnética. Sua versão mais atual (5ª edição), foi lançada em 2013. Esse sistema utiliza categorias de 0 a 6 para descrever os achados do exame e prevê recomendações de condutas. A Tabela 1 sintetiza os resultados do exame mamográfico e as principais condutas, consoante categorias do sistema BI-RADS (ELSHINAWY et al., 2011).

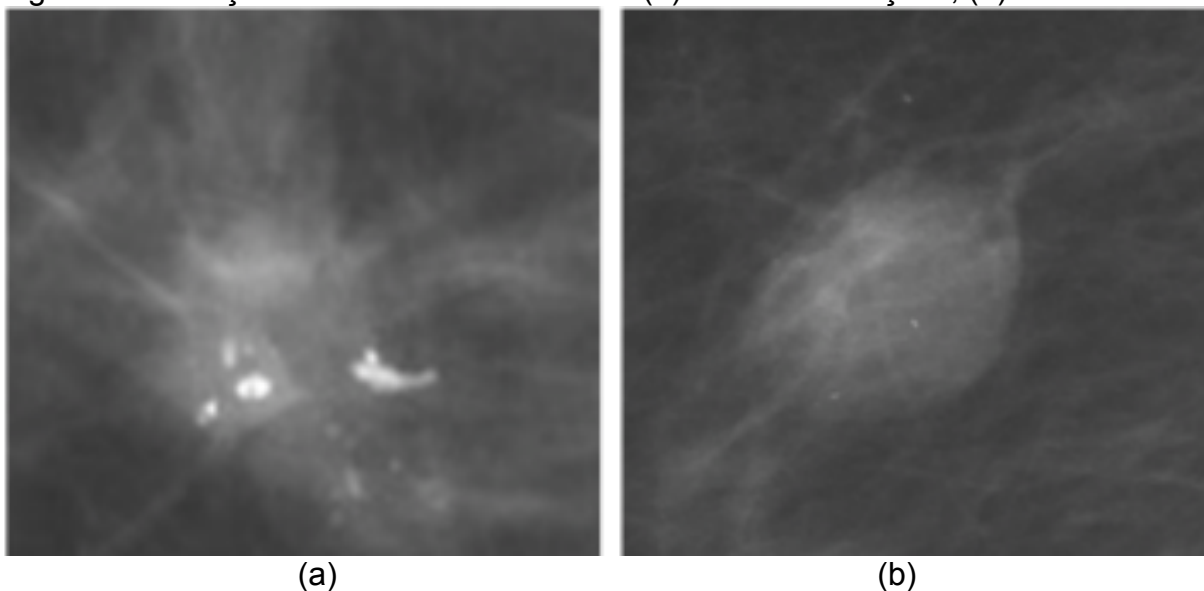
Quadro 1: Avaliação BI-RADS das lesões mamárias (ELSHINAWY et al., 2011).

Categoria	Avaliação	VPP²	Recomendação de condutas
0	incompleta		Exames adicionais
1	Discreta	0%	Exames de rotinas
2	Achado benigno	0%	Exames de rotinas
3	Achado provavelmente benigno	<2%	Acompanhamento a curto prazo
4 (A,B,C)¹	Suspeita de anormalidade	>2% e <90%	Biópsia
5	Altamente sugestivo a malignidade	>95%	Achado tem características típicas de malignidade.
6	Câncer com confirmação histológica	100%	Iniciar procedimento terapêuticos

¹(A) Menor, (B) Médio e (C) Maior ²VPP: Valor Preditivo Positivo.

O objetivo da mamografia é gerar uma imagem de alta resolução em tons de cinza das estruturas internas da mama, com o intuito de detectar anomalias. As principais anomalias visualizadas a partir de uma mamografia, em geral, são calcificações e massas, contudo as microcalcificações são as lesões mamárias não palpáveis mais frequentes (ELSHINAWY et al., 2011), como ilustra a Figura 5.

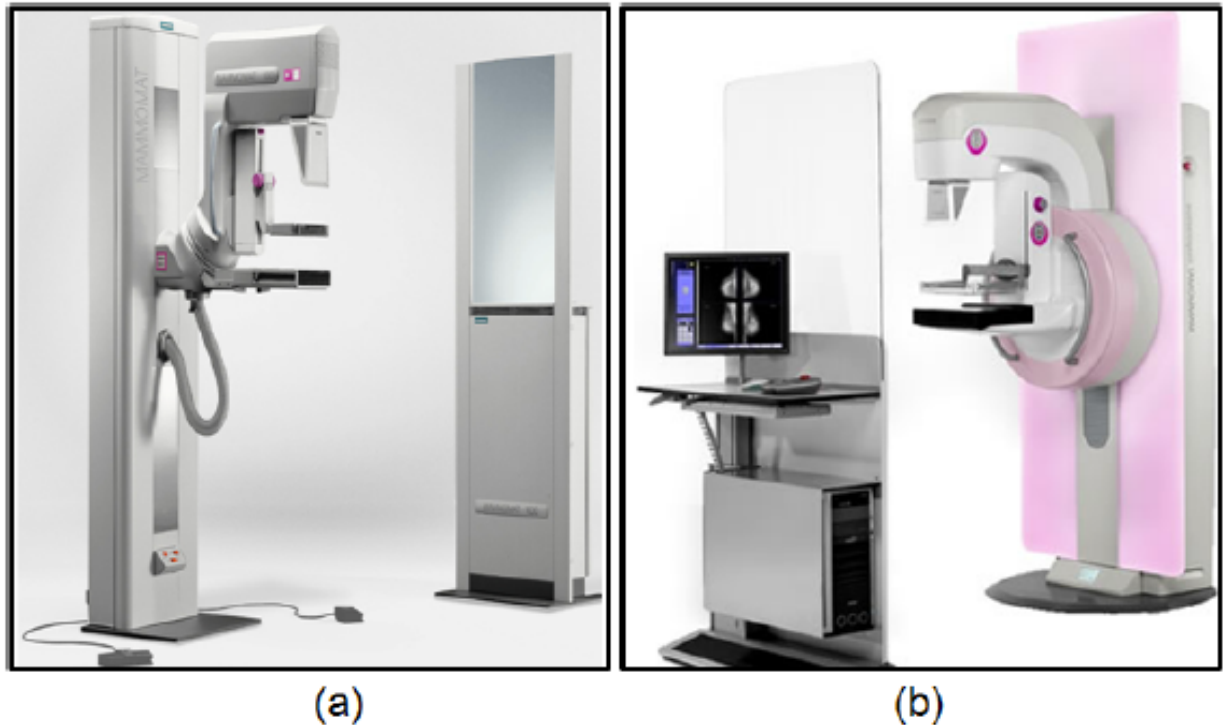
Figura 5. Ilustração de anomalias mamárias: (a) microcalcificações; (b) massas.



Fonte: (BCDR, 2012)

O exame da mama é realizado pelo mamógrafo, aparelho de raio-X que pode ser convencional (analógico) ou digital (Figura 6).

Figura 6. Mamógrafo: (a) analógico e (b) digital.



Fonte: <http://www.siemens.com>

A formação da imagem de mamografia, em ambos equipamentos (analógico e digital), depende dos diferentes graus de densidade dos tecidos mamários e, conseqüentemente, da capacidade do tecido de absorver a radiação. Assim, quanto menor a densidade mamária, menor será a exposição de raios-X, pois penetram facilmente em tecidos adiposos fazendo com que estes apareçam escuros (radiolúcido). Já nos tecidos conjuntivo e glandular que são radiologicamente densos (radiodenso), os raios-X não penetram bem, resultando em uma imagem de baixo contraste com áreas predominantemente brancas (radiodenso) (BOYD et al., 1998); (ELSHINAWY et al., 2011).

5. CAD – Computer Aided Diagnosis

Devido os raios-X não penetrarem bem nas mamas densas a mamografia apresenta algumas limitações, dificultando a interpretação dos radiologistas. A

propósito, de maneira a solucionar estes problemas e para que não seja necessária uma segunda leitura que eleva os custos, surgiu então os sistemas CADs (TANG et al., 2009).

Técnicas de análise auxiliadas por computador têm sido propostas com o objetivo de oferecer melhores parâmetros para a elaboração de um diagnóstico mais preciso, indicando áreas suspeitas, bem como anomalias. Essas técnicas têm sido desenvolvidas por diversos grupos de pesquisas, sendo integradas em sistemas CADs visando auxiliar na detecção precoce do câncer da mama (NISHIKAWA, 2010).

Os sistemas CADs foram desenvolvidos com o intuito de minimizar os casos de falsos negativos, bem como os casos falsos positivos. Tais sistemas têm proporcionado um importante auxílio ao diagnóstico médico em diversas aplicações radiológicas, sobretudo na mamografia. Estes sistemas são divididos em dois tipos diferentes: Detecção Auxiliada por Computador (*Computer Aided Detection - CADe*) e Diagnóstico Auxiliado por Computador (*CADx*). O CADe identifica regiões suspeitas numa imagem e ajuda os radiologistas a encontrar o câncer na mamografia de rastreamento, fazem uso de técnicas de visão computacional, já o Diagnóstico Auxiliado por Computador (*Computer Aided Diagnosis - CADx*) classifica as regiões suspeitas, ou seja, permite que os radiologistas decidam se a lesão é benigna ou maligna, através de técnicas de Inteligência Artificial (NISHIKAWA, 2010).

6. Processamento Digital de Imagens

O processamento digital de imagens pode ser definido com um conjunto de técnicas para capturar, representar e transformar imagens com o auxílio de processos computacionais. A utilização destas técnicas permite identificar e extrair informações das imagens, alterando a qualidade visual de determinados aspectos estruturais e facilitando a percepção humana e a interpretação automatizada por meio de máquinas (PEDRINI; SCHWARTZ, 2008).

De acordo com (GOMES; VELHO, 2002), um sistema de processamento digital de imagens admite um conjunto de imagens de entrada e retorna um conjunto de imagens de saída de forma adequada às necessidades do problema abordado.

Ao longo das duas últimas décadas, a área de processamento de imagens digitais apresentou um rápido crescimento. A evolução da tecnologia de computação

digital, bem como o desenvolvimento de novos algoritmos para lidar com sinais bidimensionais estão permitindo uma gama de aplicações cada vez maior. Na medicina, por exemplo, o processamento de imagens digitais permite a coloração ou a intensificação de contraste para facilitar a interpretação de radiografias e outras imagens biomédicas (GONZALES; WOODS; 2010).

Através desse processo consiste um conjunto de técnicas para capturar, representar e transformar imagens com auxílio do computador. De acordo com (GONZALES; WOODS; 2010), o processamento das imagens é dividido nas etapas de aquisição das imagens, pré-processamento, segmentação, pós-processamento, extração dos atributos e classificação.

6.1. Aquisição de Imagens

A aquisição de imagens é a primeira etapa no processamento digital de imagens. Nesta etapa a imagem pode ser obtida diretamente, como em um aparelho de raio-X digital, ou através de um digitalizador. O digitalizador é um aparelho que tem a função de ler sinais analógicos e convertê-los em sinais digitais. Por exemplo, a mamografia é impressa em uma folha de filme e é digitalizado por scanners especializados (GONZALES; WOODS; 2010).

6.2. Pré-Processamento das Imagens

A segunda etapa é o pré-processamento das imagens adquiridas. Esta etapa tem o objetivo de aumentar a qualidade da imagem, proporcionando uma melhora através da redução de ruídos, melhoramento do contraste e recorte do objeto de interesse da imagem de fundo, dentre outros. Esta etapa na maioria dos casos é indispensável à eficiência das etapas posteriores.

Para melhorar a visualização e interpretação das imagens por parte do observador existem inúmeras técnicas de realce, sua escolha e utilização está diretamente relacionada ao contexto e a preferência humana no que se refere a uma “imagem boa” (GONZALEZ; WOODS, 2010).

6.2.1. Filtro de Média

O filtro da média é uma técnica muito usada no processamento de imagens para reduzir o ruído, melhorando sua qualidade (GONZALES; WOODS, 2010).

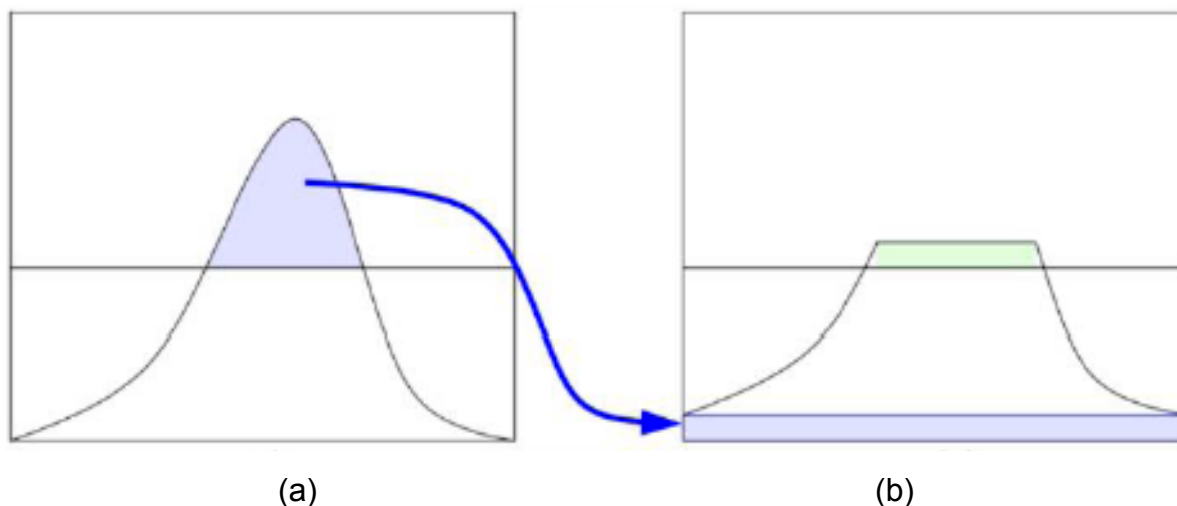
Dada uma imagem I e uma janela J de tamanho $N \times N$, centralizada no pixel p_{ij} , o resultado do filtro da média sobre p_{ij} é a média aritmética dos valores dos pixels contidos em J . Quanto maior for o valor de N , mais influência o p_{ij} transformado sofrerá, e isto pode resultar no efeito de borramento da imagem.

6.2.2. Contrast-Limited Adaptive Histogram Equalization

O realce local é uma maneira interessante de contrastar os pixels de uma imagem. O *Contrast Limited Adaptive Histogram Equalization* (CLAHE) é uma técnica de realce local de contraste que altera o tom de cinza de um pixel através da análise de sua vizinhança. O CLAHE evita o aumento de contraste em ruídos, baseando-se numa equalização de histograma adaptativa. Em resumo, cada pixel é transformado com base no histograma de um quadrado ao seu redor, onde a função de transformação é obtida através da Função de Distribuição Cumulativa (FDC), encontrada em Pedrin e Schwartz (2008) de valores de pixels na vizinhança (GONZALES; WOODS; 2010).

O CLAHE limita a amplificação por recorte do histograma em um valor pré-definido antes de calcular o FDC. Isto limita a inclinação do mesmo e, por conseguinte, a função de transformação. Este valor de corte do histograma depende do tamanho da vizinhança. Os valores de corte estão entre 3 e 4 vezes o valor médio do histograma (ZUIDERVELD, 1994). É mais interessante não descartar a parte do histograma que ultrapassa o limite de corte, então pega-se a parte que excedeu e distribui-se igualmente em todas as faixas do histograma, como mostra a Figura 7.

Figura 7. Ilustração da redistribuição do histograma de CLAHE: (a) Ponto de corte, (b) Redistribuição dos valores de corte do histograma.



Fonte: (ZUIDERVELD, 1994)

O CLAHE, possui diversas funções de transformação dos tons de cinza (ZUIDERVELD, 1994). Para este trabalho foi utilizado a função uniforme, de acordo com a Equação 1.

$$g = [g_{max} - g_{min}]P(f) + g_{min}, \quad (1)$$

onde g é o novo valor de cinza do pixel. Os valores g_{min} e g_{max} são as variáveis, menor e maior valor de cinza na vizinhança, respectivamente; e $P(f)$ é a função FDC.

Esta técnica de realce local (CLAHE), foi utilizada para realçar as estruturas internas das mamas nas mamografias digitais.

6.3. Crescimento de Região

O crescimento de região é baseado na propagação de um ponto inicial chamado semente de acordo com um critério específico de homogeneidade, iterativamente aumentando o tamanho da região.

Segundo Oliver et al. (2010) algoritmos de crescimento de região têm sido amplamente utilizados na segmentação de massas mamográficas, com o objetivo de extrair a lesão de seu fundo (HUO et al. 1995; KUPINSKI; GIGER, 1998; KINNARD et al. 2002; MARTÍ et al., 2003).

Para este trabalho, o algoritmo de crescimento de região foi utilizado na etapa de segmentação, para gerar as regiões de Interesse do Objeto (ROI) das imagem, de modo a isolar as ROIs candidatas.

6.4. Segmentação

Nesta etapa o objetivo é extrair da imagem apenas partes dela que realmente interessam para o processamento. A etapa de segmentação é extremamente importante, pois por meio dela podemos isolar os objetos de interesse, permitindo assim trabalharmos com estes elementos individualmente, proporcionando resultados mais satisfatórios na determinação destes objetos, dando prosseguimento à etapa de representação na análise da imagem.

Também pode-se afirmar que a segmentação é um processo de particionamento de imagens em regiões desconexas de maneira a buscar detectar descontinuidade ou similaridade na imagem. Podemos dizer assim, mais homogêneas possíveis entre os elementos contidos nesta e mais heterogênea possível entre as demais regiões da imagem (PEDRIN; SCHWARTZ, 2008).

Devido à grande quantidade de estruturas diferentes (músculo peitoral, rótulos identificadores, vasos sanguíneos, mamilos, nódulos, micro calcificações) que uma imagem de mamografia pode ter, a segmentação torna-se um processo indispensável para o sucesso do processamento digital de imagens, pois é através da segmentação que podemos encontrar os objetos de interesse.

6.4.1. Algoritmo de Otsu

O método de Otsu (OTSU, 1979) assume que a imagem está dividida em duas classes, objeto e fundo, e tenta encontrar o limiar que melhor divide essas classes por meio de múltiplas iterações. Para isso, calcula-se a variância entre as classes de objeto e fundo para todos os possíveis limiares. O limiar que possui maior variância inter-classe é o limiar escolhido pelo método. Segundo o artigo de Otsu, maximizar a variância inter-classe é o mesmo que minimizar a variância intra-classe, definida como a soma ponderada entre as variâncias das duas classes.

Partindo de uma imagem em tons de cinza com N pixels e L possíveis níveis de cinza, a probabilidade de ocorrência dos níveis de cinza i na imagem é mostrada na Equação 2.

$$p_i = \frac{f_i}{N} \quad (2)$$

onde, f_i representa a frequência de repetição dos níveis de cinza de i , com $i = 1, 2, \dots, L$.

Os pixels são divididos em duas classes C_1 e C_2 , com níveis de cinza $[1, 2, \dots, t]$ e $[t + 1, t + 2, \dots, L]$ respectivamente, onde as distribuições de probabilidade são mostrados nas Equações 3 e 4.

$$C_1 = \frac{P_1}{\omega_1(t)}, \dots, \frac{P_t}{\omega_1(t)} \quad (3)$$

$$C_2 = \frac{P_{t+1}}{\omega_2(t)}, \frac{P_{t+2}}{\omega_2(t)}, \dots, \frac{P_L}{\omega_2(t)} \quad (4)$$

onde as intensidades para cada uma das classes $\omega_1(t)$ e $\omega_2(t)$ são definidas como:

$$\omega_1(t) = \sum_{i=1}^t P_i \quad (5)$$

$$\omega_2(t) = \sum_{i=t+1}^L P_i \quad (6)$$

As medidas para cada uma das classes $\mu_1(t)$ e $\mu_2(t)$ são definidas com as Equações 7 e 8 a seguir:

$$\mu_1(t) = \sum_{i=1}^t \frac{i \cdot P_i}{\omega_1(t)} \quad (7)$$

$$\mu_2(t) = \sum_{i=t+1}^L \frac{i \cdot P_i}{\omega_2(t)} \quad (8)$$

A intensidade média total da imagem μ_T é definida com as Equações 9 e 10.

$$\omega_1 \cdot \mu_1 + \omega_2 \cdot \mu_2 = \mu_T \quad (9)$$

$$\omega_1 + \omega_2 = 1 \quad (10)$$

Por fim, o algoritmo de Otsu definirá a variância entre as classes por meio da Equação 11 e encontrará o número t , que representa a maior variância, onde $1 \leq t \leq L$.

$$\delta_B^2 = \omega_1.(\mu_1 - \mu T)^2 + \omega_2.(\mu_2 - \mu T)^2 \quad (11)$$

$$T^* = \text{Max} \{ \delta_B^2(t) \} \quad (12)$$

onde, t^* é o limiar ótimo encontrado pela Equação 12 que será utilizado para dar início aos centroides da primeira partícula do enxame do Particle Swarm Optimization (PSO).

6.4.2. Particle Swarm Optimization

A técnica *Particle Swarm Optimization* (PSO), é uma abordagem estocástica, baseada em população que simula o processo comportamental de interação entre os indivíduos de um grupo (MEDEIROS, 2005).

Sua teoria é fundamentada pela observação do comportamento de grupos de animais como pássaros, peixes, entre outros, que realizam interessantes tarefas de otimização na execução de atividades simples, como a busca por alimentos (KENNEDY; EBERHART, 1995). Com isso o PSO busca a simulação destas atividades naturalmente simples, porém matematicamente complexas.

Segundo Medeiros (2005), o PSO mantém uma população de partículas onde cada partícula representa uma solução potencial para o problema e está associada a uma posição em um espaço de busca multidimensional, mantendo as seguintes informações:

- x_i - Posição atual da partícula.
- v_i - Velocidade atual da partícula.
- y - Melhor posição local da partícula.
- \hat{y} - Melhor posição global da partícula.
- w - Valor de inércia da partícula.

Para cada iteração a velocidade da partícula é alterada conforme a Equação 13.

$$v_{i,k}(t+1) = wv_{ik}(t) + c_1r_{1,k}(t)(y_{i,k}(t) - x_{i,k}(t)) + c_2r_{2,k}(t)(\hat{y}_k(t) - x_{i,k}(t)) \quad (13)$$

onde $v_{i,k}$ representa a k -ésima dimensão do vetor velocidade da i -ésima partícula. Cada velocidade é atualizada separadamente para cada dimensão $k \in i \dots n$. Os valores r_1 e r_2 são gerados aleatoriamente entre 0 e 1 (KENNEDY, 2010).

As constantes c_1 e c_2 regulam a aceleração, onde $0 < c_1, c_2 \leq 1$, c_1 regula a direção da melhor posição local e c_2 regula a direção da posição global (gbest) ou da vizinhança (lbest) (KENNEDY, 2010). Os termos $c_1 r_{1,k}(t) y_{i,k}(t) - x_{i,k}(t)$ e $c_2 r_{2,k}(t) (\hat{y}_k(t) - x_{i,k}(t))$ representam as experiências passadas da partícula, sendo o primeiro associado à cognição e o segundo ao social, pois cada partícula leva em consideração a melhor solução ao seu redor (KENNEDY, 2010).

A partícula terá sua posição atualizada usando o novo vetor velocidade calculado na Equação 13 e usado na Equação 14.

$$x_i(t + 1) = x_i(t) + v_i(t + 1) \quad (14)$$

Para o PSO cada indivíduo pode ser uma possível solução para o problema investigado, sendo atribuído a cada indivíduo um valor que está relacionado a adequação da partícula com a solução do problema, denominada de *fitness*, e também uma variável velocidade que representa a direção do movimento do indivíduo (partícula). Com o passar do tempo os indivíduos vão ajustando suas velocidades em relação a melhor solução (melhor *fitness*) encontrada por ele próprio e também pela melhor solução do grupo, realizando este processo até que os indivíduos (partículas) encontrem o alimento (melhor solução global ou próxima dela).

O valor de *fitness* é calculado com base na função objetivo que se deseja maximizar (minimizar). Segundo (LOPES, 2006) o valor de *fitness* é definido pela natureza do problema de otimização e é computada por uma função objetivo que avalia um vetor solução.

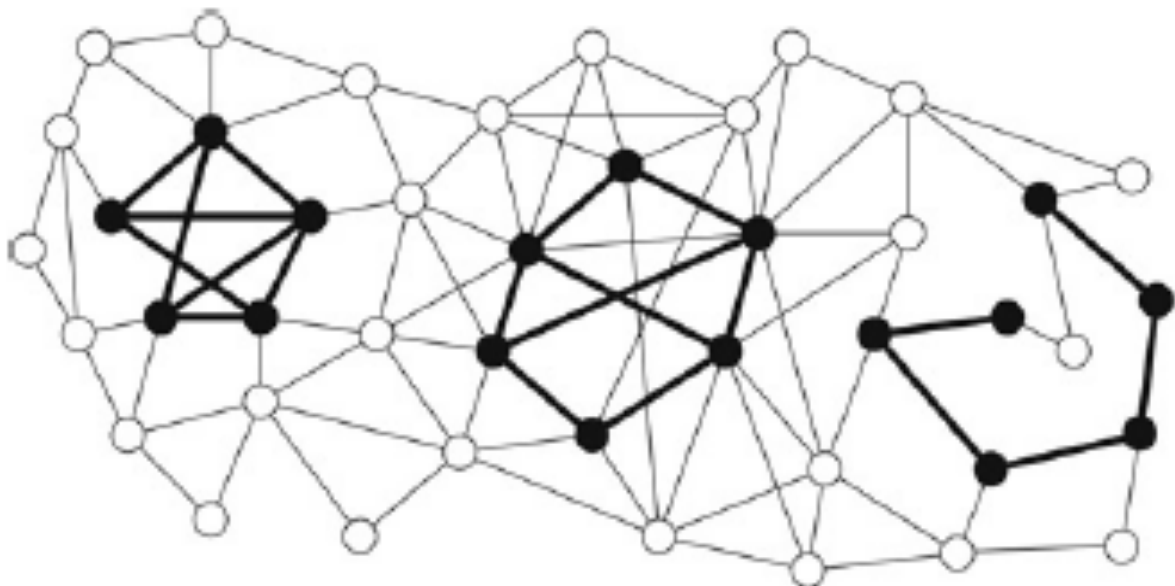
6.4.3. Graph Clustering

Algoritmos de agrupamento podem ser divididos em dois grupos principais: algoritmos hierárquicos e particionados. Estes algoritmos consistem em dividir o conjunto de dados em $k \leq n$ grupos, onde n é o número de elementos no conjunto de dados. Soluções são obtidas através de objetos em movimento entre os clusters até que um critério de parada seja satisfeita. Cada solução é avaliada por uma dada função objetivo (SCHAEFFER, 2007).

O algoritmo *Graph Clustering* é um grafo de agrupamento de vértices em grupos, com base na estrutura da extremidade do grafo. A partição do vértice

resultante deve ter a propriedade de que dentro de cada cluster os vértices são altamente conectados, sendo que só existem poucas arestas entre clusters. Esta abordagem adota explicitamente os conceitos da teoria dos grafos e pode fornecer as definições necessárias e o formalismo matemático, resultando em um importante suporte para a análise de modelos de grafos (SCHAEFFER, 2007). A Figura 8 mostra este modelo.

Figura 8. Figura ilustrativa do *Graph Clustering* com 3 subgrafos, representando cada um os clusters.



Fonte: (SCHAEFFER, 2007)

No particionamento do grafo, os problemas são definidos como sendo o estudo na dificuldade de encontrar bons agrupamentos, e nas configurações de ambos os grafos simples e ponderados respectivamente. Restringindo a atenção primeiro para a bi-partição de um gráfico, denotamos (S, S^c) uma partição do conjunto de vértices V de um grafo (V, E) e deixe $\Sigma(S, S^c)$ sendo a soma total dos pesos de arestas entre S , e S^c , que é o custo associado com (S, S^c) . O problema máximo do corte é encontrar uma partição (S, S^c) que maximiza $\Sigma(S, S^c)$, e o problema do corte do quociente mínimo é encontrar uma partição que minimiza $\min \Sigma(S, S^c) / \min(\|S\|, \|S^c\|)$ Dongen et al., (2001).

Em problemas de partição de grafos padrão, os tamanhos dos elementos da partição são prescritos, para se minimizar o peso total das arestas que ligam os nós em subconjuntos distintos na partição. Assim tamanhos $m_i > 0, i = 1, \dots, k$,

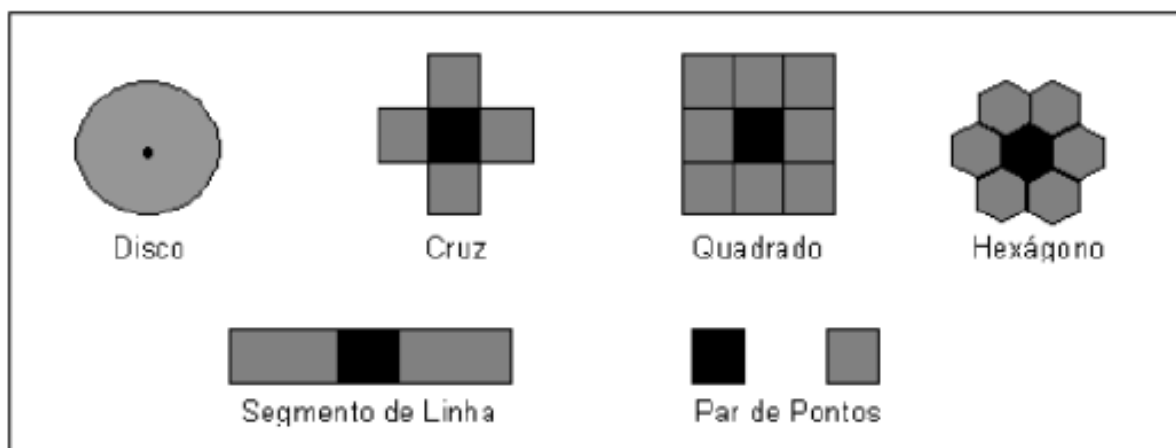
satisfazendo $k \leq n$ e $\sum m_i = n$, forem especificados, em que n é o tamanho de V , e uma partição de V em subconjuntos S_i de tamanho m_i é chamado de função que minimiza o custo. O papel de uma função de custo/rendimento é principalmente útil para a comparação de diferentes estratégias em relação a referências comuns Dongen et al., (2001).

6.5. Pós-Processamento

O intuito da etapa de pós-processamento de imagens consiste na correção de possíveis erros e na eliminação de ruídos obtidos durante a etapa de segmentação. A solução para problemas como esses seria a utilização de técnicas de Morfologia Matemática (MM), as quais podem ser aplicadas em diversas tarefas de processamento de imagens tais como restauração, realce, filtragem, segmentação, extração de medidas, entre outras (SCHAEFFER, 2007).

A MM trata do estudo das estruturas geométricas presentes em uma imagem (SCHAEFFER, 2007), tendo como objetivo a extração de informações sobre a geometria de uma imagem desconhecida, através da transformação de outra imagem bem definida por meio de um elemento estruturante como seu tamanho e sua forma (SCHAEFFER, 2007). Na Figura 9 é possível visualizar as formas de alguns elementos estruturantes. Na forma do elemento cruz, por exemplo, são considerados como vizinhos do pixel central somente os 4 pixels adjacentes a ele, lateral e verticalmente (SCHAEFFER, 2007).

Figura 9. Formas de elementos estruturantes.



Fonte: (SOILLE, 1999)

Por meio da MM podem ser realizadas operações para expandir os objetos da imagem (dilatação) e para diminuí-los (erosão) utilizando o elemento estruturante (vizinhança). Em outras palavras, na erosão os pixels que não atendem a um determinado padrão são apagados da imagem e na dilatação áreas relacionadas a um pixel são alteradas para um dado padrão (SCHAEFFER, 2007).

6.6 Extração de Atributos

A extração de atributos é a etapa do processamento digital de imagens que consiste na obtenção de atributos relevantes para representar e diferenciar os tipos de objetos presentes na imagem e posteriormente analisá-los.

Existem inúmeros atributos para representar Regiões de Interesse (ROIs) em imagens médicas, eles são distribuídos nas principais categorias: textura e geometria (morfologia).

Intuitivamente, textura descreve características da superfície de um objeto na imagem, tais como rugosidade, uniformidade e regularidade. De acordo com Gonzalez e Woods (2010), as três abordagens mais utilizadas no processamento digital de imagens para descrever a textura de uma região são:

- Estatísticas: produzem caracterizações da textura como suave, rugosa, granulada, etc;
- Estruturais: descrevem a textura como arranjos das primitivas de imagem. A primitiva é um elemento fixo que se repete numa área da imagem, como linhas paralelas espaçadas regularmente;
- Espectrais: baseadas em propriedades do espectro de Fourier, é possível detectar características como periodicidade global em uma imagem mediante a identificação de picos de alta energia no espectro.

O reconhecimento de características de textura é um desafio, visto que não apresentam um padrão regular e são dependentes de escala (FERREIRA; BORGES, 2005). Além disso, a análise de textura geralmente é um processo muito demorado, por isso não é indicado em situações onde um grande volume de imagens está envolvido (COSTA; HUMPIRE-MAMANI; TRAINA, 2012).

Em relação aos atributos geométricos, eles descrevem as propriedades morfológicas das regiões de interesse como área, perímetro e circularidade (AL-SHAMLAN; EL-ZAART, 2010). Possuem alta representatividade para identificar

objetos de interesse em imagens (AL-SHAMLAN; EL-ZAART, 2010) e com isso auxiliar no diagnóstico médico.

6.7 Classificação

Na classificação, as características obtidas de cada indivíduo são analisadas para se chegar a alguma conclusão sobre a imagem. Esta etapa pode inferir conceitos a respeito dos objetos analisados e agrupá-los em categorias.

No contexto de processamento de imagens, o objetivo da classificação é a identificação de características ou padrões das estruturas encontradas em uma imagem e de suas atribuições à determinadas classes (SOLOMON; BRECKON, 2011). Em casos em que se tem conhecimento a priori sobre os atributos e suas respectivas classes, a classificação é supervisionada, caso contrário, a classificação é não supervisionada e os atributos são agrupados por critérios de similaridades formando assim os clusters. De maneira geral, a classificação têm a finalidade de substituir a análise visual dos dados por técnicas de análise automática para a identificação de objetos na imagem (QUEIROZ, 2003).

7. Justificativa

Tendo em vista a taxa de exames mamográficos falso negativos e positivos, levando-se em consideração a agressividade do câncer de mama e mortalidade decorrente da doença, assim como a necessidade de melhorar a qualidade da imagem mamográfica, em particular por meio do diagnóstico por imagem auxiliada por computador vem se tornando uma ferramenta valiosa na melhora da acurácia diagnóstica das lesões mamárias, todavia até onde investigamos não encontramos métodos de auxílio diagnóstico (CAD) que utilizasse inteligência artificial com técnicas de redes neurais artificiais, com uma melhor taxa da acurácia de lesões mamárias obtidas neste estudo, o que nos levou à concepção do presente trabalho.

REFERÊNCIAS

1. ABREU, E.; KOIFMAN, S. Fatores prognósticos no câncer da mama feminina. **Revista Brasileira de Cancerologia**, v. 48, n. 1, p. 113-31, 2002.
2. AL-SHAMLAN, Hala; EL-ZAART, Ali. Feature extraction values for breast cancer mammography images. In: **Bioinformatics and Biomedical Technology (ICBBT), 2010 International Conference on**. IEEE. p. 335-340, 2010.
3. ALBUQUERQUE, M. P. de et al. Análise de imagens e visão computacional. **V Escola do CBPF**, 2004. Disponível em: <http://www.cbpf.br/mpa/G7-marcio.pdf>.
4. AMERICAN COLLEGE OF RADIOLOGY; AMERICAN COLLEGE OF RADIOLOGY. Breast imaging reporting and data system atlas (BI-RADS atlas). **Reston, VA: American College of Radiology**, v. 98, 2003.
5. BANDEIRA, D.; VAN DER SAND, I. C. P.; CABRAL, F. B.; FLORES, J. S.; MARON, L. C.; DOS SANTOS, M.; FANTINEL, N. M. Repercussões da mastectomia nas esferas pessoal, social e familiar para a mulher mastectomizada: uma revisão. **Revista Contexto & Saúde**, v. 11 n.20, p.473-482, 2013.
6. BERRY, D.A.; CRONIN, K.A.; PLEVITIS, S.K.; FRYBACK, D.G.; CLARKE, L.; ZELEN, M.; MANDELBLATT, J.S.; YAKOVLEV, A.Y.; HABBEMA, J.D.F.; FEUER, E.J. Effect of screening and adjuvant therapy on mortality from breast cancer. **New England Journal of Medicine**, v. 353, n. 17, p. 1784-1792, 2005.
7. BORCHARTT, T. B. Análise de imagens termográficas para a classificação de alterações na mama. **UFF**, Niterói, 2013.
8. BOYD, N.F.; LOCKWOOD, G.A.; BYNG, J.W.; TRITCHLER, D.L.; YAFFE, M.J. Mammographic densities and breast cancer risk. **Cancer Epidemiology Biomarkers & Prevention**, v. 7, n. 12, p. 1133-1144, 1998.
9. CASES IN 2012: Marked increase in breast cancers must be addressed. **International Agency for Research on Cancer, World Health Organization**, p. 2012–2014, 2013. Disponível em: https://www.iarc.fr/en/media-centre/pr/2013/pdfs/pr223_E.pdf. Acesso em: 26 nov. 2016.
10. COSTA, A.F.; HUMPIRE-MAMANI, G.; TRAINA, A.J.M. An efficient algorithm for fractal analysis of textures. In: **SIBGRAPI Conference on Graphics, Patterns and Images**. IEEE, 2012. p. 39-46, 2012.

11. DESHPANDE, D. S.; RAJURKAR, A. M.; MANTHALKAR, R. M. Mammogram classification using association rule mining. In: **Proceedings of the International Conference on Data Mining (DMIN)**. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp). p. 1, 2014.
12. DESANTIS, C.E.; LIN, C.C.; MARIOTTO, A.B.; SIEGEL, R.L.; STEIN, K.D.; KRAMER, J.L.; ALTERI, R.; ROBBINS, A.S.; JEMAL, A. Cancer treatment and survivorship statistics. **CA: a cancer journal for clinicians**, v. 64, n. 4, p. 252-271, 2014.
13. DONGEN, V.; STIJN, M. Graph clustering by flow simulation. **Dspace**, 2001
Disponível em:
<<http://dspace.library.uu.nl/bitstream/handle/1874/848/full.pdf?sequence=1>>.
Acesso em: 10/11/2015.
14. Dumard, C. H. Possíveis complicações no pós-cirúrgico de mastectomia radical e mastectomia radical modificada. **FisioWeb**, 2005. Disponível em:
<http://www.wgate.com.br/conteudo/medicinaesaudefisioterapia/cancer/cancer_mama_2.html>. Acesso: 02 maio 2016.
15. ELSHINAWY, M.; BADAWEY, A.; ABDELMAGEED, W.; CHOUIKHA, M. Effect of breast density in selecting features for normal mammogram detection. In: **2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro**, p. 141-147, 2011.
16. FERREIRA, C. B. R.; BORGES, D. L. A. selection strategy of a minimal subset of wavelet features in a multiresolution approach for the classification of tumors on mammograms. **V Workshop of Medical Informatics**, v. 1, 2005.
17. GARÓFOLO, A.; AVESANI, C.M.; CAMARGO, K.G.; BARROS, M.E.; SILVA, S.R.J.; TADDEI, J.A.D.A.C.; SIGULEM, D.M. Dieta e câncer: um enfoque epidemiológico. **Revista de Nutrição**, 2004.
18. GOMES, J.; VELHO, L. Computação gráfica: imagem. **IMP**, 2002.
19. GONZALEZ, R. C.; WOODS, R. E. Digital image processing. 3. ed. [S.l.]: **Pearson Prentice Hall**, 2010.
20. HAYKIN S. Redes Neurais. Princípios e práticas, 2a. ed, **Editora Bookman**, 2001.
21. HUO, Z.; GIGER, M.L.; VYBORNÝ, C.J.; BICK, U.; LU, P.; WOLVERTON, D.E.; SCHMIDT, R.A. Analysis of spiculation in the computerized classification of mammographic masses. **Medical Physics**, v. 22, n. 10, p. 1569-1579,

- 1995.
22. INCA, Estimativa 2012: Incidência de Câncer no Brasil. Rio de Janeiro: Divisão de Tecnologia da Informação, Editor 2011. Disponível em: <http://portal.saude.sp.gov.br/resources/ses/perfil/gestor/homepage/estimativas-de-incidencia-de-cancer-2012/estimativas_incidencia_cancer_2012.pdf>. Acesso em: 28 dez. 2016.
23. INCA, Estimativa da incidência de Câncer para 2008 no Brasil e nas cinco Regiões. Divisão de Tecnologia da Informação, Editor 2008. Disponível em: <http://www.inca.gov.br/conteudo_view.asp?id=1793>. Acesso em: 28 dez. 2016.
24. INCA, Mamografia: da prática ao controle. Recomendações para profissionais de saúde., INCA, Editor 2007, Divisão de Tecnologia da Informação, Ministério da Saúde, Brasil: Rio de Janeiro. p. 109. Disponível em: <http://www.inca.gov.br/inca/Arquivos/publicacoes/qualidade_em_mamografia.pdf>. Acesso em: 22 dez. 2016.
25. JUSTO, N.; WILKING, N.; JÖNSSON, B.; LUCIANI, S.; CAZAP, E. A review of breast cancer care and outcomes in Latin America. **Oncologist**. v.18, n.3, p. 248-256, 2013.
26. KENNEDY, J.; EBERHART, R. Particle swarm optimization. In: **Proceedings IEEE International Conference on Neural Networks**. [s.n.], 1995. v. 4, p. 1942–1948. Disponível em: <<http://dx.doi.org/10.1109/ICNN.1995.488968>>.
27. KINNARD, L.; LO, S.C.; WANG, P.; FREEDMAN, M.T.; CHOUKHA, M. Automatic segmentation of mammographic masses using fuzzy shadow and maximum-likelihood analysis. In: **Biomedical Imaging IEEE International Symposium on**. p. 241-244, 2002.
28. KUPINSKI, Matthew A.; GIGER, Maryellen L. Automated seeded lesion segmentation on digital mammograms. **IEEE Transactions on medical imaging**, v. 17, n. 4, p. 510-517, 1998.
29. KUTANZI, K.R.; KOTURBASH, I.; BRONSON, R.T.; POGRIBNY, I.P.; KOVALCHUK, O. Imbalance between apoptosis and cell proliferation during early stages of mammary gland carcinogenesis in ACI rats. **Mutation Research**. v. 694,n.1-2, p.1-6, 2010.
30. LAU, T. K.; BISCHOF, W. F. Automated detection of breast tumors using the asymmetry approach. **Computer and Biomedical Researches**, v. 24, n. 3, p. 273- 295, 1991.
31. LIU, X.; XU, X.; LIU, J.; FENG, Z. A new automatic method for mass detection

- in mammography with false positives reduction by supported vector machine. **In Biomedical Engineering and Informatics (BMEI)**, 2011 4th International Conference on, volume 1, páginas 33–37. IEEE, 2011.
32. LOPES, H. S. Fundamentos da computação evolucionária e aplicações. **XIII Escola Regional de**. p. 52-106, 2006.
33. MARQUES, F.O.; NETO, H.V. Processamento digital de imagens. **Brasport**, 1999.
34. MARTÍ, J.; FREIXENET, J.; MUNOZ, X.; OLIVER, A. Active region segmentation of mammographic masses based on texture, contour and shape features. In: **Iberian conference on pattern recognition and image analysis**. Springer Berlin Heidelberg. p. 478-485, 2003.
35. MATHENGE, E.G.; DEAN, C.A.; CLEMENTS, D.; VAGHAR-KASHANI, A.; PHOTOPOULOS, S.; COYLE, K.M.; GIACOMANTONIO, M.; MALUETH, B.; NUNOKAWA, A.; JORDAN, J.; LEWIS, J.D. Core needle biopsy of breast cancer tumors increases distant metastases in a mouse model. **Neoplasia**, v. 16, n. 11, p. 950-960, 2014.
36. MEDEIROS, J.A.C.C. Exame de partículas como ferramenta de otimização em problemas complexos de engenharia nuclear. 2005. Tese de Doutorado. UNIVERSIDADE FEDERAL DO RIO DE JANEIRO.
37. MEDEIROS, R.M.; DA SILVA, G.; CORRÊA, D.; DA LUZ, E.L.; SCHMIDT, P.C. Câncer De Mama: Análise Situacional em uma Cidade do Norte do Rio Grande do Sul. **Inova Saúde**, v. 2, n. 2, 2014.
38. MÈNDEZ, A. J.; TAHOCES, P. G.; LADO, M. J.; SOUTO, M.; VIDAL, J. J. Computer- Aided Diagnosis: automatic detection of malignant masses in digitized mammograms. **Medical Physics**, v. 25, n. 6, p. 957-964, jun. 1998.
39. MENTOR. Anatomia da mama. 1 figura JPG, color 11 x 10 cm. Disponível em: <<http://www.mentorcorp.com/global-pt/images/breast-anatomy.jpg>>. Acesso em: jun. 2016.
40. MILLER, P.; ASTLEY, S. Detection of breast asymmetry using anatomical features. In: ACHARYA, R. S.; GOLDFOG, C. B. Biomedical Image Processing and Biomedical Visualization, Proc. of SPIE 1993, v. 1905, p. 433-442, 1993.
41. MOREIRA, A. D. S.; BRAZ, G. J.; ROCHA, S. V.; SILVA, A. C.; PAIVA, A. C. Detecção de massas em imagens da mama usando índices de diversidade e algoritmos de segmentação em grafo. **Cad. Pesq., São Luís-MA.**, páginas v. 20, n. especial, julho 2013.

42. NISHIKAWA R. M. "Computer-aided detection and diagnosis", In Digital Mammography, **Springer Berlin Heidelberg**, pages 85-106, 2010.
43. OLIVER, A.; FREIXENET, J.; MARTI, J.; PÉREZ, E.; PONT, J.; DENTON, E.R.; ZWIGGELAAR, R. A review of automatic mass detection and segmentation in mammographic images. **Medical image analysis**, v. 14, n. 2, p. 87-110, 2010.
44. OTSU, N. A threshold selection method from gray-level histograms. **Automatica**, v. 11, n. 285-296, p. 23-27, 1975.
45. PAL, N.R.; Pal, S.K., A review on image segmentation techniques. **Pattern recognition**, v. 26, n. 9, p. 1277-1294, 1993.
46. PEDRINI, H.; SCHWARTZ, W.R. Análise de Imagens Digitais: Princípios. Algoritmos e Aplicações, **Thomson Learning, São Paulo**, 2008.
47. PISANO, E.D.; GATSONIS, C.; HENDRICK, E.; YAFFE, M.; BAUM, J.K.; ACHARYYA, S.; CONANT, E.F.; FAJARDO, L.L.; BASSETT, L.; D'ORSI, C.; JONG, R. Diagnostic performance of digital versus film mammography for breast-cancer screening. **New England Journal of Medicine**, v. 353, n. 17, p. 1773-1783, 2005.
48. QASEEM, A.; SNOW, V.; SHERIF, K.; ARONSON, M.; WEISS, K.B.; OWENS, D.K. Screening mammography for women 40 to 49 years of age: a clinical practice guideline from the American College of Physicians. **Annals of internal medicine**, v. 146, n. 7, p. 511-515, 2007.
49. QUEIROZ, C. J. Análise de transformações geométricas para o georreferenciamento de imagens do satélite **CBERS-I**. Dissertação (Mestrado) - UFRGS – CEPSSRM, 2003.
50. RIBAS, F. E.; MELO, M. P. Anatomia, histologia e embriologia da mama. In: BOFF, R. A.; WISINTAINER, F. (Orgs.). **Mastologia moderna**. Caxias do Sul: Mesa Redonda, p. 23-24, 2006.
51. RUSSEL, S; NORVIG, P. Inteligência Artificial. 2. **Editora Rio de Janeiro: Campos**, 2004.
52. SCHAEFFER, S. E. Graph clustering. **Computer Science Review**, p.27-64, 2007.
53. SICKLES, E.A.; D'ORSI, C.J.; BASSETT, L.W.; APPLETON, C.M.; BERG, W.A.; BURNSIDE, E.S. ACR BI-RADS® Mammography. **ACR BI-RADS Atlas, Breast Imaging Reporting and Data System**, p. 123-132, 2013.

54. SILVA-NETO, P. O. Detecção automática de massas em imagens mamográficas usando particles warm optimization e índice de diversidade funcional. Master'sthesis, **Universidade Federal do Maranhão**, São Luís, Dissertação (Mestrado em Engenharia Elétrica) 2016.
55. SOILLE, P. Morphological image analysis: principles and applications. **Springer Science & Business Media**, 2013.
56. SOLOMON, C; BRECKON, T. Fundamentals of Digital Image Processing: A practical approach with examples in Matlab. **John Wiley & Sons**, 344 p, 2011.
57. SPEECHLEY, V. Tudo sobre câncer. **Editora Andrei**, 2000.
58. SPENCE; ALEXANDER, P. Anatomia Humana Básica. 2a Edição, **Editora Manole LTDA**, 1991.
59. TANG, J.; RANGAYYAN, R. M.; XU, J.; NAQA, I. E.; YANG Y. "Computer-Aided Detection and Diagnosis of Breast Cancer With Mammography : Recent Advances", **IEEE Transactions on Information Technology in Biomedicine**, vol. 13, no. 2, pages 236–251, 2009.
60. TOSTESON, A.N.; STOUT, N.K.; FRYBACK, D.G.; ACHARYYA, S.; HERMAN, B.A.; HANNAH, L.G.; PISANO, E.D. Cost-effectiveness of digital mammography breast cancer screening. **Annals of internal medicine**, v. 148, n. 1, p. 1-10, 2008.
61. VAN DER MERWE, D. W.; ENGELBRECHT, A. P. Data clustering using particle swarm optimization. In: **Evolutionary Computation. CEC'03. The 2003 Congress on**. IEEE, 2003. p. 215-220, 2003.
62. VEISY, A.; LOTFINEJAD, S.; SALEHI, K.; ZHIAN, F. Risk of breast cancer in relation to reproductive factors in North-West of Iran, 2013-2014. **Asian Pacific journal of cancer prevention: APJCP**, v. 16, n. 2, p. 451-455, 2014.
63. WALKER, M.J.; CHIARELLI, A.M.; KNIGHT, J.A.; MIREA, L.; GLENDON, G.; RITVO, P. Perceived risk and adherence to breast cancer screening guidelines among women with a familial history of breast cancer: a review of the literature. **The Breast**, v. 22, n. 4, p. 395-404, 2013.
64. WORLD HEALTH ORGANIZATION et al. WHO position paper on mammography screening. **World Health Organization**, 2014.
65. ZUIDERVELD, K. Contrast limited adaptive histogram equalization. In: **Graphics gems IV**. Academic Press Professional, Inc. p. 474-485, 1994.

CAPÍTULO 1

Artigo publicado



Timeline Publication Pvt. Ltd. , ISSN(Online):2249-071X

Impact Factor(JIF) :2.7 , UIF:1.0909

PROSPECÇÃO DE SOFTWARE PARA PROCESSAMENTO DE IMAGEM MAMÁRIA

Francisco Adelson Alves-Ribeiro¹, Miguel de Sousa Freitas¹, Juscelino Lopes da Silva¹, Carla Solange Escórcio-Dourado¹, Fabiane Araújo Sampaio¹, Luana Mota Martins¹, Camila Maria Simplício-Revoredo¹, Benedito Borges da Silva^{1,2}

¹Pos-graduate Program of Northeast Network of Biotechnology (RENORBIO), Northeast, Brazil

²Department of Mastology, Federal University of Piauí, Teresina, Piauí, Brazil

Corresponding author:

Benedito Borges da Silva

Avenida Elias João Tajra, 1260

CEP 64049-300, Jockey

Teresina, Piauí, Brazil

Phone: +55 86 3232 5063

Fax: +55 86 3215 0470

Email: beneditoborges@globo.com

RESUMO

Diagnóstico Auxiliado por Computador (CAD) são sistemas computacionais que visam ajudar os médicos em seu diagnóstico. Software para processamento de imagens tem contribuído de várias maneiras para a medicina, tais como, no diagnóstico e tratamentos de doenças como o câncer de mama. O câncer de mama é uma das principais causas de morte das mulheres em todo o mundo. No entanto, a detecção precoce da doença aumenta em muito a possibilidade de cura. Portanto, vários tipos de sistemas computacionais baseados em processamento de imagens estão sendo desenvolvidos por vários grupos de pesquisa, a fim de auxiliar o radiologista na precisão do diagnóstico. O objetivo desta pesquisa foi a busca de sinais sobre o estado atual da tecnologia, identificando software de processamento de imagem no mercado global para auxílio no diagnóstico médico. Os resultados foram obtidos através de pesquisas feitas em publicações científicas repositório *Web of Science* usando o banco de dados *Derwent*, incluindo registros de depósitos de patentes mundiais. Os dados mensurados demonstraram apenas 107 registros de patentes. Demonstrando que estudos e pesquisas precisam ser desenvolvidos, uma vez que a palavra-chave "processamento de imagem software" mostrou apenas 3 registros de patentes. Os Estados Unidos são os líderes em número de patentes por país, com um total de 40 registros na busca por resumo. A pesquisa por títulos resultou em apenas 20 registros nos US, embora esta quantidade seja maior que a encontrada em outros países. Os resultados desta prospecção mostraram a escassez de registro de patentes dos softwares, para processamento de imagens, nos países subdesenvolvidos e em desenvolvimento. As razões podem ser a ausência de uma estratégia empreendedora, para o melhoramento técnico-científico.

Palavras-chave: neoplasias mamárias, processamento de imagens, softwares mamografia

ABSTRACT

Computer-Aided Diagnosis (CAD) systems are computational tools that aim to help medical professionals in their diagnostic decisions. Software for processing images has contributed in many ways to medicine, such as, the diagnosis and treatments of diseases such as breast cancer. Breast cancer is one of the main causes of women death all over the world. However, early detection of the disease increases greatly the possibility of cure. Therefore, several types of computer systems based on image processing are being developed by many research groups in order to aid the radiologist in the accuracy of the diagnosis. The aim of this survey was to search for signs of the current state of technology, identifying image processing software in the global marketplace for assistance in medical diagnosis. The outcomes measured were obtained through the searches done in scientific publications repository Web of Science using the Derwent database, including records of deposits of world patents. Data measurement revealed that there were only 107 patent registrations. Further studies and research need to be designed, since the keyword “image processing software” showed only 3 patent registrations. The United States is the leader in the number of patents per country, with a total of 40 registrations in the abstract search. The title search resulted in only 20 registrations, although this quantity was much larger than that found in other countries. The this survey showed that there is a paucity of patent registration in lesser-developed and developing countries. The reasons may be the absence of a strategic enterprising vision for encouragement to improve technical and scientific endeavors.

Keywords: breast neoplasms, image processing, software mammography

I. INTRODUÇÃO

O câncer de mama é o tipo de câncer que mais acomete as mulheres em todo o mundo, tanto em países em desenvolvimento quanto em países desenvolvidos. Nos Estados Unidos, estima-se que em 2014 foram diagnosticados mais de 200 mil novos casos e 40 mil mortes em mulheres [1]. No Brasil, para o ano de 2014, foram estimados 57.120 casos novos de câncer de mama, com um risco estimado de 56,09 casos para cada 100 mil mulheres [2].

A mamografia é o exame diagnóstico por imagem padrão para avaliação da glândula mamária. Esse tipo de exame pode detectar nódulos não palpáveis, microcalcificações e assimetrias das mamas. Entretanto, a taxa de falso negativo é de 10% a 15%, podendo chegar a 40% em pacientes com mamas densas [1,3]. Leitura da mamografia é um trabalho difícil para os radiologistas, suas decisões dependem da sua experiência e de critérios subjetivos, mesmo os mais experientes podem ter uma taxa de erro de 65-75%. O Computer aided diagnóstico (CAD) podem ajudar os radiologistas na interpretação da mamografia, detecção e classificação de massas. De 65-90% das biópsias suspeita de câncer revelou-se benigno, é muito importante desenvolver CADs, que podem ajudar a distinguir benigna de lesões malignas. A combinação do esquema CAD e do conhecimento dos especialistas melhoraria muito a precisão da detecção. A sensibilidade de detecção sem CAD é de 80% e com CAD até 90% [4-6]. No entanto, a utilização desse exame como método de rastreamento reduz a mortalidade em 25%. Portanto, faz-se necessário o desenvolvimento de novos sistemas de detecção assistido por computador baseado em análise das características globais da imagem mamográfica.

Dessa forma, os sistemas de Detecção Assistida por Computadores CADE, podem auxiliar os radiologistas na difícil tarefa de leitura e interpretação de imagens da mamografia, automatizando o processo de interpretação de imagens médicas [9]. De 65 a 90% das biópsias de tecidos com suspeita de câncer mostram-se benignas, o que mostra a importância do desenvolvimento de sistemas CADs que permitam a distinção entre lesões malignas e benignas [9].

A demanda crescente dos hospitais e clínicas por um diagnóstico de exames radiológicos de forma rápida e precisa, somada aos avanços computacionais e de processamento de imagens, fizeram surgir novas frentes de pesquisa relacionadas

ao diagnóstico radiológico auxiliado por computador [10]. O diagnóstico auxiliado por computador ou detecção auxiliada por computador é uma ferramenta computacional relativamente recente, que tem sido implementada para prover dupla releitura, aumentando a sensibilidade do método. A propósito, estudos clínicos têm demonstrado que o CAD aumenta a sensibilidade de detecção do câncer de mama em até 21% [10-13].

Daí a importância da análise atual da prospecção de software de imagens mamográficas. Uma prospecção é uma constatação dos estudos e registros de patentes para analisar o mercado atual em relação à tecnologia aplicada e verificar se o mesmo está saturado para um lançamento de algum futuro produto [11].

Assim, o presente artigo teve como objetivo apresentar estudo sobre o estado da técnica por meio da prospecção dos softwares para processamento de imagens mamográficas e auxílio ao diagnóstico médico.

II. METODOLOGIA

Atendendo aos procedimentos metodológicos realizamos busca na base internacional de patentes *Derwent Innovation Index* no mês de junho de 2015, fazendo uso do repositório de publicações científicas *Web of Science* [12,13]. A pesquisa foi realizada nos moldes de um estudo quantitativo, baseado em análise comparativa de resultados com visão neopositivista já que mensurações são consideradas mais ricas que descrições verbais. A base *Derwent* possui registros de depósito de patentes de vários órgãos emissores de patentes, tais como dos escritórios americanos, europeus, alemães e japoneses (14,15). Foram usadas as seguintes palavras-chave nessa prospecção: *software diagnosis breast cancer*, *Software processing images* e *software mammography*. Os campos de busca utilizados na base *Derwent* foram de título e resumo de tópicos [13]. Com o resultado da busca foi possível a construção dos gráficos via *Microsoft Office Excel 2013*. Os 107 registros encontrados foram extraídos da pesquisa por títulos e tópicos. A análise dos resultados foi feita através do código do sistema Internacional de Classificação de Patentes (IPC), permitindo analisar a situação do software em desenvolvimento, se já desenvolvido previamente ou se já apropriado por outro inventor.

III. RESULTADOS

Realizada a busca por patentes, notou-se um retorno maior de depósitos de patentes por resumos e por título, relacionados com os termos software *diagnosis breast cancer*, resultando em 72 registros, 32 *software mammography* e 3 *Image processing software*, totalizando 107 (cento e sete) patentes encontradas até 2014.

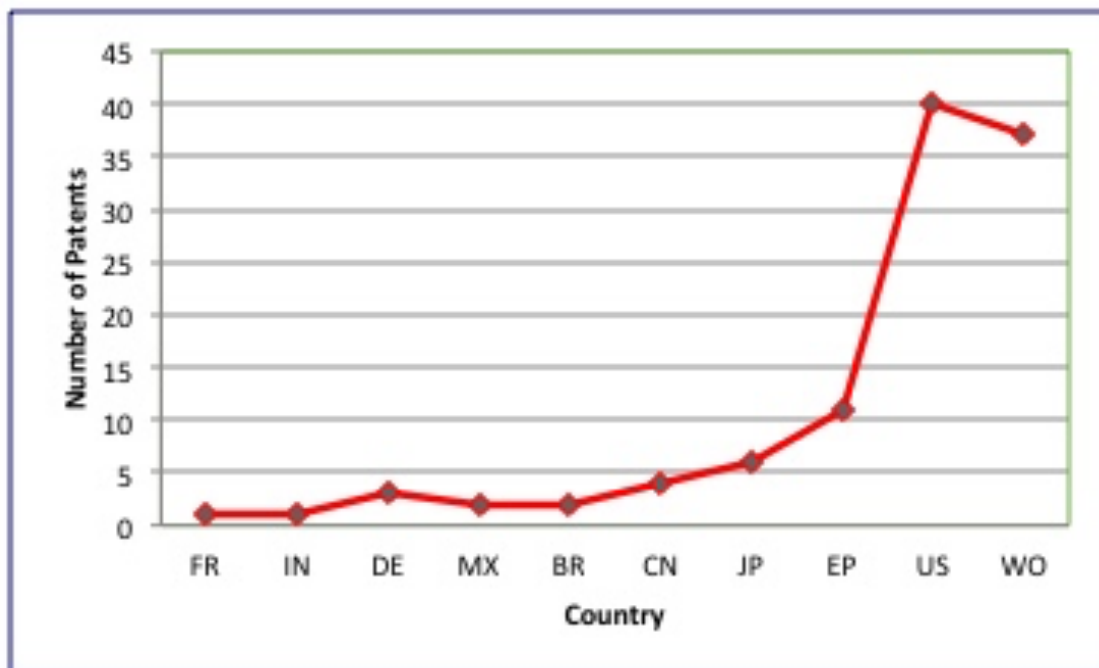
O Quadro 1 apresenta os resultados da quantidade de patentes depositadas relativas aos termos pesquisados:

Quadro 1. Quantidade de registros obtidos com os termos, por resumos e título.

Termo	No de Patentes encontradas
<i>Breast cancer diagnosis software</i>	72
<i>Mammography software</i>	32
<i>Image processing software</i>	3
TOTAL	107

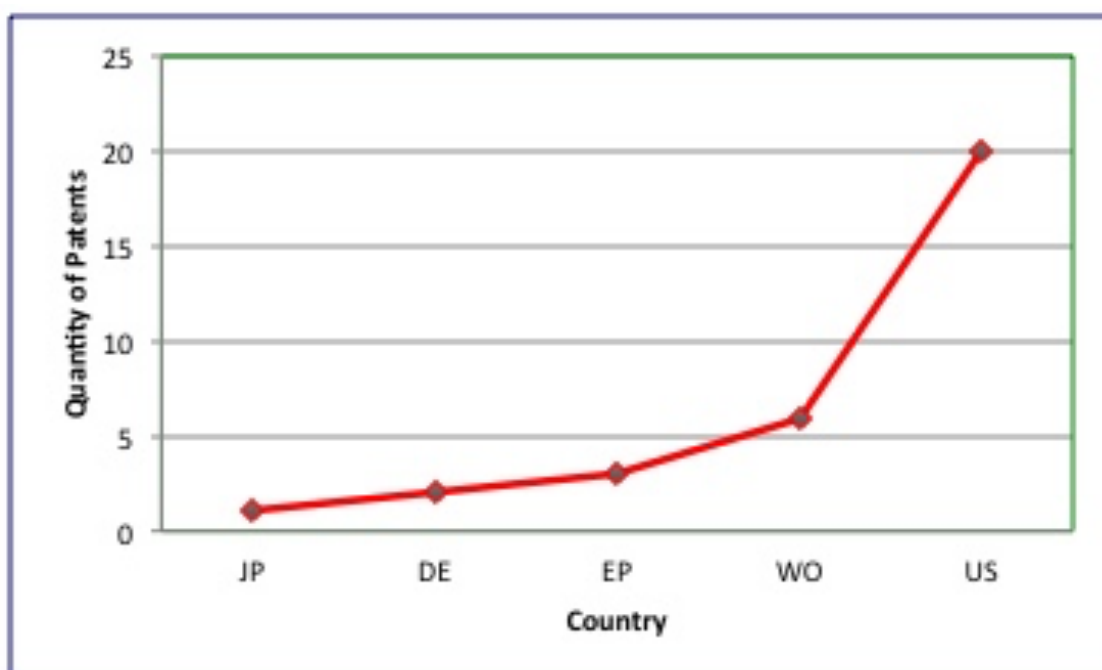
A Figura 1 apresenta o número de patentes por país, resultante da busca por resumos, onde se observa que o maior número de patentes encontra-se nos Estados Unidos (US), com 40 registros, vindo em segundo lugar as patentes registradas via base mundial (WO), mediante o Tratado de Cooperação de Patentes (PCT), com 37 patentes depositadas, em seguida, terceiro lugar, a Organização Européia de Patentes (EP) com 11 patentes e em quarto lugar vem o Japão com seis registros. Por fim, vêm os países com menor número de patentes, França (FR), Índia (IN), Alemanha (DE), México (MX), China (CN) e Brasil (BR), que juntos somam 13 patentes.

Figura 1. Quantidade de Patentes por Países (Resumo).



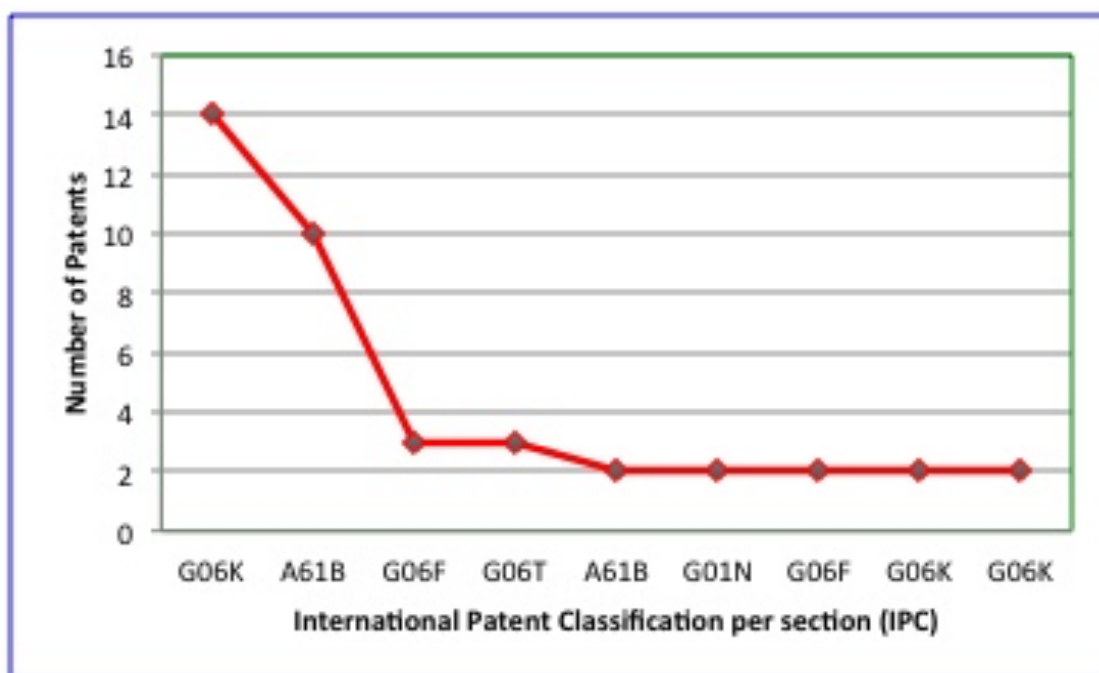
A Figura 2 apresenta a quantidade de patentes depositadas por país de origem, resultante da busca por títulos, onde foi verificada a maior quantidade de registros depositados pelos US. Tal fato pode ser devido não só ao maior número de patentes depositado pelos Estados Unidos como também a corroboração da palavra chave *software diagnosis breast cancer* utilizada na busca.

Figura 2. Quantidade de patentes depositadas por País de origem.



Nesta prospecção foi feito também um levantamento dos registros de patente baseados na Classificação Internacional de Patentes (CIPs/IPCs). Como não há uma seção específica para Computação ou Engenharia, na Seção que representa a Física denominada G que contém uma subclasse G06K, de interesse neste estudo, que se encontra compartimentado nos subgrupos 019/00 (equipamentos ou métodos de computação digital), e 019/20 (processamentos de dados). Portanto é nesta seção G que se verifica a maior quantidade de ocorrências de registro de patentes, com 14 casos. Já na Seção A (Necessidade Humanas), representante da subclasse que abrange diagnósticos, cirurgia e identificação da referida classificação (CIPs/IPC), que totalizou 10 ocorrências, consoante o realçado na figura 3.

Figura 3. Classificação Internacional de Patentes por CIP



IX. CONCLUSÃO

A prospecção foi feita utilizando o repositório de publicações científicas *Web of Science*, fazendo uso da base de dados *Derwent*, por apresentar o maior número de registros de softwares na área de processamento de imagens. O banco da *Derwent* inclui registros de depósitos de patentes provenientes de escritórios renomados, tais como: Europeu, Americano, Alemão e Japonês. Os dados mensurados demonstram que com apenas 107 registros de patentes, há

necessidade desenvolvimento de estudos e pesquisas, pois as palavras-chaves “*Image processing software*”, mostrou apenas 3 registros de patentes.

Os Estados Unidos estão a frente na quantidade de patentes por país com o total de 40 registros, na busca por resumos, enquanto a busca por títulos foi de apenas 20 registros, todavia bem superior aos outros países. Os resultados desta prospecção mostraram nos países subdesenvolvidos e em desenvolvimento o registro de patentes de software é escasso, por falta de uma visão empreendedora estratégica e fomentadora do aprimoramento técnico-científico e da pesquisa voltada o desenvolvimento de softwares para o auxílio ao diagnóstico médico, que é uma tendência mundial e uma carência para os países em questão.

REFERÊNCIAS

- [1] INCA - National Cancer Institute. Breast Cancer-Patient Version, 2014. Available in: <http://www.cancer.gov/cancertopics/types/breast>. Accessed: 22 Jun 2016.
- [2] INCA - National Cancer Institute. Estimativa 2014 – Incidência de Câncer no Brasil. Available in: http://www.inca.gov.br/rbc/n_60/v01/pdf/11-resenha-estimativa-2014-incidencia-de-cancer-no-brasil.pdf. Accessed 20 May 2016.
- [3] TANG, J.; RANJAYYAN, R. M.; NAQA, R. I. EI, YANG, Y. Computer-aided detection and diagnosis of breast cancer with mammogram: recent advances, IEEE Trans Inf Technol Biomed. Vol 12, p:236–251, 2006.
- [4] GANESAN, K.; ACHARYA, U. R.; CHUA, C. K.; MIN, L. C.; ABRAHAM, KT. Computer-aided breast cancer detection using mammograms: a review. IEEE Rev Biomed Eng. Vol 6, p.77-98, 2009.
- [5] CHENG, H, D.; SHI X J.; MIN R.; HU L. M., CAI, X. P., DU, H.N. Approaches for automated detection and classification of masses in mammograms. Pattern Recognition.” Vol 39, 2006.
- [6] DAVIES, D. H.; DANCE, D. R.; JONES, C. H. Automatic detection of clusters of calcifications, SPIE Med. Imaging IV: ImageProcess. Vol 1233, p.185–191, 1990.
- [7] GURCAN, M.N.; CHANG H. P.; SAHINER, B.; HADJIISKIL, L.; PETRICK, N.; HELVIE, M.A. Optimal neural network architecture selection: improvement in computerized detection of microcalcifications, Acad Radiol, p. 420–429, 2002.
- [8] TAN, M.; QIAN W.; PU, J.; LIU, H.; ZHENG, B. A new approach to develop computer-aided detection schemes of digital mammograms. Phys Med Biol. Vol. 60, p.4413-4427, 2015.

- [9] TANG, J.; RANGAYYAN, R.M.; XU J.; NAQA, I. E.; YANG, Y. Computer-aided detection and diagnosis of breast cancer with mammography: recent advances, *IEEE Trans. Inf.Technol. Biomed.* Vol. 13, 2009.
- [10] CALAS, M. J. G.; GUTFILEN, B.; PEREIRA, W. C. A. CAD and mammography: why use this tool?, *Radiol Bras.* Vol. 45, p. 46-52, 2012.
- [11] Derwent Innovation Index. Web of Knowledge. Available in: <http://isiknowledge.com/diidw>. Accessed 04 May 2015.
- [12] Web of Science. 2015. [http://apps.webofknowledge.com/DIIDW_GeneralSearch_input.do?product=DIIDW&SID=1CeBUyvDHUgYyLKM\[11\]MBN&search_mode=GeneralSearch](http://apps.webofknowledge.com/DIIDW_GeneralSearch_input.do?product=DIIDW&SID=1CeBUyvDHUgYyLKM[11]MBN&search_mode=GeneralSearch). Accessed 09 April 2015.
- [13] Derwent Help. 2013. Abrangência do país da citação no Derwent. Available in: http://images.webofknowledge.com/WOKRS511B5/help/pt_BR/DII/hs_citation_country_coverage.html. Accessed 14 April 2015.
- [14] Derwent. Derwent Innovations Index. On the Thomson Reuters Web of Science research platform. Available in: <http://thomsonreuters.com/derwent-innovations-index/>. Accessed 20 May 2015.
- [15] BIRDWELL, R. L.; BANDODKAR; P.; IKEDA D. M. Computer-aided detection with screening mammography in a university hospital setting, *Radiology.* Vol. 236, p.451–7, 2005.
- [16] KO, J. M.; NICHOLAS, M. J.; MENDEL, J. B.; SLANETZ, P. J. Prospective assessment of computer-aided detection in interpretation of screening mammography. *American Journal of Roentgenology.* Vol. 187, p.1483–91, 2006.
- [17] Open Letter to the National Institute of Industrial Property on Patent. Available in: <http://ccsl.ime.usp.br/pt-br/news/12/03/20/ccsl-usp-divulga-carta-aberta-ao-inpi-sobre-patents-de-software>. Accessed 12 July 2015.
- [18] World Intellectual Property Organization (WIPO). The World Intellectual Property Organization (WIPO) is the global forum for intellectual property services, policy, information and cooperation. Available in: <http://www.wipo.int/>. Accessed 20 May 2015.
- [19] CHENG, H. D.; SHI, X. J.; MIN, R.; HU, L. M.; CAI, X. P.; DU, H N. Approaches for automated detection and classification of masses in mammograms. *Pattern Recognition.* Vol.39, p.10-12, 2003.
- [20] International patent classification. 8th Edition. INPI 2006. Available in: <http://ipc.inpi.gov.br/IPCpubPrep/Full-BR/guide/br/guide.pdf> classification online <http://ipc.inpi.gov.br/ipcpub/#refresh=page>. Accessed 12 May 2015.
- [21] European Patent Office. 2013. Available in: <http://www.epo.org/>. Accessed 20 April 2015.

- [22] “A comprehensive overview of the coverage of the worldwide patent database managed by the EPO”. In: Global Patent Data Coverage 2011. Available in: [http://documents.epo.org/projects/babylon/eponet.nsf/0/204052542F7D7507C12572EB00455D35/\\$File/comprehensive_overview.pdf](http://documents.epo.org/projects/babylon/eponet.nsf/0/204052542F7D7507C12572EB00455D35/$File/comprehensive_overview.pdf). Accessed 07 July 2015.
- [23] HOWLADER, N.; NOONE, A. M.; KRAPCHO, M. Cancer Statistics Review, 1975-2010, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2010/. Accessed 20 April 2015.
- [24] LAW OF AUTHOR'S RIGHTS ACT - LAW N° 9.610, from February, 19, 1998. http://www.planalto.gov.br/ccivil_03/leis/l9610.htm. Accessed 10 June 2015.
- [25] LAW OF SOFTWARE - LAW N° 9.609, from February 19, 1998. Available in: http://www.planalto.gov.br/ccivil_03/leis/L9609.htm. Accessed 10 July 2015.
- [26] LAW OF INNOVATION, n. ° 10.973, de 2 de dezembro de 2004. It provides for incentives for innovation and scientific and technological research in the productive environment and other matters. Official Gazette of [the] Federative Republic of Brazil, Brasilia, DF 3 December 2004. Available in: http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/lei/l10.973.htm. Accessed 13 July 2015.
- [27] NUNES, M. A.; CAZELLA, S. C.; PIRES, E. A.; RUSSO, S. L. Discussion About Academic Scientific Production & Production Technology: Changing Paradigms. GEINTEC. Vol.3, p. 205-220, 2013.
- [28] RUSSO, S. L.; SILVA, G. F.; OLIVEIRA, L. B.; SANTOS, M. M.; NUNES, M. A.; VASCONCELOS, J. S.. Training in Technological Innovation for Entrepreneurs. Saint Kitts: Publisher UFS, p. 55-90, 2012.
- [29] TANG J.; RANGAYYAN, R. M.; I. NAQA, EL; YANG, Y. Computer-aided detection and diagnosis of breast cancer with mammography: recent advances, IEEE Trans InfTechnol Biomed, Vol.13, p. 236–51, 2009.
- [30] RASHED, E. A.; ISMAIL, I. A.; ZAKI, S. I. Multiresolution mammogram analysis in multi-level decomposition. Pattern Recognit Lett. Vol.28, p.286–92, 2007.
- [31] VARELA, C.; TAHOSES, P. G.; MNDEZ, A. J.; SOUTO, M.; VIDAL, J J. Computerized detection of breast masses in digitized mammograms. Comput Biol Med. Vol. 37, p.214–26, 2007.
- [32] GUO, X. C.; YANG, J.; WU, G.; WANG, C. Y.; LIANG, Y. A novel LS-SVMs hyper-parameter selection based on particle swarm optimization. Neurocomputing. Vol.71, p. 3211–5, 2008.
- [33] SHI, Y.; Eberhart, R. Modified particle swarm optimizer. In: Proceedings of the IEEE international conference on evolutionary computation, p.69–73, 1998.
- [34] SUCKLING. J.; PARKER, J.; DANCE, D.; ASTLEY, S.; HUTT, I.; BOGGIS, C. The mammographic image analysis society digital mammogram database. Exerpta Med Int CongrSer. Vol. 1069, p.375–8, 1994.

- [35] JONG-BAE, K.; HANG-JOON, K. Multiresolution-based watersheds for efficient image segmentation. *Pattern Recognition Letters*. v. 24, p. 473–488, 2003.
- [36] LI, J.; GAO, L.; SHEN, Z. Extraction and analysis of digital images feature of three kinds of wheat diseases. *3rd International Congress on Image and Signal Processing (CISP)*. Yantai (China). v.6. p. 2543–8, 2010.
- [37] MERWE, V.D.W.; ANDRIES, P.E. Data clustering using particle swarm optimization. *Proceedings of the Congress on Evolutionary Computation: Vol. 1*, 2003. <http://dx.doi.org/10.1109/cec.2003.1299577>.
- [38] NETO, O.P.S.; CARVALHO, O.; SAMPAIO, W.; CORREA, A.; PAIVA, A. Automatic segmentation of masses in digital mammograms using particle swarm optimization and graph clustering.; *Institute of Electrical & Electronics Engineers (IEEE)* 2, 2015. <http://dx.doi.org/10.1109/iwSSIP.2015.7314189>.
- [39] SMITH, R.A.; COKKINIDES, V.; EYRE, H.J. American Cancer Society guidelines for the early detection of cancer. *CA: a cancer journal for clinicians*. v. 53, p. 27-43.

CAPÍTULO 2

Artigo aceito



ISSN: 1546-1955 (Print): EISSN: 1546-1963 (Online)

JCR: 1.665

DETECÇÃO DE MASSAS EM IMAGENS MAMOGRÁFICAS UTILIZANDO EXTRAÇÃO DE CARACTERÍSTICAS DE TEXTURAS E REDES NEURAIS

Francisco Adelson Alves-Ribeiro^{1,}, Danylo Rafael Costa-Silva¹, Carla Solange Escórcio-Dourado¹, Otilio Paulo da Silva-Neto¹, Marcelo Eugenio de Castro-Gonçalves¹, Thiara Lorena Bezerra da Silva¹, Wilson de Oliveira Júnior¹, Vanessa de Oliveira Ribeiro¹, Luana Mota Martins¹, Elmo de Jesus Nery-Júnior¹, Fabiane Araújo Sampaio¹, Benedito Borges da Silva^{1,2}*

¹ Postgraduate Program of the Northeast Network of Biotechnology (RENORBIO), Northeast, Brazil

² Breast Unit, Getúlio Vargas Hospital, Federal University of Piauí, Teresina, Piauí, Brazil.

***Corresponding author**

Francisco Adelson Alves Ribeiro
Rua Amazonas, 1747
Zip code 64009-130, Marquês de Paranagua
Teresina, Piauí, Brasil
Phone: +55 86 3223 0619
Email: adelton@me.com

RESUMO

Sistemas computacionais são pervasivos em várias áreas de conhecimento, na saúde auxiliam diagnósticos e tratamentos de doenças como o câncer. As contribuições nessa linha podem ser pelo desenvolvimento de softwares como ferramentas para detecção de massas em imagens mamográficas. O objetivo desta pesquisa é apresentar os resultados obtidos nas técnicas de treinamento e aprendizado de máquina com uso de Redes Neurais Artificiais, na tarefa de detecção de massas a partir de imagens mamográficas. Os resultados comprovaram que o algoritmo Random Forest obteve o melhor índice de acerto, no que diz respeito à sensibilidade, à especificidade e à acurácia, atingindo 92.10%, 93.50% e 92.90%, respectivamente.

Palavras-chave: cancer de mama, mamografia, processamento de imagens, redes neurais.

ABSTRACT

Computer systems are widely used in various fields of knowledge. In healthcare settings, computer technology contributes to the diagnosis and management of diseases such as cancer, including breast cancer which is the most common malignancy among women. Software developed for masse detection in mammographic images may contribute as valuable tools. The aim of the current article was to present the results obtained during training and machine-learning techniques with the use of Artificial Neural Networks for masse detection using mammographic images. Results showed that the Random Forest algorithm obtained the best correct classification rate, relative to sensitivity, specificity and accuracy, achieving 92.10%, 93.50% and 92.90%, respectively.

Key words: breast cancer, breast masses, mammography, image processing, neural networks.

I. INTRODUÇÃO

O câncer de mama é a doença maligna que mais comumente as mulheres em todo o mundo, tanto em países desenvolvidos em desenvolvimento. Nos Estados Unidos, estima-se que uma em cada oito mulheres desenvolverá o câncer de mama (FEUER et al., 1993). No Brasil, para o ano de 2016, foram estimados 57.960 casos novos de câncer de mama e cerca de 14.206 mortes em mulheres pela doença, com um risco estimado de 56,20 casos novos para cada 100 mil mulheres (INCA, 2016).

O mecanismo mais comum e padrão para o rastreamento e diagnóstico do câncer de mama é a mamografia, um exame radiológico que gera uma imagem em tons de coloração cinza da mama, que o especialista analisa e identifica de forma visual onde se encontra a lesão ou as lesões (GIGER, 2000). Com o uso do rastreamento mamográfico foi observada uma redução na taxa de mortalidade ocasionada por essa doença (SMITH, 2003). Entretanto, a taxa de falso negativo é de 10% a 15%, podendo chegar a 40% em mulheres com mamas densas (INCA, 2014). De forma geral, a leitura e interpretação de imagens da mamografia são tarefas que exigem experiência do radiologista, e seu julgamento depende de treinamento, experiência e critérios subjetivos. Mesmo experts bem treinados apresentam uma taxa de acertos de classificação com variação de 65 a 75% entre observadores (CHENG et al., 2006). Todavia, a utilização da mamografia como método de rastreamento é estimada reduzir a mortalidade em cerca de 25%, mesmo com uma taxa de falso negativo que pode chegar até 15% (TANG, 2009).

Daí a busca de métodos mais avançados, tais como os imagens computadorizadas, em particular os sistemas para auxílio e detecção assistida por computadores, “*Computer-Aided Detection*” (CAD) e diagnóstico auxiliado por computador “*Computer Aided Diagnosis*” (CADx). Estes sistemas podem auxiliar os radiologistas na difícil tarefa de leitura e interpretação de imagens da mamografia, automatizando o processo de interpretação de imagens médicas. De 65 a 90% das biopsias de tecidos com suspeita de câncer mostram-se benignas, com traumas físicos e emocionais para as paciente, o que mostra a importância do desenvolvimento de sistemas de detecção e diagnóstico mais precisos como os CADs, que permitam aprimorar a distinção entre lesões benignas e malignas (TANG et al., 2009).

A propósito, o avanço dos hospitais e clínicas por um diagnóstico de exames radiológicos de forma rápida e precisa, somada aos avanços computacionais e de processamento de imagens, fizeram surgir novas frentes de pesquisa relacionadas a detecção e diagnóstico de imagens auxiliados por computador (CADs) (TANG et al., 2015). Assim, no presente estudo do artigo foram abordadas cinco técnicas de treinamento e aprendizado de máquina com uso de Redes Neurais Artificiais (RNAs), na tarefa de detecção de tumores a partir de imagens mamográficas, com o intuito de avaliar os classificadores para o uso em um sistema CADs para a detecção e diagnóstico de câncer de mama. Os algoritmos utilizados foram o *Random Forests* (RF), o *Multi-Layer Perceptron* (MLP), o *Support Vector Machines* (SVM), o *Radial-Basis Function Networks* (RBF Network) e o *Naive Bayes* (NB), que se encontram estruturadas neste levantamento bibliográfico na sequência supracitada.

II. MATERIAIS E MÉTODOS

O presente trabalho é um estudo transversal sobre detecção de tumores em imagens mamográficas auxiliadas por computador, que foi aprovado pelo Comitê de Ética da Universidade Federal do Piauí, Brasil, sob o nº 40474114.8.0000.5214. O Estudo incluiu 100 imagens, divididas em dois grupos benignas e malignas, com 50 casos cada. As imagens foram adquiridas na base de dados de domínio público DDSM (*Digital Database for Screening Mammography*), Universidade do Sul da Flórida, EUA. Em seguida, foi realizado o pré-processamento da imagem para remoção artefatos, seguido pela segmentação. Esta etapa separa as regiões suspeitas do tecido mamário, músculo e demais estruturas presentes na imagem mamográfica, para então extrair as regiões de interesse e localizar os tumores (FERNANDES et al., 2014), que se apresentam mais transparentes e brilhantes com características na imagem compatíveis com o tipo de tumor (CHENG et al., 2006; Fernandes et al., 2014). A abordagem de segmentação dividida em duas etapas, na primeira foi utilizado o método de (Otsu, 1979) e na segunda o método PSO (*Particle Swarm Optimization*) proposto por (MERWE et al., 2003). Com o intuito de analisar individualmente cada região segmentada, foi utilizada a técnica de crescimento de Região de Interesse do Objeto (ROI), seguida de técnicas de redução de falsos positivos. Por fim, as técnicas de treinamento e aprendizado de máquina com uso de Redes Neurais Artificiais (Artificial Neural Networks, ANN).

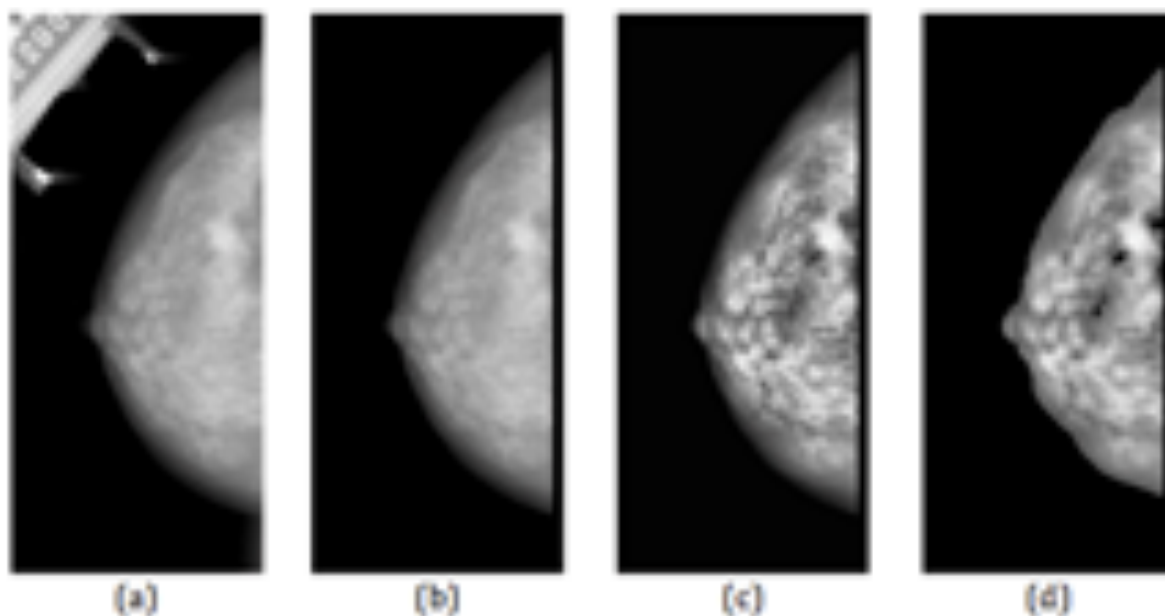
Aquisição das imagens

A base de dados DDSM dispõe de aproximadamente 2.600 casos, fornecidos gratuitamente na internet. Cada caso mamográfico possui duas imagens, obtidas da mama direita e esquerda, nas visões Médio-Lateral Oblíqua (MLO) e Crânio-Caudal (CC). Estas imagens são compactadas no padrão LJPEG e representadas em níveis de cinza com resolução de 12 ou 16 bits por pixels. Os laudos obtidos por especialistas acompanham as imagens, fornecendo informações como o código de cada paciente, a idade na época do estudo, a classificação na categoria e na composição da densidade mamária de acordo com o padrão BI-RADS (*Breast Imaging Report and Data System*), tipo de anormalidade, a localização e o laudo dado por especialistas (TOURASSI et. al., 2003). O conjunto de 50 casos, com de 100 imagens, foram selecionados para avaliar o desempenho dos algoritmos propostos.

Pré-processamento

Antes de iniciar a etapa de segmentação das mamografias, faz-se necessário remover as estruturas indesejadas, tais como ruídos, bordas, marcações e músculo peitoral, estes oriundos na aquisição das imagens. Para remover estas estruturas, utilizou-se a metodologia desenvolvida por (SAMPAIO et al., 2011). Após a remoção destas estruturas, foi aplicada a técnica de realce baseada no histograma e o filtro da média, com objetivo de realçar as estruturas internas das mamografias (Figura 1).

Figura 1. Resultado das etapas de melhoramento da imagem mamográfica: (a) Imagem original; (b) Imagem sem as bordas e marcações; (c) Realce local; (d) Filtro da média.



Segmentação

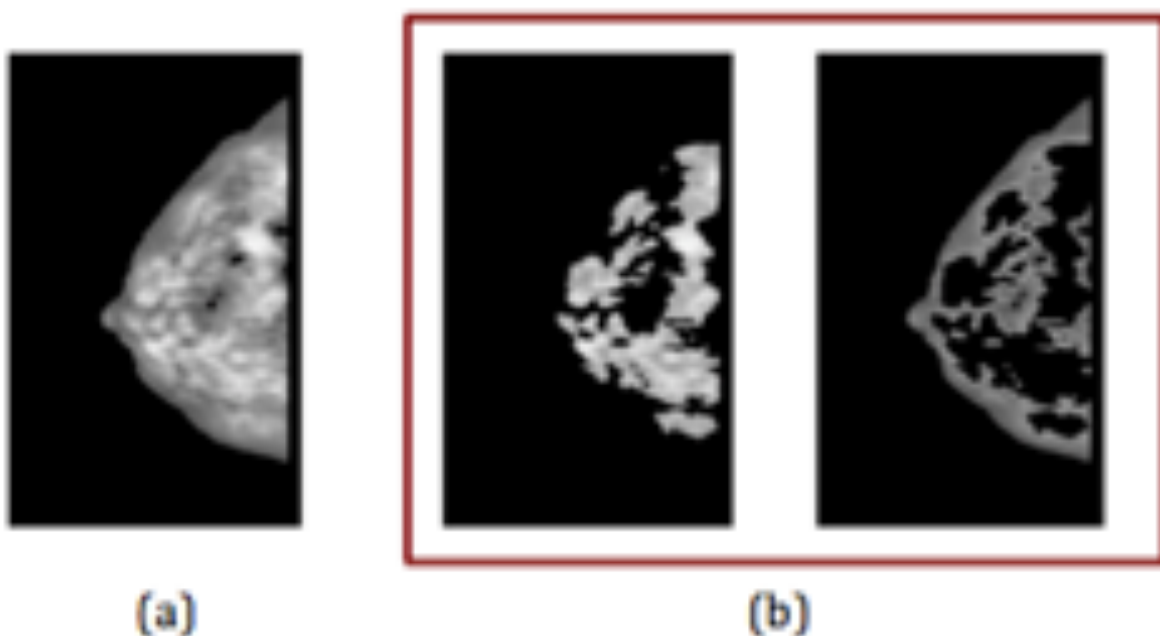
As imagens geradas após a etapa de pré-processamento, foram submetidas ao processo de segmentação responsável pela separação das regiões mais homogêneas da imagem, resultando nos objetos contendo artefatos suspeitos como nódulo.

Nesse artigo propomos a abordagem de segmentação dividida em duas etapas, na primeira utilizamos o método (OTSU, 1979) e a segunda o método PSO proposto por (MERWE et al., 2003).

Agrupamento de Otsu

O algoritmo de (OTSU, 1979) assume que a imagem está dividida em duas classes, objeto e fundo, e tenta encontrar o limiar que melhor divide essas classes por meio de múltiplas iterações, calculando a variância entre as classes de objeto e fundo para todos os possíveis limiares. Uma vez definido, todos os pixels da imagem cujos valores de nível de cinza são inferiores ao limiar escolhido, esses serão selecionados para o primeiro cluster e, conseqüentemente, aqueles cujos valores de nível de cinza superarem ou igualam o limiar, são selecionados como pertencentes ao segundo cluster. A (figura 2) mostra um exemplo de imagem mamográfica binarizada pelo método de Otsu.

Figura 2: Imagem mamográfica binarizada: (a) Imagem de entrada; (b) processamento Otsu.



Algoritmo Particle Swarm Optimization

No método proposto por (MERWE et al., 2003), cada partícula do PSO é composta por um vetor de centróides de tamanho N_c , onde N_c é o número de grupos a serem criados, e representa uma solução completa para o problema. A cada iteração, os dados são atribuídos ao centróide ao qual estão mais próximos, as soluções são avaliadas e então atualizadas de acordo com sua melhor posição no espaço de busca e a melhor posição já encontrada por alguma partícula.

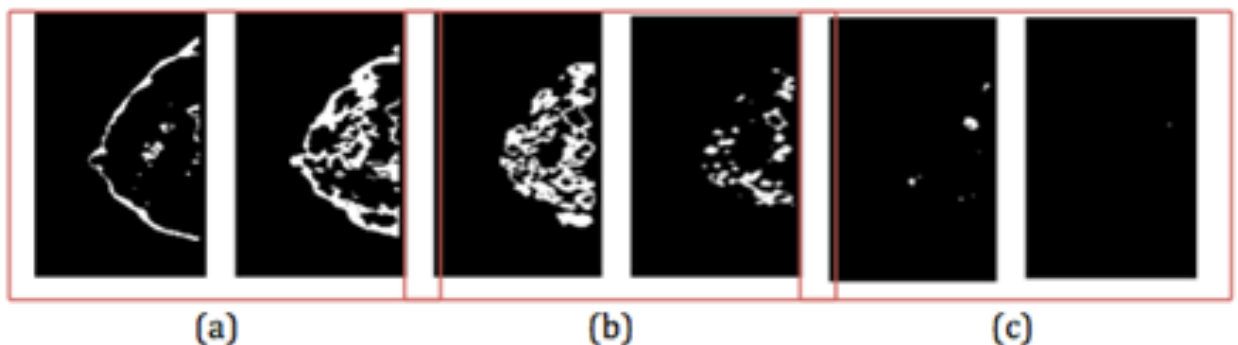
De acordo com (MERWE et al., 2003), cada partícula é tido como um conjunto de clusters, pois só assim, cada partícula representaria uma resposta completa da Clusterização. Em outras palavras, no enxame do PSO, cada partícula possui um vetor de centróides que são construídos da seguinte forma:

$$x_i = (m_{i1}m_{i2}, \dots, m_{ij}m_{iN_c}) \quad [\text{Equação 1}]$$

Onde N_c é o número de clusters a serem criados em m_{ij} corresponde ao j -ésimo centróide da i -ésima partícula. Assim, uma única partícula representa um candidato à solução do problema de Clusterização.

Gerado o vetor inicial X_i da primeira partícula do enxame, através do algoritmo de Otsu, o mesmo é submetido ao PSO para que estes valores sejam otimizados. A figura 3 mostra o resultado dos clusters gerados.

Figura 3: Resultado dos clusters gerados: (a) 10 cluster gerado; (b) 20 cluster gerado a partir do anterior; (c) 30 cluster gerado a partir do segundo.



Ao fim das iterações dos métodos de Otsu e PSO são geradas as regiões suspeitas de lesão. Objetivando analisar individualmente cada região segmentada,

utilizamos a técnica de crescimento de Região de Interesse do Objeto (ROI), seguida de duas técnicas de redução de falsos positivos.

Primeira Redução de Falso Positivo

Objetivando analisar individualmente cada região segmentada, utilizamos a técnica ROI, seguida da técnica filtro de área e do filtro de distância.

Redução com filtro de área

A técnica filtro de área é soma de todos os pixels pertencente a região, caracterizando a área da mesma (JONG-BAE et al., 2003; NETO et al., 2015), a qual foi utilizada nesta pesquisa para localizar as regiões de interesses menores que 300x300 e maiores que 10x10, em relação à imagem original, não são candidatas a tumores as regiões que possuem área fora deste intervalo.

Redução com filtro de distância

O filtro de distância calcula a distância euclidiana do primeiro pixel ao último pixel da região da imagem, objetivando remover as regiões que tiverem distância maior que 60% em relação à imagem original (JONG-BAE et al., 2003; BERBER et al., 2013). O percentual de 60% foi escolhido de forma empírica, pois foi o que apresentou melhores resultados.

As imagens resultantes do processo de segmentação, serão submetidas ao procedimento de redução por altura da imagem, que é a distância euclidiana entre o primeiro ponto (x_1, y_1) até o último ponto (x_2, y_2) da ROI, de acordo com a [Equação 2].

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad \text{[Equação 2]}$$

Redução com Graph Clustering

O Graph Clustering é método de agrupamento dos vértices do grafo em clusters levando em consideração a estrutura das arestas dos grafos (SCHAEFFER, 2007). Neste trabalho, definimos alguns parâmetros: a vizinhança será definida em 3x3, o grafo será formado a partir das ROI de uma imagem, verificando sua vizinhança com todas as ROIs existentes do cluster original, e por fim, o grafo será direcionado (JONG-BAE et al., 2003).

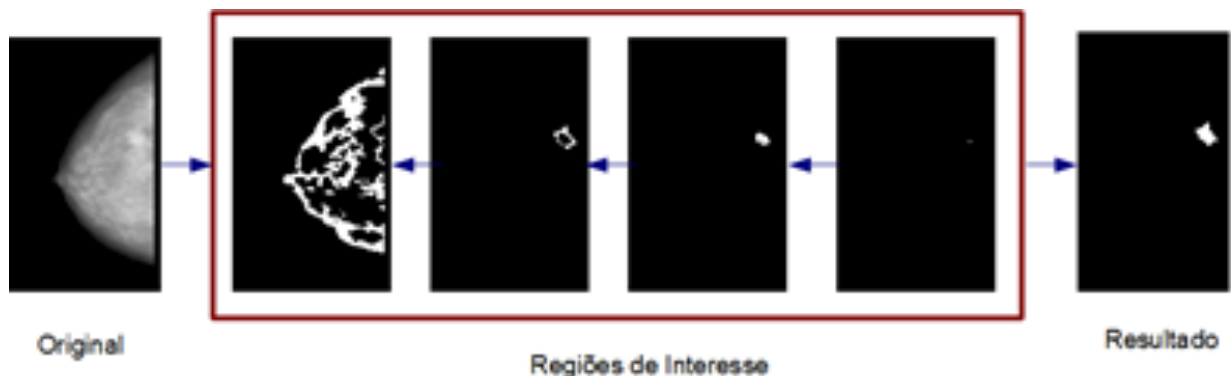
Após o término do processo, os nós do grafo que possuírem mais de duas ligações, serão removidos e todos os nós que tenham nenhuma ou no máximo duas ligações devem permanecer. Cada nó do grafo resultante desse processo, representa uma ROI (Figura 3). A partir de um nó qualquer do grafo, é calculado seu valor de Fator de Forma Circular (FFC) (Equação FFC), da seguinte forma: i) se FFC for menor que 10%: o nó será descartado e será escolhido outro nó, e o processo se repete; ii) se FFC for maior que 10%: verifica-se suas adjacências, realizando as uniões do nós; iii) após cada união, o FFC é calculado novamente, caso este, seja maior que 10% a união é válida e será verificado (caso exista) a próxima adjacência; iv) se o FFC, após cada união, resultar em um valor menor que 10%, esta união não será válida, e o nó que foi unido será descartado.

$$FFC = \frac{4 \cdot \pi \cdot A}{P^2}, \quad [\text{Equação 3}]$$

onde A correspondem a área e P ao perímetro de cada ROI.

O percentual da FFC foi adotado em 10%, em decorrência dos testes, sendo este o valor que apresentou melhores resultados. A Figura 4 ilustra o resultado.

Figura 04: Resultado do algoritmo Graph Clustering.



Segunda Redução de Falso Positivo

Após a primeira redução de falso positivo, aplicamos em cada região que será individualmente caracterizada, os descritores de textura de Haralick para quantificar a textura das imagens, medindo os atributos de suavidade, rugosidade e regularidade (HARALICK et al., 1973).

Extração de Características

Para analisar os padrões de textura das ROIs, para os descritores de Haralick a matriz de coocorrência foi computada verificando os pixels com distância 1 e foram utilizadas nove descritores propostos por (HARALICK et al., 1973): entropia, energia, contraste, homogeneidade, momento diferença inverso, momento diferença de ordem, probabilidade máxima, variância inversa e correlação.

III. EXPERIMENTOS

Uma vez que os atributos relacionados às lesões foram extraídos, os mesmos são utilizados como entrada para um classificador que deverá realizar a tarefa de classificação dos dados em dois tipos de classes (tumores e não tumores). Classificadores como discriminantes lineares e redes neurais artificiais possuem boa performance em classificação de tumores (CHENG, 2006). As ANNs utilizadas possuem uma camada de entrada com 216 atributos, extraídas de 100 nódulos mamários, sendo 50 benignos e 50 malignos e uma camada de saída com duas classes (tumores e não tumores).

Com o conjunto de dados resultantes, o software Weka 3.6 foi utilizado nas classificações com algoritmos de cada técnica de classificação prevista no projeto do software. Os algoritmos utilizados foram o *Random Forests* (RF), o *Multi-Layer Perceptron* (MLP), o *Support Vector Machines* (SVM), o *Radial-Basis Function Networks* (RBF Network) e o *Naive Bayes* (NB). Para cada algoritmo foi aplicado sobre os dados o método de validação cruzada com 10 grupos, objetivando classificar os achados em tumores ou não tumores. Para avaliar os classificadores, foram adotadas as seguintes medidas proporcionais: taxa de acerto, sensibilidade, especificidade e acurácia (ZHANG et al., 2013).

Todos os dados foram normalizados entre -1 e 1 e a técnica SMOTE (CHAWLA, et al., 2002) foi aplicada aos dados originais, na fase de treinamento, para fazer o balanceamento de dados que cria novos indivíduos sinteticamente com o objetivo de melhorar a proporção entre as classes tumores e não tumores obtidas durante o processo de extração das características.

IV. RESULTADO E DISCUSSÃO

Para a realização dos testes foram utilizadas 100 imagens mamográficas selecionadas do banco DDSM, onde 50 são benignas e 50 malignas. Sobre essas imagens são executadas as etapas de pré-processamento, segmentação, primeira e segunda redução de falso positivo e extração das características.

No Quadro 1, mostra o resultado obtido com cada algoritmo utilizado na etapa de classificação. Observando-se que o melhor classificador, no que diz respeito à a sensibilidade, à especificidade e à acurácia, foi o *Random Forest*, atingindo 92.10%, 93.50% e 92.90% de sensibilidade, especificidade e acurácia, respectivamente.

Quadro 1. Etapa de classificação.

Classificador	Sensibilidade %	Especificidade %	Acurácia %
<i>Random Forest</i>	92.10	93.50	92.90
<i>Multi-Layer Perceptron</i>	77.20	79.70	78.60
<i>Support Vector Machines</i>	65.20	76.80	74.50
<i>RBF Network</i>	74.45	75.08	75.09
<i>NaiveBayes</i>	76.05	77.90	76.96

Dentre classificadores de redes neurais utilizados no presente estudo o que mostrou maior porcentagem na análise das proporções amostrais foi o *Random Forest* com uma sensibilidade, especificidade e acurácia de 92.10%, 93.50% e 92.90%, respectivamente. Já os classificadores *Support Vector Machines* e *RBF Network* foram os que apresentaram menor porcentagem nos índices supracitados. Os resultado mencionados podem ser alterados se modificarmos a metodologia utilizada na etapa de pré-processamento para aumentar o contraste das imagens e melhorar o percentual de acerto dos classificadores, pois uma detecção adequada do contorno da mama, em todas as imagens, deve ser alcançada na etapa de segmentação para posterior extração dos atributos das imagens, influenciando no resultado de todas as etapas posteriores e na porcentagens das proporções amostrais dos classificadores utilizados.

V. CONCLUSÃO

Verificamos que metodologia utilizada para detecção de tumores em imagens mamográficas, realizada de forma independente para o método de determinação de ROIs, segmentação de Otsu e PSO, somados a extração de atributos pelos descritores de textura de *Haralick* e a classificação por redes *Random Forest*, resultou numa taxa de classificação superior à 93%. Este estudo permitiu verificar o potencial da rede *Random Forest* para classificação de tumores em mamografia e a vantagem na aplicação do classificador, independente do tipo de mama. Esta rede pode ser otimizada para melhorar seu poder de classificação e acurácia e, assim, estudos futuros são necessários para fazer a classificação automática das massas de acordo com a morfologia definida pelo sistema *Breast Image Reporting and Data System* (BI-RADS).

REFERÊNCIAS

1. ANDERSON, K.N.; SCHWAB, R.B.; MARTINEZ, M.E. Reproductive risk factors and breast cancer subtypes: a review of the literature. **Breast cancer research and treatment**. p. 144:1–10, 2014.
2. AZEVEDO-MARQUES, P.M.; Diagnóstico auxiliado por computador na radiologia. **Radiologia Brasileira**. p. 34 :285–293, 2001.
3. BERBER, T.; ALPKOCAK, A.; BALCI, P.; DICLE, O. Breast mass contour segmentation algorithm in digital mammograms. **Computer Methods and Programs in Biomedicine**. Computer methods and programs in biomedicine p. 110:150–159, 2013. <http://dx.doi.org/10.1016/j.cmpb.2012.11.003>.
4. CALAS, M.J.G.; GUTFILEN, B.; PEREIRA, W.C.A. CAD and mammography: why use this tool?. **Radiol Bras**. p. 45:46-52, 2012. <http://dx.doi.org/10.1590/s0100-39842012000100011>.
5. CHAWLA, N.V.; BOWYER, K.W. HALL, L.O.; KEGELMEYER, W.P. Smote: synthetic minority over-sampling technique. **Journal of artificial intelligence research**. p. 321–357, 2002. http://dx.doi.org/10.1007/3-540-45428-4_5.
6. CHENG, H.D.; SHI, X.J.; MIN, R.H.; CAI, X.P.; DU, H.N. Approaches for automated detection and classification of masses in mammograms. **Pattern Recognition**. v 39, p. 646-668, 2006. <http://dx.doi.org/10.1016/j.patcog.2005.07.006>.

7. FEUER, E.J.; WUN, L.M.; BORING, C.C.; FLANDERS, W.D.; TIMMEL, M.J.; TONG T. **The lifetime risk of developing breast cancer**. Journal of the National Cancer Institute 1993. 85:892-7.
<http://dx.doi.org/10.1093/jnci/85.11.892>.
8. GANESAN, K.; ACHARYA, U.R.; CHUA, C.K.; MIN, LC.; ABRAHAM, K.T. **Computer-aided breast cancer detection using mammograms: A Review**. IEEE Reviews in Biomedical Engineering [Internet]. Institute of Electrical and Electronics Engineers (IEEE); 2013. 6:77-98.
<http://dx.doi.org/10.1109/rbme.2012.2232289>.
9. GIGER, M.L. Computer-Aided Diagnosis of Breast Lesions in Medical Images. **Computing in Science & Engineering**. v. 2, p. 39-45, 2000.
<http://dx.doi.org/10.1109/5992.877391>.
10. HARALICK, R.M.; SHANMUGAM, K.; DINSTEN, H. Textural features for image classification. **Systems, Man and Cybernetics, IEEE Transactions on**. v. 6, p. 610–621, 1973. <http://dx.doi.org/10.1109/tsmc.1973.4309314>.
11. INCA - INSTITUTO NACIONAL DO CÂNCER. Detecção precoce do câncer de mama (2014). Available in:
http://www.inca.gov.br/conteudo_view.asp?id=1932. Accessed 10 July 2016.
12. INCA - INSTITUTO NACIONAL DO CÂNCER (2016). Estimativa do câncer no Brasil, Available in: <http://www.inca.gov.br/estimativa/2016/>. Accessed 10 may 2016.
13. JONG-BAE, K.; HANG-JOON, K. Multiresolution-based watersheds for efficient image segmentation. **Pattern Recognition Letters**. v. 24, p. 473–488, 2003. [http://dx.doi.org/10.1016/s0167-8655\(02\)00270-2](http://dx.doi.org/10.1016/s0167-8655(02)00270-2).
14. LI, J.; GAO, L.; SHEN, Z. Extraction and analysis of digital images feature of three kinds of wheat diseases. **3rd International Congress on Image and Signal Processing (CISP)**. Yantai (China). v.6. p. 2543–8, 2010.
<http://dx.doi.org/10.1109/CISP.2010.5646912>.
15. MERWE, V.D.W.; ANDRIES, P.E. Data clustering using particle swarm optimization. **Proceedings of the Congress on Evolutionary Computation**: Vol. 1, 2003. <http://dx.doi.org/10.1109/cec.2003.1299577>.
16. NETO, O.P.S.; CARVALHO, O.; SAMPAIO, W.; CORREA, A.; PAIVA, A. Automatic segmentation of masses in digital mammograms using particle swarm optimization and graph clustering.; **Institute of Electrical & Electronics Engineers (IEEE) 2**, 2015.
<http://dx.doi.org/10.1109/iwSSIP.2015.7314189>.
17. OTSU, K. A threshold selection method from gray-level histograms. **Automatica**. v.11, p. 23–27, 1979.
<http://dx.doi.org/10.1109/tsmc.1979.4310076>.

18. SAMPAIO, W.B.; DINIZ, E.M.; SILVA, A.C.; DE PAIVA, A.C.; GATTASS, M. Detection of masses in mammogram images using CNN, geostatistic functions and SVM. **Computers in Biology and Medicine**. v41, p. 653-64, 2011. <http://dx.doi.org/10.1016/j.compbimed.2011.05.017>.
19. SCHAEFFER, S.E. Graph clustering. **Computer Science Review**. v. 1, p. 27–64, 2007. <http://dx.doi.org/10.1016/j.cosrev.2007.05.001>.
20. SMITH, R.A.; COKKINIDES, V.; EYRE, H.J. American Cancer Society guidelines for the early detection of cancer. **CA: a cancer journal for clinicians**. v. 53, p. 27-43, 2003. <http://dx.doi.org/10.3322/canjclin.53.1.27>.
21. TAN, M.; QIAN, W.; PU, J.; LIU, H.; ZHENG, B. A new approach to develop computer-aided detection schemes of digital mammograms. **Physics in medicine and biology**. v. 60, p. 4413, 2015. <http://dx.doi.org/10.1088/0031-9155/60/11/4413>.
22. TANG, J.; RANGAYYAN, R.M.; XU, J.; EL NAQA, I.; YANG, Y. Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. **IEEE Transactions on Information Technology in Biomedicine**. v. 13, p. 236-51, 2009. <http://dx.doi.org/10.1109/titb.2008.2009441>.
23. TOURASSI, G.D.; CATARIOUS, D.M.; FLOYD, C. Computer-assisted detection of mammographic masses: A template matching schema based on mutual information. **Medical Physical**. v. 30, p. 2123-2130, 2003. <http://dx.doi.org/10.1118/1.1589494>.
24. ZHANG, M.; XU, C.; YU, J.; WU, Y.; LI, C.; JIN, Z.; LI, Z. Differentiation of Pancreatic Cancer and Chronic Pancreatitis Using Computer-Aided Diagnosis of Endoscopic Ultrasound (EUS) Images: A Diagnostic Test. **Plos One**. v. 8, p. 1-6, 2013. <http://dx.doi.org/10.1371/journal.pone.0063820>.

CAPÍTULO 3

Artigo submetido



ISSN: ISSN 0001-3765 (Print): EISSN: 1678-2690 (Online)

Impact Factor: 1.10

MAPEAMENTO SISTEMÁTICO DE ALGORITMOS UTILIZADOS EM PROCESSAMENTO DE IMAGENS MAMOMOGRÁFICAS

Francisco A. Alves-Ribeiro¹³, Thiara L. B. Da Silva², Wilson de Oliveira Júnior²,
Marcelo Eugenio de Castro-Gonçalves², Vanessa de Oliveira Ribeiro², Raimundo
Santos Moura, Francisco das Chagas Alves-Lima¹, Elmo de Jesus Nery Júnior²,
Benedito B. da Silva¹²

¹Postgraduate Program of the Northeast Network of Biotechnology (RENORBIO),
Northeast, Brazil

² Breast Unit, Getúlio Vargas Hospital, Federal University of Piauí, Teresina, Piauí,
Brazil

³Federal University of Piauí, Teresina, Piauí, Brazil.

RESUMO

A importância da informática aplicada a medicina tem adquirido juntamente com a utilização de técnicas de inteligência artificial em suas aplicações, principalmente no que diz respeito à engenharia do conhecimento em Oncologia, faz-se necessário mapear a evolução desta área de maneira completa, imparcial e sistemática. Para tanto, foi construído um mapeamento sistemático sobre os algoritmos utilizados nas etapas de processamento de imagens mamomográficas, a partir de buscas nas bases de dados *Web Of Science*, *Engineering Village* e *Scopus*, de modo a evidenciar e preencher as necessidades da área. Após a realização de uma seleção rigorosa de três especialistas em processamento de imagens, dentre 199 artigos retornados das bases de dados, 14 foram aceitos, sobre os quais foram realizadas análises consubstanciadas para a construção do mapeamento sistemático proposto.

Palavras-chave: câncer de mama, detecção de massas, imagens mamograficas, extração de texturas, redes neurais.

ABSTRACT

Computer science applied to medicine has acquired an important role, along with the use of artificial intelligence techniques. In particular, Knowledge Engineering in Oncology, requires a more thorough, impartial and systematic mapping of the progress in this area. For this purpose, systematic mapping of algorithms used in stages of mamographic image processing was constructed, starting from a search in the Web of Science, Engineering Village and Scopus databases. Evidence was sought and requirements in the area were met. A strict selection was carried out by three specialists in image processing. Fourteen out of 199 articles retrieved from databases were accepted. Consolidated analyses of the 14 accepted articles were conducted for the construction of the systematic study proposed.

Keywords: breast cancer, mass detection, mammographic images, texture feature extraction, neural networks.

I. INTRODUÇÃO

O mapeamento sistemático é um tipo de verificação da literatura, onde se realiza uma revisão mais ampla dos estudos primários, em busca de identificar quais evidências estão disponíveis, bem como identificar lacunas no conjunto dos estudos primários onde seja direcionado o foco de revisões sistemáticas futuras e identificar áreas onde mais estudos primários precisam ser conduzidos (KITCHENHAM, 2004).

De acordo com Bailey et al (2007), o mapeamento sistemático é um método usado frequentemente nas pesquisas médicas, mas que vem sendo negligenciado em Engenharia de Software. A propósito, de forma resumida, os mapeamentos sistemáticos podem ser entendidos como modalidades de pesquisa que possuem os mesmos passos iniciais e formalidades de uma verificação da literatura, possuindo, no entanto, uma pergunta de pesquisa mais ampla e geral devido seu propósito exploratório (KITCHENHAM; CHARTERS, 2007). O mapeamento sistemático com origem nas pesquisas médicas, com o passar dos anos e a disseminação das revisões literárias no contexto da Engenharia de Software, foi ganhando um papel cada vez maior, sendo destacado por Zhang e Ali Babar (2013) como um estudo de grande crescimento nos últimos anos.

O mapeamento sistemático é geralmente composto de três estágios, sendo o primeiro aquele em que ocorre uma identificação de estudos primários que possam conter resultados relevantes. A seguir, a partir dos estudos identificados, é realizada uma triagem sobre os documentos para encontrar aqueles que sofrerão uma análise mais aprofundada. E, por fim, são realizadas atividades para aferir a qualidade dos estudos selecionados na segunda etapa. Em uma revisão sistemática os estágios descritos seriam acompanhados por fases de extração e análise de dados (KITCHENHAM et al., 2004).

Partindo do pressuposto que uma das tarefas mais complexas em um mapeamento sistemático é a criação de um modelo que represente o conhecimento de uma determinada área de interesse, este trabalho propõe o estudo de técnicas e ferramentas existentes na área de processamento de imagens mamográficas com o uso de redes neurais artificiais.

II. MATERIAIS E MÉTODOS

Atendendo aos procedimentos metodológicos realizamos busca nas bases de dados *Web Of Science*, *Engineering Village* e *Scopus*, de junho a setembro de 2016. O mapeamento sistemático foi realizado nos moldes de um estudo quantitativo, baseado na análise da quantidade de artigos com palavras chaves, abstracts e títulos relacionados ao desenvolvimento de softwares como ferramentas para detecção de tumores em imagens mamográficas com uso de redes neurais artificiais.

A partir da busca nas base de dados, foram retornados 200 artigos relacionados com as *strings mamographic images and neural networks*, entre os anos de 2004 a 2016. Após a realização da pesquisa foi gerado um arquivos no formato *BibTex*, das bases de dados, os quais foram submetidos a ferramenta *TheEnd - Systematic Mapping Tool*, disponível em <http://easii.ufpi.br/theend/home>, para a construção do Mapeamento Sistemático proposto, nas etapas que se segue.

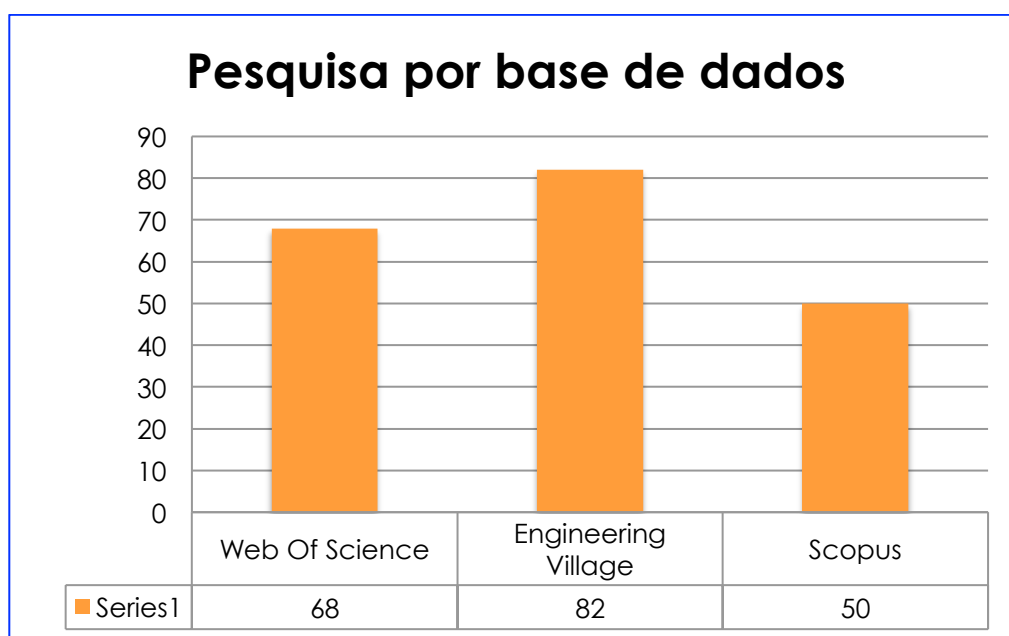
- I. Planejamento
 - a. Objetivo: Identificar algoritmos e técnicas de treinamento aprendizado de máquina, utilizadas nas etapas de processamento de imagens para detecção de tumores em mamográficas.
 - b. Questões da pesquisa
 - i. Quais os algoritmos são mais utilizados para segmentar imagens;
 - ii. Quais os algoritmos são mais utilizados para extrair características de imagens;
 - iii. Quais os classificadores de redes neurais possuem boa performance em classificação de tumores mamários.
- II. *Strings* de busca: A construção da *string* de busca, utilizada nas bibliotecas digitais selecionadas, seguiu a estratégia definida por Kitchenham (2006) que identifica as principais palavras-chaves a partir das perguntas de pesquisa, e utiliza o conector OR para combinar sinônimos e termos alternativos de cada palavra-chave e o conector AND para combinar as palavras-chave. As *strings* de busca utilizadas foram: *mamographic images and neural networks*.
- III. Critérios de inclusão e exclusão: O critério se deu pela sua relevância em relação às questões investigadas.

- a. Inclusão: *Masse detection; mammographic images; texture feature extraction e neural networks.*
- b. Exclusão: *microcalcifications e ultrasonography.*

III. RESULTADOS

Após a busca nas base de dados *Web Of Science, Engineering Village e Scopus*, foram retornados 200 artigos, conforme figura 1.

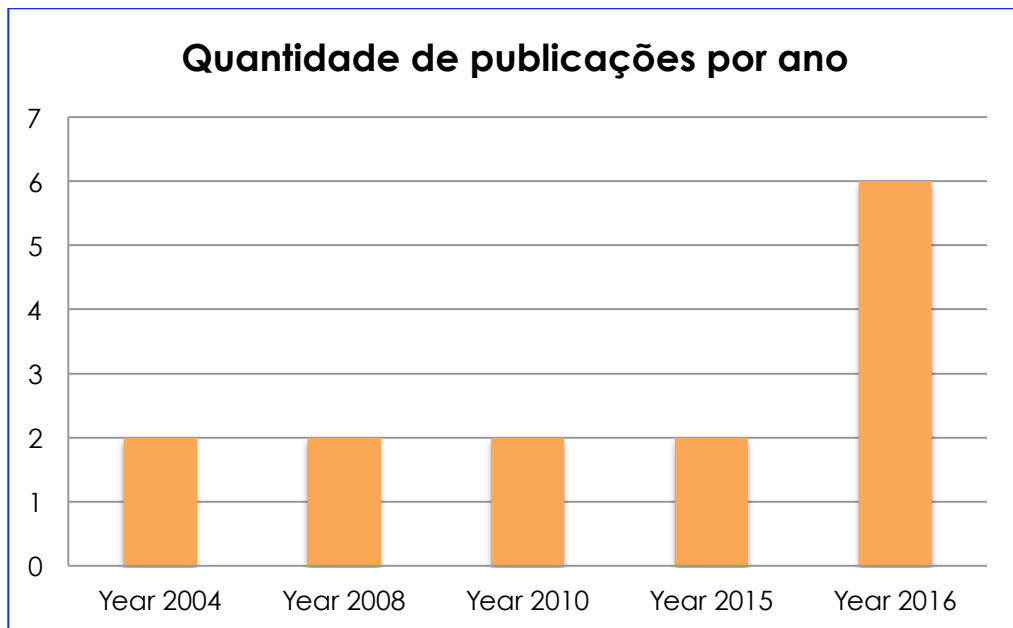
Figura 1. Artigos obtidos por base de busca.



Na etapa de refinamento dos artigos, foi encontrado um arquivo repetido o qual foi excluído do mapeamento, restando 199 artigos a serem avaliados por três especialistas na área de processamento de imagens.

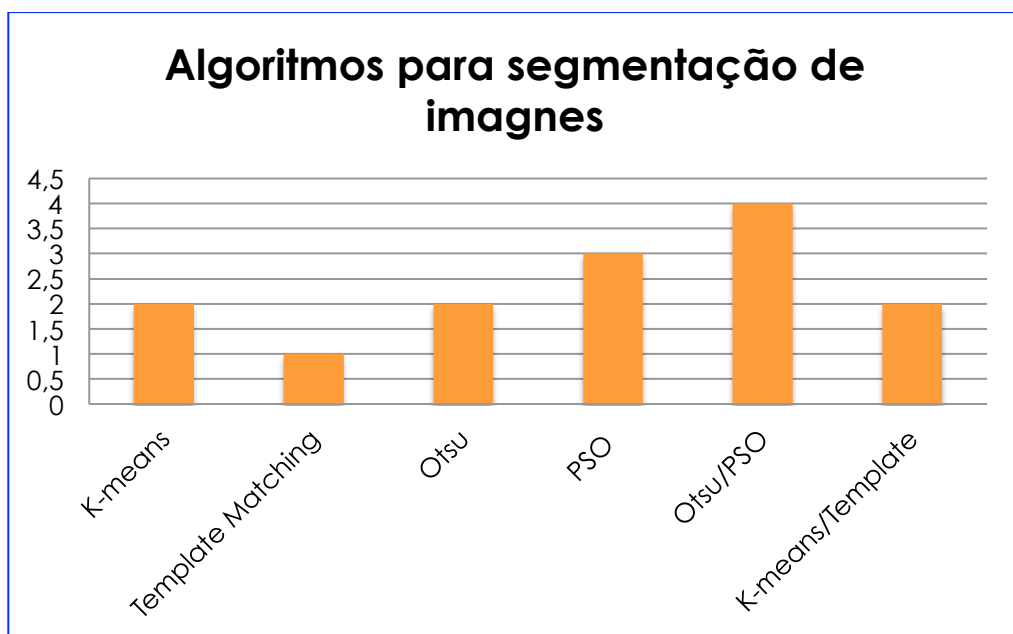
Seguindo os critérios de inclusão e exclusão, os avaliadores fizeram a leitura de todos os abstracts dos artigos, objetivando selecionar conteúdos relevantes ao mapeamento, resultando em 185 artigos rejeitados e 14 aceitos. Após a etapa de avaliação, foi possível perceber que durante os anos de 2004 a 2015 a quantidade de artigos publicados com temas em questão foram poucas em torno de duas publicações nos anos de 2004, 2008, 2010 e 2015, obtendo um aumento significativo no ano de 2016, conforme figura 2.

Figura 2. Publicações por ano.



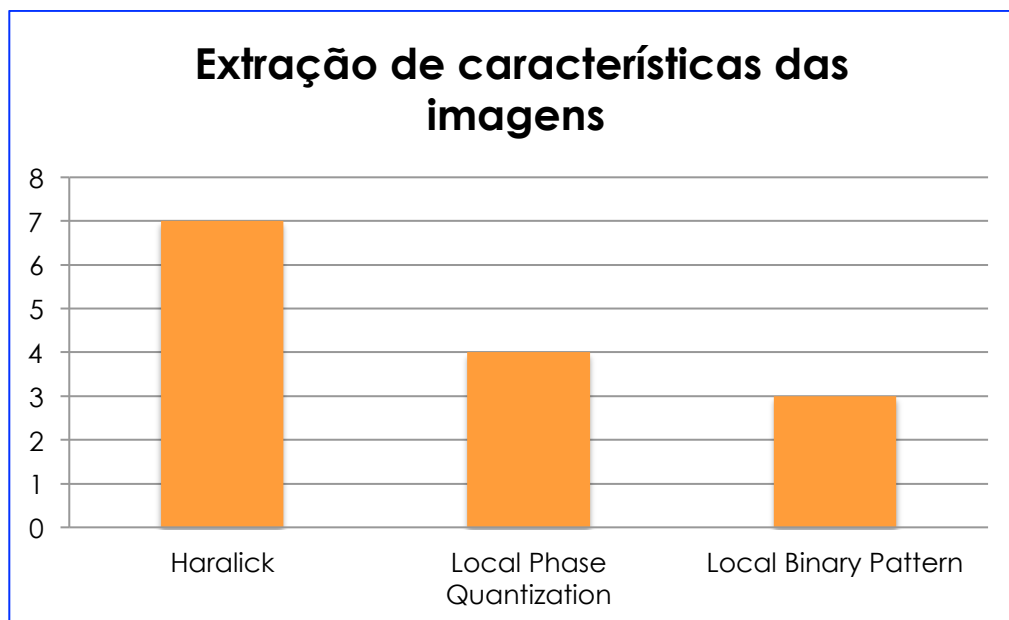
Dentre os 14 artigos aceitos, observamos que os algoritmos de agrupamento *K-means*, a técnica de *Template Matching*, método de Otsu (OTSU, 1979) e o método PSO (*Particle Swarm Optimization*), foram os mais utilizados para segmentar as regiões suspeitas da mama, conforme figura 3.

Figura 3. Algoritmos mais utilizados na segmentação de imagens.



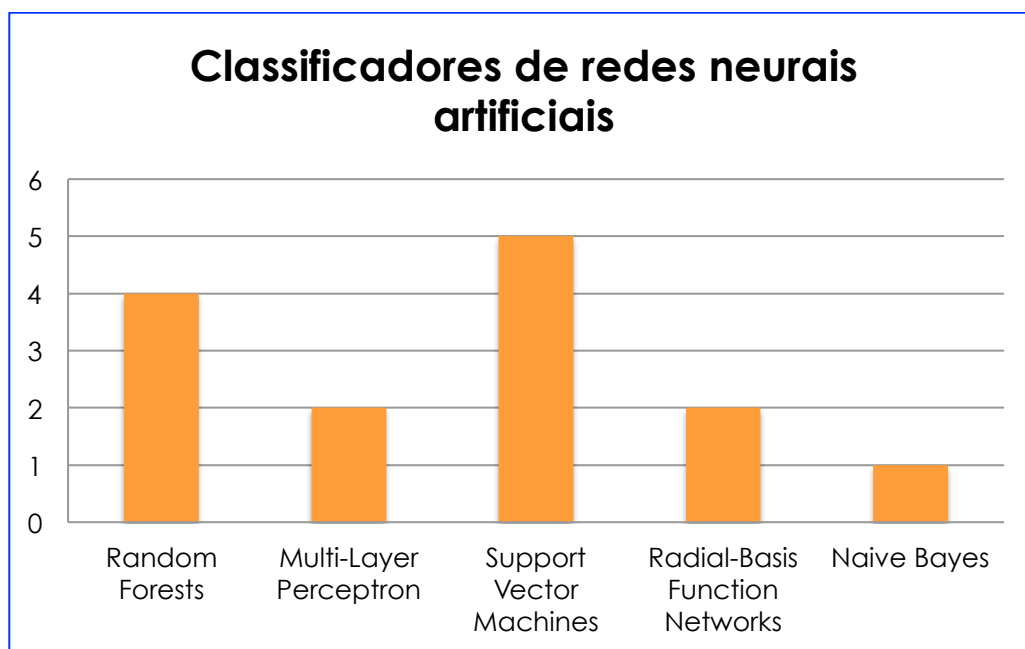
Na etapa de extração de características das imagens obtiveram destaques entre os artigos aceitos os algoritmos de *Haralick*, *Local Phase Quantization* e *Local Binary Pattern*, conforme figura 4.

Figura 4. Algoritmos utilizados na extração de características das imagens.



Segundo os artigos aprovados, os melhores classificadores de redes neurais artificiais que obtiveram melhor performance foram algoritmos de *Random Forests*, *Multi-Layer Perceptron*, *Support Vector Machines*, *Radial-Basis Function Networks* e *Naive Bayes*, destacando-se os algoritmos de *Support Vector Machines* e *Random Forests* com 5 e 4 utilizações, respectivamente, nos artigos aceitos, conforme figura 5.

Figura 5. Classificadores de redes neurais artificiais



IV. CONCLUSÃO

Os resultados desta pesquisa apresentaram um mapeamento sistemático que visou encontrar mecanismos para guiar estudos empíricos em processamento de imagens mamográficas com uso de redes neurais artificiais. No conjunto de 199 artigos analisado, após análise criteriosa de três especialistas em processamento de imagens, 14 artigos foram indicados como estudos relevantes para o mapeamento. A partir dos artigos selecionados foram levantadas hipóteses para se investigar quais são os algoritmos mais utilizados para segmentar, extrair características e obter melhor performance como classificadores em redes neurais artificiais, para o processamento de imagens mamográficas. Os algoritmos Otsu e o PSO destacaram-se na etapa de segmentação, os quais foram utilizados em 4 dos 14 artigos aceitos, na etapa de extração das características obteve destaque o algoritmo de *Haralick* o qual foi utilizado em 7 das 14 publicações e como melhor classificador de redes neurais artificiais o algoritmo *Support Vector Machines* destacou-se em 5 dos 14 artigos aceitos.

REFERÊNCIAS

1. ARKSEY, H. A.; O'MALLEY, L. Scoping studies: towards a methodological framework. **International Journal of Social Research Methodology**. v. 8, p. 19–32, 2005.
2. AUDY, J.; PRIKLADNICKI, R. Desenvolvimento Distribuído de Software: Desenvolvimento de software com equipes distribuídas. Rio de Janeiro: **Elsevier**, 2007.
3. BAILEY, J.; BUDGEN, D.; TURNER, M.; KITCHENHAM, B.; BRERETON, P.; LINKMAN, S.. Evidence relating to object-oriented software design: A survey, in 'Proc. of the 1st Int. Symp. **On Empirical Software Engineering and Measurement (ESEM 2007)**. p. 482–484, 2007.
4. DYBÅ, T.; DINGSØYR, T.; HANSSEN, G.K. Applying Systematic Reviews to Diverse Study Types: An Experience Report. **First International Symposium on Empirical Software Engineering and Measurement, ESEM**, 2007.
5. FREIMUT, B.; PUNTER, T.; BIFFL, S.; E CIOLKOWSKI, M. State-of-the-art in empirical studies authors. Relatório Técnico VISEK/007/E, **Fraunhofer IESE**, 2002.
6. HERBSLEB, J. D.; MOITRA, D. Global Software Development. IEEE Software Magazine, **IEEE Computer Society**, EUA, 2001.
7. KITCHENHAM, B. Empirical paradigm - the role of experiments. **In Proceedings of the 2006 international conference on Empirical software engineering issues: critical assessment and future directions**. Berlin, Heidelberg. Springer-Verlag. p. 25–32, 2006.
8. KITCHENHAM, B. A.; CHARTERS, S. **Guidelines for performing systematic literature reviews in software engineering**. Relatório Técnico **EBSE 2007-001**, Keele University and Durham University Joint Report, 2007.
9. KITCHENHAM, B. A.; DYBA, T.; JORGENSEN, M. Evidence-based software engineering. **In ICSE**. Washington, DC, USA. IEEE Computer Society. v.04, p. 273–281, 2004.
10. MACGREGOR, E.; HSIEH, Y.; KRUCHTEN, P. The impact of intercultural factors on global software development. Proc. **Canadian Conference on Electrical and Computer Engineering**, p. 920- 926, 2005.
11. PETTICREW, M. E ROBERTS, H. Systematic Reviews in the Social Sciences: A Practical Guide. **Wiley-Blackwell**, 2005.
12. PRIKLADNICKI, R. MuNDDoS - Um modelo de referência para desenvolvimento distribuído de software. Dissertação de Mestrado, **PUC/RS**, Porto Alegre, RS, Brasil, 2003.

13. RAMALHO, R. A. S. Representação do conhecimento e ontologias. **reflexões interdisciplinares**, 2009.
14. SENGUPTA, B.; CHANDRA, S.; SINHA, V. Test-driven global software development. In **Proceedings of The 3rd International Workshop on Global Software Development**, co-located with ICSE 2004.
15. VAVASSORI, F. B. Metodologia para o Gerenciamento Distribuído de Projetos e Métricas de Software. Dissertação de Mestrado, **UFSC**, Florianópolis, SC, Brasil, 2002.
16. WOHLIN, C.; RUNESON, P.; HÖST, M.; OHLSSON, M.C.; REGNELL, B.; WESSLÉN, A. Experimentation in software engineering: an introduction. **Kluwer Academic Publishers**, Norwell, MA, USA, 2000.
17. ZHANG, J.; ALI BABAR, M. Systematic reviews in software engineering: An empirical investigation. *Information and Software Technology*, **Elsevier BV**, v. 55, n. 7, 2013.

CAPÍTULO 4

REGISTRO DE PROGRAMA DE COMPUTADOR



INPI INSTITUTO
NACIONAL
DA PROPRIEDADE
INDUSTRIAL

REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA E COMÉRCIO EXTERIOR
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
DIRETORIA DE CONTRATOS, INDICAÇÕES GEOGRÁFICAS E REGISTROS

CERTIFICADO DE REGISTRO DE PROGRAMA DE COMPUTADOR

Processo: BR 51 2015 001494-3

O INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL expede o presente Certificado de Registro de Programa de Computador, **válido por 50 anos** a partir de 1º de janeiro subsequente à data de criação indicada, em conformidade com o art. 3º da Lei Nº 9.609, de 19 de Fevereiro de 1998, e arts. 1º e 2º do Decreto 2.556 de 20 de Abril de 1998.

Título: **MEDMAMO**
Criação: 01 de outubro de 2015
Titular(es): UNIVERSIDADE FEDERAL DO PIAUÍ (06.517.387/0001-34)
Autor(es): BENEDITO BORGES DA SILVA (027.232.633-04)
FRANCISCO ADELTON ALVES RIBEIRO (092.183.428-43)
Linguagem: C++
Aplicação: SD-05, SD-06, SD-08, SD-09
Tipo Prog.: GI-08, IA-01, TC-04

DOCUMENTAÇÃO TÉCNICA EM DEPÓSITO SOB SIGILO ATÉ 25/11/2025.

A exclusividade de comercialização deste programa de computador não tem a abrangência relativa à exclusividade de fornecimento estatuida pelo art.25, I, da Lei nº8.666, de 21 de Junho de 1993, para fins de inexigibilidade de licitação para compras pelo poder público.

Expedido em 26 de julho de 2016

Assinado digitalmente por:

Breno Bello de Almeida Neves
Diretor de Contratos, Indicações Geográficas e Registros

CONSIDERAÇÕES FINAIS

Concluimos este trabalho, mostrando que a metodologia proposta atingiu seu objetivo de otimizar imagens mamográficas para auxílio ao diagnóstico médico, usando técnicas de treinamento e aprendizado de máquina para detectar massas em imagens mamográficas. O primeiro capítulo fez a prospecção sobre os softwares para processamento de imagem, buscando indícios do estado atual da técnica via cenário de patentes depositadas, no Brasil e no mundo, relacionadas softwares de diagnóstico e processamento de imagens mamográficas, demonstrando a necessidade de desenvolvimento do software devido a pequena quantidade de patentes de software para processamento de imagens mamográficas. O segundo capítulo mostrou que as técnicas de aprendizado de máquinas utilizadas para desenvolvimento do software foram eficazes, destacando-se os métodos de Otsu, PSO, os descritores de textura de *Haralick* e o algoritmo *Random Forest*, que foi o melhor classificador de redes neurais artificiais com uma taxa de classificação superior à 92%. O terceiro capítulo apresentou um mapeamento sistemático sobre algoritmos utilizados nas etapas de processamento de imagens mamográficas. Destacando-se na etapa de segmentação os algoritmos Otsu e o PSO, na etapa de classificação o algoritmo *Support Vector Machines* com o melhor classificador de aprendizagem supervisionada, seguido do algoritmo *Random Forest*, o que permitiu verificar que o software foi desenvolvido com os algoritmos mais utilizados atualmente nos programas de processamento de imagens mamográficas. Finalizando o quarto capítulo com o registro do software MedMamo. Acreditamos melhorar o poder de classificação e acurácia do software, testando outros classificadores de redes neurais, havendo necessidade de futuros estudos para fazer a classificação automática das massas de acordo com a morfologia definida pelo sistema BI-RADS.

ANEXOS

ANEXO A - Aprovação pelo Comitê de Ética e Pesquisa



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Otimização de imagens mamográficas para auxílio ao diagnóstico médico

Pesquisador: benedito borges da silva

Área Temática:

Versão: 3

CAAE: 40474114.8.0000.5214

Instituição Proponente: FUNDAÇÃO UNIVERSIDADE FEDERAL DO PIAUÍ

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 1.055.235

Data da Relatoria: 29/05/2015

Apresentação do Projeto:

O projeto apresenta uma proposta de pesquisa de Mestrado intitulada: "Sistema Computacional de otimização da imagem mamográfica para prever o diagnóstico de tumores da mama". Justifica a relevância da investigação devido a escassez de estudos na literatura utilizando a segmentação em textura, que é uma nova técnica que consiste em dividir a imagem em vários segmentos, uma mamografia, por exemplo, em várias regiões, particionadas por textura, inspirada na operação multicanal do Sistema Visual

Humano (HVS – Human Visual System), o que nos levou à concepção do presente estudo.

Objetivo da Pesquisa:

Objetivo Primário:

Desenvolver sistema computadorizado para processamento de imagens, baseado na segmentação em textura, para auxiliar o diagnóstico de tumores da mama, utilizando redes neurais artificiais.

Objetivo Secundário:

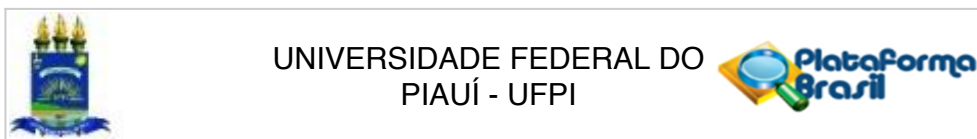
Avaliar o sistema proposto através da comparação de seus resultados com o diagnóstico convencional e visualização gráfica.

Avaliação dos Riscos e Benefícios:

Riscos:

Como durante a pesquisa será aplicado um questionário, este poderá ocasionar um risco de

Endereço: Campus Universitário Ministro Petronio Portella - Pró-Reitoria de Pesquisa
Bairro: Ininga **CEP:** 64.049-550
UF: PI **Município:** TERESINA
Telefone: (86)3237-2332 **Fax:** (86)3237-2332 **E-mail:** cep.ufpi@ufpi.edu.br



Continuação do Parecer: 1.055.235

constrangimento, mas o pesquisador responsável, adotará procedimentos éticos no decorrer dessa aplicação conforme Resolução 466/2012 para evitar esse risco. Ao participar da pesquisa, o voluntário não sofrerá nenhum prejuízo. Não há riscos diretos aos participantes.

Benefícios:

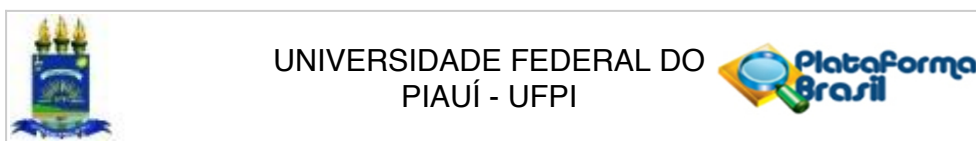
As participantes do estudo terão como benefício o uso de sistema computacional para apoio ao diagnóstico médico; detecção precoce do câncer de mama e a redução do risco de radiação.

Comentários e Considerações sobre a Pesquisa:

O estudo apresenta uma metodologia computacional para detecção e classificação de nódulos em imagens digitalizadas de mamografias auxiliando o especialista em sua análise. O método proposto utiliza a técnica de Análise de Componentes Independentes – ICA e algoritmo de cluster “K-means” para identificar as regiões suspeitas. Em seguida, medidas de geometria e textura são extraídas de cada uma dessas regiões, para análise e classificação, a textura será descrita através do algoritmo “Forward selection”,

caracterizando o tecido normal e com blastoma benigno ou maligno, utilizando redes Neurais. Define a amostra de 10 participantes. A proposta do projeto está dividida em três etapas: 1 - aquisição da mamografia digital: As imagens serão adquiridas em laboratórios, no formato digital, de pacientes do Hospital Getúlio Vargas, setor de Ginecologia, submetidas a exames mamográficos. Serão incluídas apenas imagens cuja película não possua artefatos e excluídas imagens com artefatos que prejudiquem o diagnóstico médico. Após seu melhoramento será feito estudo comparativo com a imagem anterior para verificar se existe ganho na sensibilidade diagnóstica. 2 - Segmentação das Regiões Suspeitas (RI): A segmentação será realizada da seguinte forma: primeiramente será realizada uma filtragem multi-canal na mamografia, utilizando o banco de filtros bionspirados do HVS com a técnica de análise de componentes independentes. Após a filtragem, será gerada uma imagem multibanda, que será clusterizada através do algoritmo de segmentação K-means. A imagem segmentada vai apontar regiões suspeitas de câncer, e essas regiões suspeitas serão analisadas e classificadas pelos métodos a serem desenvolvidos. 3 - Codificação do software: Será utilizada a linguagem de programação JAVA para codificação do software, o qual poderá ser utilizado em máquinas com configuração simples, sem a necessidade de um software robusto, ou uma super máquina. Análise de Componentes Independentes - ICA: Devemos assumir que uma imagem pode ser expressa como uma combinação linear de imagens bases a_1, a_2, \dots, a_n ponderadas por coeficientes mutua e estatisticamente independentes entre si s_1, s_2, \dots, s_n , conforme a equação: $x_n = a_1s_1 + a_2s_2 + \dots +$

Endereço: Campus Universitário Ministro Petronio Portella - Pró-Reitoria de Pesquisa
Bairro: Ininga **CEP:** 64.049-550
UF: PI **Município:** TERESINA
Telefone: (86)3237-2332 **Fax:** (86)3237-2332 **E-mail:** cep.ufpi@ufpi.edu.br



Continuação do Parecer: 1.055.235

ansn (1) Sendo: x_n = Sinal aleatório; s_n = Componente independente aleatório; a_n = Coeficiente de mistura. Utilizando notação matricial, podemos reescrever esta equação da seguinte forma: $X=A.S$ (2) O modelo apresentado na equação (2) é chamado de Análise de Componentes Independentes - ICA (HYVÄRINEN, OJA, 2000), que descreve como os dados são gerados a partir do processo de mistura com as componentes independentes.

Considerações sobre os Termos de apresentação obrigatória:

A proposta apresenta os componentes básicos exigidos por uma pesquisa acadêmica, referencial teórico que dará sustentação ao estudo, bem como os aspectos éticos do estudo, cronograma e orçamento afirmando ser financiada com recursos próprios. Os objetivos estão coerentes com a proposta de estudo. O coordenador é docente da UFPI com experiência na temática evidenciada e se compromete cumprir os termos da Resolução CNS nº 466/12 - e zelar pela privacidade e confidencialidade dos dados.

Recomendações:

Sem recomendações

Conclusões ou Pendências e Lista de Inadequações:

Cumprida as pendências recomendo sua aprovação

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Considerações Finais a critério do CEP:

TERESINA, 08 de Maio de 2015

Assinado por:
Adrianna de Alencar Setubal Santos
(Coordenador)

Endereço: Campus Universitário Ministro Petronio Portella - Pró-Reitoria de Pesquisa
Bairro: Ininga **CEP:** 64.049-550
UF: PI **Município:** TERESINA
Telefone: (86)3237-2332 **Fax:** (86)3237-2332 **E-mail:** cep.ufpi@ufpi.edu.br

ANEXO B – Notificação de Aceite – Software Surver for Breast Image Processing



**International Journal of Electronics Communication
and Computer Engineering (IJECCCE)**

ISSN (Online Journal) : 2249 - 071X, ISSN (Print): 2278 – 4209

www.ijeccce.org

Notification of Acceptance

Date: September 26, 2016

Dear **Francisco Adeltton Alves-Ribeiro, Miguel de Sousa Freitas, Thiara Lorena Bezerra da Silva, Wilson de Oliveira Júnior, Marcelo Eugenio de Castro-Gonçalves, Vanessa de Oliveira Ribeiro, Benedito Borges da Silva**

Paper ID : IJECCCE-3973

Paper Title: **SOFTWARE SURVEY FOR BREAST IMAGE PROCESSING**

The review process for International Journal of Electronics Communication and Computer Engineering (IJECCCE) has been completed. Based on the recommendations of the reviewers and the Technical Program Committee, we are pleased to inform you that your paper identified above has been accepted for online publication. You have to **update your paper as per reviewer's comments**, so that it can be published. You are requested to complete the formalities. This notification email serves as our formal acceptance of your paper as well as consideration for online publication provided you complete the formalities and update the paper as per reviewer's comments.

(Important) So in order to publish your paper successfully, you must finish following steps.

1. **Send your final paper (updated according to reviewers comments)** in .pdf and .doc format. (Format your paper according to the [Rules](#), [Template carefully](#)), *mail it to* submit2ijeccce@yahoo.in, submit@ijeccce.org
2. **Mail signed and completely filled copyright** (in .jpg format), at submit2ijeccce@yahoo.in, submit@ijeccce.org
3. **Finish the payment of Registration fee at the Bank or online**, scan "payment slip" in jpg format and mail it to submit2ijeccce@yahoo.in, submit@ijeccce.org
4. **Send your filled scanned payment form** (in jpg format) to us at: submit2ijeccce@yahoo.in, submit@ijeccce.org

You can find all the required formats and details from [Download](#).

If the above requirements are met by 04/10/2016, the paper will be published in the IJECCCE.

Finally, we would like to further extend our congratulations to you.

Feel free to contact us for any query...

With regards,

Editor in Chief,

www.ijeccce.org

IJECCCE Copyright © 2016. All Rights Reserved

ANEXO C - Certificado de Publicação de Artigo



ANEXO D - Certificado de Apresentação

CERTIFICATE

We hereby certify that the work entitled **SOFTWARE SURVEY FOR BREAST IMAGE PROCESSING** authored by *Francisco Adelson Alves-Ribeiro, Miguel de Sousa Freitas, Benedito Borges da Silva, Francisco das Chagas Alves-Lima, Carla Solange Escórcio-Dourado, Fabiane Araújo Sampaio, Luana Mota Martins,* was presented during the poster session of the "X-Meeting 2015 - 11th International Conference of th AB3C + Brazilian Symposium of Bioinformatics" held in Sao Paulo - Brazil from November 3 to 6 2015.

Glória Regina Franco

Glória R Franco - AB3C President

Mainá Bitar

Mainá Bitar - Poster Chair

Nicole Scherer

Nicole Scherer - Poster Chair

 **X - meeting + BSB 2015**
USP - SÃO PAULO - SP - BRAZIL | NOVEMBER 03 to 06 2015
11th International Conference of the AB3C + Brazilian Symposium of Bioinformatics



ANEXO E – Notificação de Aceite – Masse detection in mammographic images using texture feature extraction and neural networks

gsw.schommers@t-online.de 

Para: adelton@me.com

Responder A: gsw.schommers@t-online.de

WG: WG: Journal of Computational and Theoretical Nanoscience (CTN)

Journal of Computational and Theoretical Nanoscience (CTN)

Dear authors,

this is to acknowledge receipt of the article

Masse detection in mammographic images using texture feature extraction and neural networks
3631

for publication in the **Journal of Computational and Theoretical Nanoscience (CTN)**.

The paper has been refereed by two of our experts; the judges are positive.

The acceptance of your paper becomes automatically effective after your Institute has transferred the publication fee of 980 US dollars.

The publisher will send an invoice soon.

There are some minor formal points which have to be changed and added, respectively:

- we need your signed **Copyright Transfer Agreement**. Please fax the attached form directly to W. Schommers (+49 7247 82 4972), or scan the signed document and send it by E-mail to the Editor-in-Chief.


ANEXO F – Submissão – Systematic Mapping of Algorithms Used in Mammographic Image Processing

Annals of the Brazilian Academy of Sciences - Manuscript ID AABC-2017-0003



Entrada x

 **Anais da Academia Brasileira de Ciências** <onbehalfof+dlopes+abc.org.br@m: para beneditoborges, mim, fdcalima, danylorafhael, vanessaribeir0, beneditoborges ▾ 13:37 (Há 20 horas) ☆  

 inglês ▾ > português ▾ Traduzir mensagem Desativar para: inglês x

03-Jan-2017

Dear Prof. da Silva:

Your manuscript entitled "Systematic Mapping of Algorithms Used in Mammographic Image Processing" has been successfully submitted online and is presently being given full consideration for publication in the Annals of the Brazilian Academy of Sciences.

Your manuscript ID is AABC-2017-0003.

Please mention the above manuscript ID in all future correspondence or when calling the office for questions. If there are any changes in your street address or e-mail address, please log in to ScholarOne Manuscripts at <https://mc04.manuscriptcentral.com/aabc-scielo> and edit your user information as appropriate.

You can also view the status of your manuscript at any time by checking your Author Center after logging in to <https://mc04.manuscriptcentral.com/aabc-scielo>.

Thank you for submitting your manuscript to the Annals of the Brazilian Academy of Sciences.

Sincerely,
Annals of the Brazilian Academy of Sciences Editorial Office

ANEXO G - Publicação de Concessão do Registro de Software

INPI INSTITUTO
NACIONAL DA
PROPRIEDADE
INDUSTRIAL

Revista da Propriedade Industrial

Seção I

**Nº 2377
26 de Julho de 2016**

**Patentes
Desenhos Industriais
Contratos de Tecnologia
Programas de Computador
Indicações Geográficas
Topografias de Circuitos
Integrados**



RPI 2377 de 26/07/2016

DICIG - Diretoria de Contratos, Indicações Geográficas e Registros 377

<p>Procurador: Não informado ou inexistente</p> <p>Processo: BR 51 2015 001356-4 120 Título: SOFTWARE BUSCA CLIENTES Titular: JOSE PAULO PEREIRA SILVA Criador: DAVID ANDERSON DA SILVA ROCHA; OSMAR DE SOUZA QUEIROZ Linguagem: PHP Campo de Aplicação: AD-10; CO-04; IF-01; IF-02; IF-04 Tipo de Programa: FA-02; GI-01; GI-06; LG-01; LG-02 Data da Criação: 01/01/2014 Regime de Guarda: SIGILO ATÉ 17/11/2025 Procurador: Não informado ou inexistente</p> <p>Processo: BR 51 2015 001358-0 120 Título: PLATAFORMA BOADFACTOR Titular: BROADFACTOR SISTEMAS LTDA EPP Criador: RICARDO MARTINS CURY Linguagem: AMAZON AWS; ASP.NET; CSS 2.0; CSS3; HTML 5.0; MICROSOFT.NET; SAP.NET MVC; SQL - SERVER Campo de Aplicação: AD-05; FN-03; FN-05; FN-06 Tipo de Programa: GI-01; GI-03; GI-06; SO-02; SO-05; SO-07 Data da Criação: 14/01/2015 Regime de Guarda: SIGILO ATÉ 17/11/2025 Procurador: A PROVINCIA MARCAS E PATENTES LTDA</p> <p>Processo: BR 51 2015 001359-9 120 Título: COMIN STANDARD CONTATOS Titular: PEDRO ANTONIO ROSSATO COMIN Criador: PEDRO ANTONIO ROSSATO COMIN Linguagem: C++2010; COMIN XASP SYSTEMS; VISUAL Campo de Aplicação: AD-11 Tipo de Programa: AT-02 Data da Criação: 02/01/2015 Regime de Guarda: SIGILO ATÉ 18/11/2025 Procurador: Não informado ou inexistente</p> <p>Processo: BR 51 2015 001371-8 120 Título: CPQD3222. MODELO DE CIRCUITO SINTETIZÁVEL EM FPGA PARA CONTROLE DE FLUXO E MAPEAMENTO DOS CANAIS FÍSICOS DO DOWNLINK DO SISTEMA LTE RELEASE 8-V.1.0 Titular: FUNDAÇÃO CPQD - CENTRO DE PESQUISA E DESENVOLVIMENTO EM TELECOMUNICAÇÕES Criador: ANDERSON FERNANDO TORQUATO; FABBRYCCIO AKKAZZHA CHAVES MACHADO CARDOSO; FABIANO SILVA MATHILDE; FELIPE AUGUSTO PEREIRA DE FIGUEIREDO; ORLANDO DOS REIS PEREIRA; RAFAEL MENDES VILELA Linguagem: VHDL Campo de Aplicação: TC-03 Tipo de Programa: SM-01 Data da Criação: 06/10/2012 Regime de Guarda: SIGILO ATÉ 18/11/2025 Procurador: ANA LÚCIA FORNI POPPI</p> <p>Processo: BR 51 2015 001401-3 120 Título: SACEE - SISTEMA DE ANÁLISE DE CONTAS DE ENERGIA ELÉTRICA Titular: UNIVERSIDADE FEDERAL DE SANTA MARIA Criador: MÁRCIO DALCUL DEPEXE; MATIAS AMÉRICO BORTOLUZZI; RENATO MACHADO; THIAGO CATTANI NAIDON Linguagem: MATLAB Campo de Aplicação: EC-09; EN-04 Tipo de Programa: DS-05; FA-01</p>	<p>Data da Criação: 01/07/2014 Regime de Guarda: SIGILO ATÉ 20/11/2025 Procurador: Não informado ou inexistente</p> <p>Processo: BR 51 2015 001403-0 120 Título: AGAPES Titular: FUNDAÇÃO UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL; PSG TECNOLOGIA APLICADA LTDA Criador: EDSON TAKASHI MATSUBARA; JOÃO ALEXANDRE VENTORINI Linguagem: PYTHON Campo de Aplicação: AG-01 Tipo de Programa: TC-04 Data da Criação: 01/03/2015 Regime de Guarda: SIGILO ATÉ 10/11/2025 Procurador: Não informado ou inexistente</p> <p>Processo: BR 51 2015 001406-4 120 Título: GPS GEOLÓGICO GEOTEC Titular: UNIVERSIDADE FEDERAL DO PARANÁ Criador: ALBERTO PIO FIORI; CLAUDINEI TABORDA DA SILVEIRA; CYNTHIA ROBERTI LIMA; PAULA ROBERTA APARECIDA SILVA; RAFAEL ROMUALDO WANDRESEN; TIAGO GEANEZINI Linguagem: ANDROID; JAVA Campo de Aplicação: CO-02; GL-01; GL-03; GL-05; GL-06 Tipo de Programa: AP-01; FA-04; GI-06; SO-02 Data da Criação: 12/08/2015 Regime de Guarda: SIGILO ATÉ 23/11/2025 Procurador: Não informado ou inexistente</p> <p>Processo: BR 51 2015 001410-2 120 Título: SISDAGRI Titular: PETRÓLEO BRASILEIRO S/A - PETROBRAS Criador: RAPHAEL RIEMKE DE CAMPOS CESAR LEÃO Linguagem: JAVA; JAVASCRIPT; PHP Campo de Aplicação: AD-05; AD-06; AG-01; AG-05; IF-02 Tipo de Programa: GI-01; GI-02; GI-03; GI-04; SO-02 Data da Criação: 01/12/2009 Regime de Guarda: SIGILO ATÉ 25/11/2025 Procurador: Não informado ou inexistente</p> <p>Processo: BR 51 2015 001432-3 120 Título: HOTWELDWISE@ Titular: SERVIÇO NACIONAL DE APRENDIZAGEM INDUSTRIAL- SENAI/DR-BAHIA Criador: ALEX ALISSON BANDEIRA SANTOS; IVO ANDREI DE OLIVEIRA LINO Linguagem: APDL Campo de Aplicação: IN-01; IN-03 Tipo de Programa: IA-02; SM-01; SM-04 Data da Criação: 01/10/2014 Regime de Guarda: SIGILO ATÉ 17/11/2025 Procurador: Não informado ou inexistente</p> <p>Processo: BR 51 2015 001436-6 120 Título: OLART - SMD0-ONE-LAYER ANISOTROPIC RADIATIVE TRNASFER - SINGLE MESH Titular: UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO Criador: ANTÔNIO JOSÉ DA SILVA NETO; DIEGO CAMPOS KNUPP; JOÃO VITOR MUSSEL CANATO; LEONARDO TAVARES STUTZ Linguagem: FORTRAN; MATEMÁTICA Campo de Aplicação: FQ-07; IN-03 Tipo de Programa: SM-01; TC-01</p>	<p>Data da Criação: 24/08/2015 Regime de Guarda: SIGILO ATÉ 27/11/2025 Procurador: JOSÉ CARLOS VAZ E DIAS</p> <p>Processo: BR 51 2015 001471-4 120 Título: SIGAJ - SISTEMA DE GESTÃO DE AÇÕES JUDICIAIS Titular: DELPHOS SERVIÇOS TÉCNICOS S/A. Criador: LEANDRO KLING Linguagem: DEVELOPER; ORACLE; ORACLE 9i; PLSQL Campo de Aplicação: IF-02; IF-04; IF-08 Tipo de Programa: AP-03; AT-02; AT-06; GI-01; GI-06 Data da Criação: 04/01/2010 Regime de Guarda: SIGILO ATÉ 02/12/2025 Procurador: JOÃO LUIZ DE MAGALHÃES CASTRO</p> <p>Processo: BR 51 2015 001483-8 120 Título: SIMULADOR DE GUINDASTE SIMCRANE Titular: FUNDAÇÃO PARQUE TECNOLÓGICO DE SÃO CARLOS Criador: JONAS DE CARVALHO; SYLVIO GOULART ROSA JUNIOR Linguagem: C#; UNITY 3D Campo de Aplicação: IN-02; IN-03; TP-04 Tipo de Programa: AT-01; ET-04; SM-02; SO-01; SO-09 Data da Criação: 18/12/2014 Regime de Guarda: SIGILO ATÉ 27/11/2025 Procurador: ROGERIA MARIA DA SILVA MHIRDAUI</p> <p>Processo: BR 51 2015 001491-9 120 Título: XC - ERP Titular: XC LTDA - ME Criador: ELISIANA PODGORSKI; KARLA DOS SANTOS LENCINA; TIAGO SIQUEIRA BRUM Linguagem: ANGULAR; CSS; HTML; JAVA; PHP Campo de Aplicação: AD-05; AD-11; EC-05; FN-05; FN-06 Tipo de Programa: AP-01; AP-02; AP-03; AP-05; DS-04 Data da Criação: 20/01/2014 Regime de Guarda: SIGILO ATÉ 08/12/2025 Procurador: MARCA BRAZIL MARCAS & PATENTES</p> <p>Processo: BR 51 2015 001494-3 120 Título: MEDMAMO Titular: UNIVERSIDADE FEDERAL DO PIAUÍ Criador: BENEDITO BORGES DA SILVA; FRANCISCO ADELTON ALVES RIBEIRO Linguagem: C++ Campo de Aplicação: SD-05; SD-06; SD-08; SD-09 Tipo de Programa: GI-08; IA-01; TC-04 Data da Criação: 01/10/2015 Regime de Guarda: SIGILO ATÉ 25/11/2025 Procurador: Não informado ou inexistente</p>
--	--	--

APÊNDICES

APÊNDICE A - Ficha de entrevista para coleta de dados com as pacientes

Questionário sobre a História Familiar

Entrevista

Data da entrevista: ____/____/____ Horário de início: ____:____

O entrevistado é o próprio indivíduo?

1. Sim ()
2. Não () especifique grau de relacionamento _____

Entrevistador: _____

Identificação

Iniciais do indivíduo: _____ Número: _____

Número do CPF: _____._____._____ - _____ RG: _____

Endereço: _____

Bairro: _____ CEP: _____

Cidade: _____ UF : _____

Tel. residencial: _____ Celular: _____

Dados do Paciente

1. Qual é a sua data de nascimento? ____/____/____

2. Qual a sua idade? ____ anos

3. Sexo:

1. () Masculino
2. () Feminino

4. O(a) Sr.(a) nasceu no Brasil?

1. () Sim **passé 6**
2. () Não

5. Em que país o(a) Sr.(a) nasceu? _____ **Passé 8**

6. Em que Estado o(a) Sr.(a) nasceu? _____

7. Em que cidade o(a) Sr.(a) nasceu? _____

8. Qual é o seu grau de instrução?

Entrevistador: Caso tenha pós-graduação assinale superior completo.

1. () Analfabeto
2. () Educação infantil
3. () Ensino fundamental incompleto
4. () Ensino fundamental completo
5. () Técnico / ensino médio incompleto
6. () Técnico / ensino médio completo

7. () Superior incompleto
8. () Superior completo
8. () Não sabe
9. Estado Civil
1. () Casado(a) / união consensual
 2. () Separado(a) / divorciado(a) / desquitado(a)
 3. () Solteiro(a)
 4. () Viúvo(a)
 5. () Ignorado
10. Por quem o(a) Sr.(a) foi encaminhado(a) para o aconselhamento genético?
1. () Profissional de saúde deste Centro
 2. () Profissional de saúde de outro Centro
 3. () Veio por conta própria (autoreferido)
 4. () Outro (especifique _____)
11. O(a) Sr.(a) tem ou já teve algum diagnóstico de câncer ou tumor maligno dado pelo médico?
1. () Sim
 2. () Não *passa 27*
 3. () Não sabe *passa 27*
12. Quantos diagnósticos de câncer ou tumor maligno dados pelo médico o(a) Sr.(a) já teve?
() Não sabe *passa 27*
13. Em que parte do corpo o câncer ou tumor maligno começou?
Entrevistador: Se houver mais de um primário, descreva o mais recente.
Parte do corpo: _____
14. Qual era o tipo deste câncer ou tumor maligno? Por exemplo, melanoma, sarcoma, carcinoma, adenocarcinoma etc.
Tipo: _____ CID-10: _____
15. Em que ano ele foi diagnosticado? _____ (*não sabe = 9999*)
Entrevistador: Se o entrevistado sabe o ano, passe 17; caso contrário siga para 16)
16. Quantos anos o(a) Sr.(a) tinha quando ele foi diagnosticado?
_____ (*não sabe = 9999*)
17. O(a) Sr.(a) teve algum outro tipo de câncer ou tumor maligno, que tenha começado em outra parte do corpo, e sem estar relacionado com o anterior?
1. () Sim
 2. () Não *passa 27*
 3. () Não sabe *passa 27*
18. Em que parte do corpo este câncer ou tumor maligno começou?
Entrevistador: Descreva o segundo mais recente.

Parte do corpo: _____

19. Qual era o tipo deste câncer ou tumor maligno? Por exemplo, melanoma, sarcoma, carcinoma, adenocarcinoma etc.

Tipo: _____ CID-10: _____

20. Em que ano ele foi diagnosticado? _____ (*não sabe = 9999*)
Entrevistador: Se o entrevistado sabe o ano, passe 22; caso contrário siga para 21)

21. Quantos anos o(a) Sr.(a) tinha quando ele foi diagnosticado? _____ anos
(*não sabe = 9999*)

22. O(a) Sr.(a) teve algum outro tipo de câncer ou tumor maligno, que tenha começado em outra parte do corpo, e sem estar relacionado com o anterior?

1. Sim
2. Não *passe 27*
3. Não sabe *passe 27*

23. Em que parte do corpo este câncer ou tumor maligno começou?

Entrevistador: Descreva o terceiro mais recente.

Parte do corpo: _____

24. Qual era o tipo deste câncer ou tumor maligno? Por exemplo, melanoma, sarcoma, carcinoma, adenocarcinoma etc.

Tipo: _____ CID-10: _____

25. Em que ano ele foi diagnosticado? _____ (*não sabe = 9999*)
Entrevistador: Se o entrevistado sabe o ano, passe 27; caso contrário siga para 26)

26. Quantos anos o(a) Sr.(a) tinha quando ele foi diagnosticado? _____ anos
(*não sabe = 9999*)

27. O(a) Sr.(a) teve algum tipo de tumor benigno diagnosticado pelo médico?

1. Sim
2. Não
3. Não sabe

28. Qual o tipo, quantos e em que ano ou com que idade os tumores benignos foram diagnosticados?

Entrevistador: Pergunte primeiro o ano de diagnóstico da primeira lesão de cada tipo de tumor benigno; se o entrevistado não souber o ano pergunte a idade do diagnóstico.

1. Cisto no rim	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
2. Pólipo intestinal	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
3. Tumor da glândula suprarrenal	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
4. Tumor de pâncreas	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
5. Tumor da hipófise	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
6. Tumor do osso	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
7. Tumor de tireóide	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
8. Meningeoma	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
9. Mioma	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
10. Lipoma	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
11. Lesão benigna de mama	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
12. Tumor de cerebelo	1. <input type="checkbox"/> Sim 2. <input type="checkbox"/> Não	Quantos. __ __ Ano. __ __ __ __ Idade. __ __
13. Outros. Descreva os tipos:		
1. _____		Quantos. __ __ Ano. __ __ __ __ Idade. __ __
2. _____		Quantos. __ __ Ano. __ __ __ __ Idade. __ __
3. _____		Quantos. __ __ Ano. __ __ __ __ Idade. __ __

APÊNDICE B - Termo de Consentimento Livre e Esclarecido



MINISTÉRIO DA EDUCAÇÃO E CULTURA
UNIVERSIDADE FEDERAL DO PIAUÍ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA
RENORBIO/UFPI

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

A Resolução CNS 466/12, item II.2 considera como pesquisa em seres humanos as realizadas em qualquer área do conhecimento e que, de modo direto ou indireto, envolva pessoas ou coletividades, em sua totalidade ou partes, incluindo o manejo de informações e materiais. Assim, também são consideradas pesquisas envolvendo seres humanos as entrevistas, aplicações de questionários, utilização de bancos de dados e revisões de prontuários. Portanto, em consonância com esta resolução este projeto será submetido ao parecer da Comissão de Ética da UFPI, como também serão respeitados todos os direitos dos pacientes ao anonimato e à autonomia. Será utilizado um termo de consentimento livre e esclarecido, no qual o responsável pelo paciente autorizará ou não a sua participação no projeto.

Assim, você está sendo convidada como voluntária a participar da pesquisa: Otimização de imagens mamográficas para auxílio ao diagnóstico médico.

A JUSTIFICATIVA, OS OBJETIVOS E OS PROCEDIMENTOS:

Justificativa: As neoplasias mamárias estão entre as maiores causas de morte de mulheres no Brasil. Esquemas computacionais de auxílio ao diagnóstico (CAD - “Computer-Aided Diagnosis”) são desenvolvidos para tentar fornecer ferramentas que ajudem na detecção precoce da doença. Tais esquemas utilizam primariamente a imagem mamográfica como fonte de dados. Com esse fim, muitas pesquisas têm sido feitas para possibilitar o desenvolvimento de técnicas para esquemas CAD. O **objetivo** desse projeto é a implementação de um sistema de processamento de imagens para a detecção de massas ou nódulos em imagens mamográficas, fazendo uso de redes neurais, com intuito de fornecer imagens para o diagnóstico mais preciso.

Procedimentos: As imagens serão adquiridas de pacientes do Hospital Getulio Vargas, setor de Ginecologia, submetidas a exames mamográficos. Não será realizada



**MINISTÉRIO DA EDUCAÇÃO E CULTURA
UNIVERSIDADE FEDERAL DO PIAUÍ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA
RENORBIO/UFPI**

entrevista gravada ou filmada, aceitando participar da pesquisa você terá garantia de manutenção do sigilo e da privacidade dos resultados durante todas as fases da pesquisa.

RISCOS E BENEFÍCIOS: Riscos: Como durante a pesquisa será aplicado um questionário, este poderá ocasionar um risco de constrangimento, mas o pesquisador responsável, Francisco Adelson Alves Ribeiro adotar procedimentos éticos no decorrer dessa aplicação conforme Resolução 466/2012 para evitar esse risco. Ao participar da pesquisa, a voluntária não sofrerá nenhum prejuízo. Não há riscos diretos aos participantes. **Benefícios:** As participantes do estudo terão como benefício o uso de sistema computacional para apoio ao diagnóstico médico; detecção precoce do câncer de mama e a redução do risco de radiação. **GARANTIA DE ESCLARECIMENTO, LIBERDADE DE RECUSA E GARANTIA DE SIGILO:** Você será esclarecida sobre a pesquisa em qualquer aspecto que desejar. Você é livre para recusar-se a participar, retirar seu consentimento ou interromper a participação a qualquer momento. A sua participação é voluntária e a recusa em participar não irá acarretar qualquer penalidade ou perda de benefícios.

Os pesquisadores irão tratar a sua identidade com padrões profissionais de sigilo. Os resultados serão enviados para você e permanecerão confidenciais. Seu nome ou o material que indique a sua participação não será liberado sem a sua permissão. Você não será identificada em nenhuma publicação que possa resultar deste estudo. Uma cópia deste consentimento informado será arquivada na Universidade Federal do Piauí e outra será fornecida a você.

CUSTOS DA PARTICIPAÇÃO, RESSARCIMENTO E INDENIZAÇÃO POR EVENTUAIS DANOS: A participação no estudo não acarretará custos para você e não será disponível nenhuma compensação financeira adicional.

DECLARAÇÃO DA PARTICIPANTE OU DO RESPONSÁVEL PELA PARTICIPANTE: Eu, _____ fui informada dos objetivos da pesquisa acima de maneira clara e detalhada e esclareci minhas dúvidas. Sei que em qualquer momento poderei solicitar novas informações e motivar



**MINISTÉRIO DA EDUCAÇÃO E CULTURA
UNIVERSIDADE FEDERAL DO PIAUÍ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA
RENORBIO/UFPI**

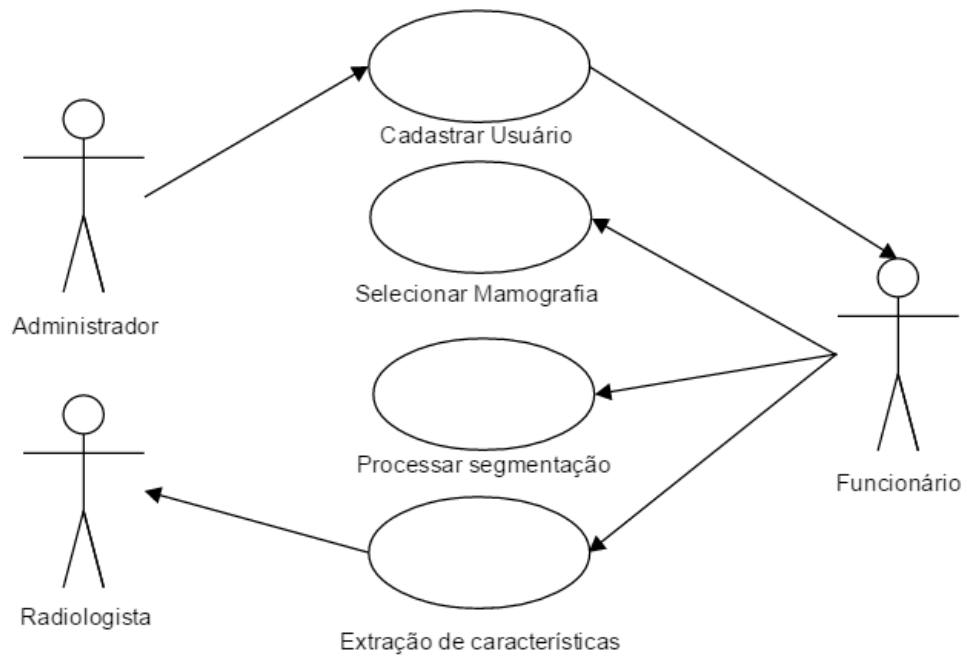
minha decisão se assim o desejar. O professor orientador certifica-me de que todos os dados desta pesquisa serão confidenciais.

Também sei que caso existam gastos adicionais, estes serão absorvidos pelo orçamento da pesquisa. Em caso de dúvidas poderei localizar o professor orientador **Benedito Borges da Silva** no seguinte endereço: Universidade Federal do Piauí, Centro de Ciências da Saúde. Campus Universitário Ministro Petrônio Portela, Bairro Ininga, CEP 64049-550 Teresina, PI. Telefone: (86) 3215-5160 ou o Comitê de Ética em Pesquisa se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato: Comitê de Ética em Pesquisa/UFPI, Campus Universitário Ministro Petrônio Portela, Bairro Ininga, CEP 64049-550

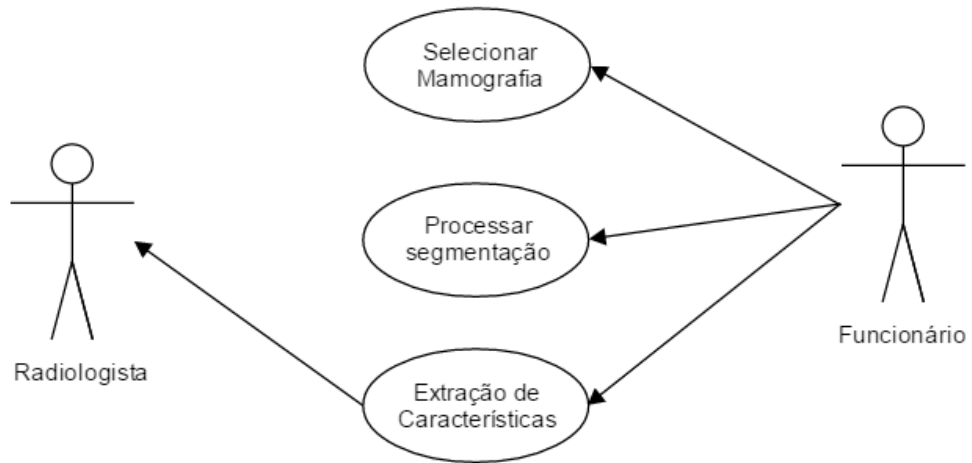
Declaro que concordo em participar desse estudo. Recebi uma cópia deste termo de consentimento livre e esclarecido e me foi dada a oportunidade de ler e esclarecer as minhas dúvidas.

Nome	Assinatura do Participante	Data
Nome	Assinatura do Pesquisador	Data
Nome	Assinatura da Testemunha	Data

APÊNDICE C - Diagrama de Contexto - MedMamo



APÊNDICE D - Diagrama de casos de uso – MedMamo



APÊNDICE E - Diagrama de Classes do Projeto – MedMamo

