



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA
MESTRADO ACADÊMICO EM ENGENHARIA DE TELEINFORMÁTICA

DANIEL FREITAS COLAÇO

CLASSIFICAÇÃO AUTOMÁTICA DA DENSIDADE DA MAMA BASEADA EM
MODELOS GRAVITACIONAIS SIMPLIFICADOS

FORTALEZA

2017

DANIEL FREITAS COLAÇO

CLASSIFICAÇÃO AUTOMÁTICA DA DENSIDADE DA MAMA BASEADA EM
MODELOS GRAVITACIONAIS SIMPLIFICADOS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Ciências e Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Orientador: Prof. Dr. Paulo César Cortez

FORTALEZA

2017

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

C642c Colaço, Daniel Freitas.
Classificação automática da densidade da mama baseada em modelos gravitacionais simplificados /
Daniel Freitas Colaço. – 2017.
68 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-
Graduação em Engenharia de Teleinformática, Fortaleza, 2017.
Orientação: Prof. Dr. Paulo César Cortez.

1. Mamografia. 2. CAD. 3. Lacunaridade. 4. MIAS. I. Título.

CDD 621.38

DANIEL FREITAS COLAÇO

CLASSIFICAÇÃO AUTOMÁTICA DA DENSIDADE DA MAMA BASEADA EM
MODELOS GRAVITACIONAIS SIMPLIFICADOS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Ciências e Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Aprovada em: 31 de Julho de 2017

BANCA EXAMINADORA

Prof. Dr. Paulo César Cortez (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Wagner Coelho de Albuquerque Pereira
Universidade Federal do Rio de Janeiro (UFRJ)

Prof. Dr. Tarcísio Ferreira Maciel
Universidade Federal do Ceará (UFC)

Prof. Dr. José Marques Soares
Universidade Federal do Ceará (UFC)

A DEUS, pelo presente da vida e por nos guiar
nessa caminhada. À minha esposa e filhos, pela
paciência, amor e incentivos. Aos meus pais,
pela dedicação e amor. Ao meu sogro e sogra,
pelo apoio. Às minhas irmãs e cunhado, pela
amizade e carinho.

AGRADECIMENTOS

Ao amigo Prof. Dr. Paulo César Cortez por me orientar durante toda a minha vida acadêmica, indo muito além da função de professor.

Ao amigo Ricardo Jardel e família pela amizade incondicional, incentivos e apoio para terminar este trabalho.

Aos amigos Luiz Alves e Edmilson Moreira pela amizade e apoio.

Aos amigos de laboratório e de universidade pelas discussões acadêmicas e políticas que muito me engrandeceram.

Aos professores Jarbas Joaci de Mesquita Sá Júnior e André R. Backes pela ajuda no desenvolvimento deste trabalho.

A todos os professores por compartilharem pacientemente os seus conhecimentos, fazendo-me aprender e preparando-me para a minha vida profissional.

Ao Prof. Dr. Tobias Rafael Fernandes Neto, coordenador do Laboratório de Sistemas Motrizes (LAMOTRIZ) onde este *template* foi desenvolvido.

Ao Doutorando em Engenharia Elétrica, Ednardo Moreira Rodrigues, e seu assistente, Alan Batista de Oliveira, aluno de graduação em Engenharia Elétrica, pela adequação do *template* utilizado neste trabalho para que o mesmo ficasse de acordo com as normas da biblioteca da Universidade Federal do Ceará (UFC).

Aos bibliotecários da Universidade Federal do Ceará: Francisco Edvander Pires Santos, Juliana Soares Lima, Izabel Lima dos Santos, Kalline Yasmin Soares Feitosa e Eliene Maria Vieira de Moura, pela revisão e discussão da formatação utilizada neste *template*.

Agradeço a todos os professores por me proporcionar o conhecimento não apenas racional, mas a manifestação do caráter e afetividade da educação no processo de formação profissional, por tanto que se dedicaram a mim, não somente por terem me ensinado, mas por terem me feito aprender.

“A força não provém da capacidade física, e sim de uma vontade indomável.”

(Mahatma Gandhi)

RESUMO

O câncer de mama permanece como o segundo tipo de câncer mais frequente no mundo. Atualmente a mamografia é o principal exame para o diagnóstico precoce da doença, fator fundamental para o sucesso do tratamento. Entretanto, a composição do tecido mamário pode dificultar a detecção de lesões devido à alta densidade do tecido fibroglandular atenuar muito a energia efetiva do feixe de raios X da mamografia. Assim, mamas densas apresentam um risco de 4 a 6 vezes maior de não se detectar lesões no estágio inicial. Esta dissertação apresenta uma pesquisa dos modelos gravitacionais simplificados aplicados na análise de textura de imagens mamográficas para posterior classificação automática da densidade do tecido mamário em sistemas de diagnóstico auxiliado por computador. O modelo gravitacional simplificado em estudo foi aplicado em 300 mamografias do banco de imagens *Mammographic Image Analysis Society (MIAS)* e, através de descritores de lacunaridade, combinado com os classificadores Máquina de Vetor Suporte (*Support Vector Machine – SVM*) e K vizinhos mais próximos (*K-nearest neighbors - Knn*), obteve performance de 76,7% de acurácia com especificidade média de 87,27%, indicando que o método de modelos gravitacionais simplificados pode ser utilizado na análise de textura de imagens mamográficas, combinado com outros métodos de classificação e de pré-processamento.

Palavras-chave: Mamografia. CAD. Lacunaridade. MIAS.

ABSTRACT

Breast cancer remains the second most common cancer in the world. Currently, mammography is the main test for early diagnosis of the disease, a fundamental factor to the success of the treatment. However, the composition of the mammary tissue can make it difficult to detect lesions as the high density of the fibroglandular tissue greatly attenuates the effective energy of the X-ray beam of mammography. As a result, dense breasts present 4 to 6-fold greater risk of not detecting lesions in the initial stage. This master thesis presents a research of the simplified gravitational model applied to the analysis of mammographic images textures for later automatic density classification of the mammary tissue in computer-aided diagnostic systems. The simplified gravitational model under study was applied to 300 mammograms from the MIAS image database and, through lacunarity descriptors, combined with the Support Vector Machine (SVM) and K-nearest neighbors (Knn), obtained a performance of 76.7% of Accuracy with average Specificity of 87.27%, indicating that the method of simplified gravitational models can be used in the analysis of mammographic images texture, combined with other methods of classification and preprocessing.

Keywords: Mammography. CAD. Lacunarity. MIAS.

LISTA DE FIGURAS

Figura 1 – Mamografias com diferentes densidades de tecidos. Esquerda: Mama esquerda com tecido fibroglandular. Direita: Mama direita com tecido adiposo.	18
Figura 2 – Etapas do processamento de imagens digitais.	20
Figura 3 – Estruturas mamárias.	22
Figura 4 – Mamógrafo e seus componentes principais.	23
Figura 5 – Diagrama esquemático da produção de radiação X por Bremsstrahlung, mostrando a liberação de um fóton de raios X devido à perda de energia do elétron com a desaceleração.	24
Figura 6 – Descrição da produção de radiação característica, quando um elétron incidente retira um elétron orbital, causando uma vacância. Um elétron de uma camada mais externa ocupa esta vacância e o excesso de energia é emitido em forma de fóton de raios X.	25
Figura 7 – Leitora CR.	26
Figura 8 – Mamógrafo digital.	27
Figura 9 – Mamografia com incidência Médio Lateral Oblíqua - MLO.	28
Figura 10 – Mamografia com incidência Médio Lateral Oblíqua - MLO.	29
Figura 11 – Imagens de exame mamográfico com incidência Médio Lateral Oblíqua - MLO.	29
Figura 12 – Mamografia com incidência Crânio Caudal - CC.	30
Figura 13 – Imagens de exame mamográfico com incidência Crânio Caudal - CC.	31
Figura 14 – Ilustração de força gravitacional entre duas massas.	31
Figura 15 – Ilustração de uma partícula em movimento ao redor de outra partícula.	32
Figura 16 – Modelo gravitacional simplificado onde o pixel p colapsa. D_1 e D_2 determinam a nova posição do pixel.	33
Figura 17 – Demonstração do processo de cálculo de lacunaridade em uma imagem.	35
Figura 18 – Apresentação de um hiperplano e margens gerados por SVM treinado com amostras de duas classes de objetos. As amostras nas margens são chamadas de vetores suporte.	37
Figura 19 – Kernel mais comuns.	37

Figura 20 – Uma observação x é testada em um conjunto treinamento com $k = 3$, formado pelas observações com menor distância euclidiana do ponto x^* . No caso apresentado, a observação x é alocada à categoria dos pontos representados pelos triângulos azuis.	38
Figura 21 – Metodologia proposta para classificação da densidade mamária.	40
Figura 22 – Imagem de mama densa com área de amostra destacada e apresentada ao lado.	41
Figura 23 – Imagem de mama fibroglandular com área de amostra destacada e apresentada ao lado.	41
Figura 24 – Imagem de mama adiposa com área de amostra destacada e apresentada ao lado.	42
Figura 25 – Etapa de Extração de Atributos.	43
Figura 26 – Determinação dos parâmetros de configuração do Modelo Gravitacional Simplificado.	46
Figura 27 – Lacunaridades para uma amostra de textura de imagem mamográfica usando $t = 1, 2, \dots, 20$ e tamanho de janelas $r = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$	47
Figura 28 – Lacunaridades de uma amostra de tecido de imagem mamográfica usando diferentes tamanhos de janela r	48
Figura 29 – Imagem original; amostra 200x200, imagens colapsadas com $M=500$, $G=1$ e $t = 1, 3, 6, 12$, respectivamente.	49
Figura 30 – Imagem original; amostra 200x200, imagens colapsadas com $M=750$, $G=1$ e $t = 1, 3, 6, 12$, respectivamente.	49
Figura 31 – Classificador <i>Support Vector Machine</i> (SVM): (a) Tecido Denso classificado corretamente. (b) Tecido Denso classificado como Fibroglandular. (c) Tecido Denso classificado como Gorduroso.	51
Figura 32 – Classificador SVM: (a) Tecido Gorduroso classificado corretamente. (b) Tecido Gorduroso classificado como Denso. (c) Tecido Gorduroso classificado como Fibroglandular.	51
Figura 33 – Classificador SVM: (a) Tecido Fibroglandular classificado corretamente. (b) Tecido Fibroglandular classificado como Denso. (c) Tecido Fibroglandular classificado como Gorduroso.	51

Figura 34 – Classificador Knn: (a) Tecido Denso classificado corretamente. (b) Tecido Denso classificado como Gorduroso. (c) Tecido Denso classificado como Fibroglandular.	52
Figura 35 – Classificador Knn: (a) Tecido Gorduroso classificado corretamente. (b) Tecido Gorduroso classificado como Denso. (c) Tecido Gorduroso classificado como Fibroglandular.	52
Figura 36 – Classificador Knn: (a) Tecido Gorduroso classificado corretamente. (b) Tecido Gorduroso classificado como Denso. (c) Tecido Gorduroso classificado como Fibroglandular.	53

LISTA DE TABELAS

Tabela 1 – Modelos gravitacionais avaliados	48
Tabela 2 – Acurácia do classificador SVM de acordo com os modelos gravitacionais simplificados avaliados.	50
Tabela 3 – Acurácia do classificador Knn de acordo com os modelos gravitacionais simplificados avaliados e considerando diferentes tamanhos de vizinhança (2, 10, 15, 18 e 20).	50
Tabela 4 – Performance dos algoritmos de classificação para cada tipo de densidade de tecido de acordo com a acurácia e a especificidade.	54
Tabela 5 – Matriz de Confusão para o classificador SVM.	55
Tabela 6 – Matriz de Confusão para o classificador Knn.	55

LISTA DE ABREVIATURAS E SIGLAS

<i>MIAS</i>	<i>Mammographic Image Analysis Society</i>
<i>SVM</i>	<i>Support Vector Machine</i>
<i>INCA</i>	Instituto Nacional do Câncer
<i>BI-RADS</i>	<i>Breast Imaging Reporting and Data System</i>
<i>CAD</i>	<i>Computer-Aided Diagnosis</i>
<i>MGC</i>	Mamografia Convencional
<i>CR</i>	Radiologia Computadorizada
<i>DR</i>	Radiologia Digital
<i>CC</i>	Crânio Caudal
<i>MLO</i>	Médio Lateral Oblíquo
<i>RBF</i>	<i>Radial-Basis Function</i>
<i>Knn</i>	<i>k-Nearest Neighbor</i>
<i>2DPCA</i>	<i>Two-Dimensional Principal Component Analysis</i>
<i>ROI</i>	<i>Region of Interest</i>
<i>PGM</i>	<i>Portable Gray Map</i>
<i>DDSM</i>	<i>Digital Database for Screening Mammography</i>

LISTA DE SÍMBOLOS

\mathbf{a}_c	Aceleração centrípeta
Be	Berílio
\mathbf{f}_a	Força gravitacional
\mathbf{f}_c	Força centrípeta
G	Constante Gravitacional
Mo	Molibdênio
\mathbf{r}	Distância entre duas partículas
\mathbf{r}_{max}	Maior distância entre o pixel e o centro da imagem
Rh	Ródio
t	Tempo
\mathbf{v}	Velocidade tangencial
\mathbf{v}_{max}	Velocidade máxima do pixel
W	Tungstênio
Λ	Lacunaridade

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Objetivos	19
1.2	Organização da Dissertação	19
2	CONCEITOS BÁSICOS	20
2.1	Processamento de Imagens Digitais	20
2.2	Densidade dos tecidos mamários	21
2.3	Mamografia	23
2.3.1	<i>Processo de Produção de raios X</i>	24
2.3.2	<i>Formação da Imagem</i>	25
2.3.3	<i>Incidências de Rotina</i>	27
2.3.3.1	<i>Médio Lateral Oblíqua - MLO</i>	27
2.3.3.2	<i>Crânio Caudal - CC</i>	28
2.4	Sistema Gravitacional Simplificado e Análise de Textura	30
2.5	Lacunaridade	34
2.6	Conceitos Básicos de Aprendizado de Máquina	34
2.6.1	<i>Support Vector Machine - SVM</i>	36
2.6.2	<i>Algoritmo K-Nearest Neighbor</i>	36
2.7	Revisão Bibliográfica	38
3	MATERIAIS E MÉTODOS	40
3.1	Aquisição das Imagens	40
3.2	Extração de Atributos	41
3.2.1	<i>Análise de textura com Modelo Gravitacional Simplificado</i>	41
3.3	Classificação / Reconhecimento da Densidade da Mama	43
3.4	Determinação dos Parâmetros de Configuração	44
4	RESULTADOS	47
5	CONCLUSÕES E TRABALHOS FUTUROS	54
5.1	Contribuições	56
5.2	Sugestões para trabalhos futuros	56
	REFERÊNCIAS	58
	APÊNDICES	61

ANEXOS	61
ANEXO A-E-MAIL ACEITE DO ARTIGO NO CLAIB 2016	61

1 INTRODUÇÃO

O câncer da mama permanece como o segundo tipo de câncer mais frequente no mundo. De acordo com estudos realizados pelo Instituto Nacional do Câncer (INCA), no Brasil, para o biênio 2016-2017, estimou-se o surgimento de 57.960 novos casos de câncer de mama por ano, com um risco estimado de 56,20 casos a cada 100 mil mulheres. No Ceará estima-se a ocorrência de 2.160 novos casos por ano. Sem considerar os tumores de pele não melanoma, é o tipo que possui maior incidência e a maior mortalidade na população feminina em todo o mundo, tanto em países em desenvolvimento quanto em países desenvolvidos (SCHLITZ *et al.*, 2015).

Apesar de ser considerado um câncer de relativamente bom prognóstico, se diagnosticado e tratado oportunamente, as taxas de mortalidade por câncer da mama continuam elevadas no Brasil, muito provavelmente porque a doença ainda seja diagnosticada em estágios avançados. A sobrevida média após cinco anos na população de países desenvolvidos tem apresentado um aumento, cerca de 85%. Entretanto, nos países em desenvolvimento, a sobrevida média após cinco anos fica em torno de 60%.

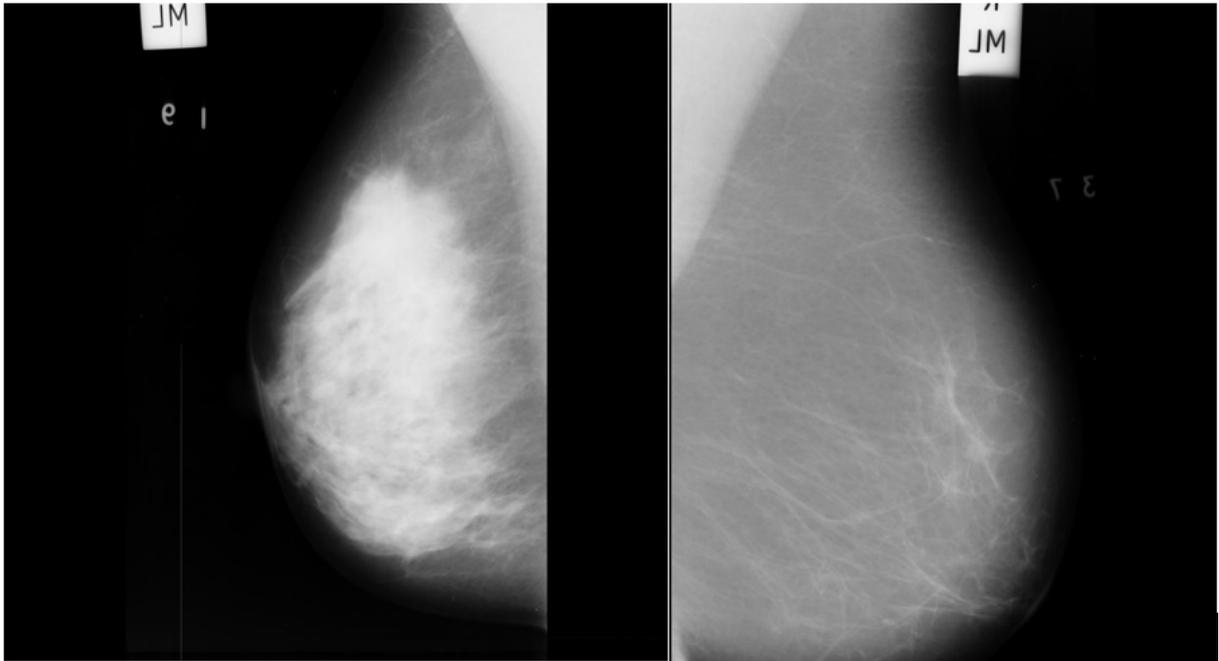
No Brasil, o exame clínico anual das mamas e o rastreamento são as estratégias recomendadas para controle do câncer da mama. As recomendações do Ministério da Saúde para detecção precoce e diagnóstico desse câncer considera, como principais estratégias de rastreamento, o exame clínico anual das mamas a partir dos 40 anos e um exame mamográfico, a cada dois anos, para mulheres de 50 a 69 anos.

A mamografia é o exame mais indicado para a detecção precoce do câncer de mama, uma vez que permite a visualização de possíveis artefatos nas mamas que evidenciem a presença ou ausência deste tipo de câncer. Entretanto, a avaliação da imagem mamográfica é subjetiva, dependente da experiência, concentração e habilidade do radiologista e bastante difícil. Essa dificuldade ocorre devido a vários fatores, tais como a grande variedade dos formatos das lesões, variação na densidade do tecido mamário e a alta taxa de ruído presente nas mamografias (MELLO-THOMS, 2003). Além disso, a própria natureza da imagem é um complicador, já que as mamografias são representações bidimensionais de estruturas tridimensionais, resultando em sobreposição de regiões. Tudo isso, resulta em um número alto de lesões não identificadas.

A densidade do tecido mamário pode dificultar a detecção de lesões. O tecido fibroglandular tem alta densidade e atenua muito a energia efetiva do feixe de raios X do exame mamográfico. Em contraposição, o tecido gorduroso absorve pouco o feixe de raios X. Mamas densas apresentam um risco de 4 a 6 vezes maior de não se detectar lesões no estágio inicial.

Diante desses riscos, a densidade é incorporada ao laudo das mamografias, o qual foi padronizado pelo Colégio Americano de Radiologia dando origem à classificação *Breast Imaging Reporting and Data System* (BI-RADS). A Figura 1 apresenta imagens mamográficas de mamas com diferentes composições de tecidos, fibroglandular e adiposo, respectivamente.

Figura 1 – Mamografias com diferentes densidades de tecidos. Esquerda: Mama esquerda com tecido fibroglandular. Direita: Mama direita com tecido adiposo.



Fonte: *Mammographic Image Analysis Society-MIAS* (SUCKLING, 1994)

Nos últimos anos, técnicas de processamento de imagens e de inteligência artificial vêm sendo desenvolvidas com o propósito de auxiliar no diagnóstico médico, incrementando a taxa de diagnóstico atempado, compondo soluções conhecidas por sistemas *Computer-Aided Diagnosis* (CAD).

No contexto da visão computacional, entre os vários atributos relevantes para uma análise de imagens, a textura é um dos mais importantes por ser uma fonte rica de informações a qual permite a caracterização de estruturas e regiões em uma imagem digital.

Diante do exposto, esta dissertação visa a pesquisa e o desenvolvimento de técnicas e algoritmos para classificação automática da densidade do tecido mamário por análise de texturas de imagens mamográficas.

1.1 Objetivos

O objetivo geral dessa dissertação consiste em desenvolver e implementar algoritmos para classificar automaticamente a densidade das mamas em mamografias, tomando como base o modelo gravitacional simplificado. Durante o desenvolvimento desta dissertação outros objetivos específicos devem ser alcançados:

1. obtenção de um novo método de classificação da densidade do tecido mamário;
2. implementação e comparação de algoritmos descritos na literatura para compor um novo método de identificação da densidade de tecidos mamários;

1.2 Organização da Dissertação

Esta dissertação está estruturada em 5 capítulos. Além do capítulo 1, o qual apresenta a problemática do câncer de mama no Brasil, no capítulo 2 são apresentados alguns conceitos importantes relativos ao processamento de imagens digitais, mamografia, densidade dos tecidos mamários, análise de texturas e lacunaridade.

No capítulo 3 são descritos os materiais e métodos utilizados para a implementação de cada algoritmo avaliado nesta dissertação. No capítulo 4 são apresentados os resultados obtidos no decorrer deste projeto e, por fim, no capítulo 5 são apresentadas as conclusões e contribuições obtidas, assim como, os trabalhos futuros que podem ser desenvolvidos, tomando como base esta dissertação.

2 CONCEITOS BÁSICOS

Neste capítulo são apresentados alguns conceitos básicos inerentes ao desenvolvimento da dissertação.

2.1 Processamento de Imagens Digitais

O processamento de imagens digitais tem um papel importante na análise de imagens médicas. Geralmente um sistema para processamento de imagens digitais é composto pelas seguintes etapas: aquisição, pré-processamento, segmentação, extração de atributos e reconhecimento/classificação. A Figura 2 apresenta as etapas de um sistema padrão de processamento de imagens digitais.

Figura 2 – Etapas do processamento de imagens digitais.



Em um sistema de processamento de imagens médicas, a etapa de aquisição da imagem é realizada através de equipamentos como o mamógrafo, raios X, ultrassom, tomógrafo, dentre outros. Em seguida, as imagens coletadas são pré-processadas, corrigindo distorções oriundas da coleta ou realizando ajustes de formatação, cor, brilho, contraste e filtragem de ruídos. Na etapa seguinte, as áreas de interesse são segmentadas, permitindo uma análise mais específica e eliminando as regiões e informações não relevantes. Segmentada as áreas de interesse, é realizada a extração dos atributos, os quais caracterizam as áreas analisadas. Através dos atributos, é possível realizar a etapa de classificação e reconhecimento, a qual tem como resultado a classificação das áreas segmentadas usando diferentes técnicas de inteligência artificial e reconhecimento de padrões.

Os sistemas de visão computacional podem ser compostos ou não por todas as etapas de processamento de imagens digitais descritas anteriormente. A necessidade da presença de cada etapa é determinada de acordo com o objetivo do sistema de visão computacional, com a qualidade das imagens e com as técnicas utilizadas.

2.2 Densidade dos tecidos mamários

As glândulas mamárias, que têm como principal função a secreção do leite, estão situadas na parede anterior do tórax e são compostas por:

1. Ácino - menor parte da glândula e responsável pela produção do leite no período de lactação;
2. Lóbulo mamário - conjunto de ácinos;
3. Lobo mamário - conjunto de lóbulos mamários, ligados à papila através de um ducto;
4. Ductos mamários - composto por 15 a 20 canais, através dos quais o leite é conduzido à papila;
5. Tecido glandular - conjunto de lobos e ductos;
6. Papila - protuberância elástica onde desembocam os ductos mamários;
7. Aréola - estrutura central da mama onde se projeta a papila;
8. Tecido Adiposo - todo o restante da mama é preenchido por tecido adiposo, cuja quantidade varia de acordo com as características físicas, estado nutricional e idade da mulher.

A Figura 3 apresenta a imagem das estruturas da mama.

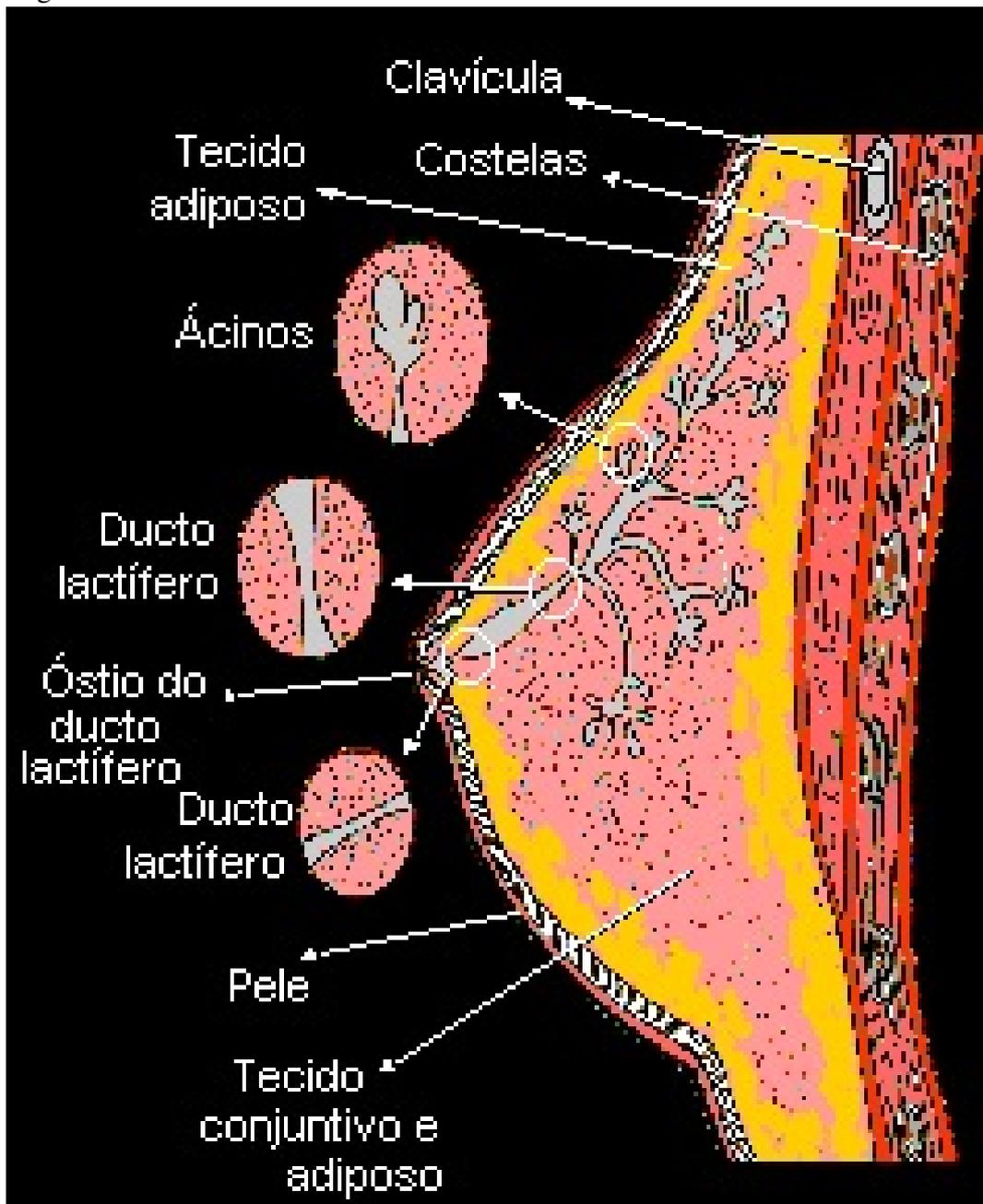
A densidade do tecido da mama não está relacionada ao exame de palpação, mas sim de como os seus tecidos aparecem na mamografia. Mamas com alta densidade indicam uma alta quantidade de tecidos glandulares e conjuntivos em relação à gordura. Já no caso de tecidos com baixa densidade indicam uma grande quantidade de gordura se comparada aos outros tecidos.

A densidade da mama varia com a idade e com o peso da mulher. Mamas densas são mais comuns em mulheres jovens e em mulheres magras. Cerca de 50 a 60 % das mulheres com idades entre 40 e 44 anos de vida possuem seios densos, comparado com somente 20 a 30% das mulheres com idade entre 70 e 74 anos de vida. As mesmas relações percentuais se aplicam às mulheres com peso dentro dos limites considerados normais para uma pessoa saudável com relação às mulheres obesas (HO *et al.*, 2014).

Medicamentos que possuem hormônios também podem afetar a densidade da mama. Mulheres na menopausa, que fazem terapia de reposição hormonal, podem ter a densidade da mama aumentada. Com a idade, as mamas ficam menos densas e mais gordurosas. Com o tratamento hormonal na menopausa, este processo pode ser desacelerado (KHAN, 2014).

A preocupação médica com as mamas densas é devida à grande dificuldade de

Figura 3 – Estruturas mamárias.



Fonte: Faculdade de Medicina de São José do Rio Preto-Famerp (ARANTES *et al.*, 2004)

deteção de tumores em estágio precoce neste tipo de mama. Tumores encontrados em mamas mais densas são difíceis de detectar e podem estar em um estágio mais avançado quando diagnosticados. Por isso, é aconselhável um intervalo mais curto de rastreamento para mulheres com este tipo de mama, através da execução de exames preventivos mais frequentes a partir de determinada idade. Com isso, é possível acompanhar a mudança do padrão mamográfico, devido principalmente ao aumento da idade da paciente (SANTOS, 2002). Mulheres com mamas densas apresentam um risco 4 a 5 vezes maior de desenvolver câncer de mama do que mulheres com mamas não densas (YAGHJYAN *et al.*, 2011).

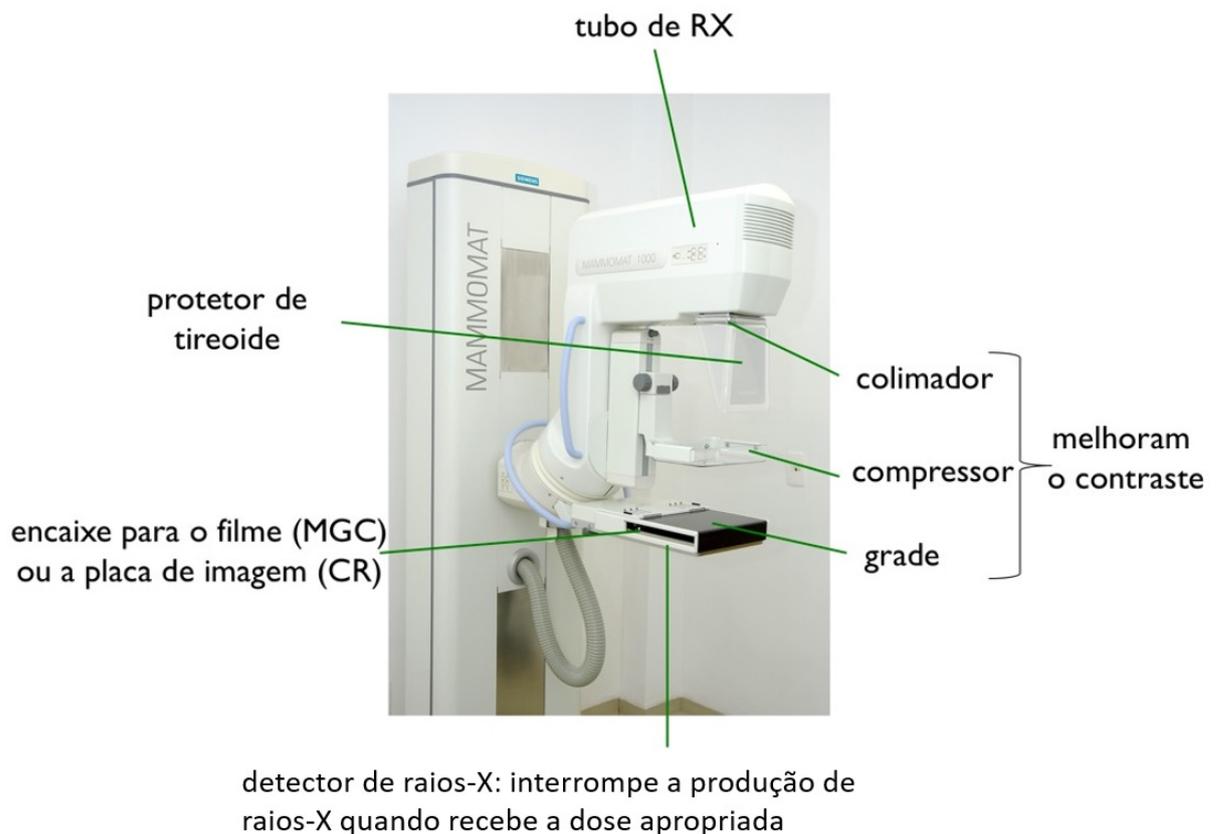
2.3 Mamografia

A mamografia é uma forma particular de radiografia, que trabalha com níveis de tensões e correntes em intervalos específicos, voltada para a coleta de imagens da mama, com o intuito de evidenciar a presença de microcalcificações e ou nódulos que possam indicar uma anormalidade a ser investigada.

O exame mamográfico é uma técnica simples, tolerada pelas pacientes, reproduzível, de baixo custo e com uma alta sensibilidade e especificidade. Apropriadamente executada e interpretada, a mamografia possui uma sensibilidade acima de 90%.

No caso da mamografia, a coleta da imagem é realizada por um aparelho de raios X específico para esta aplicação, conhecido por mamógrafo. A Figura 4, apresenta a imagem de um mamógrafo com a identificação dos seus elementos principais.

Figura 4 – Mamógrafo e seus componentes principais.

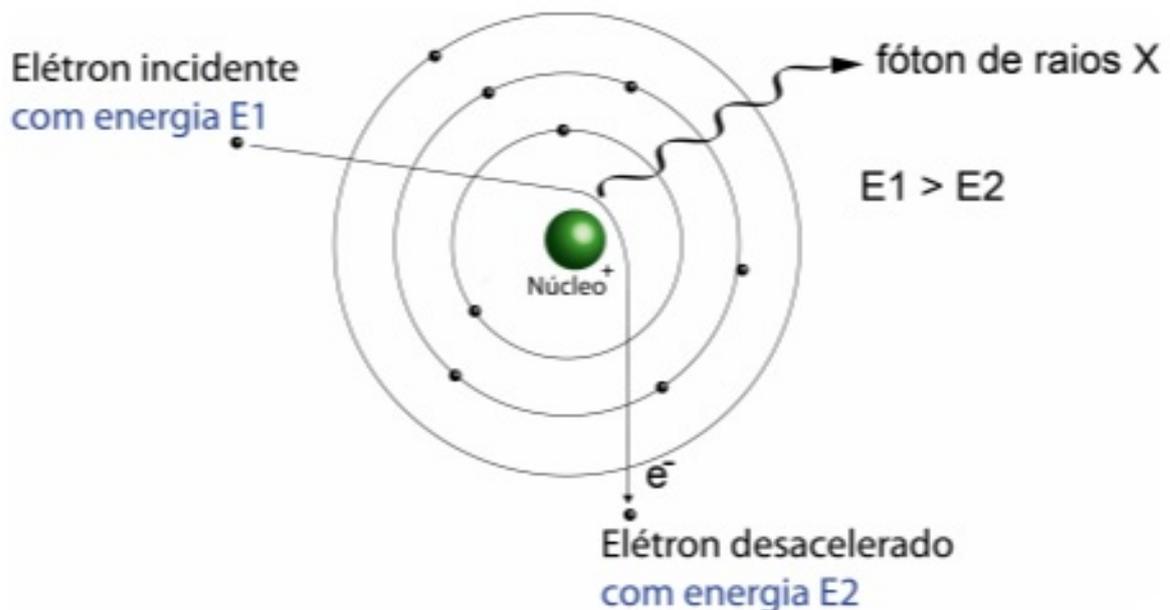


Fonte: Faculdade de Ciências Médicas-UNICAMP (JALES, 2017)

2.3.1 Processo de Produção de raios X

Os raios X são produzidos no tubo de raios X através do seguinte processo: Um catodo é aquecido liberando um feixe de elétrons (processo de emissão termoiônica). Esses elétrons são acelerados até um anodo, devido a uma diferença de potencial, formando uma corrente. Ao se chocarem com o anodo, os elétrons penetram no material e passam muito próximo dos núcleos dos átomos, causando uma desaceleração, fazendo com que os elétrons percam energia, a qual é emitida em forma de raios X (fenômeno conhecido como *Bremsstrahlung* ou radiação de freamento). Além desse processo, os elétrons podem colidir com elétrons orbitais dos átomos do material do anodo e retirá-los, caso a energia cinética destes seja maior ou igual à energia de ligação dos elétrons no átomo. Ao ser removido do átomo, o elétron deixa uma vacância naquela órbita. Isso faz com que elétrons de camadas mais externas ocupem essa vacância. Como em órbitas mais internas há uma menor energia de ligação, o excesso é emitido em forma de fótons de raios X (fenômeno conhecido como radiação característica)(KOPANS, 1998). As Figuras 5 e 6 apresentam os fenômenos de radiação de freamento e a característica, respectivamente.

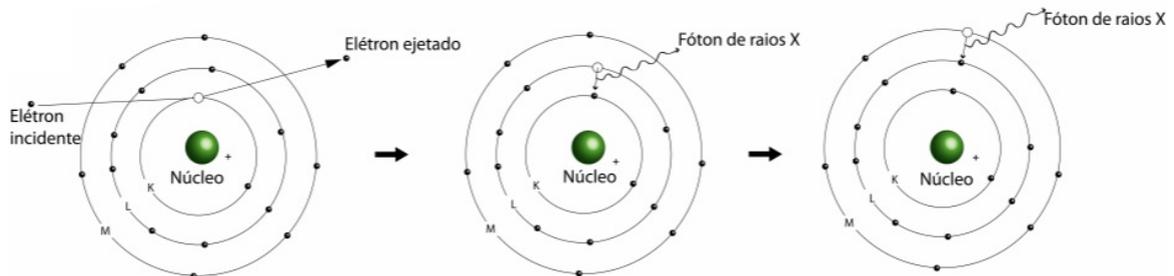
Figura 5 – Diagrama esquemático da produção de radiação X por Bremsstrahlung, mostrando a liberação de um fóton de raios X devido à perda de energia do elétron com a desaceleração.



Fonte: Instituto de Física da Universidade de São Paulo-USP (FURQUIM, 2017)

Depois de formados, os raios X saem do tubo passando através de uma janela de berílio (Be) e passam por um filtro que pode ser de molibdênio (Mo), ródio (Rh), tungstênio(W),

Figura 6 – Descrição da produção de radiação característica, quando um elétron incidente retira um elétron orbital, causando uma vacância. Um elétron de uma camada mais externa ocupa esta vacância e o excesso de energia é emitido em forma de fóton de raios X.



Fonte: Instituto de Física da Universidade de São Paulo-USP (FURQUIM, 2017)

cobre ou alumínio os quais eliminam os componentes de baixa energia do espectro de raios x que não são necessários para criar a imagem, reduzindo não somente a dose na pele do paciente, como também a radiação espalhada para o operador.

Em seguida o feixe de raios X passa pelo colimador, o qual é construído a partir de um material que absorve radiação e é usado para direcionar e suavizar feixes de radiação. Na maioria dos tubos de raios X, utilizam-se colimadores de abertura variável, constituídos de dois conjuntos de lâminas de chumbo que podem ser ajustadas para a obtenção de campos de incidência retangulares de tamanhos variáveis.

2.3.2 Formação da Imagem

Os raios X, produzidos por poucos segundos e dirigidos para a mama, podem atravessá-la ou ser absorvidos ou desviados, de acordo com a espessura da mama, composição do parênquima e a presença de doenças. Os raios X que atravessam a mama são detectados pelo filme - Mamografia Convencional (MGC) ou pela placa de fósforo - Radiologia Computadorizada (CR) ou pelo receptor digital - Radiologia Digital (DR).

No mamógrafo convencional, as imagens são gravadas em filme através de reações de agentes químicos. O filme é analisado pelo radiologista no negatoscópio. Quando a imagem não é satisfatória, o exame tem que ser refeito e a mulher será exposta a uma nova carga de radiação.

No caso da radiologia computadorizada CR, a imagem da mama é formada em uma placa de fósforo inserida em um cassete. Após a exposição da mama, as informações contidas na placa de fósforo são digitalizadas pela leitora CR. Com isso, as imagens podem ser ajustadas digitalmente, permitindo que os exames possam ser interpretados e arquivados em alta resolução. A Figura 7 apresenta a imagem de uma leitora CR de um dos fabricantes mundiais.

Figura 7 – Leitora CR.



Fonte: Carestream (CARESTREAM, 2017)

Já na tecnologia de radiologia digital DR, a radiação é processada diretamente para o formato digital, sem a necessidade de encaixe ou suporte para o cassete com o filme ou com a placa de fósforo. A Figura 8 apresenta um mamógrafo digital. Neste caso, as imagens também são geradas através de raios X, porém o sistema é equipado por um receptor eletrônico o qual converte a imagem em uma foto digital. Esta imagem pode ser analisada pelo radiologista em um computador, fazendo uso de ferramentas que realcem áreas da imagem permitindo uma análise mais precisa da mama. Nos procedimentos com mamógrafo digital, o paciente recebe uma dose menor de raios X, são mais rápidos, as imagens são de melhor qualidade e permite um pós-processamento, em comparação com os mamógrafos convencionais.

Figura 8 – Mamógrafo digital.



Fonte: *Medical Expo* (MEDICALEXPO, 2017)

2.3.3 Incidências de Rotina

Segundo o INCA (SCHILITZ *et al.*, 2015), o exame mamográfico é composto por duas incidências de cada mama: uma incidência Crânio Caudal (CC) e uma incidência Médio Lateral Oblíquo (MLO).

2.3.3.1 Médio Lateral Oblíqua - MLO

A incidência 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.

Para a aquisição desta incidência o mamógrafo deve ser angulado de 35 a 45 graus, dependendo do biotipo da paciente. Dessa forma, o feixe de raios X será da região medial para

a lateral oblíqua. A paciente deve estar de frente para o aparelho, elevar o braço para que o canto do bucky esteja na região axilar, para que assim seja visualizado a músculo peitoral. As Figuras 9 e 10 demonstram a aquisição de imagens com incidência MLO e a figura 11 apresenta as imagens resultantes dessa incidência nas mamas direita e esquerda, respectivamente.

Figura 9 – Mamografia com incidência Médio Lateral Oblíqua - MLO.



Fonte: *Centro de Educação Profissional Integrado - CIE (CIE, 2017)*

2.3.3.2 Crânio Caudal - CC

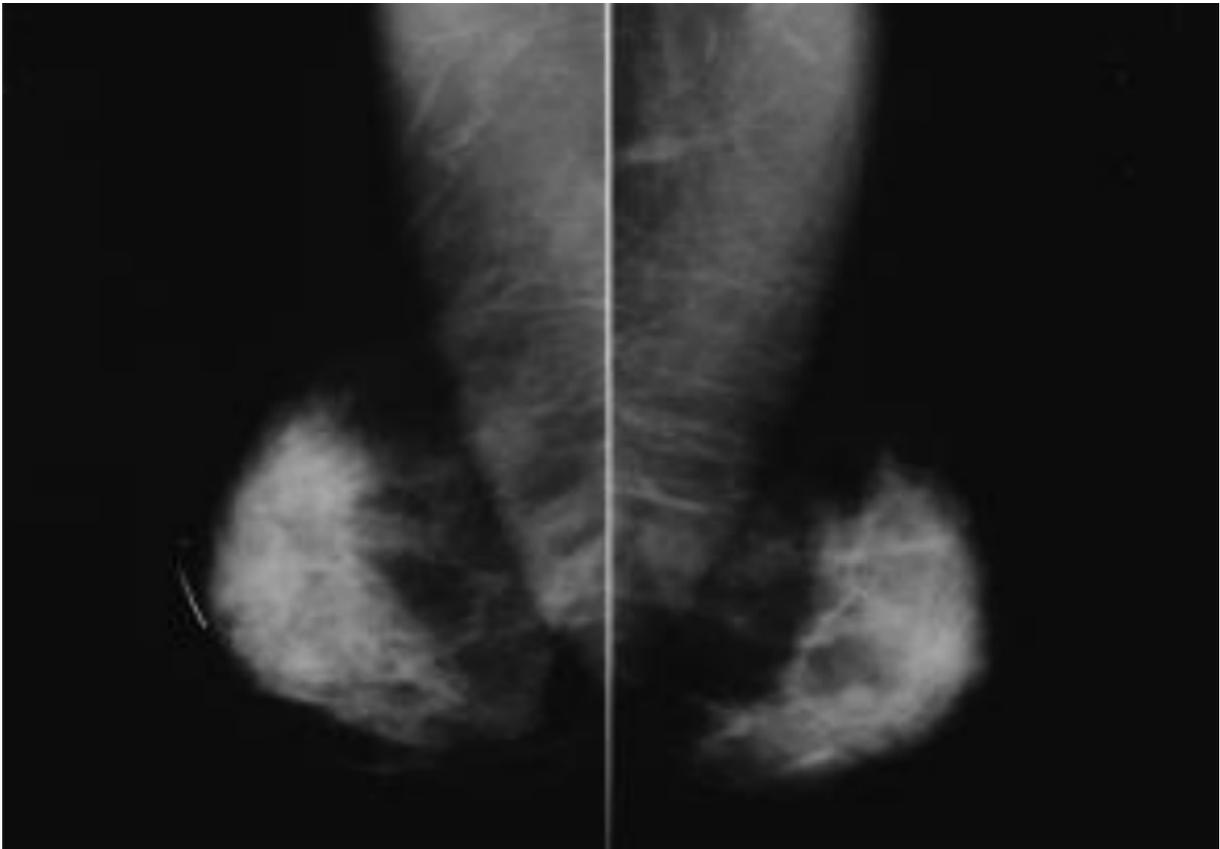
A incidência CC tem como objetivo complementar a MLO incluindo todo o material pósteromedial. Neste caso, o paciente deve estar de frente para o mamógrafo, com a cabeça virada para o lado oposto ao exame, o tubo deve estar na vertical e o feixe perpendicular à mama. Deve-se centralizar a mama com o mamilo paralelo ao filme, com a parte inferior sobre o bucky. Além disso, deve-se tracionar a mama o máximo possível antes de aplicar a compressão. A Figura 12 ilustra o exame com incidência crânio caudal e a Figura 13 apresenta as imagens resultantes dessa incidência nas mamas direita e esquerda, respectivamente.

Figura 10 – Mamografia com incidência Médio Lateral Oblíqua - MLO.



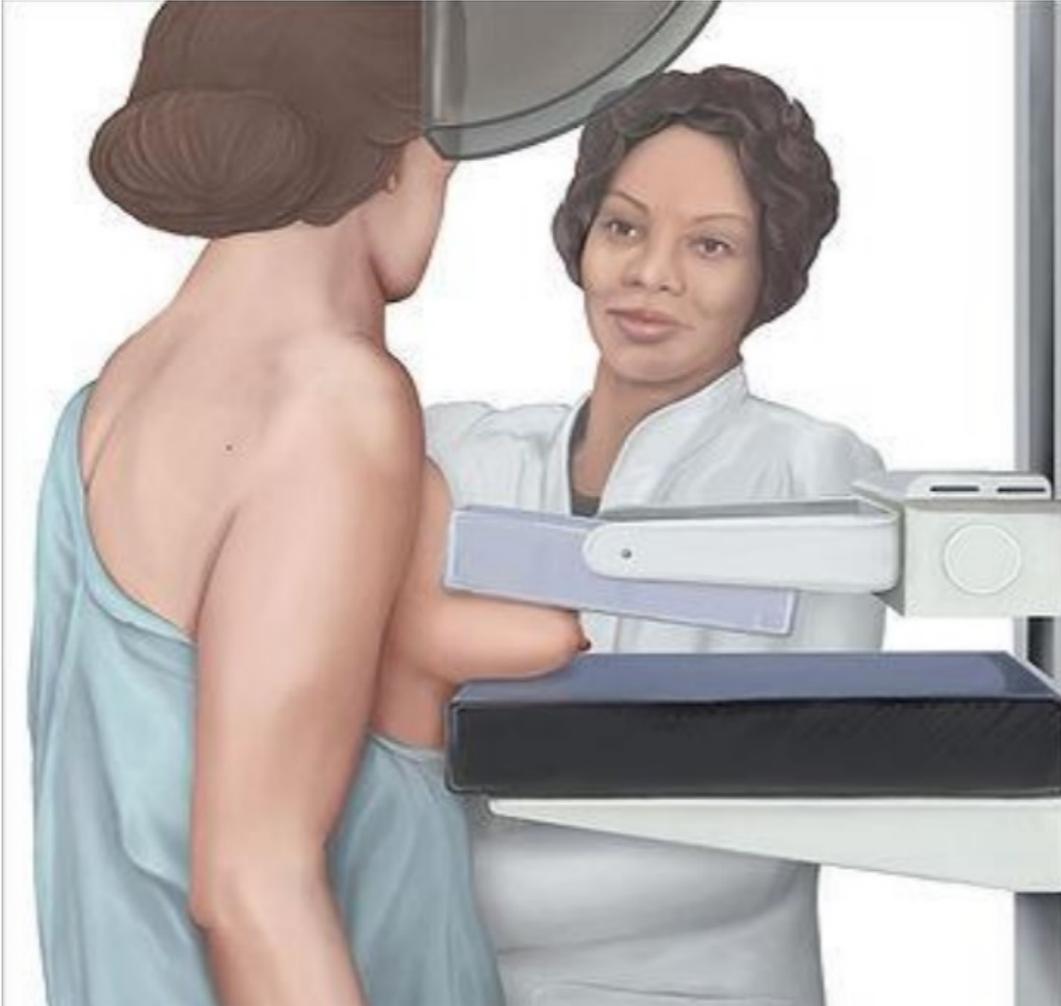
Fonte: *Centro de Educação Profissional Integrado - CIE* (CIE, 2017)

Figura 11 – Imagens de exame mamográfico com incidência Médio Lateral Oblíqua - MLO.



Fonte: *Centro de Educação Profissional Integrado - CIE* (CIE, 2017)

Figura 12 – Mamografia com incidência Crânio Caudal - CC.



Fonte: *Centro de Educação Profissional Integrado - CIE (CIE, 2017)*

2.4 Sistema Gravitacional Simplificado e Análise de Textura

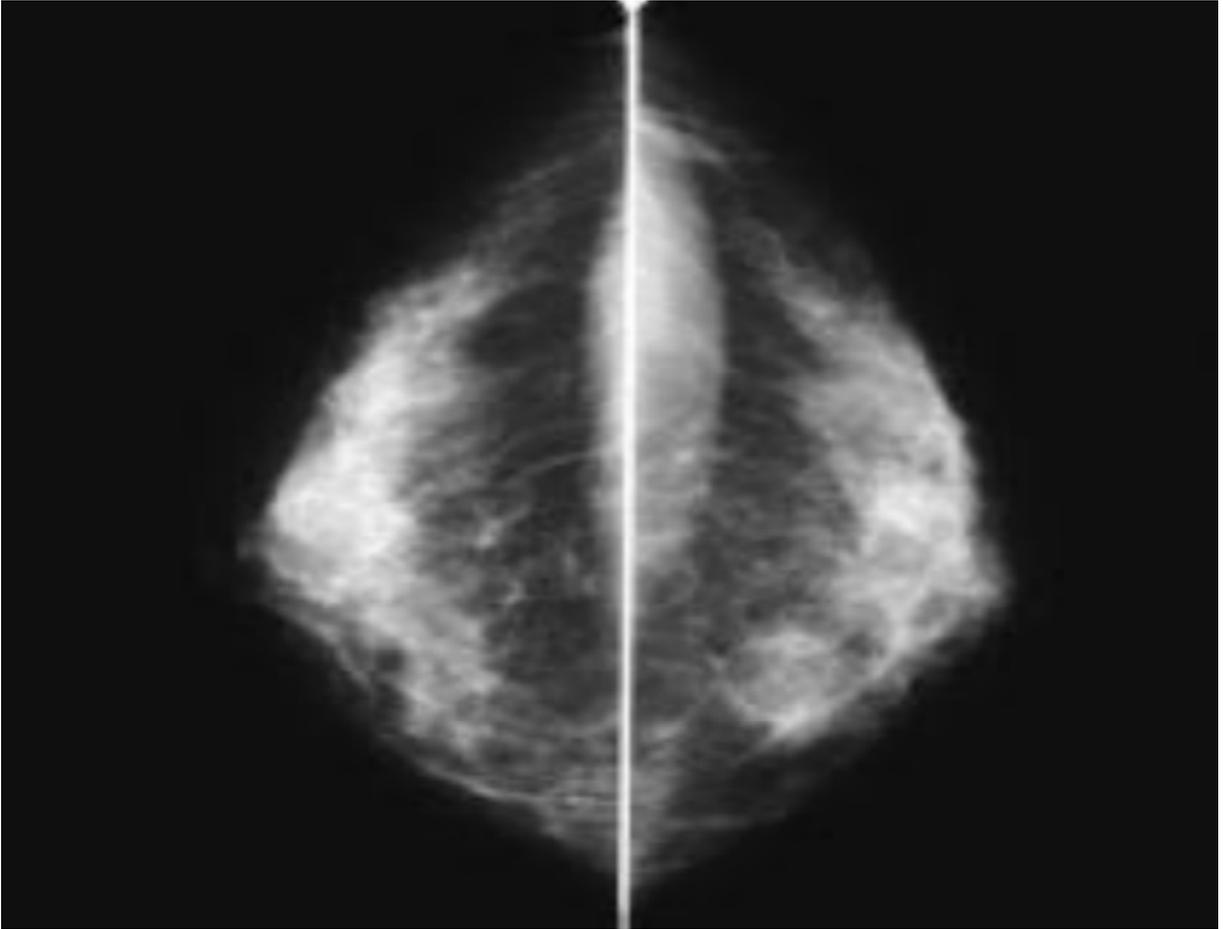
Em um sistema gravitacional, duas partículas no universo que possuem massas atraem-se com uma força diretamente proporcional ao produto de suas massas e inversamente proporcional ao quadrado da distância que as separa (NEWTON, 1999). Esta força é chamada de força gravitacional, calculada pela Equação 2.1.

$$\mathbf{f}_a = \frac{G.m_1.m_2}{\|\mathbf{r}\|^2} \cdot \frac{\mathbf{r}}{\|\mathbf{r}\|} \quad (2.1)$$

onde \mathbf{f}_a é a força gravitacional existente entre duas partículas com massas m_1 e m_2 , conectadas por um vetor \mathbf{r} . G é chamada de constante gravitacional. A Figura 14 exemplifica a força gravitacional entre duas partículas massivas.

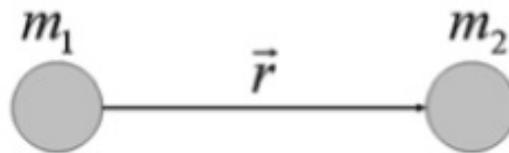
Entre partículas em movimento circular existe uma segunda força, além da força gravitacional, chamada de força centrípeta, a qual é diretamente proporcional à velocidade

Figura 13 – Imagens de exame mamográfico com incidência Crânio Caudal - CC.



Fonte: *Centro de Educação Profissional Integrado - CIE (CIE, 2017)*

Figura 14 – Ilustração de força gravitacional entre duas massas.



Fonte: (SA JUNIOR, 2013)

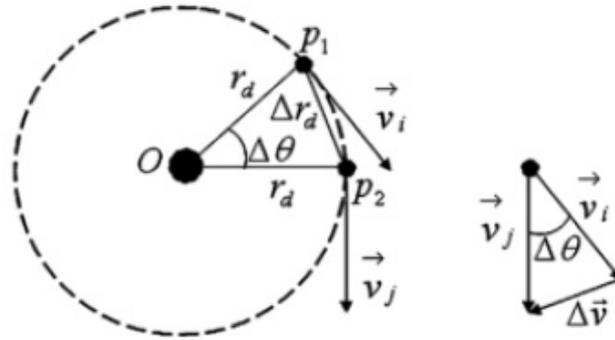
tangencial e aponta para o centro da trajetória circular descrito por um objeto. A Figura 15 ilustra a força centrípeta de uma partícula se movendo do ponto p_1 para p_2 ao redor de uma partícula O.

A força centrípeta é obtida pela Equação 2.2, onde f_c é a força centrípeta em uma partícula com massa m que se move com uma aceleração centrípeta a_c , velocidade tangencial v e raio da trajetória circular r .

$$\mathbf{f}_c = m\mathbf{a}_c = m \frac{\mathbf{v}^2}{\|\mathbf{r}\|}, \quad (2.2)$$

Para que esse conceito possa ser utilizado na análise de uma imagem com textura

Figura 15 – Ilustração de uma partícula em movimento ao redor de outra partícula.



Fonte: (SA JUNIOR, 2013)

I são necessárias algumas considerações. Cada pixel $I(x, y)$ compõe uma partícula no sistema gravitacional não equilibrado, cujo nível de cinza do pixel corresponde à massa da partícula e ele interage somente com um objeto localizado no centro da imagem de textura com massa M . A interação desse pixel com cada pixel da imagem não é levada em consideração devido aos custos computacionais (SA JUNIOR *et al.*, 2013b).

A força gravitacional para cada pixel é obtida de acordo com a Equação 2.1, substituindo a massa m_1 por M e m_2 por $I(x, y)$ (nível de cinza do pixel). A velocidade tangencial é obtida pela Equação 2.2 substituindo m por $I(x, y)$. Na determinação da velocidade deve-se considerar que velocidades muito baixas causam um colapso rápido do sistema, assim como velocidades muito altas implicam em não colapso do sistema. Dessa forma, para encontrar o intervalo de velocidades em que todos os pixels irão colapsar lentamente, estabelece-se que $\mathbf{f}_a = \mathbf{f}_c$. Com isso, a maior velocidade tangencial pode ser calculada pela Equação 2.3, onde \mathbf{v}_{max} é a velocidade máxima do pixel e \mathbf{r}_{max} a maior distância entre o pixel e o centro da imagem (massa M) que garantem que o pixel mais distante do centro da imagem irá colapsar, ou seja, todos os pixels da imagem da textura irão se aproximar do centro da imagem.

$$\mathbf{v}_{max} = \sqrt{\frac{GM}{\mathbf{r}_{max}}}, \quad (2.3)$$

Cada pixel terá sua velocidade própria, que permita a extração de informações considerando a distância e a intensidade do nível de cinza, dada pela Equação 2.4, onde \mathbf{v}_{pix} é a velocidade tangencial do pixel, $I(x, y)$ é o nível de cinza do pixel e \mathbf{v}_{max} sua velocidade máxima. Dessa forma, cada pixel tem sua trajetória particular determinada pela sua distância e

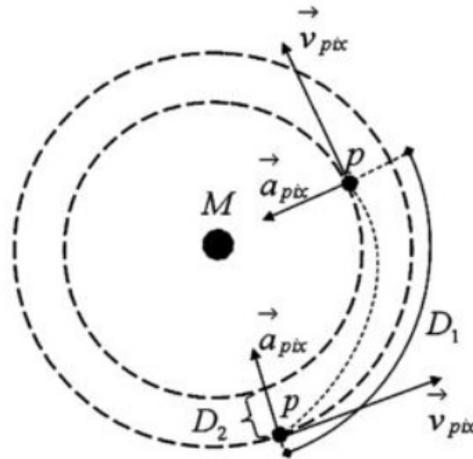
intensidade, dando à imagem uma assinatura própria (SA JUNIOR *et al.*, 2013a).

$$\mathbf{v}_{pix} = \left(1 + \frac{I(x,y)}{255}\right) \frac{\mathbf{v}_{max}}{2}, \quad (2.4)$$

Considerando as duas forças e as associações anteriores em um modelo gravitacional não equilibrado, cada pixel deverá apresentar dois tipos de movimentos. O primeiro é constante, circular, anti-horário e calculado por $S_1 = \mathbf{v}_{pix} \cdot t$, onde D_1 é a distância do deslocamento do pixel em um tempo t . O segundo movimento é retilíneo, acelerado, orientado para o centro da imagem e calculado pela equação $S_2 = (1/2) \cdot \mathbf{a}_{pix} \cdot t^2$, onde \mathbf{a}_{pix} é a aceleração do pixel orientado para o centro da imagem e dado por 2.5 e D_2 corresponde à distância percorrida pelo pixel em um tempo t . A Figura 16 apresenta o movimento de um pixel em um determinado tempo t .

$$\mathbf{a}_{pix} = \begin{cases} \frac{\|\mathbf{f}_a - \mathbf{f}_c\|}{I(x,y)}, & \text{if } I(x,y) \neq 0 \\ 0, & \text{if } I(x,y) = 0 \end{cases} \quad (2.5)$$

Figura 16 – Modelo gravitacional simplificado onde o pixel p colapsa. D_1 e D_2 determinam a nova posição do pixel.



Fonte: (SA JUNIOR; BACKES, 2012)

Durante o processo de colapso dos pixels em uma imagem de textura, no modelo gravitacional simplificado, é possível que dois ou mais pixels tentem ocupar a mesma posição. Neste caso, para reduzir a complexidade do método e preservar a informação da imagem, a posição receberá a média dos níveis de cinza dos pixels.

Em seguida, a partir das imagens colapsadas, são extraídas características que descrevem a assinatura própria de cada padrão de textura.

Vale salientar que, entre os vários atributos relevantes, a textura é um dos mais importantes por ser uma fonte rica de informações, sendo base para muitas aplicações na análise de imagens mamográficas.

Atualmente, o modelo gravitacional simplificado vem sendo utilizado para análise de texturas em taxonomia vegetal (SA JUNIOR *et al.*, 2013b) e (SA JUNIOR *et al.*, 2011), para análise de rugosidade de texturas (SA JUNIOR; BACKES, 2012) e classificação de texturas coloridas (SA JUNIOR *et al.*, 2013a).

Em (SA JUNIOR *et al.*, 2013b), o modelo gravitacional simplificado é utilizado para ampliar a capacidade de discriminação de um descritor de textura tradicional, a lacunaridade, aplicado em imagens de texturas da base Brodatz (BRODATZ, 1966).

2.5 Lacunaridade

O conceito de lacunaridade foi introduzida por Mandelbrot (MANDELBROT, 1982) com o intuito de caracterizar padrões de textura diferentes mas que apresentavam a mesma dimensão fractal. A Lacunaridade descreve uma textura de acordo com a distribuição de lacunas dispersadas ao longo da imagem, tornando-a uma medida multiescalar da heterogeneidade da textura. O cálculo da lacunaridade em imagens de níveis de cinza consiste em deslizar uma janela de tamanho r sobre a imagem de textura e determinar o mínimo u e o máximo v valor de nível de cinza dentro da janela. Considerando os valores anteriores, uma coluna com mais de uma cúbica pode ser necessária para cobrir as coordenadas de uma imagem. O tamanho relativo desta coluna é determinado por $S = [v/r] - [u/r]$. Considerando a posição possível de cada janela, calcula-se a distribuição de probabilidade $Q(S, r)$ de uma altura relativa para uma janela de tamanho r (SA JUNIOR *et al.*, 2013b). Então, a lacunaridade é obtida através da Equação 2.6.

$$\Lambda(r) = \frac{\sum S^2 \cdot Q(S, r)}{[\sum S \cdot Q(S, r)]^2} \quad (2.6)$$

A figura 17 apresenta o processo de cálculo de lacunaridade em uma imagem de textura utilizando uma janela 3x3.

2.6 Conceitos Básicos de Aprendizado de Máquina

As técnicas de aprendizado de máquina empregam um princípio de inferência denominado indução, no qual obtém-se conclusões genéricas a partir de um conjunto particular de

Figura 17 – Demonstração do processo de cálculo de lacunaridade em uma imagem.

110	117	121	119	120	122	$v = 124$ $u = 110$ $r = 3$ $S = [v/r] - [u/r]$ $= 41 - 37 = 4$
119	118	123	120	119	121	
120	124	120	125	118	121	
124	120	117	116	114	118	
123	121	125	117	118	121	
123	120	115	117	119	120	
110	117	121	119	120	122	$v = 125$ $u = 117$ $r = 3$ $S = [v/r] - [u/r]$ $= 42 - 39 = 3$
119	118	123	120	119	121	
120	124	120	125	118	121	
124	120	117	116	114	118	
123	121	125	117	118	121	
123	120	115	117	119	120	

exemplos. O aprendizado indutivo pode ser dividido em dois tipos principais: supervisionado e não-supervisionado (LORENA; CARVALHO, 2007).

No aprendizado supervisionado tem-se a figura de um professor externo, o qual apresenta o conhecimento do ambiente por conjuntos de exemplos na forma: entrada, saída desejada. O algoritmo de aprendizado de máquina extrai a representação do conhecimento a partir desses exemplos. O objetivo é que a representação gerada seja capaz de produzir saídas corretas para novas entradas não apresentadas previamente. Já no aprendizado não-supervisionado não há a presença de um professor, ou seja, não existem exemplos rotulados. O algoritmo de aprendizado de máquina aprende a representar (ou agrupar) as entradas submetidas segundo uma medida de qualidade. Essas técnicas são utilizadas principalmente quando o objetivo for encontrar padrões ou tendências que auxiliem no entendimento dos dados (LORENA; CARVALHO, 2007).

O tipo de aprendizado que será utilizado neste trabalho é o supervisionado. Nesta caso, dado um conjunto de exemplos rotulados na forma $(\mathbf{x}_i, \mathbf{y}_i)$, em que \mathbf{x}_i representa um exemplo e \mathbf{y}_i denota o seu rótulo, deve-se produzir um classificador, também denominado modelo, preditor ou hipótese, capaz de prever precisamente o rótulo de novos dados. Esse processo de indução de um classificador a partir de uma amostra de dados é denominado treinamento. O classificador obtido também pode ser visto como uma função f , a qual recebe um dado \mathbf{x} e fornece uma predição \mathbf{y} (LORENA; CARVALHO, 2007).

Os rótulos ou classes representam o fenômeno de interesse sobre o qual se deseja fazer previsões. Neste trabalho, considera-se o caso em que os rótulos (classes) assumem os valores discretos 1, 2 ou 3, que representam respectivamente as classes denso, gorduroso e

fibroglandular, constituindo-se um problema de classificação. Como o conjunto de amostras possuem mais de duas classes, configura-se um problema de multiclases. Caso contrário, quando a quantidade de classes é igual a 2, a classificação é binária.

2.6.1 *Support Vector Machine - SVM*

Em aprendizado de máquina, uma SVM são modelos de aprendizado supervisionado que analisam dados e reconhecem padrões e é utilizado para classificação e análise de regressão. O SVM padrão toma como entrada um conjunto de dados e prediz, para cada entrada dada, qual de duas possíveis classes a entrada faz parte, o que faz do SVM um classificador linear binário não probabilístico. Entretanto, em problemas multiclases, pode-se considerar uma técnica de "um-contra-um" onde cada classificador é treinado para separar as classes duas a duas. Dessa forma, a classificação é realizada tomando como base o voto majoritário (THEODORIDIS; KOUTROUMBAS, 2009).

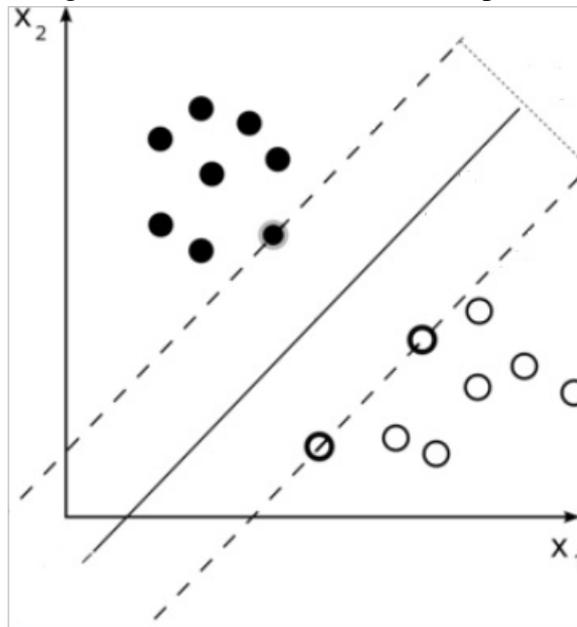
Considerando um conjunto de amostras de treinamento, cada uma marcada como pertencente a uma de duas classes, um algoritmo de treinamento do SVM constrói um modelo que atribui novas amostras a uma classe ou outra. Um modelo SVM determina uma linha de separação entre dados de duas ou mais classes, chamada de hiperplano. Essa linha busca maximizar a distância entre os pontos mais próximos em relação a cada uma das classes. A figura 18 apresenta os hiperplanos construídos pelo SVM para separação de duas classes de amostras.

Para manter o custo computacional razoável, os mapeamentos utilizados pelo SVM são projetados para garantir que os produtos escalares possam ser facilmente computados em termos das variáveis no espaço original, definindo-os em termos de uma função *Kernel* $k(\mathbf{x}, \mathbf{y})$ selecionada para se adequar ao problema (PRESS *et al.*, 2011). Alguns dos *Kernels* mais utilizados na prática são os Polinomiais, os Gaussianos ou *Radial-Basis Function* (RBF) e os Sigmoidais. Cada um deles possuem parâmetros que devem ser configurados pelo usuário (LORENA; CARVALHO, 2007). A figura 19 apresenta as funções *Kernels* mais comuns nos modelos SVM.

2.6.2 *Algoritmo K-Nearest Neighbor*

O *k-Nearest Neighbor* (KNN) é um método supervisionado de classificação de dados baseado na proximidade de seus vizinhos em um espaço amostral. Seu objetivo é formar uma

Figura 18 – Apresentação de um hiperplano e margens gerados por SVM treinado com amostras de duas classes de objetos. As amostras nas margens são chamadas de vetores suporte.



Fonte: (DECOSTE, 2002)

Figura 19 – Kernel mais comuns.

Tipo de Kernel	Função $K(\mathbf{x}_i, \mathbf{x}_j)$	Parâmetros
Polinomial	$(\delta (\mathbf{x}_i \cdot \mathbf{x}_j) + \kappa)^d$	δ, κ e d
Gaussiano	$\exp(-\sigma \ \mathbf{x}_i - \mathbf{x}_j\ ^2)$	σ
Sigmoidal	$\tanh(\delta (\mathbf{x}_i \cdot \mathbf{x}_j) + \kappa)$	δ e κ

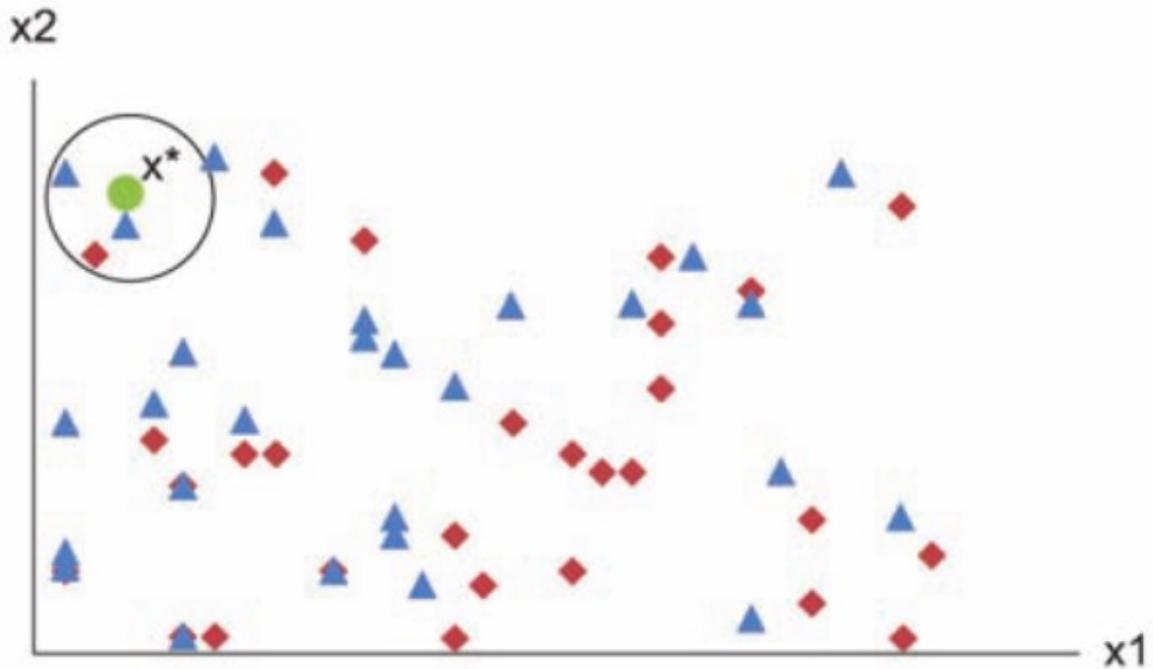
Fonte: (LORENA; CARVALHO, 2007)

generalização com base em um conjunto de treinamento, maximizando a acurácia da classificação de novos dados. O algoritmo pressupõe que o conjunto de treinamento é composto pelas variáveis descritivas e pela sua classificação; o *k-Nearest Neighbor* (Knn) então utiliza tais variáveis para classificar um novo item (MAZZILO; ANZANELLO, 2015).

Com base em um conjunto de treinamento, formado por n observações previamente classificadas, como $X = x_1, x_2, \dots, x_n$, e uma observação a ser classificada, formada por $Y = y_1, y_2, \dots, y_n$ em um espaço n -dimensional, tem-se que a distância pode ser calculada para cada ponto do conjunto de treinamento. Com base nesta distância, os k vizinhos mais próximos são identificados e, com base em um k escolhido, uma nova observação é atribuída à classe com maior número de observações em k , conforme ilustrado na Figura 20 (MAZZILO; ANZANELLO, 2015).

As distâncias entre os pontos dos conjuntos de amostras podem ser calculadas utilizando diferentes funções como por exemplo: distância euclideana, de Mahalanobis, Chebychev, Minkowski e distância euclideana padronizada.

Figura 20 – Uma observação x é testada em um conjunto treinamento com $k = 3$, formado pelas observações com menor distância euclidiana do ponto x^* . No caso apresentado, a observação x é alocada à categoria dos pontos representados pelos triângulos azuis.



Fonte: (MAZZILO; ANZANELLO, 2015)

2.7 Revisão Bibliográfica

Nos últimos anos, algumas pesquisas vêm sendo realizadas com o objetivo de classificar automaticamente a densidade do tecido mamário utilizando diferentes algoritmos de pré-processamento, segmentação, atributos e classificadores. No trabalho de Sheshadri e Kandaswamy (SHESHADRI; KANDASWAMY, 2007) foram utilizados 6 medidas estatísticas baseadas no histograma de 320 imagens do banco de dados *MIAS* (SUCKLING, 1994) conseguindo uma acurácia de 80%. Já no trabalho de Subashini *et al* (SHUBASHINI *et al.*, 2010) foram utilizadas 9 medidas estatísticas de histograma em somente 43 imagens do *MIAS*, obtendo uma acurácia média de 95,44%. Kumar *et al* (KUMAR *et al.*, 2015) obtiveram uma acurácia de 73,7% através do cálculo de descritores de textura (média, desvio padrão e energia) em imagens obtidas aplicando wavelets com filtros Haar e no trabalho de Oliveira *et al* (OLIVEIRA

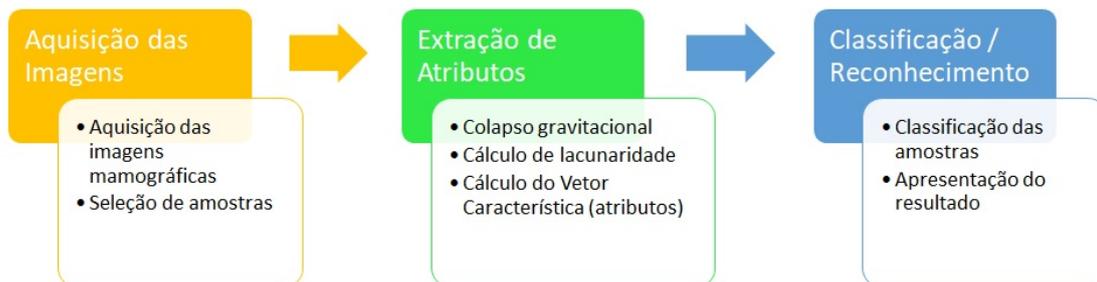
et al., 2010) foi obtida uma acurácia média na classificação de 83% utilizando *Two-Dimensional Principal Component Analysis* (2DPCA) sobre *Region of Interest* (ROI) de 300x300 pixels com validação cruzada de 10 folds.

3 MATERIAIS E MÉTODOS

A presente dissertação tem por objetivo o processamento de imagens mamográficas visando a classificação automática da densidade do tecido mamário através de técnicas de análise de textura baseadas em modelos gravitacionais simplificados, cálculo de lacunaridade e classificadores.

O processo proposto para classificação da densidade do tecido mamário é constituído por 3 etapas: aquisição das imagens, extração de atributos e classificação das densidades. Neste sentido, de posse das técnicas descritas no capítulo anterior, é realizada neste capítulo a descrição da metodologia proposta para a classificação, assim como a determinação dos parâmetros de configuração dos algoritmos utilizados. A Figura 21 apresenta a metodologia proposta com as suas etapas e atividades específicas.

Figura 21 – Metodologia proposta para classificação da densidade mamária.



3.1 Aquisição das Imagens

Os mamogramas utilizados foram obtidos do banco de imagens mamográficas *MIAS* (SUCKLING, 1994) composto por 322 imagens de 1024x1024 pixels, digitalizadas com resolução de 50 microns no formato *Portable Gray Map (PGM)* e divididas entre 112 imagens com tecido denso glandular, 104 imagens com tecido fibroglandular e 106 imagens com tecido adiposo. Para o estudo foram separadas aleatoriamente 100 imagens de cada tipo de tecido.

Inicialmente foram selecionadas manualmente amostras de imagens, com tamanho 200x200 pixels, da parte central das mamografias. As Figuras 22, 23 e 24 apresentam imagens mamográficas da base de imagens *MIAS*, ressaltando a área de seleção da amostra, e as suas respectivas amostras selecionadas para a análise.

Figura 22 – Imagem de mama densa com área de amostra destacada e apresentada ao lado.

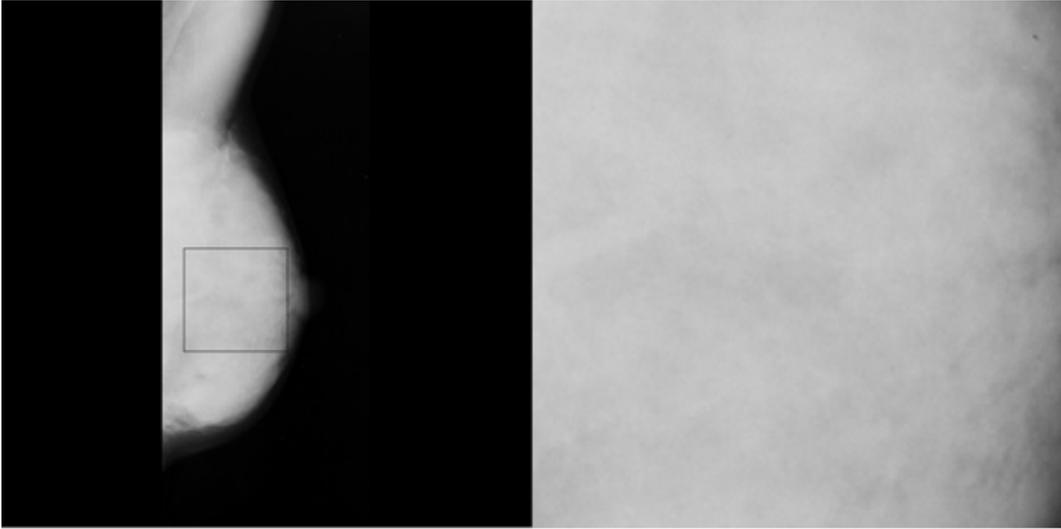
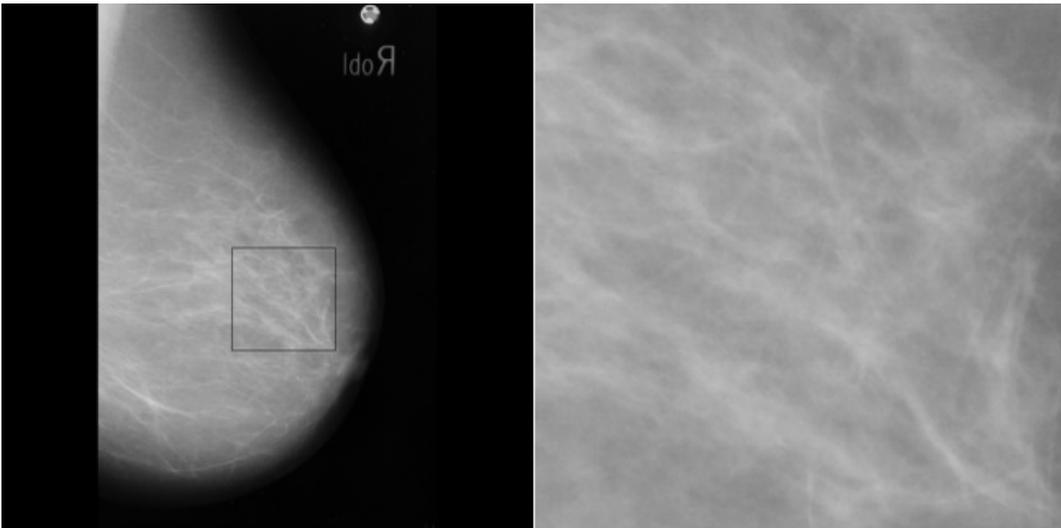


Figura 23 – Imagem de mama fibroglandular com área de amostra destacada e apresentada ao lado.



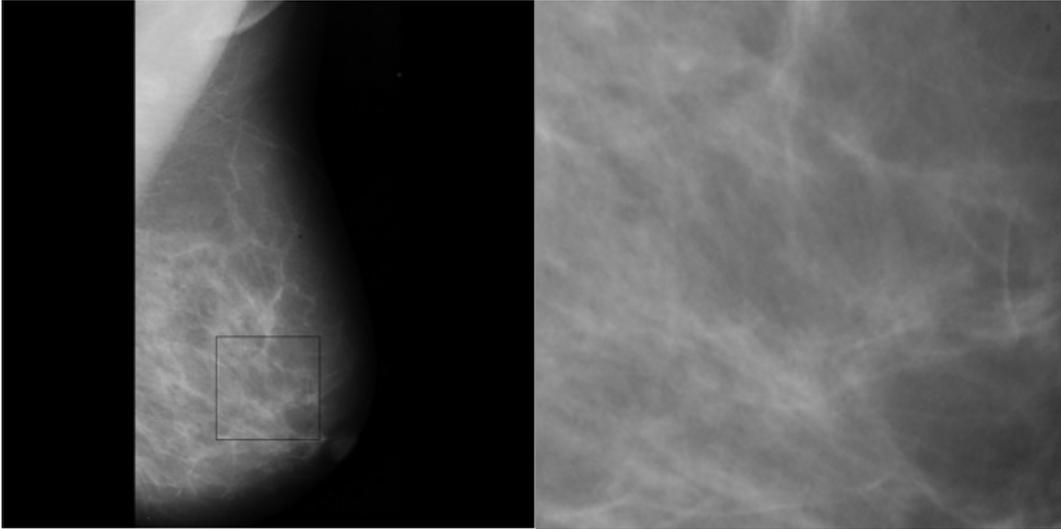
3.2 Extração de Atributos

Após a etapa de aquisição das imagens, com as amostras selecionadas, inicia-se a etapa de extração de atributos, a qual é baseada na análise de textura utilizando modelo gravitacional simplificado e cálculo recursivo de lacunaridade em um processo de colapso gravitacional.

3.2.1 *Análise de textura com Modelo Gravitacional Simplificado*

Durante o processo de colapso gravitacional o padrão de textura muda, assim como a sua rugosidade, alterando também a sua lacunaridade para cada tempo t do colapso, ou seja,

Figura 24 – Imagem de mama adiposa com área de amostra destacada e apresentada ao lado.



para cada imagem de textura que representa uma simples etapa do processo de colapso. Dessa forma, o processo de colapso permite que o padrão de textura seja caracterizado pelas variações dos seus valores de lacunaridade. Com isso, o padrão de textura é caracterizado por um vetor formado por um conjunto de medidas de lacunaridade de cada etapa do processo de colapso para uma dada janela de tamanho r , de acordo com a Equação 3.1.

$$\Psi_{t_1, t_2, \dots, t_M}(r) = [\Lambda_{t_1}(r), \Lambda_{t_2}(r), \dots, \Lambda_{t_M}(r)]. \quad (3.1)$$

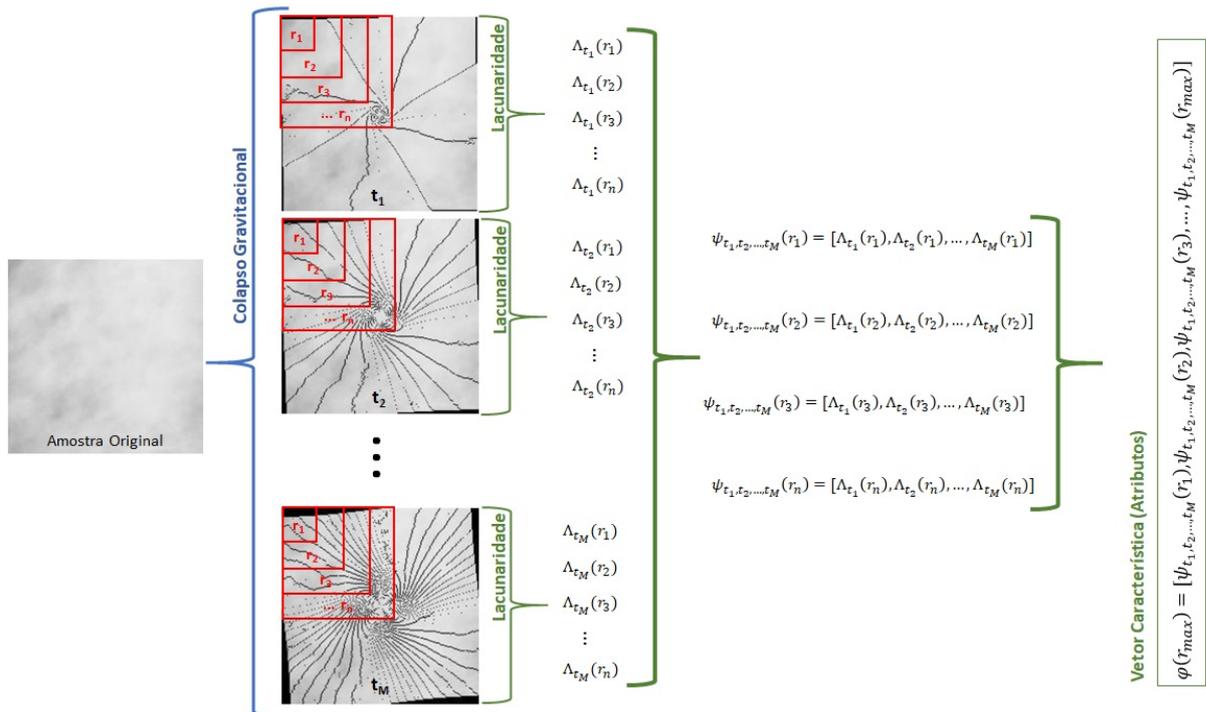
Como a lacunaridade é uma medida multi-escalar, ou seja, depende do tamanho da janela r usada no seu cálculo, é possível considerar um segundo vetor de medidas que explora essa característica. Dessa forma, pode-se calcular esse vetor de características concatenando os diferentes vetores de lacunaridade obtidos para diferentes tamanhos r de janela, ou seja:

$$\Phi(r_{max}) = [\Psi_{t_1, t_2, \dots, t_M}(2), \dots, \Psi_{t_1, t_2, \dots, t_M}(r_{max})], \quad (3.2)$$

onde r_{max} é o máximo tamanho de janela permitido (SA JUNIOR *et al.*, 2013b).

A Figura 25 ilustra a análise de textura com o cálculo de lacunaridade em cada etapa do processo de colapso gravitacional. Vale salientar que está sendo utilizado o modelo gravitacional simplificado para o processo de colapso gravitacional, ou seja, não é considerado o efeito gravitacional dos pixels vizinhos, somente da massa do centro da imagem, e é calculada a média dos pixels que colapsam para a mesma posição em um tempo de colapso.

Figura 25 – Etapa de Extração de Atributos.



3.3 Classificação / Reconhecimento da Densidade da Mama

Após o cálculo dos descritores de cada imagem, foram utilizados os algoritmos de classificação Máquina de Vetores de Suporte SVM e o K vizinhos mais próximos Knn para avaliar a capacidade de descrição das medidas de lacunaridades, obtidas através do processo de colapso gravitacional, de acordo com o modelo gravitacional simplificado adequado para as imagens mamográficas.

Para tal, foram utilizadas 60 amostras (imagens de 200x200 pixels) de cada tipo de densidade de tecido para realizar o treinamento supervisionado dos classificadores. Em seguida, foram realizados testes com outras 40 amostras de cada tipo de densidade de tecido, não pertencentes ao conjunto de imagens utilizadas no treinamento.

Na classificação utilizando SVM foram testadas diferentes configurações. Neste caso, variaram-se os parâmetros de constante gravitacional (G), massa central (M) e manteve-se fixo o parâmetro de tempo (t).

No caso do classificador Knn, diferentes parâmetros de configuração do modelo gravitacional foram testados, como foi feito anteriormente no SVM, e testaram-se diferentes configurações do classificador como: número de vizinhos e a função do cálculo de distância (Euclidiana, Mahalanobis, Chebychev, Minkowski) do vetor de característica da amostra para os

vetores características de cada classe de densidade de tecido.

3.4 Determinação dos Parâmetros de Configuração

A determinação dos parâmetros de configuração da metodologia proposta, que mais se adequam para a classificação automática da densidade de tecidos mamários pela análise da textura desse tipo de tecido, ou seja, os parâmetros que produzem a melhor classificação, consiste na definição das variáveis do modelo gravitacional simplificado e dos classificadores.

A validação de um algoritmo computacional é uma etapa importante, principalmente no caso de imagens médicas onde a importância dessa avaliação é ressaltada por se tratar de uma imagem que caracteriza o estado de saúde de um ser humano.

Para validar os parâmetros do algoritmo proposto foi realizada uma análise quantitativa dos resultados finais da classificação baseada em duas métricas, a acurácia e a especificidade. A acurácia indica a proporção de tecidos classificados corretamente. Já a especificidade mostra a proporção de negativos de tecidos identificados corretamente (DHEEBA *et al.*, 2014). As Equações 3.3 e 3.4 apresentam as fórmulas para o cálculo da acurácia e da especificidade respectivamente.

$$Acuracia = \frac{TP}{(TP + FN)} \quad (3.3)$$

$$Especificidade = \frac{TN}{(TN + FP)} \quad (3.4)$$

Onde, TP é quantidade de tecidos classificados positivamente corretamente (*True Positive*); FN a quantidade de tecidos classificados negativos erroneamente (*False Negative*); TN a quantidade de tecidos classificados negativamente corretamente (*True Negative*) e FP a quantidade de tecidos classificados positivamente erroneamente (*False Positive*).

O processo de determinação dos parâmetros de configuração consiste em: separar dois conjuntos de amostras, um para treinamento e outro para teste; definir os valores inerentes ao modelo gravitacional simplificado; calcular os atributos (vetor característica) de todas as amostras; realizar um treinamento supervisionado do classificador com os vetores de características referentes ao conjunto de amostras de treinamento; testar os classificadores com os vetores de características referentes ao conjunto de amostras de testes; calcular a acurácia e a

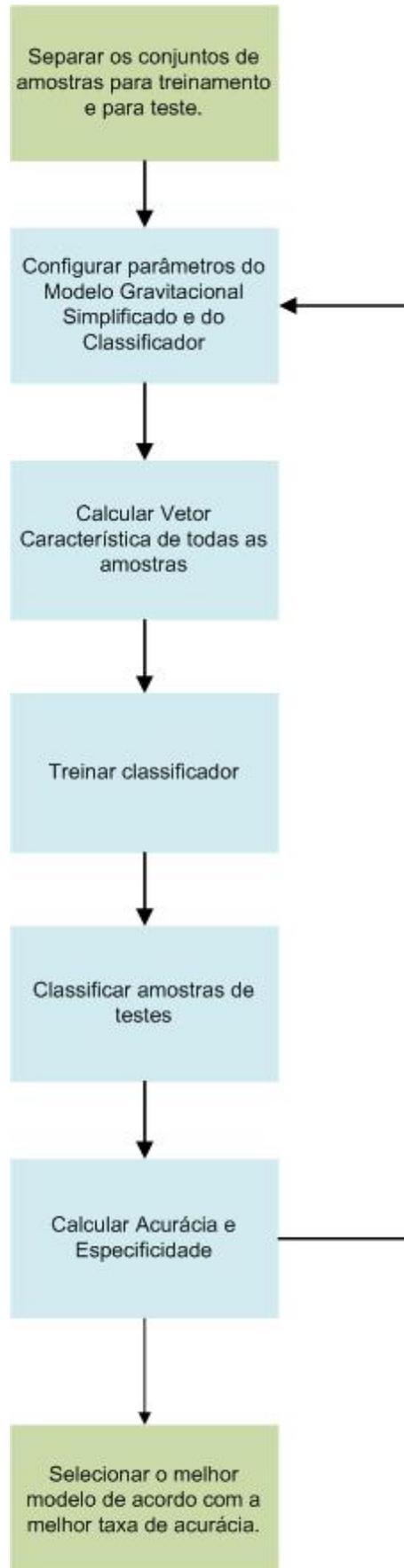
especificidade do sistema; definir novos parâmetros de configuração do modelo gravitacional, repetindo o resto das etapas até encontrar um valor máximo de acurácia. A Figura 26 ilustra o processo de determinação dos parâmetros de configuração exposto anteriormente.

Para definir o modelo gravitacional simplificado para este tipo de imagens mamográficas, precisa-se determinar os parâmetros da constante gravitacional G , o tamanho da janela r_{max} , os tempos t de cada etapa do processo de colapso e o valor da massa central M .

Depois de determinado o modelo gravitacional simplificado mais adequado para as imagens mamográficas, é realizado o processo de colapso gravitacional de cada imagem do banco *MIAS*, durante o qual foram calculados os vetores de características concatenando os vetores de lacunaridades obtidos em cada etapa do processo de colapso.

Dessa forma, cada imagem é descrita (representada) por vetores de características, oriundos da análise de textura pelo cálculo da lacunaridade em cada etapa do processo de colapso gravitacional.

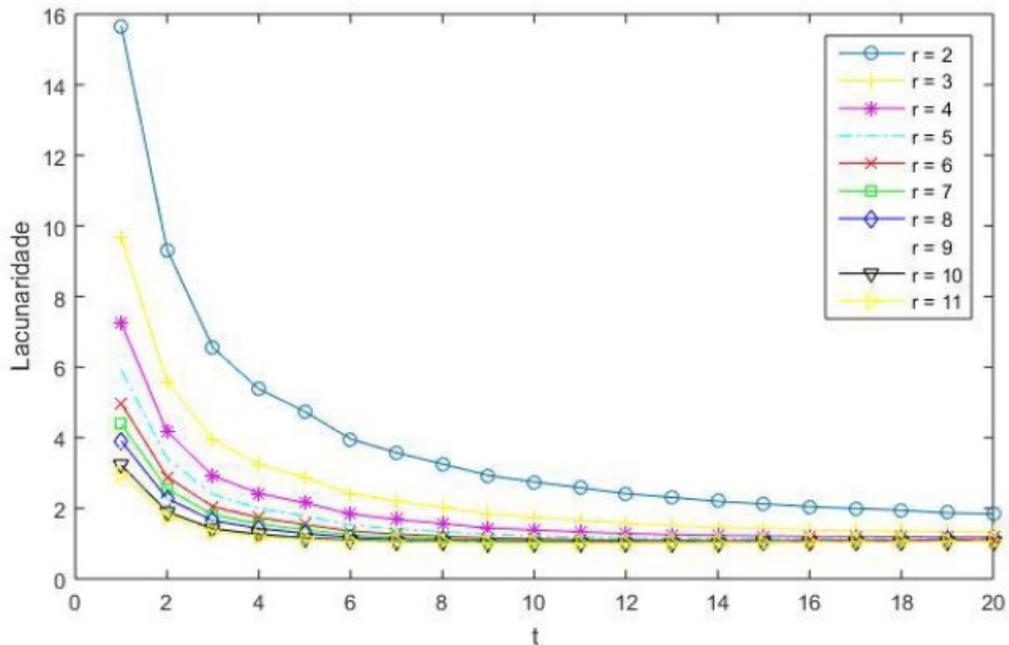
Figura 26 – Determinação dos parâmetros de configuração do Modelo Gravitacional Simplificado.



4 RESULTADOS

Inicialmente considera-se $M = 500$, constante gravitacional $G = 1$ e calcula-se a lacunaridade usando diferentes tamanhos de janela r nas imagens colapsadas. O valor de $G = 1$ não limita o passo do deslocamento da partícula no processo de colapso gravitacional. A figura 27 apresenta o gráfico dos valores de lacunaridade obtidos.

Figura 27 – Lacunaridades para uma amostra de textura de imagem mamográfica usando $t = 1, 2, \dots, 20$ e tamanho de janelas $r = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$.

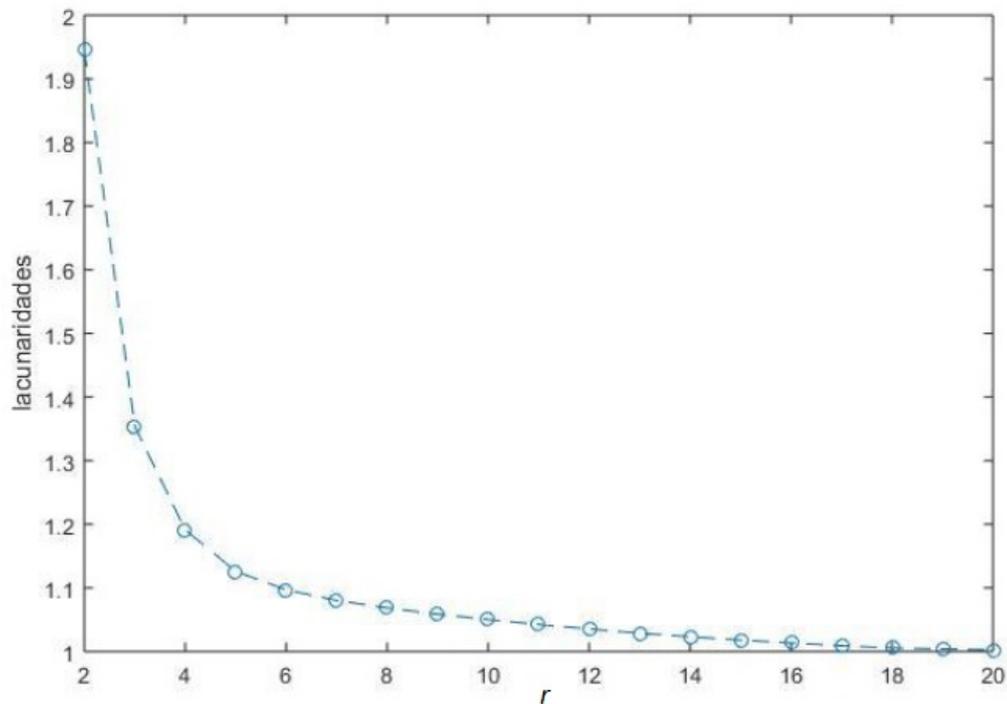


Observando o Gráfico (27), nota-se que para um tamanho de janela r o valor de lacunaridade pouco muda para cada tempo t . Dessa forma, pode-se definir o melhor conjunto de tempo t do modelo gravitacional simplificado para definir o vetor de lacunaridades, evitando uma grande quantidade de informações redundantes que compromete a habilidade de descrição das texturas. De acordo com o Gráfico 27 observa-se que o melhor conjunto é $t = 1, 3, 6, 12$ devido às maiores variações dos valores de lacunaridade. A partir do tempo $t = 12$, os valores de lacunaridade pouco variam, trazendo informações redundantes.

Considerando que a medida de lacunaridade é dependente do tamanho da janela r é necessário definir também o tamanho de janela máximo r_{max} para evitar informações redundantes. Para tal, calculam-se os valores de lacunaridade de uma amostra considerando diferentes tamanhos de janela r . A Figura 28 apresenta o gráfico obtido.

A partir do Gráfico 28, observa-se que, para $r \geq 12$, os valores de lacunaridade são

Figura 28 – Lacunaridades de uma amostra de tecido de imagem mamográfica usando diferentes tamanhos de janela r .



similares. A inclusão destes valores no vetor de características iria comprometer a habilidade de descrição do padrão de textura devido à uma grande quantidade de informações redundantes. Com isso, os tamanhos de janela r a serem considerados no vetor de lacunaridades é $r = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

Considerando os tempos $t = 1, 3, 6, 12$, $G = 1$ e uma amostra de 200×200 pixels retirada do centro de uma imagem mamográfica da base de dados *MIAS* (SUCKLING, 1994), as Figuras 29 e 30 apresentam as imagens colapsadas para um centro de massa $M = 500$ e $M = 750$, respectivamente. O centro de massa maior acelera o processo de colapso gravitacional.

Foram avaliados diferentes modelos gravitacionais, de acordo com as variações dos parâmetros apresentados na Tabela 1.

Tabela 1 – Modelos gravitacionais avaliados

Modelo Gravitacional	G	M	t
GRAV1	1	500	[1 3 6 12]
GRAV2	1	750	[1 3 6 12]
GRAV3	1	1000	[1 3 6 12]
GRAV4	1	900	[1 3 6 12]
GRAV5	1	800	[1 3 6 12]
GRAV6	2	750	[1 3 6 12]

Figura 29 – Imagem original; amostra 200x200, imagens colapsadas com $M=500$, $G=1$ e $t = 1, 3, 6, 12$, respectivamente.

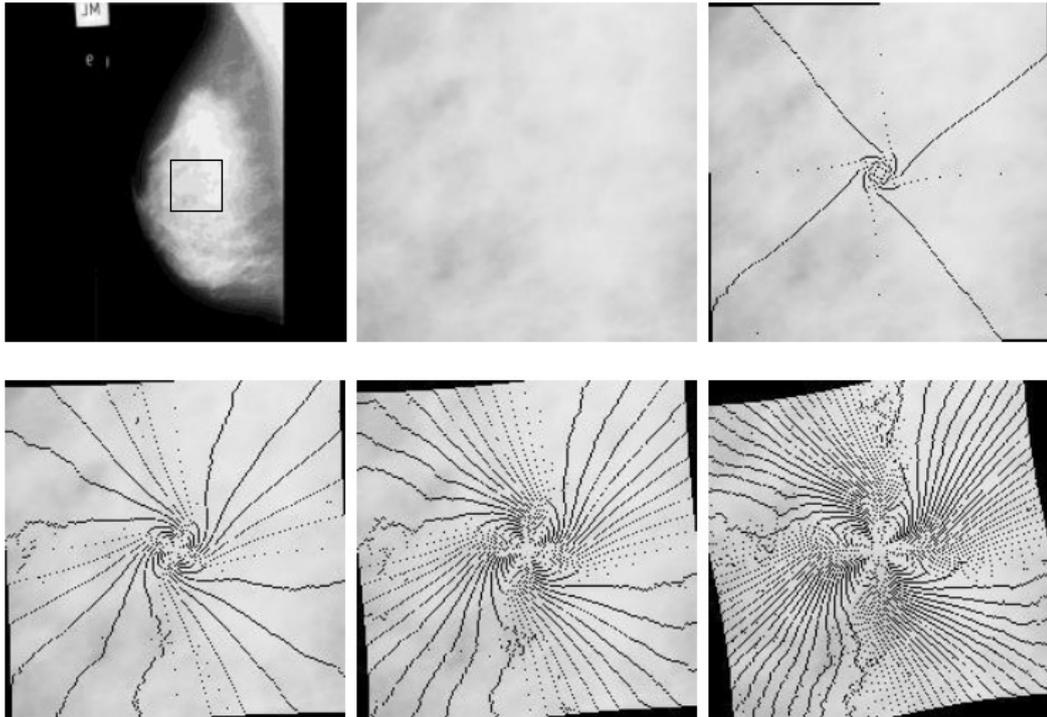
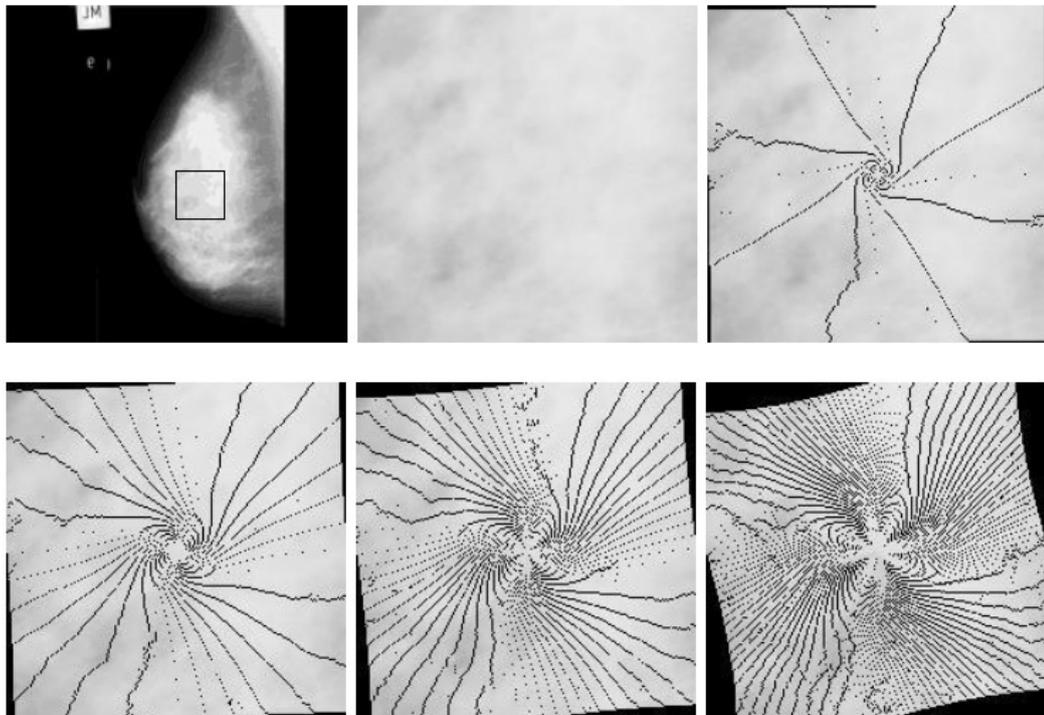


Figura 30 – Imagem original; amostra 200x200, imagens colapsadas com $M=750$, $G=1$ e $t = 1, 3, 6, 12$, respectivamente.



Considerando os modelos gravitacionais apresentados na Tabela 1, utiliza-se o classificador SVM para avaliar a capacidade de descrição de cada um dos vetores características de lacunaridades desses modelos, através da acurácia. A Tabela 2 apresenta a acurácia do

classificador SVM de acordo com as variações dos modelos gravitacionais simplificados.

Tabela 2 – Acurácia do classificador SVM de acordo com os modelos gravitacionais simplificados avaliados.

Modelo Gravitacional	Acurácia %
GRAV1	75,00
GRAV2	76,70
GRAV3	69,17
GRAV4	75,83
GRAV5	73,33
GRAV6	69,17

Além do classificador SVM, foi utilizado também o classificador Knn para classificar a densidade dos tecidos da mama considerando os vetores características de lacunaridades obtidos dos modelos gravitacionais apresentado na Tabela 1. Neste caso, foram testados diferentes parâmetros de número de vizinhos e função de cálculo de distância. A Tabela 3 apresenta as acurácias obtidas para cada modelo gravitacional e parâmetros de configuração do classificador Knn.

Tabela 3 – Acurácia do classificador Knn de acordo com os modelos gravitacionais simplificados avaliados e considerando diferentes tamanhos de vizinhança (2, 10, 15, 18 e 20).

Modelo Gravitacional	Euclidean			Mahalanobis			Chebychev			Minkowski			Seuclidean				
	2	10	15	2	10	15	2	10	15	2	10	15	2	10	15	18	
GRAV1	62,50	65,83	64,17	62,50	39,17	47,50	49,17	59,17	60,00	63,33	62,50	65,83	64,17	66,67	70,83	71,67	66,67
GRAV2	65,00	70,00	65,83	69,17	48,33	53,33	43,33	58,33	64,17	61,67	65,00	70,00	65,83	61,67	70,83	71,67	75,00
GRAV3	65,00	62,50	66,67	62,50	44,17	47,50	44,17	59,17	59,17	65,83	65,00	62,50	66,67	58,33	68,33	71,67	73,33
GRAV4	60,00	64,17	65,00	67,50	45,00	45,00	55,83	58,33	64,17	63,33	60,00	64,17	65,00	61,67	70,83	67,50	69,17
GRAV5	64,17	68,33	70,83	71,67	44,17	57,50	59,17	64,17	67,50	65,00	64,17	68,33	70,83	61,67	70,00	66,67	68,33
GRAV6	61,67	65,00	63,33	66,67	50,83	55,83	60,83	58,33	61,67	60,00	61,67	65,00	63,33	59,17	65,83	70,00	73,33

A figura 31 apresenta exemplos de amostras de tecido denso com as suas respectivas classificações obtidas utilizando o classificador SVM combinado com o modelo gravitacional GRAV2.

A figura 32 apresenta exemplos de amostras de tecido gorduroso com as suas respectivas classificações obtidas utilizando o classificador SVM combinado com o modelo gravitacional GRAV2.

A figura 33 apresenta exemplos de amostras de tecido fibroglandular com as suas respectivas classificações obtidas utilizando o classificador SVM combinado com o modelo gravitacional GRAV2.

A figura 34 apresenta exemplos de amostras de tecido denso com as suas respectivas classificações obtidas utilizando o classificador Knn combinado com o modelo gravitacional

Figura 31 – Classificador SVM: (a) Tecido Denso classificado corretamente. (b) Tecido Denso classificado como Fibroglandular. (c) Tecido Denso classificado como Gorduroso.

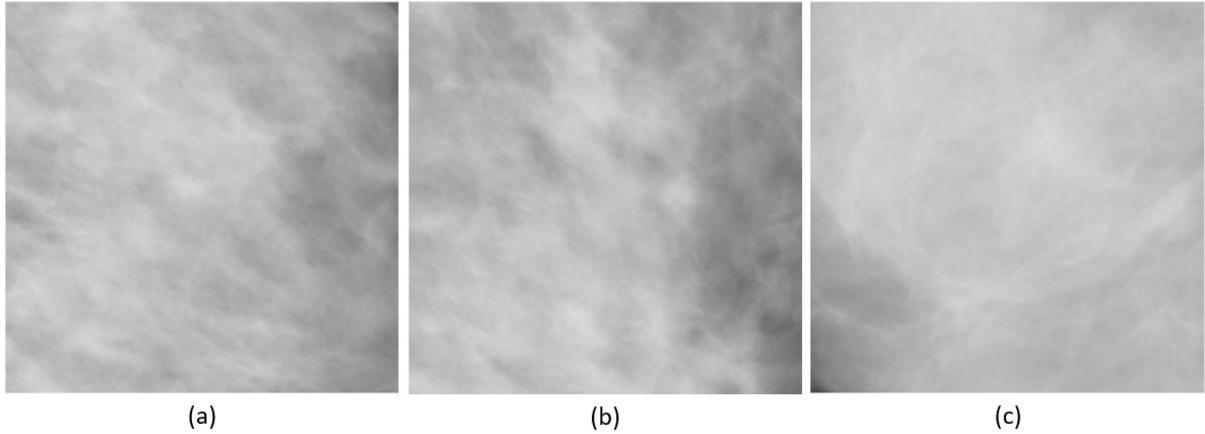


Figura 32 – Classificador SVM: (a) Tecido Gorduroso classificado corretamente. (b) Tecido Gorduroso classificado como Denso. (c) Tecido Gorduroso classificado como Fibroglandular.

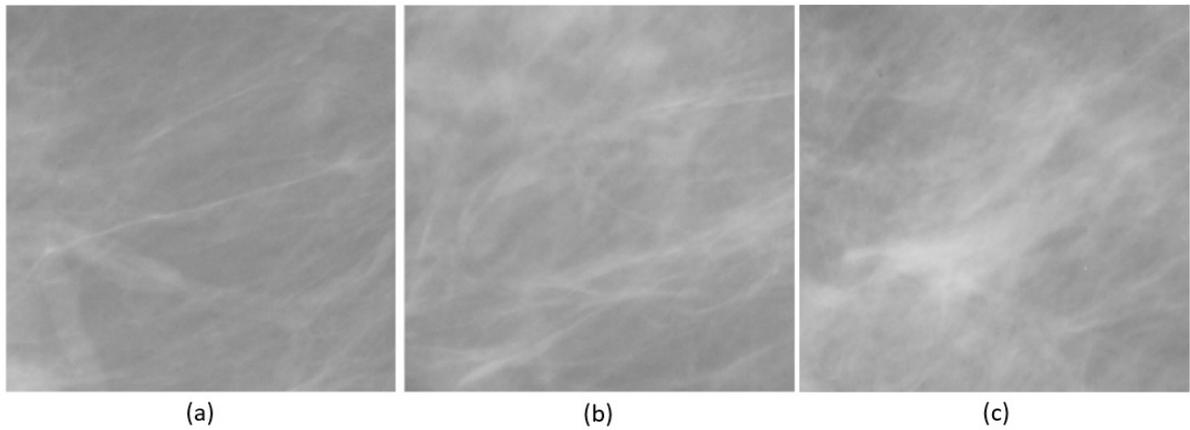
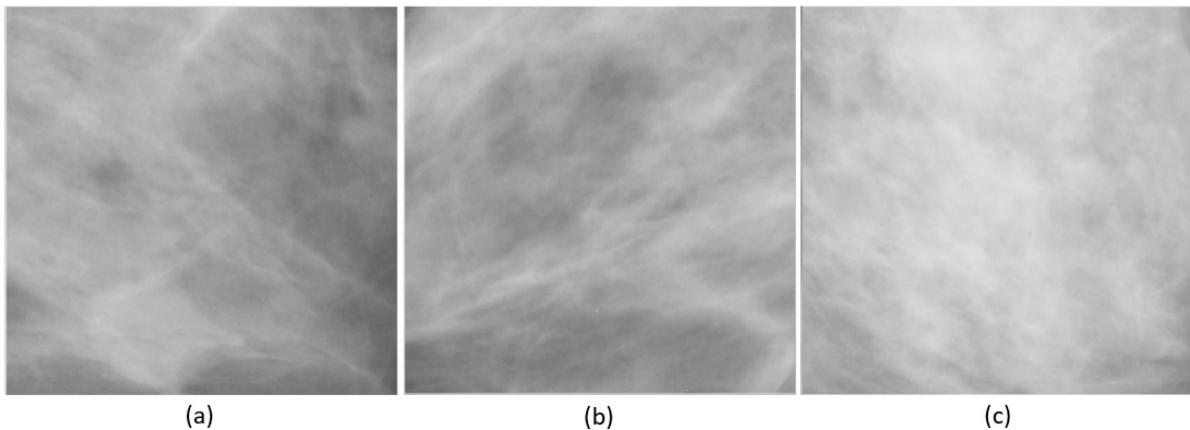


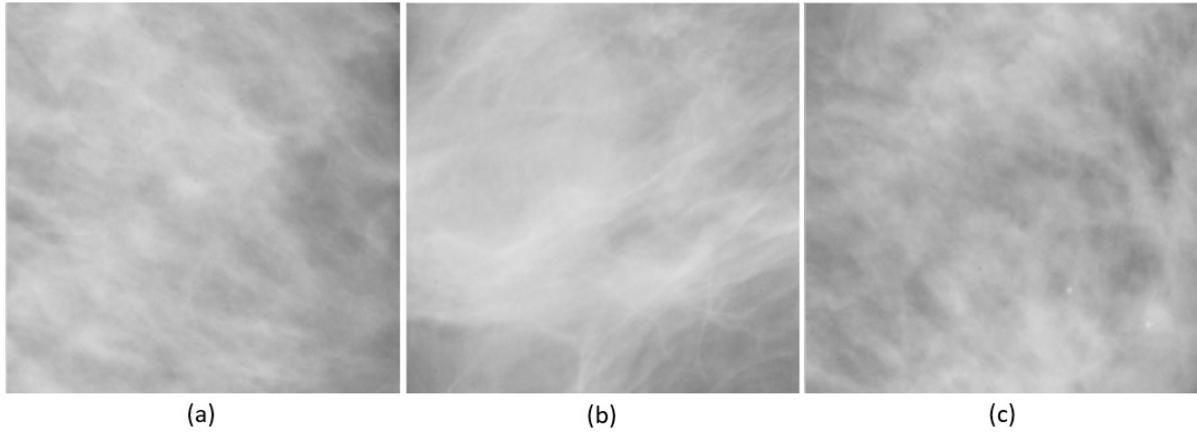
Figura 33 – Classificador SVM: (a) Tecido Fibroglandular classificado corretamente. (b) Tecido Fibroglandular classificado como Denso. (c) Tecido Fibroglandular classificado como Gorduroso.



GRAV2.

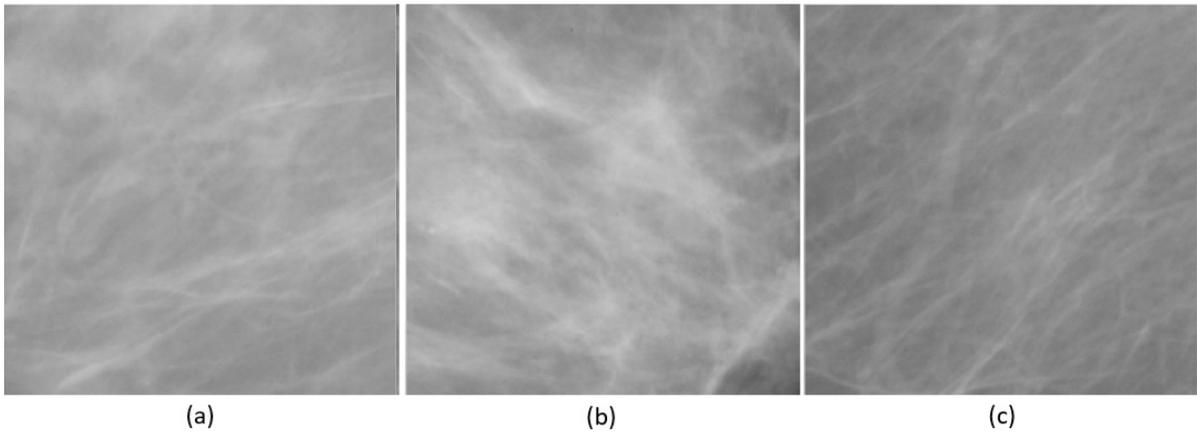
A figura 35 apresenta exemplos de amostras de tecido gorduroso com as suas respectivas classificações obtidas utilizando o classificador Knn combinado com o modelo

Figura 34 – Classificador Knn: (a) Tecido Denso classificado corretamente. (b) Tecido Denso classificado como Gorduroso. (c) Tecido Denso classificado como Fibroglandular.



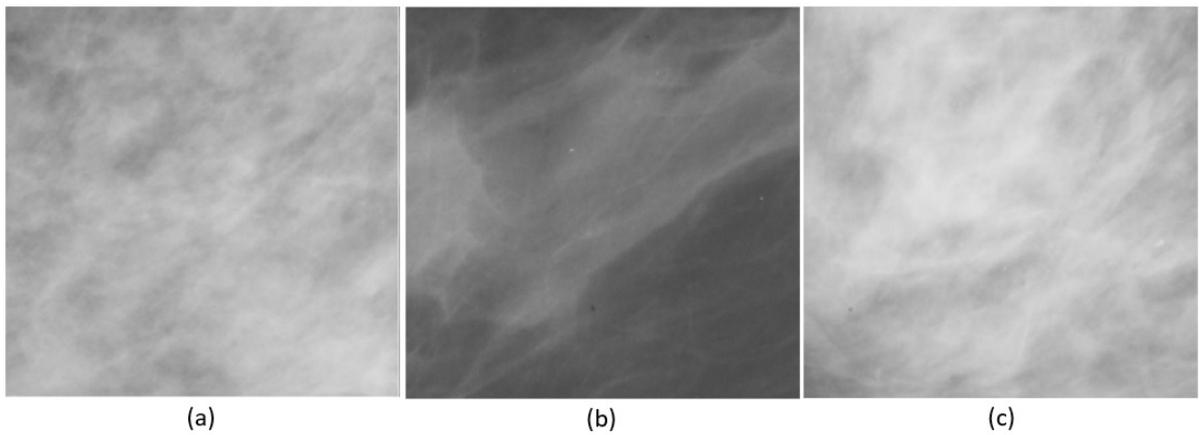
gravitacional GRAV2.

Figura 35 – Classificador Knn: (a) Tecido Gorduroso classificado corretamente. (b) Tecido Gorduroso classificado como Denso. (c) Tecido Gorduroso classificado como Fibroglandular.



A figura 36 apresenta exemplos de amostras de tecido fibroglandular com as suas respectivas classificações obtidas utilizando o classificador Knn combinado com o modelo gravitacional GRAV2.

Figura 36 – Classificador Knn: (a) Tecido Gorduroso classificado corretamente. (b) Tecido Gorduroso classificado como Denso. (c) Tecido Gorduroso classificado como Fibroglandular.



5 CONCLUSÕES E TRABALHOS FUTUROS

Nesta dissertação desenvolve-se uma aplicação do modelo gravitacional simplificado para a análise de textura de imagens mamográficas voltada para a classificação automática da densidade do tecido mamário. Para tal, foram calculadas medidas de lacunaridade durante o processo de colapso gravitacional para compor um vetor de atributos capaz de descrever os diferentes tipos de textura das amostras. Em seguida, para classificar os três diferentes padrões de textura presentes no banco de dados *MIAS* foram testados dois classificadores o SVM e o Knn.

Com o intuito de analisar a performance dos sistemas de classificação automática da densidade dos tecidos mamários foi calculado a acurácia e a especificidade para cada tipo de densidade de tecido. A Tabela 4 apresenta a acurácia e a especificidade obtidas para cada tipo de tecido, considerando somente as duas configurações de classificadores com as melhores acurácias, o modelo gravitacional 2 (GRAV2) combinado com SVM e o modelo gravitacional 2 (GRAV2) com o classificador Knn configurado com vizinhança 18 e distância euclideana padrão.

A partir da análise da Tabela 4 pode ser observado que os classificadores SVM e Knn propostos possuem melhores acurácia e especificidade para a densidade de tecido gorduroso, com a diminuição da performance na distinção dos tecidos denso e fibroglandular. Tal fenômeno acontece, devido à semelhança das imagens de mamografia referentes às duas densidades. Nas imagens onde há tecido fibroglandular, as texturas associadas às glândulas se assemelham à textura de um tecido mais denso, dificultando a sua classificação.

Através da matriz de confusão para o classificador SVM, apresentada na tabela 5, é possível observar que o sistema de classificação tem dificuldade para distinguir entre tecidos densos e fibroglandulares, mas pode realizar a distinção entre o tecido gorduroso e os outros tipos de tecido de forma conveniente.

Observando a matriz de confusão para o classificador Knn, apresentada na tabela 6, é possível constatar que o sistema de classificação consegue distinguir melhor as amostras

Tabela 4 – Performance dos algoritmos de classificação para cada tipo de densidade de tecido de acordo com a acurácia e a especificidade.

Tecido	SVM		Knn	
	Acurácia	Especificidade	Acurácia	Especificidade
Denso	57,50	95,83	80,00	82,86
Gorduroso	90,00	90,32	75,00	93,75
Fibroglandular	82,50	75,64	70,00	81,58

Tabela 5 – Matriz de Confusão para o classificador SVM.

		Classes Previstas		
		Denso	Gorduroso	Fibroglandular
Classes Reais	Denso	23	1	16
	Gorduroso	1	36	3
	Fibroglandular	2	5	33

Tabela 6 – Matriz de Confusão para o classificador Knn.

		Classes Previstas		
		Denso	Gorduroso	Fibroglandular
Classes Reais	Denso	32	1	7
	Gorduroso	3	30	7
	Fibroglandular	9	3	28

com tecido denso, entretanto apresenta uma maior dificuldade para identificar as amostras com tecidos gordurosos e fibroglandulares, o qual reflete em uma menor acurácia se comparado ao classificador SVM.

Observando as Tabelas de acurácia 2 e 3, referentes aos resultados do classificador SVM e do Knn, respectivamente, além da tabela de performance 4, vê-se que as melhores taxas de acerto 76,70% e 75,00% foram obtidas utilizando o modelo gravitacional GRAV2. Dessa forma, pode-se concluir que este é o melhor modelo gravitacional para ser utilizado no processo de colapso e cálculo de lacunaridades, ou seja, $G = 1$, $M = 750$, tempo $t = [1 \ 3 \ 6 \ 12]$ e tamanhos de janela $r = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

Com base nos resultados descritos no capítulo anterior, é possível concluir que o método proposto, com uma taxa de acurácia de cerca de 77% e especificidade média de 87,27%, obteve uma taxa de acerto próxima às obtidas nos trabalhos de Sheshadri e Kandaswamy (SHESHADRI; KANDASWAMY, 2007) e Oliveira *et al* (OLIVEIRA *et al.*, 2010), de 80% e 83%, respectivamente; superior à acurácia obtida no trabalho de Kumar (KUMAR *et al.*, 2015) de 73,7% e muito inferior ao trabalho de Subashini *et al* (SHUBASHINI *et al.*, 2010), a qual obteve uma acurácia média de 95,44%. Entretanto, neste último trabalho, foram analisadas somente 43 imagens do *MIAS*.

Vale ressaltar também que esta é a primeira abordagem realizada utilizando modelo gravitacional simplificado para análise de textura de tecidos presentes em imagens mamográficas, havendo perspectivas de realizar mudanças no modelo para adaptá-lo melhor a essa finalidade através do uso de outros descritores, combinados ou não com o descritor de lacunaridade.

5.1 Contribuições

Esta dissertação gerou como contribuição, de acordo com um levantamento bibliográfico prévio, uma aplicação inédita do modelo gravitacional simplificado aplicado para a análise de texturas de imagens mamográficas, até então utilizado principalmente em imagens para análise de taxonomia vegetal.

Além disso, a dissertação teve como contribuição o artigo **Modelo Gravitacional Simplificado na Análise de Textura de Imagens Mamográficas**, aceito no VII Congresso Latino-Americano de Engenharia Biomédica - CLAIB 2016 em Bucaramanga - Colômbia. Segue no anexo A a carta da organização do evento informando o aceite do artigo no referido congresso.

5.2 Sugestões para trabalhos futuros

Com as técnicas apresentadas nos capítulos anteriores obtivemos bons resultados. No entanto, ainda temos várias propostas e técnicas a serem implementadas, como trabalhos futuros:

1. Aplicar técnicas de pré-processamento de imagens como a equalização adaptativa de histograma com contraste limitado, utilizado por Costa em (COSTA *et al.*, 2008);
2. Utilizar outros descritores para análise de textura combinados à lacunaridade como medidas de energia da textura de Laws, energia de textura estatística, energia espectral e entropia, dentre outras;
3. Utilizar imagens de outros bancos de imagens como o do *Digital Database for Screening Mammography* (DDSM);
4. Aplicar modelos gravitacionais simplificados para a detecção e classificação de nódulos e microcalcificações;
5. Pesquisar e desenvolver algoritmos específicos de detecção de nódulos para cada tipo de tecido, a partir da definição automática da densidade do tecido mamário, potencializando o diagnóstico.
6. Pesquisar e desenvolver algoritmo otimizado para definição automática do passo, do centro de massa (M) e do tamanho de janela r .

Vale ressaltar que os sistemas de auxílio ao diagnóstico computadorizados objetivam

melhorar a interpretação das imagens, reduzindo o número de falsos negativos. Entretanto, em todos os casos, a decisão do diagnóstico final e do tratamento mais adequado é da equipe clínica.

REFERÊNCIAS

- ARANTES, C. K.; FILIPE, M. S.; MUNIZ, A. K. M.; WAKABAYASHI, W.; OLIVEIRA, M. S.; TAJARA, H. H.; RONDINA, J. M. **Projeto Internet e Saúde - Câncer de Mama**. São José do Rio Preto: [S. n.], 2004. Disponível em: <http://www.famerp.br/projis/grp08/estrutura.html>. Acesso em: 01 jul. 2016.
- BRODATZ, P. **Textures: a photographic album for artists and designers**. [S. l.]: Dover, New York, 1966.
- CARESTREAM. **Fabricante de Equipamentos Médicos**. 2017. Disponível em: <http://www.carestream.com.br>. Acesso em: 01 mai. 2017.
- CIE. **Centro de Educação Profissional Integrado**. 2017. Disponível em: <http://www.cintegrado.com.br>. Acesso em: 01 mai. 2017.
- COSTA, G. G.; PEREIRA, W. C. A.; ALVARENGA, A. V.; INFANTOSI, A. F. C. Detecção automática de nódulos em mamografias utilizando diferença de gaussianas e morfologia matemática. In: **21 Congresso Brasileiro de Engenharia Biomédica**. BA, BR: Sociedade Brasileira de Engenharia Biomédica, 2008. p. 1447–1450.
- DECOSTE, D. Training invariant support vector machines. **Machine Learning**, v. 46, p. 161–190, 2002.
- DHEEBA, J.; SINGH, N. A.; SELVI, S. T. Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. **Journal of Biomedical Informatics**, v. 49, p. 45–52, 2014.
- FURQUIM, T. A. C. **Processo de Produção de raios X**. 2017. Disponível em: <http://rle.dainf.ct.utfpr.edu.br/hipermidia/images/fisica/rx.pdf>. Acesso em: 01 mai. 2017.
- HO, J.; JAFFERJEE, N.; COVARRUBIAS, G.; GHESANI, M.; HANDLER, B. Dense breasts: a review of reporting legislation and available supplemental screening options. **American Journal of Roentgenology**, v. 203, n. 2, p. 449–456, 2014.
- JALES, R. M. **Mamografia - Bases Físicas**. Campinas, SP: [S. n.], 2017. Disponível em: <http://www.fcm.unicamp.br/drpixel/pt-br/metodos-de-imagem/mamografia-bases-fisicas>. Acesso em: 01 mai. 2017.
- KHAN, S. A. **Diseases of the breast**. Filadélfia: Lippincot Williams and Wilkins, 2014.
- KOPANS, D. B. **Breast Imaging**. Filadélfia: Lippincott-Raven Publishers, Philadelphia, 1998.
- KUMAR, I.; BHADAURIA, H. S.; VIRMANI, J. Wavelet packet texture descriptors based four-class birads breast tissue density classification. **Procedia Computer Science**, v. 70, p. 76–84, 2015.
- LORENA, A. C.; CARVALHO, A. C. P. L. F. Uma introdução às support vector machines. **Revista de Informática Teórica Aplicada**, v. 14, p. 43–67, 2007.
- MANDELBROT, B. **The Fractal Geometry of Nature**. [S. l.]: Freeman, San Francisco, 1982.

MAZZILO, C. A.; ANZANELLO, M. J. Sistemática de seleção de variáveis para classificação de produtos em categorias de modelos de reposição. **Gestão e Produção**, v. 22, n. 1, p. 201–212, 2015.

MEDICALEXPO. **Equipamentos Médicos**. 2017. Disponível em: <http://www.medicalexpo.com>. Acesso em: 01 mai. 2017.

MELLO-THOMS, C. Perception of breast cancer: Eye-position analysis of mammogram interpretation. **Academic Radiology**, Elsevier, v. 10, 2003.

NEWTON, I. **Philosophiae Naturalis Principia Mathematica**. [S. l.]: University of California, 1999.

OLIVEIRA, J. E. E. D.; MACHADO, A. M. C.; CHAVEZ, G. C.; LOPES, A. P. B.; DESERNO, T. M.; ARAÚJO, A. Mammosys: A content-based image retrieval system using breast density patterns. **Computer Methods and Programs in Biomedicine**, v. 99, n. 3, p. 289–297, 2010.

PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W. T.; FLANNERY, B. P. **Numerical Recipes: The Art of Scientific Computing**. [S. l.]: Cambridge University Press, New York, 2011.

SA JUNIOR, J. J. M. **Novos métodos de análise de texturas baseadas em modelos gravitacionais simplificados e caminhos mais curtos em grafos**. 2013. Dissertação (Doutorado em Engenharia de Teleinformática) – Departamento de Engenharia de Teleinformática da UFC, Fortaleza, 2013.

SA JUNIOR, J. J. M.; BACKES, A. R. A simplified gravitational model to analyze texture roughness. **Pattern Recognition**, Elsevier, v. 45, n. 2, p. 732–741, 2012.

SA JUNIOR, J. J. M.; BACKES, A. R.; CORTEZ, P. C. Color texture classification based on gravitational collapse. **Pattern Recognition**, Elsevier, v. 46, n. 6, p. 1628–1637, 2013.

SA JUNIOR, J. J. M.; BACKES, A. R.; CORTEZ, P. C. A simplified gravitational model for texture analysis. **Journal of Mathematical Imaging and Vision**, Springer, v. 47, p. 70–78, 2013.

SA JUNIOR, J. J. M.; BACKES, A. R.; ROSSATO, D. R.; KOLB, R. M.; BRUNO, O. M. Measuring and analyzing color and texture information in anatomical leaf cross sections: an approach using computer vision to aid plants species identification. **Botany**, v. 89, n. 7, p. 467–479, 2011.

SANTOS, V. T. **Segmentação de imagens mamográficas para detecção de nódulos em mamas densas**. Dissertação (Mestrado em Engenharia Elétrica) – Escola de Engenharia de São Carlos - USP, São Carlos, 2002.

SCHILITZ, A.; LIMA, F. Silva de; ANDRADE, J.; OLIVEIRA, J.; SANTOS, M.; REBELO, M.; CUNHA, R.; REIS, R. **Estimativa 2016 - Incidência de Câncer no Brasil**. Rio de Janeiro: Instituto Nacional de Câncer José Alencar Gomes da Silva, 2015.

SHESHADRI, H. S.; KANDASWAMY, A. Experimental investigation on breast tissue classification based on statistical feature extraction of mammograms. **Computerized Medical Imaging and Graphics**, v. 31, n. 1, p. 46–48, 2007.

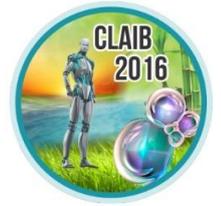
SHUBASHINI, T.; RAMALINGAM, V.; PALANIVEL, S. Automated assesment of breast tissue density in digital mammograms. **Computer Vision and image Understanding**, v. 114, n. 1, p. 33–43, 2010.

SUCKLING, J. **The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica**. [S. l.]: International Congress Series, 1994.

THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern Recognition**. [S. l.]: Academic Press, Burlington, 2009.

YAGHJYAN, L.; COLDITZ, G. A.; COLLINS, L. C. Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to tumor characteristics. **National Cancer Institute**, v. 103, n. 15, p. 1179–1189, 2011.

ANEXO A – E-MAIL ACEITE DO ARTIGO NO CLAIB 2016



Bucaramanga, Colômbia, Septiembre 1 de 2016

Autores

Daniel Freitas Colaco
Paulo Cesar Cortez
Jarbas Joaci de Mesquita Sá Júnior
André Ricardo Backes

Assunto: Aceitação de trabalho no 332 Congresso CLAIB 2016

Ilustres Autores

Por meio desta declaração nós permitimos informar-lhes que seu trabalho: "Modelo Gravitacional Simplificado na Análise de Textura de Imagens Mamográficas" tem sido aceitado para ser apresentado no Congresso Latino-americano de Engenharia Biomédica - CLAIB 2016, que será realizado desde 26 de Outubro até 28 de outubro 2016, nas instalações da Universidade Autônoma de Bucaramanga, Colômbia.

Sua aceitação confirma a inclusão de seu trabalho no evento e nas memórias com ISBN. Sim em agosto 30 de 2016, temos o escrito que atenda todas as regras associadas ao Comitê editorial Verlag Springer decidirá sobre sua inclusão nas memórias indexadas.

O modo de apresentação será informado via e-mail dentro de um prazo razoável antes do evento.

Em nome do coral, ABIOIN e com o apoio da IFMBE e IEEE-EMBS, nos queremos parabenizá-los por sua aceitação neste evento científico.

Senhores Autores, com consideração e apreço, estamos ansiosos por a sua presença no CLAIB 2016 em Bucaramanga, Colombia.

ISNARDO TORRES
Chair CLAIB 2016

DANIEL A. SIERRA
Director Publicaciones y
Documentación CLAIB 2016

JOHN BUSTAMANTE
Vice-Chair CLAIB 2016