

UNIVERSIDADE FEDERAL FLUMINENSE

Rodrigo Carvalho Serrano

**ANÁLISE DA VIABILIDADE DO USO DO
COEFICIENTE DE HURST E DA LACUNARIDADE
NO AUXÍLIO AO DIAGNÓSTICO PRECOCE DE
PATOLOGIAS DA MAMA**

NITERÓI

2010

UNIVERSIDADE FEDERAL FLUMINENSE

Rodrigo Carvalho Serrano

**ANÁLISE DA VIABILIDADE DO USO DO
COEFICIENTE DE HURST E DA LACUNARIDADE
NO AUXÍLIO AO DIAGNÓSTICO PRECOCE DE
PATOLOGIAS DA MAMA**

Dissertação de Mestrado submetida ao Programa
de Pós-Graduação em Computação da
Universidade Federal Fluminense como requisito
parcial para a obtenção do título de
Mestre. Área de concentração: Computação
Visual e Interfaces

Orientadora:

Prof^a. Aura Conci, D.Sc.

Co-Orientadora:

Prof^a. Regina Leal, D.Sc.

NITERÓI

12 de março de 2010

Ficha Catalográfica elaborada pela Biblioteca da Escola de Engenharia e Instituto de Computação da UFF

S487 Serrano, Rodrigo Carvalho.
Análise da viabilidade do uso do coeficiente de Hurst e da lacunaridade no auxílio ao diagnóstico precoce de patologias da mama / Rodrigo Carvalho Serrano. – Niterói, RJ : [s.n.], 2010.
76 f.

Dissertação (Mestrado em Computação) - Universidade Federal Fluminense, 2010.
Orientadores: Aura Conci, Regina Lea, Regina Leal.

1. Processamento de imagem. 2. Imagem infravermelha. 3. Diagnóstico por imagem. 4. Coeficiente de Hurst. 5. Geometria fractal. 6. Doenças da mama. I. Título.

CDD 006.37

Análise da viabilidade do uso do coeficiente de Hurst e da Lacunaridade no auxílio ao diagnóstico precoce de patologias da mama

Rodrigo Carvalho Serrano

Dissertação de Mestrado submetida ao Programa
de Pós-Graduação em Computação da
Universidade Federal Fluminense como requisito
parcial para a obtenção do título de
Mestre. Área de concentração: Computação Visual.

Aprovada por:

Profa. Dra. Aura Conci / IC-UFF
(Orientadora)

Profa. Dra. Regina Célia Paula Leal Toledo / IC-UFF
(Co-orientadora)

Profa. Dra. Rita de Cássia Fernandes de Lima / UFPE

Prof. Dr. Aristófanes Corrêa Silva / UFMA

Prof. Dr. Raul Queiroz Feitosa / PUC-Rio

Niterói, 12 de março de 2010

Agradecimentos

A Deus por sempre iluminar os meus caminhos.

À minha mãe, Maria da Guia, por ser um exemplo de vitória, por todos os seus ensinamentos e por não ter medido esforços para me dar uma boa educação.

À minha madrinha, Raquel, por sempre incentivar os meus estudos.

À Bruna por todo apoio, ajuda, incentivo e compreensão.

À minha orientadora, Prof^a. Aura, por toda a contribuição para o meu desenvolvimento ao longo deste curso, pela paciência, atenção e incentivo.

Aos amigos Carlos Roberto e Flávio Mendonça pelos conselhos em momentos difíceis.

Aos amigos Leonardo Motta, Cintia Caetano e Mônica Batista pelas parcerias nas disciplinas que cursamos juntos.

Ao Prof. Dr. Jorge Araújo e a Prof^a Dra. Clícia Valladares pelas orientações durante minha graduação.

À Prof^a Regina pela ajuda ao longo do mestrado.

Ao Rafael Melo e ao Marcus Araújo (UFPE) pelas dúvidas esclarecidas por e-mail.

Ao CNPq pela bolsa concedida para a realização deste trabalho.

Resumo

A aplicação da termografia na mastologia origina-se da constatação de que na região onde há um câncer de mama a temperatura da pele pode ser maior do que a temperatura do tecido normal. Também o sangue venoso responsável por drenar o câncer possui temperatura superior a do arterial. Com isso, se há grande assimetria na distribuição das temperaturas das mamas, então geralmente uma delas pode apresentar algum tipo de patologia. Esta dissertação apresenta uma análise da viabilidade do uso do Coeficiente de Hurst e da Lacunaridade no auxílio ao diagnóstico precoce de patologias da mama. A primeira etapa da metodologia proposta é a segmentação da região de interesse e reamostragem desta. Esta etapa dá origem a duas imagens: mama direita (MD) e mama esquerda (ME). A partir destas uma nova imagem (imagem Subtração (S)) é gerada pelo módulo da subtração pixel a pixel da imagem da mama esquerda com a imagem da mama direita rotacionada em 180° em relação a um eixo vertical. Janelas móveis de tamanhos variáveis ($w = 5, 7, 9, 11, 13$ e 15) percorrem pixel a pixel cada uma das três imagens criadas computando os valores do Coeficiente de Hurst (H). Para cada tamanho de janela é calculada para cada uma dessas imagens (MD, ME e S) a média e o desvio padrão de H. Sendo assim, trinta e seis características são extraídas utilizando o coeficiente de Hurst. A partir da avaliação da Lacunaridade é feita a extração de outras noventa e sete características.

Através de “gliding boxes” de tamanho s (com $2 \leq s \leq 25$) são calculadas a lacunaridade da mama direita ($\Lambda R(s)$), mama esquerda ($\Lambda L(s)$) e da imagem Subtração ($\Lambda S(s)$) dando origem a setenta e duas características. Vinte e quatro características são obtidas a partir de $|\Lambda R(s) - \Lambda L(s)|$ com $2 \leq s \leq 25$. O desvio padrão dessas últimas fornece mais uma característica. Finalmente, as cento e trinta e três características extraídas (divididas em quatorze grupos) foram analisadas por técnicas de aprendizado de máquina para a classificação das imagens. A metodologia apresentou excelentes resultados para alguns grupos, enquanto que para outros não. Por exemplo, o classificador “Naive Bayes” obteve área ROC “receiver operating characteristic” de 0.958.

Palavras chaves: Termografia, Imagens térmicas, Imagens médicas, Geometria fractal, Coeficiente de Hurst, Lacunaridade, Reconhecimento de padrões.

Abstract

The use of the thermography in the mastology is related with the fact that region with breast cancer can present temperatures higher than that of the normal tissues. Moreover, the blood responsible for draining the cancer tissues has temperature higher than other veins or arteries. Then, if there is great asymmetry in the distribution of temperatures of the breasts, probably the patient has some type of pathology. This dissertation presents a study about the feasibility of using the Hurst coefficient and Lacunarity to diagnosis breast diseases in early stages. The first step is the segmentation and resampling of the thermal image. This step gives us two images: right (RB) and left breast (LB). From these images is generated a new image (Subtraction image (S)) by the module of the subtraction pixel by pixel in the image of the left breast with the image of the right breast rotated 180 ° on a vertical axis. A square pixel-by-pixel movable window is used to compute the values of Hurst coefficient. Such computation is accomplished of various window sizes ($w = 5, 7, 9, 11, 13$ and 15). For each image (RB, LB and S) is calculated the average and standard deviation of Hurst computed values in each window w . Then, thirty six features are extracted using Hurst Coefficient. From the evaluation of Lacunarity is extracted ninety-seven features. Using gliding boxes of size s (with $2 \leq s \leq 25$) the lacunarity of right breast ($\Lambda R(s)$), left breast ($\Lambda L(s)$), and Subtraction image ($\Lambda S(s)$) are calculated generating seventy-two features. Twenty four features are obtained by $|\Lambda R(s) - \Lambda L(s)|$ with $2 \leq s \leq 25$. Additional feature is provided by the standard deviation of these values. Then, the one hundred and thirty-three features extracted (divided in fourteen groups) are used by machine learning techniques to classify the images. It presented excellent results for some groups, but not all. For example, a group using Naive Bayes obtained an incredible high receiver operating characteristic area of 0.958.

Key words: Thermography, Thermal imaging, Medical imaging, Fractal geometry, Pattern Recognition, Hurst Coefficient, Lacunarity.

Lista de Figuras

Figura 2.1 Estrutura da mama-----	3
Figura 2.2 Termografia paciente com cisto-----	5
Figura 2.3 Ultrassom de uma paciente com cisto-----	5
Figura 2.4 Mamografia de uma paciente com cisto-----	6
Figura 2.5 Mamografia de uma paciente com fibroadenoma-----	7
Figura 2.6 Ultrassom de uma paciente com fibroadenoma-----	7
Figura 2.7 Termografia de uma paciente com fibroadenoma-----	8
Figura 2.8 Carcinoma em mamografia crânio caudal-----	9
Figura 2.9 Carcinoma em um ultrassom-----	9
Figura 2.10 Carcinoma em termografia-----	10
Figura 2.11 Paciente sendo submetida a um exame de mamografia-----	15
Figura 2.12 Exemplos de mamografias crânio caudais e médio-laterais--	16
Figura 2.13 Paciente sendo submetida a um exame de ultrassom-----	17
Figura 2.14 Exame de ressonância magnética-----	17
Figura 2.15 Paciente saudável-----	21
Figura 2.16 Paciente com patologia (carcinoma) -----	21
Figura 2.17 Mamografia de mama menos densa (a) e mais densa (b)---	22
Figura 3.1 Etapas do reconhecimento de padrões-----	24
Figura 3.2 Segmentação da região de interesse da paciente saudável---	25
Figura 3.3 Segmentação da região de interesse da paciente com patologia-----	25
Figura 3.4 Região de interesse da Figura 3.2 com 100 x 100 pixels-----	25
Figura 3.5 Região de interesse da Figura 3.3 com 100 x 100 pixels-----	26
Figura 3.6 Imagem Subtração da paciente saudável e da paciente com patologia-----	26
Figura 4.1 Janela de 5 x 5 pixels com parte de uma imagem em níveis de cinza-----	31
Figura 4.2 Distância euclidiana-----	31
Figura 4.3 Distância “city block” -----	31
Figura 4.4 Distância “sup” -----	32
Figura 4.5 Gráfico de $\ln \Delta g \times \ln d$ -----	33

Figura 4.6 Reta obtida através do método dos mínimos quadrados-----	33
Figura 4.7 Esquema da primeira análise-----	34
Figura 4.8 Esquema da segunda análise-----	35
Figura 4.9 Imagem de 3 x 3 pixels em tons de cinza-----	36
Figura 4.10 Imagem da Figura 4.9 em 3D-----	37
Figura 4.11 Exemplo de AABB-----	37
Figura 4.12 Passos para a extração de 48 características-----	39
Figura 4.13 Passos para a extração de 24 características da imagem Subtração-----	40
Figura 4.14 Gráfico de Lacunaridade típico de uma paciente saudável---	40
Figura 4.15 Gráfico de Lacunaridade típico de uma paciente com patologia-----	41
Figura 4.16 Passos para a extração de 24 características-----	42
Figura 4.17 Passos para a extração da característica DP-----	42
Figura 4.18 Todas as características obtidas através da Lacunaridade---	43
Figura 5.1 Grupos de características-----	45

Lista de tabelas

Tabela 4.1 Elementos para o cálculo de H-----	32
Tabela 4.2 Cálculo de $\Lambda(2)$ -----	38
Tabela 5.1 Resultados das áreas abaixo das curvas ROC para os grupos 1 ao 7 -----	46
Tabela 5.2 Resultados das áreas ROC para os grupos 8 ao 14 -----	49

Sumário

1 – Introdução -----	1
1.1 Objetivos-----	2
1.2 Organização do presente trabalho-----	2
2 – Patologias da mama e exames auxiliares ao diagnóstico -----	3
2.1 Estrutura da mama-----	3
2.2 Patologias da mama-----	4
2.2.1 Cisto-----	4
2.2.2 Fibroadenoma-----	6
2.2.3 Carcinomas-----	8
2.2.3.1 Carcinoma in situ-----	10
2.2.3.2 Carcinoma ductal in situ-----	11
2.2.3.3 Carcinoma lobular in situ-----	11
2.2.3.4 Carcinoma ductal infiltrante ou invasivo-----	11
2.2.3.5 Carcinoma lobular infiltrante ou invasivo-----	11
2.3 Fatores de risco-----	12
2.4 Exames auxiliares ao diagnóstico-----	15
2.4.1 Mamografia-----	15
2.4.2 Ultrassom-----	16
2.4.3 Ressonância magnética-----	17
2.4.4 Biópsia -----	18
2.4.5 Termografia-----	19
2.4.5.1 Aplicações-----	19
2.4.5.2 Imagens térmicas-----	20
2.5 Desvantagens dos exames usuais-----	22
3. Visão geral da metodologia proposta -----	24
3.1 Segmentação-----	24
3.2 Extração de características-----	26
3.3 Classificação-----	27
4. Características utilizadas-----	30
4.1 Coeficiente de Hurst-----	30
4.1.1 Proposta do uso do Coeficiente de Hurst -----	34

4.2 Lacunaridade-----	36
4.2.1 Proposta do uso da Lacunaridade -----	39
5. Avaliação dos resultados obtidos -----	44
6. Conclusões e trabalhos futuros-----	53
6.1 Trabalhos Futuros-----	54
Referências bibliográficas -----	57
Apêndice-----	62

1 – Introdução

Essa dissertação é parte das pesquisas em Termografia que vêm sendo desenvolvidas em parceria com a Universidade Federal de Pernambuco. Essa pesquisa se originou de um projeto intitulado “Análise da viabilidade do uso de câmera termográfica como ferramenta auxiliar no diagnóstico de câncer de mama em hospital público localizado em clima tropical” aprovado, em novembro de 2005, pelo Comitê de Ética da Universidade Federal de Pernambuco (UFPE) e registrado no Ministério da Saúde sob o número CEP/CCS/UFPE Nº279/05. O projeto buscava inicialmente capturar imagens termográficas de aproximadamente trezentas mulheres (até o presente foram examinadas cerca de cento e noventa mulheres) maiores de trinta e cinco anos, todas pacientes do Ambulatório de Mastologia do Hospital das Clínicas da UFPE. A partir de 2008, com o financiamento do projeto CAPES PRÓ-ENGENHARIAS passou-se a pensar em usar as imagens adquiridas das pacientes para a geração de um banco de imagens e desenvolvimento de diversos programas para modelagem e auxílio ao diagnóstico de doenças mastológicas. O objetivo final é analisar a viabilidade da utilização de uma câmera termográfica no auxílio ao diagnóstico de distúrbios mamários, principalmente para uso de imagens por infravermelho (termografia) em hospital público situado em clima tropical. Assim, à medida que esse projeto se desenvolve, um banco de dados vem sendo implementado para o cadastramento das pacientes, a fim de ser capaz de gerar análises estatísticas e cruzamento posterior de dados. Na Universidade Federal Fluminense, a linha de pesquisa em diagnóstico de doenças mastológicas iniciou em 1998 com a dissertação “Auxílio ao Diagnóstico de Mamografias por Computação Visual” de Luciana Marinho Soares. Em seguida, em 1999, houve a dissertação “Abordagens Difuso-Neurais para Análise de Mamogramas” de Marcelle de Sá Guimarães. Outra linha de pesquisa, sobre geometria fractal, vem sendo desenvolvida desde a mesma época. Nesta linha foram elaboradas as dissertações: “Emprego da Dimensão Fractal para Análise de Texturas em Imagens Multiespectrais” de Éldman de Oliveira Nunes em 2001, e também a dissertação “Using Fractal Characteristics such as Fractal Dimension,

Lacunarity and Succolarity to Characterize Texture Patterns on Images” de Rafael Heitor Correia de Melo em 2007. A presente dissertação é um elo entre as duas linhas de pesquisa, pois utiliza medidas fractais para auxiliar o diagnóstico de doenças mastológicas. Além disso, apresenta uma abordagem nova para o diagnóstico de patologias da mama: combina imagens térmicas com geometria fractal.

1.1 Objetivos

Esta dissertação tem como objetivo analisar a viabilidade do uso do Coeficiente de Hurst e da Lacunaridade no auxílio ao diagnóstico precoce de patologias da mama.

1.2 Organização do presente trabalho

Esta dissertação está organizada conforme é descrito a seguir. No Capítulo 2 são apresentadas inicialmente a estrutura da mama e suas patologias mais comuns (cisto, fibroadenoma e carcinoma). Em seguida são mencionados os fatores de risco da mais séria das patologias: os tumores malignos. Após isso, é feita uma descrição de exames auxiliares ao diagnóstico (mamografia, ultrassom, ressonância magnética, biópsia e termografia) e uma análise das desvantagens dos exames usuais.

O Capítulo 3 mostra a visão geral da metodologia proposta (segmentação, extração de características e classificação).

Os algoritmos do coeficiente de Hurst e da Lacunaridade, assim como suas aplicações às imagens térmicas são apresentados no Capítulo 4.

No Capítulo 5 são apresentados os resultados obtidos e suas análises.

No Capítulo 6 apresentam-se as conclusões desta dissertação e os trabalhos futuros, seguidos das referências bibliográficas.

Por fim, no Apêndice são mostradas todas as imagens térmicas utilizadas.

2 – Patologias da mama e exames auxiliares ao diagnóstico

2.1 Estrutura da mama

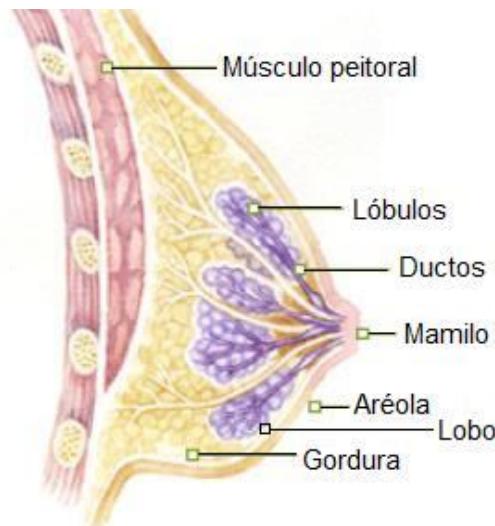


Figura 2.1 Estrutura da mama (Fonte: <http://www.pop.eu.com/uploads/images/>)

A função principal das mamas é a produção de leite. Como mostrado na Figura 2.1, as mamas são órgãos glandulares localizados em frente aos músculos peitorais. Essas glândulas dividem-se de 15 a 20 secções, chamados de lobos mamários (Figura. 2.1). Os lobos são conjuntos de lóbulos que são os elementos responsáveis pela produção de leite. Os ductos mamários (por volta de 15 a 20 canais finos) fazem a condução do leite até o mamilo (saliência elástica onde convergem os ductos mamários) que é localizado na parte central da aréola (região da mama que possui cor diferenciada da pele). Na mama também se encontram os vasos sanguíneos (responsáveis por transportar o sangue) e os vasos linfáticos (que têm a função de transportar a linfa). As demais regiões são compostas por estroma. Este se constitui por tecido gorduroso (variável conforme a idade, características físicas e alimentação) e conjuntivo suportando e ficando em torno dos lóbulos, ductos, vasos linfáticos e sanguíneos. De um modo geral, a maior parte das mamas de mulheres jovens é constituída de tecido glandular, o que faz com que as mamas sejam mais densas e firmes. Com a diminuição do nível de hormônios, as glândulas que constituem as mamas tendem a se atrofiar e passam a ser substituídas

gradativamente por tecido gorduroso. Após a menopausa a mama é composta por tecido gorduroso e por sobras insignificantes de tecido glandular. Essa substituição de tecido acarreta uma diferença significativa entre a constituição e densidade da mama de uma paciente jovem e de uma paciente após a menopausa (Bezerra, 2007).

2.2 Patologias da mama

Dentre as patologias mais freqüentes das mamas pode-se citar: os cistos, os fibroadenomas e os carcinomas (carcinoma in situ, carcinoma ductal in situ, carcinoma lobular in situ, carcinoma ductal infiltrante ou invasivo, carcinoma lobular infiltrante ou invasivo).

2.2.1 Cisto

Trata-se de uma modificação benigna da estrutura da mama. Apresenta-se como uma massa de tamanho variável de 1 milímetro até vários centímetros de diâmetro, podendo ser redonda ou ovalada, múltipla ou solitária. Normalmente se constituem por um único compartimento, que se reveste por epitélio e é preenchido por líquido. Geralmente a substância líquida que o preenche é parda e opaca. O tratamento, se necessário, para esse tipo de patologia é a aspiração por agulha (punção).

De maneira geral, mulheres entre 40 e 50 anos de idade apresentam com maior freqüência esse tipo de distúrbio mamário. O cisto pode ser assintomático, ou seja, a paciente pode ter cisto e nem saber disso. A paciente ou o médico podem detectar o cisto através da palpação das mamas, isso dependerá do seu tamanho e do grau de tecido conjuntivo que o cerca. Em alguns casos, a paciente pode sentir dor na região do cisto, ou quando a mama é comprimida no local onde este está localizado. Na termografia, o cisto se apresenta geralmente como uma região que possui temperatura mais baixa que as demais regiões da mama (Araújo et al., 2006). A Figura 2.2 mostra a imagem termográfica de uma paciente com vários cistos. Os maiores estão localizados às 3 horas, na mama esquerda e medindo 0,4cm. Como pode se

perceber essa região apresenta cor azul, sendo esta associada a temperaturas mais baixas, conforme se vê na escala de temperaturas localizada ao lado direito da imagem.

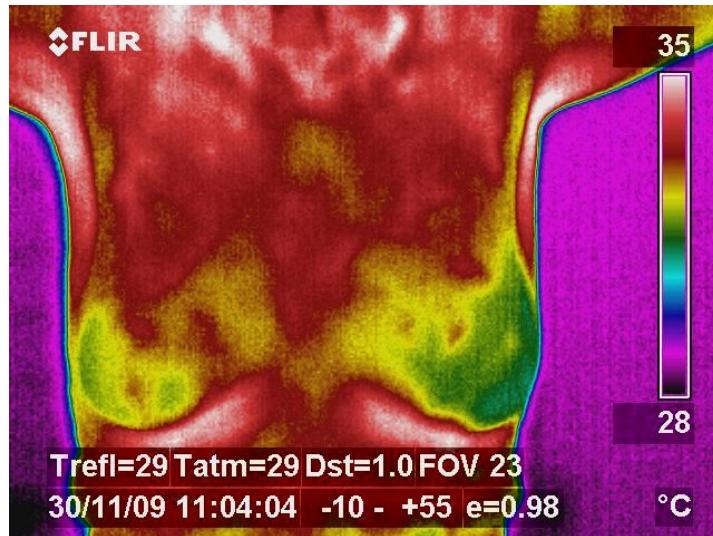


Figura 2.2 Termografia paciente com cisto (Fonte:

<http://200.20.11.171/proeng/>

No ultrassom esse tipo de distúrbio mamário se mostra como um nódulo cujas paredes são bem definidas e não possui ecos no seu interior. A Figura 2.3 mostra um cisto no ultrassom. Nesta figura o cisto é a região preta com formato oval no centro da imagem.



Figura 2.3 Ultrassom de uma paciente com cisto, (Bui e Taira, 2010).

Uma vez que identificado neste exame, não há a necessidade de exames complementares, pois este diagnóstico é bastante preciso, com acerto próximo de 100%. Quando o ultrassom diagnostica cistos pequenos que não apresentam sintomas, é possível realizar o acompanhamento sem que seja necessário o tratamento. Quando indicada a punção do cisto e a aspiração do líquido contido no seu interior tem-se o diagnóstico do distúrbio mamário e também o seu tratamento. O fato de a mulher apresentar um cisto não significa que terá chances de desenvolver o câncer de mama. Ocorre a necessidade de cirurgia em cistos recorrentes, ou caso o mesmo, após a aspiração, permaneça como um nódulo residual. A Figura 2.4 mostra uma mamografia médio-lateral da mama esquerda de uma paciente com cisto. A região que apresenta a patologia é identificada na figura por uma seta.

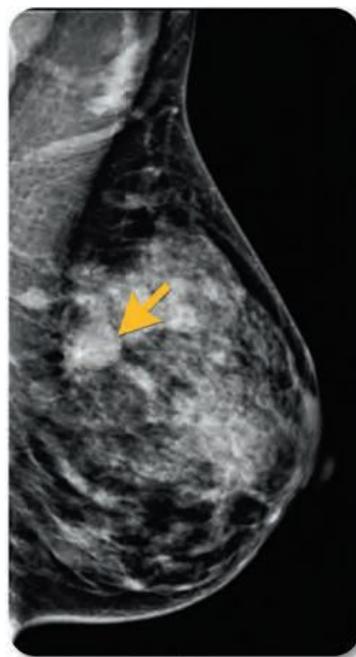


Figura 2.4 Mamografia de uma paciente com cisto, (Bui e Taira, 2010).

2.2.2 Fibroadenoma

É um tumor benigno cujo contorno é bem definido, e em geral, é palpável. Na mamografia apresenta-se com formato de pipoca arredondado ou

oval, com limites bem definidos. A Figura 2.5 mostra a imagem mamográfica de um fibroadenoma. Ele se caracteriza por tons claros devido à sua maior densidade ao raio-X.



Figura 2.5 Mamografia de uma paciente com fibroadenoma, (Bui e Taira, 2010).

No ultrassom esse tipo de tumor apresenta formato oval, com a altura menor do que a largura e ecos fracos no seu interior. A Figura 2.6 mostra a imagem ultrassonográfica de um fibroadenoma, sendo a região mais escura próxima ao centro da imagem.

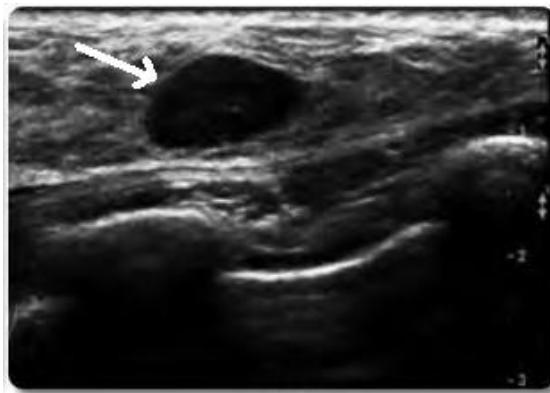


Figura 2.6 Ultrassom de uma paciente com fibroadenoma, (Bui e Taira, 2010).

Na termografia esse tipo de tumor apresenta-se como uma região com formato arredondado, cuja temperatura é diferente das demais regiões da mama da paciente. A Figura 2.7 mostra uma paciente com fibroadenoma na mama esquerda. Nesta imagem, a região do tumor possui cor vermelha, sendo esta cor associada a temperaturas mais elevadas, conforme pode ser observado através da escala com temperaturas (lado direito da imagem).

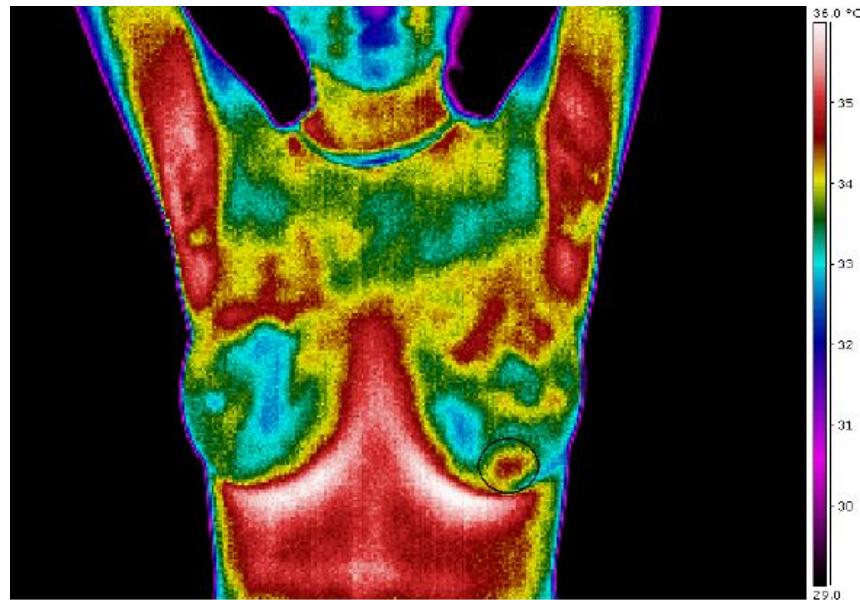


Figura 2.7 Termografia de uma paciente com fibroadenoma, (Fonte: <http://200.20.11.171/proeng/>).

2.2.3 Carcinomas

O câncer de mama se caracteriza pelo crescimento rápido e desordenado das células, que por sua vez podem se espalhar para outras partes do corpo. Nas imagens das mamas, estes diferentemente do cisto e do fibroadenoma, apresentam formato irregular e limites mal-definidos. A Figura 2.8 mostra a imagem de um carcinoma em mamografia crânio caudal (CC).



Figura 2.8 Carcinoma em mamografia crânio caudal, (Bui e Taira, 2010).

A Figura 2.9 mostra a imagem ultrassonográfica de uma mama com carcinoma. Este se localiza no centro da imagem. É região mais escura de contornos mal definidos.



Figura 2.9 Carcinoma em um ultrassom, (Bui e Taira, 2010).

Nas imagens térmicas, o carcinoma é caracterizado pela irregularidade na distribuição de temperaturas das mamas. A Figura 2.10 mostra uma termografia de uma paciente com um carcinoma (de Paget). Pode-se notar nesta termografia uma grande alteração da vascularização da mama esquerda. Não há informação visível de tumor, porém um exame clínico constatou a destruição do mamilo, devido ao carcinoma de Paget. Houve uma confirmação do carcinoma após uma biópsia. (Bezerra, 2007).

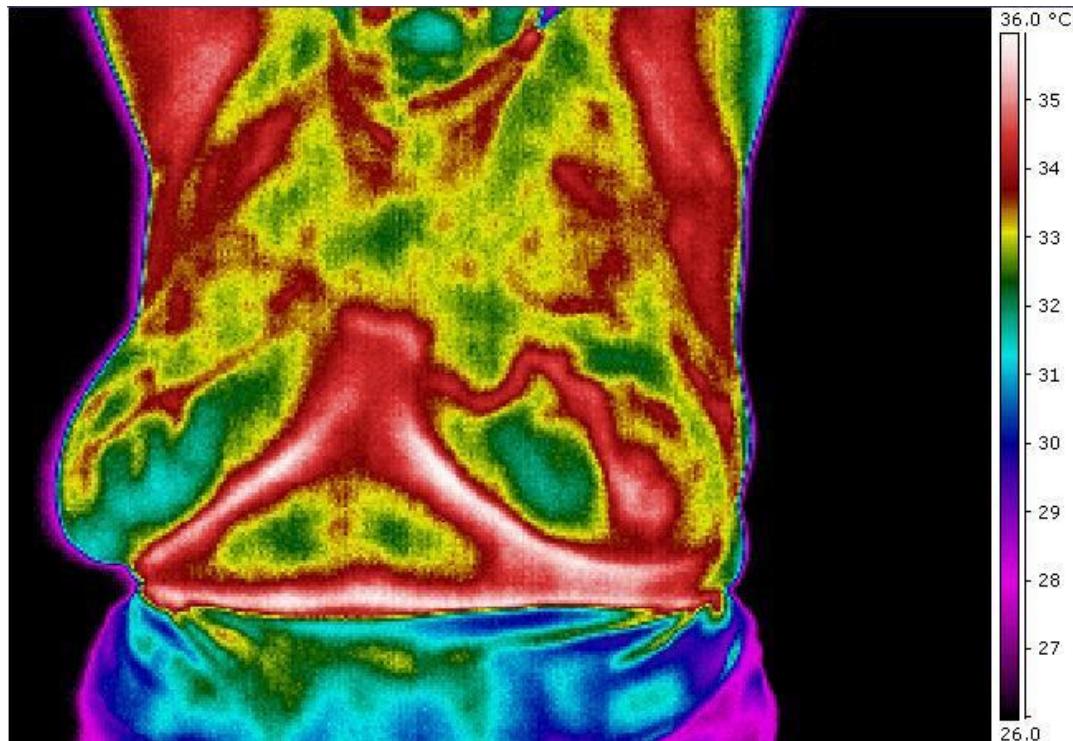


Figura 2.10 Carcinoma em termografia (Fonte:

<http://200.20.11.171/proeng/>).

O carcinoma se apresenta de diversas formas, que são descritas a seguir.

2.2.3.1 Carcinoma in situ

Essa denominação caracteriza o câncer em estágio inicial, que se limita ao local de sua aparição, ducto ou lóbulo, não afetando os tecidos gordurosos adjacentes nem outros órgãos.

2.2.3.2 Carcinoma ductal in situ

Trata-se de um tipo de câncer de mama não-invasivo bastante comum, no qual as chances de cura são elevadíssimas. Confina-se aos ductos, não chega a se espalhar através de suas paredes. Além disso, não alcança o tecido gorduroso. A doença de Paget mamária é uma variedade de carcinoma ductal in situ. Essa doença é um tipo de câncer muito raro que causa mudanças na pele da mama e em torno do mamilo. Pacientes que apresentam esse tipo de câncer, em geral, possuem outro câncer interno. Devido ao fato de a doença de Paget ocasionar mudanças na pele, inicialmente, por desconhecerem ser um problema interno da mama, os pacientes buscam tratamento em dermatologistas.

2.2.3.3 Carcinoma lobular in situ

Este tipo de câncer tem início nas glândulas produtoras de leite (lóbulos) e se restringe a elas, não atravessando a parede dos lóbulos.

2.2.3.4 Carcinoma ductal infiltrante ou invasivo

Este se inicia nos ductos atravessando sua parede e invadindo o tecido gorduroso da mama, havendo a possibilidade de se espalhar para as demais partes do corpo.

2.2.3.5 Carcinoma lobular infiltrante ou invasivo

Este tipo de câncer tem início nas glândulas produtoras de leite (lóbulos) podendo se espalhar para as demais partes do corpo.

2.3 Fatores de risco

Para complementar o entendimento sobre as patologias da mama é interessante conhecer quais são os fatores de risco da mais séria delas: os tumores malignos. Fator de risco é tudo que aumente a probabilidade de surgimento de uma doença. Alguns dos diversos fatores existentes podem ser controlados (consumo de álcool, alimentação) e outros não (histórico familiar e faixa etária). Dentre os fatores de risco, podemos citar: gênero, idade, raça, consumo de álcool, alimentação, reposição hormonal, não ter filhos, sedentarismo, pílula anticoncepcional, tratamento com dietiletibestrol, histórico pessoal de câncer de mama, mutações genéticas, biópsia anterior anormal, radioterapia anterior do tórax, idade de início e fim da fase menstrual.

- Gênero: o desenvolvimento do câncer de mama ocorre principalmente em mulheres.

- Idade: as chances de se desenvolver a doença aumentam com o passar dos anos. Estatísticas globais indicam que em torno de 18% dos cânceres diagnosticados são referentes a mulheres com idade próxima dos 40 anos. Mulheres com idade igual ou superior a 50 anos correspondem a 77%, (www.accamargo.org.br).

- Raça: há incidência do câncer de mama com maior freqüência nas mulheres brancas. Porém quando atingem mulheres negras, estas apresentam tumores mais agressivos, fazendo com que tenham maiores chances de morrer em função da doença. A incidência é menor em mulheres indígenas e asiáticas.

- Consumo de álcool: a ingestão de bebida alcoólica aumenta o risco de se ter câncer de mama. Mesmo a ingestão de uma pequena dose da bebida por dia acarreta aumento do risco. Mulheres que ingerem de 2 a 5 doses por dia tem uma vez e meia mais chance de ter a doença do que mulheres que não ingerem.

- Alimentação: a obesidade está relacionada com maiores chances de se ter o câncer de mama, principalmente se o excesso de peso for adquirido durante ou depois da menopausa. Além disso, o risco aumenta se o excesso de células adiposas estiver concentrado em torno da cintura. Os especialistas

recomendam uma alimentação equilibrada, com pouca gordura e rica em fibras, devendo-se evitar principalmente carnes vermelhas.

- Reposição hormonal: a utilização da reposição hormonal combinada (estrogênio mais progesterona) por diversos anos aumenta as chances de ter câncer de mama, derrame, coágulos e doenças cardíacas. Nestes casos quando diagnosticados os tumores se apresentam em estágios mais avançados, pois a reposição hormonal aparenta diminuir a eficácia da mamografia.

- Não ter filhos: o risco de ter câncer de mama é maior para mulheres que tiveram filhos somente após os 30 anos de idade ou que não tiveram filhos. Mulheres que tiveram mais de um filho quando jovens têm menor risco de ter câncer de mama.

- Sedentarismo: mulheres que praticam exercícios físicos regularmente reduzem a possibilidade de ter câncer de mama, no entanto o que ainda não se sabe é a quantidade de atividade física necessária.

- Pílula anticoncepcional: até o momento não se sabe precisamente se há alguma relação entre o câncer de mama e o uso de pílulas anticoncepcionais.

- Tratamento com dietiletilbestrol: há alguns anos, as mulheres grávidas usaram esse medicamento com o intuito de diminuir a possibilidade de aborto espontâneo. Tempos depois se descobriu que o medicamento possuía efeitos teratogênicos (acarretando má-formação no feto) e carcinogênicos (www.accamargo.org.br).

- Histórico familiar: o fato de possuir filha, irmã ou mãe com câncer de mama aumenta a probabilidade de a mulher vir a ter a doença, porém não se sabe precisar com exatidão qual é esse risco. Sabe-se também que a mulher que possui pai ou irmão com câncer de mama tem maiores chances de ter a doença.

- Histórico pessoal de câncer de mama: a possibilidade de vir a ter câncer de mama é maior em mulheres que já apresentaram a doença anteriormente (cerca de 3 a 4 vezes maior), podendo vir a ter novamente na outra mama ou em alguma outra região da primeira. Não cabe dizer que houve reincidência, mas sim a aparição de um novo tumor.

- Mutações genéticas: certas mutações genéticas podem ocasionar o desenvolvimento cancerígeno nas mamas, sendo que a mais comum é a que ocorre nos genes BRCA1 e BRCA2.

- Biópsia anterior anormal: a possibilidade de câncer de mama é aumentada quando alguns tipos de anomalias foram encontradas em biópsias anteriores.

- Radioterapia anterior do tórax: mulheres que realizaram radioterapia na região do tórax (quando crianças ou jovens) aumentam de forma considerável as chances de desenvolvimento do câncer de mama.

- Grande tempo de período menstrual: mulheres que tiveram a sua primeira menstruação cedo, por volta dos 12 anos de idade, ou mulheres que tenham entrado na menopausa depois dos 55 anos têm uma possibilidade um pouco maior de desenvolver o câncer de mama (www.accamargo.org.br).

O fato de a mulher pertencer a um ou mais grupos de fatores de risco não implica necessariamente no aparecimento da doença, mas aumenta a possibilidade de desenvolvê-la. Conforme foi exposto anteriormente, alguns fatores de risco podem ser controlados (consumo de álcool, alimentação), porém outros não (idade, histórico familiar, período menstrual, entre outros), sendo assim, é fundamental o acompanhamento médico para que seja possível a detecção precoce do câncer de mama. Essa detecção precoce é extremamente importante, pois alguns estudos têm revelado que se a doença for identificada em estágios iniciais (tumores de tamanho menor que 10mm) a possibilidade de cura é de 85%, enquanto que se for detectado mais tarde, as chances de cura são de 10% (Ng e Sudarshan, 2001-b).

2.4 Exames auxiliares ao diagnóstico

2.4.1 Mamografia

Atualmente, a principal forma de diagnosticar distúrbios mamários é através da mamografia. A mamografia é uma espécie de radiografia da mama. A Figura 2.11 mostra uma paciente sendo submetida a um exame de mamografia.



Figura 2.11 Paciente sendo submetida a um exame de mamografia
(Fonte: <http://pt.wikipedia.org>).

Neste procedimento, a mama da paciente é comprimida por um aparelho chamado mamógrafo. Simultaneamente à compressão são incididos raios-X sobre a mama da paciente em duas direções: paralelamente à cintura da paciente (como mostrado na Figura 2.11) ou em direção inclinada, de modo a estar presente a região axilar. No primeiro caso tem-se a exposição crânio caudal (CC). No segundo tem-se a exposição médio-lateral (ML). A Figura 2.12(a) e a Figura 2.12(c) mostram exemplos de mamografias crânio caudais de uma mama direita e de uma mama esquerda, respectivamente. A Figura 2.12(b) e a Figura 2.12(d) mostram exemplos de mamografias médio-laterais de uma mama direita e de uma mama esquerda, respectivamente.

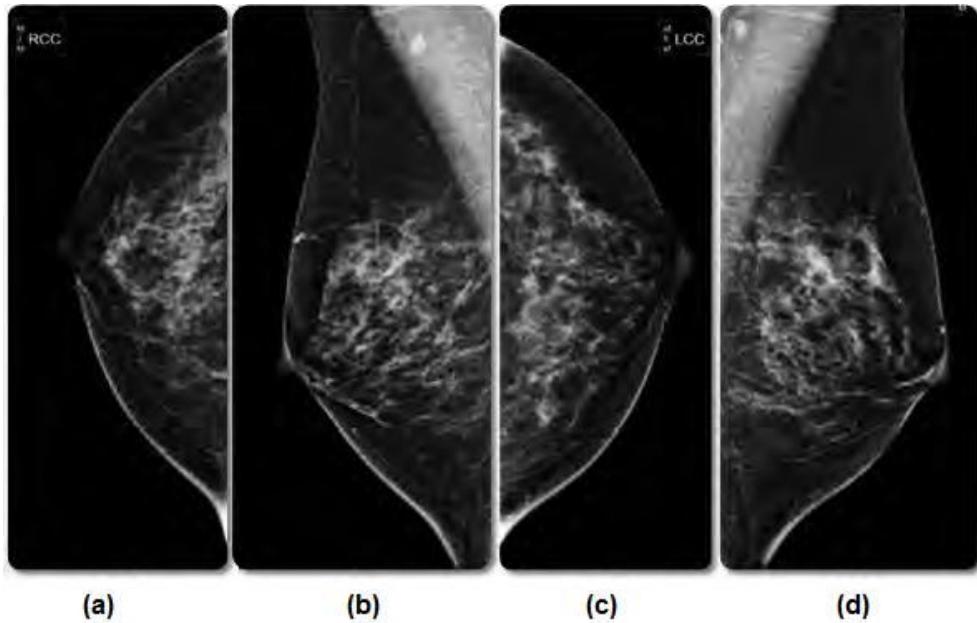


Figura 2.12 Exemplos de mamografias crânio caudais ((a) e (c)) e médio-laterais ((b) e (d)), (Bui e Taira, 2010).

Concluída a mamografia, um especialista (radiologista) analisa as imagens obtidas com o objetivo de saber se existem regiões suspeitas. A mamografia é recomendada em dois casos:

- 1) Exame preventivo, quando a paciente tem mais de 40 anos (www.siemens.com.br). Neste caso, o objetivo da mamografia é identificar possíveis nódulos que não sejam palpáveis.
- 2) Quando é encontrado algum nódulo, seja por exame clínico ou auto-exame.

Após a análise da mamografia, pode ser solicitado à paciente um ultrassom das mamas, visando identificar na região suspeita a presença de cisto (líquido) ou de algum tumor (sólido). Também podem ser indicados outros exames como punção por agulha ou biópsia.

2.4.2 Ultrassom

O ultrassom consiste na utilização do eco produzido pelo som para a geração de uma imagem. A Figura 2.13 mostra uma paciente sendo submetida a esse tipo de exame.



Figura 2.13 Paciente sendo submetida a um exame de ultrassom (Fonte:
www.noticiasnaspec.com)

2.4.3 Ressonância magnética

Além da mamografia e do ultrassom também pode ser solicitado à paciente um exame de ressonância magnética. Este tipo de exame não utiliza radiação ionizante.



Figura 2.14 Exame de ressonância magnética (Fonte:
www.advanceweb.com)

Quando avaliações anteriores (mamografia, ultrassom, ressonância magnética) indicam que determinada região pode estar com câncer de mama, a paciente é submetida a uma biópsia.

2.4.4 Biópsia

A biópsia é um procedimento no qual se realiza a retirada de parte da região suspeita. Existem diversos modos de fazer uma biópsia, cabendo ao médico avaliar a forma mais adequada para a paciente:

- Biópsia estereotáxica: este tipo de biópsia é utilizado com o objetivo de fazer a remoção de diversos cilindros de tecido.

- Biópsia por aspiração: Neste processo é utilizada uma agulha fina (podendo ser orientada por ultrassom) para se fazer a retirada de líquido do nódulo. Caso haja um nódulo sólido, é feita uma remoção de pequenas partes dele e, em seguida, é feita uma análise do material coletado no microscópio.

- Biópsia cirúrgica: Neste tipo de biópsia é feita uma cirurgia para a retirada de um pedaço do nódulo ou ele por inteiro e, após isso, é realizada sua análise no microscópio. Em diversos casos também é feita a retirada de tecido sadio localizado em torno do nódulo.

- Biópsia percutânea assistida a vácuo (mamotomia): É um tipo de biópsia que utiliza agulha grossa, sendo guiada por mamografia ou ultrassom.

Concluída a biópsia, é feita a sua análise originando dois possíveis resultados: o material coletado é cancerígeno ou não. Os tumores são relacionados a uma escala que varia de 1 a 3. Tumores que são relacionados com valores mais altos são aqueles que possuem maior tendência de crescimento, ou seja, são os tumores mais agressivos. Já os que são relacionados ao grau mais baixo são aqueles que possuem menor taxa de crescimento e mais parecidos com o tecido normal. Quando o material coletado não é cancerígeno, não é necessário que a paciente seja submetida a um tratamento.

Há a possibilidade de a amostra coletada ser avaliada com o objetivo de saber se a mesma possui receptores para alguns hormônios, tais como: estrogênio e progesterona. Caso possua receptores há indicativo de que ocorrerá resposta ao tratamento hormonal para esses tipos de câncer.

2.4.5 Termografia

Objetiva-se através da termografia obter informações relativas à condição funcional de um componente, processo ou equipamento. Trata-se de uma técnica que utiliza a radiação infravermelha que é emitida pelos corpos a fim de medir temperaturas ou mesmo observar certos padrões diferenciais de distribuição de térmica, sendo uma técnica não-destrutiva e não-invasiva.

2.4.5.1 Aplicações

A termografia possui diversas aplicações em: instalações elétricas, equipamentos rotativos, equipamentos estáticos, artes, indústria química, engenharia civil, siderurgia, indústria automobilística e medicina.

Sua aplicação na medicina iniciou-se após Lawson ter descoberto, em 1956, que na região onde há um câncer de mama a temperatura da pele pode ser maior do que a temperatura do tecido normal. Descobriu também que o sangue venoso responsável por drenar o câncer possui temperatura superior a do arterial (Lawson, (1984); in Bronzino, (2006)). Segundo Gautherie (1989), a distribuição de temperaturas em mamas saudáveis se dá de forma simétrica e as mudanças de temperatura estão relacionadas com uma possível gravidez e com o ciclo menstrual. Este estudo destacou que essas mudanças atingem de alguma forma a ambas as mamas. Se há grande assimetria na distribuição de temperaturas das mamas é porque uma delas apresenta algum tipo de patologia. Pacientes com tumores mamários deixam de possuir tal simetria, pois as células cancerígenas produzem excessivamente o óxido nítrico (NO), que acarreta uma nova vascularização (angiogênese) em torno do tumor (Yahara et al., 2003). Essa nova vascularização é responsável por aumentar o fluxo sanguíneo causando uma elevação da temperatura na região do tumor. Essa variação de temperatura na superfície da pele entre uma região normal e uma que apresenta algum tumor pode ser de 2 a 3 °C (Hu, (2004)). Por isso a importância da termografia, pois esta consegue captar esse aumento de temperatura na superfície de mamas que possuam patologias. Sendo assim,

enquanto a mamografia e ultrassom são testes anatômicos, a termografia é um teste fisiológico (Love, 1985).

Além disso, a termografia é capaz de detectar tumores menores que a mamografia (o tamanho médio do tumor detectado em uma mamografia é de 1,66 cm, já em uma termografia é de 1,28 cm (Keyserlingk et al., 2000)) e de diagnosticar tumores cerca de 8 a 10 anos antes que uma mamografia (Gautherie, 1989; Ng et al., 2001). Devido à incidência do câncer de mama em mulheres cada vez mais jovens, alguns médicos têm recomendado o exame de mamografia a partir dos 35 anos. Mas, como as mamas mais jovens são densas torna-se difícil o diagnóstico através da mamografia. Por outro lado, como a termografia capta a temperatura da mama, esta pode ser utilizada, sem restrição de espécie alguma, em pacientes jovens possibilitando diagnosticar precocemente o câncer de mama (Ng e Sudharsan, 2001-a).

2.4.5.2 Imagens térmicas

Nas imagens térmicas cada pixel corresponde a uma temperatura da cena capturada. Neste estudo foram usadas vinte e oito imagens térmicas (640 x 480 pixels) disponíveis em <http://200.20.11.171/proeng/> e no apêndice desta dissertação (Araújo, 2009; Araújo et al., 2008; Santos et al., 2008). Essas imagens pertencem a dois grupos de pacientes: saudáveis e com patologia. O grupo de pacientes com patologias é composto por imagens de mulheres com cisto, fibroadenoma (tumor benigno) e carcinoma (tumor maligno). As imagens térmicas podem ser representadas por até quinze paletas diferentes de cores: Yellow, Rain, Rain10, Rain900, RainHi, Grey10, Grey, Greyred, Glowbow, Iron, Iron10, IronHi, Medical, Midgreen, Midgrey (Menezes et al., 2009). Nesta dissertação foram usadas imagens em tom de cinza (Grey). Sendo assim, o espaço das temperaturas está associado com o espaço dos tons de cinza (0 a 255). Esta é a forma mais simples de representar uma imagem térmica. Como exemplos, a Figura 2.15 e a Figura 2.16 mostram imagens térmicas (em tamanhos reduzidos para caberem na página) de duas pacientes. Na Figura 2.15 é mostrada a imagem térmica em tom de cinza de uma paciente saudável.

Nesta figura cada ponto da imagem entre 25,2 e 35,2 °C será associado com um tom de cinza entre 0 e 255.

Na Figura 2.16 é mostrada a imagem térmica de uma paciente com patologia (carcinoma). Nesta imagem não há informação visível de tumor. Do mesmo modo, cada ponto da imagem entre 28,3 e 36,2 °C será associado com um tom de cinza entre 0 e 255.



Figura 2.15 Paciente saudável

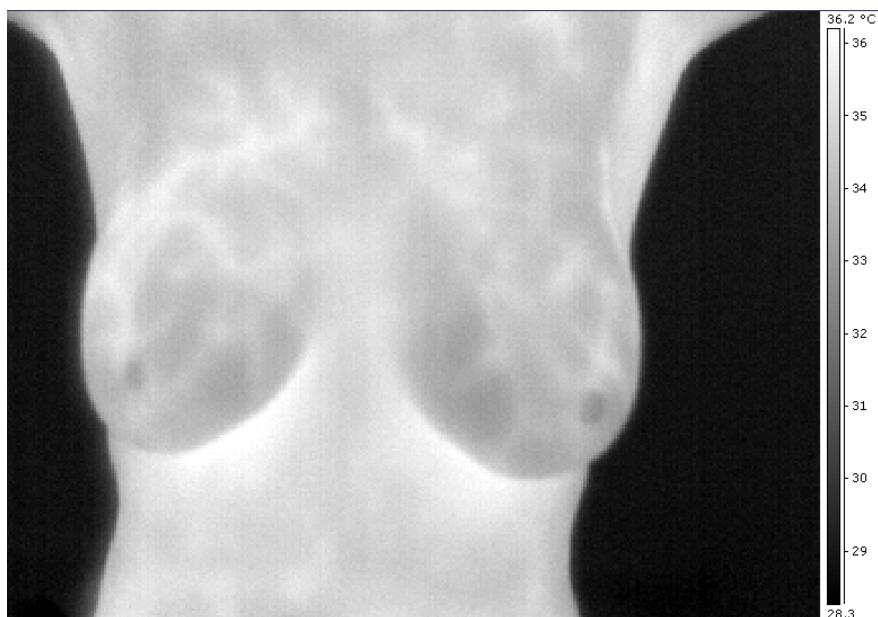


Figura 2.16 Paciente com patologia (carcinoma)

2.5 Desvantagens dos exames usuais

Cada um dos exames utilizados atualmente (mamografia, ultrassom, ressonância magnética, biópsia) apresenta características negativas.

O fato de a mamografia utilizar radiação eletromagnética pode ser considerado como um fator que desencadeie o crescimento de tumores, ou seja, com isso, a inspeção (recomendada para mulheres acima de 40 anos) teria efeitos negativos sobre a saúde da paciente.

O exame de rotina acaba agravando o caso da paciente, causando a necessidade de tratamentos mais agressivos, aumentando o número de mastectomias (retirada da mama) por cerca de 20% e o número de mastectomias e tumorectomias (procedimento cirúrgico que ocasiona a retirada do tumor e parte do tecido saudável em torno dele) por cerca de 30% (Olsen e Gotzsche, 2001).

A mamografia apresenta dificuldade de análise quando se tratam de mamas densas. Próximo à idade da menopausa e depois dessa, o tecido mamário começa a ser substituído por tecido adiposo, sendo que esta substituição ocorre de modo progressivo. Algum tempo após a menopausa a mama é constituída praticamente por gordura e vestígios de tecido glandular. Sendo assim, mamas de pacientes na fase da menopausa, ou após ela, são constituídas em quase sua totalidade por tecido gorduroso (pouco densas), enquanto que mamas de pacientes jovens em fase reprodutiva são densas.

A Figura 2.17(a) e 2.17(b) representam mamografias de duas pacientes: a primeira referente a uma paciente com mama constituída por tecido adiposo (menos densa) e a segunda de uma paciente com mama densa.

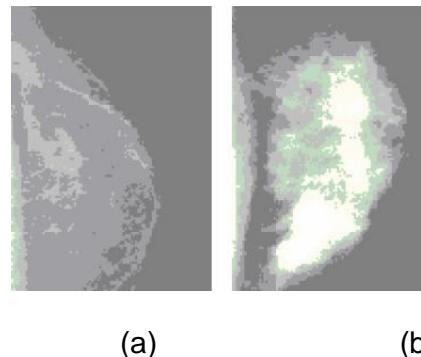


Figura 2.17 Mamografia de mama menos densa (a) e mais densa (b)

Como se pode perceber, comparando as imagens da Figura 2.17, a imagem correspondente a mamografia de uma mama densa tem a capacidade de análise prejudicada, pelas estruturas da mama, o que impossibilita a identificação de regiões com microcalcificações. O aparecimento dessas regiões com muita textura na imagem (Figura 2.17b) é devido ao fato dos raios-X serem mais absorvidos pelo estroma (tecido responsável pela sustentação das células funcionais de um órgão). De modo análogo ao estroma de mamas densas, as microcalcificações absorvem uma quantidade maior de raios-X. Com isso, as grandes regiões claras na Figura 2.17b (ocasionadas pelo fato do estroma ser mais denso) se confundem com as microcalcificações e nódulos em estágios iniciais, pois estes também se mostram na mamografia como regiões ou pontos claros. Em geral, pacientes mais jovens apresentam as mamas densas, com isso, torna-se inviável a utilização da mamografia nessas pacientes. Entretanto, conforme citado anteriormente, mulheres cada vez mais jovens têm apresentado câncer de mama e, por isso, alguns médicos têm recomendado a mamografia para as mulheres com idade a partir dos 35 anos, apesar de geralmente se obter um resultado pouco esclarecedor (Santos, 2009).

Tamanho do tumor identificado: o tamanho médio do tumor detectado em uma mamografia é de 1,66 cm e o tumor detectado por uma termografia é de 1,28 cm (Keyserlingk et al.,2000). Além disso, a termografia pode identificar um tumor cerca de 8 a 10 anos antes que uma mamografia possa identificar alguma região suspeita (Gautherie, 1989; Ng et al., 2001).

A desvantagem da ultra-sonografia é o fato de depender muito da habilidade do operador. Já o alto custo da ressonância magnética muitas vezes inviabiliza o uso desta. A biópsia tem como desvantagem o fato de ser uma técnica invasiva, mas é necessária para fechar o diagnóstico.

3. Visão geral da metodologia proposta

O reconhecimento de padrões tem como objetivo a classificação de regiões de interesse (padrões) a partir de informações extraídas destas (Sonka et al., 2008). A importância desta área está relacionada à busca de métodos automatizados e informatizados para tarefas humanas repetitivas, exaustivas e muito sujeitas a falhas.

A Figura 3.1 resume os passos necessários para o reconhecimento de padrões desta metodologia.

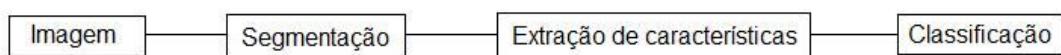


Figura 3.1 Etapas do reconhecimento de padrões desta metodologia

A seguir cada uma das etapas (segmentação, extração de características e classificação) do reconhecimento de padrões é descrita na forma proposta nesta dissertação.

3.1 Segmentação

Em geral a segmentação é a primeira etapa de um processo de reconhecimento de padrões. Esta etapa consiste na seleção de regiões de interesse na imagem. Para a aplicação considerada no presente trabalho as regiões de interesse são a mama direita e a mama esquerda. Para isso, janelas quadradas de mesmo tamanho são usadas para selecionar cada mama da paciente. O tamanho adequado da janela depende do tamanho da mama. Para as imagens utilizadas neste estudo, janelas entre 170 x 170 e 250 x 250 pixels foram usadas. A Figura 3.2 e a Figura 3.3 mostram o resultado da seleção da região de interesse das imagens térmicas representadas pela Figura 2.15 (imagem térmica em tamanho reduzido) e da Figura 2.16 (imagem térmica em tamanho reduzido) respectivamente. Neste trabalho esta seleção é feita manualmente, mas vem sendo testado um método de segmentação (Serrano et al., 2009).

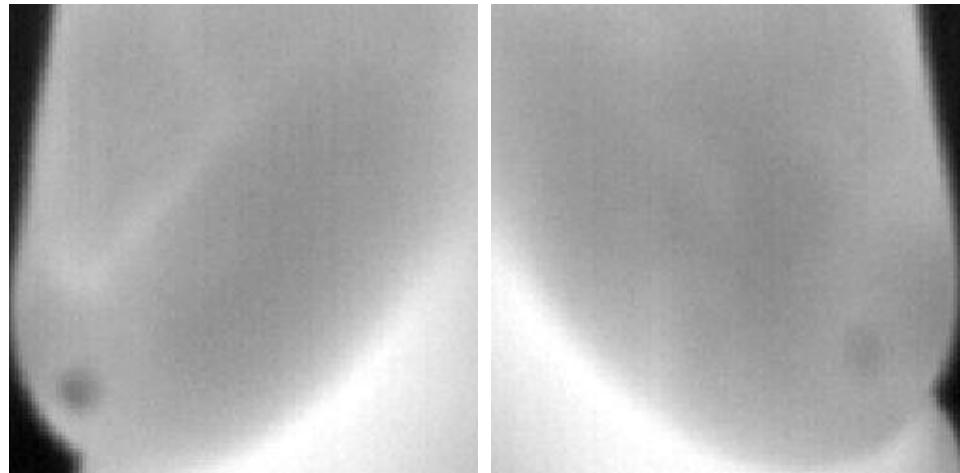


Figura 3.2 Segmentação da região de interesse da paciente saudável

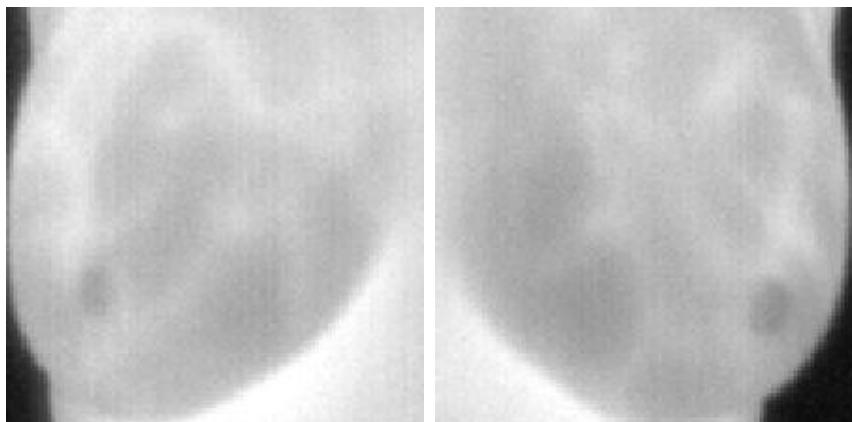


Figura 3.3 Segmentação da região de interesse da paciente com patologia

Com o intuito de reduzir o custo computacional da extração de características (cálculo da Lacunaridade) e ter todas as imagens resultantes da segmentação com o mesmo tamanho, as imagens foram reamostradas para que passassem a ter 100 x 100 pixels. Para ilustrar, a Figura 3.4 e a Figura 3.5 mostram as imagens da Figura 3.2 e da Figura 3.3 após a reamostragem através da ferramenta Photoshop 7.0.

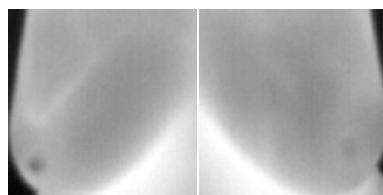


Figura 3.4 Região de interesse da Figura 3.2 com 100 x 100 pixels

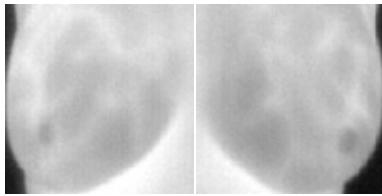


Figura 3.5 Região de interesse da Figura 3.3 com 100 x 100 pixels

Com isso, para cada imagem térmica, após a seleção da região de interesse e reamostragem têm-se duas imagens (mama direita e mama esquerda) de tamanhos iguais.

A partir destas uma nova imagem (imagem Subtração) é gerada pelo módulo da subtração pixel a pixel da imagem da mama esquerda com a imagem da mama direita rotacionada em 180º em relação a um eixo vertical.

A Figura 3.6 mostra a imagem Subtração obtida através das imagens da Figura 3.4 e da Figura 3.5 respectivamente.



Figura 3.6 Imagem Subtração da paciente saudável e da paciente com patologia

Portanto, para cada imagem térmica, são geradas três imagens: mama direita (MD), mama esquerda (ME) e imagem Subtração (S). Essas três imagens serão utilizadas na extração de características.

3.2 Extração de características

A extração de características consiste em obter atributos ou dados importantes das regiões de interesse. Para isso é necessário conhecer quais são as informações relevantes para o reconhecimento de padrões e, baseado nisso, escolher as características que deverão ser extraídas da imagem capturada. Os principais tipos de características são: geometria, dimensões, luminosidade e textura. Como exemplos de características geométricas podem-se citar: concavidade, circularidade, retilineidade, eixos principais, etc. Para

características dimensionais têm-se os seguintes exemplos: centro de gravidade, perímetro, área, comprimentos, larguras mínima e máxima, etc. Para a luminosidade o principal exemplo é a avaliação das cores.

Dentre as características de textura têm-se: matrizes de co-ocorrência, funções de autocorrelação, histograma, média, desvio padrão e medidas fractais (Coeficiente de Hurst, Dimensão Fractal, Lacunaridade e Sucolaridade). Para a aplicação dessa dissertação foram escolhidos o Coeficiente de Hurst e a Lacunaridade como características a serem extraídas das imagens resultantes da segmentação.

Objetos naturais tais como: nuvens, folhagens, montanhas e arbustos não são bem caracterizados pela geometria euclidiana tradicional. Neste aspecto, a geometria fractal consegue criar modelos naturais, pois é capaz de descrever as irregularidades da natureza. Esse foi o principal motivo da escolha da geometria fractal para a caracterização das imagens térmicas. Diversas aplicações têm utilizado a geometria fractal para extração de características: Barros e Sobreira, 2008; Conci e Nunes, 2001; Conci e Proença, 1998; Melo e Conci, 2007; Batty e Longley, 1994; Block e Schellnhuber, 1990; Barros Filho e Sobreira, 2005; Barros Filho e Sobreira, 2005-a; Fernandes e Conci, 2004; Keller e Chen 1889; Frazer, 2005; Valous et al., 2009; Plotnick et al., 1993; Müssigmann, 1992; Dong 2000, Einstein et al., 1998. A forma detalhada de como a extração de características fractais é utilizada será alvo do próximo capítulo.

3.3 Classificação

Após a conclusão da etapa de extração de características, passa-se à etapa de classificação. Esta etapa é responsável pela atribuição de um rótulo (para esta aplicação os rótulos que serão atribuídos serão “paciente saudável” e “paciente com patologia”) a um objeto (neste caso o objeto em questão é uma imagem térmica) com base nas características extraídas do objeto.

É importante ressaltar que a distribuição de temperaturas em mamas saudáveis se dá de forma simétrica e que se há grande assimetria da distribuição de temperaturas nas mamas é porque uma delas apresenta algum

tipo de patologia. É baseado nesta análise da simetria da distribuição de temperaturas que as abordagens de classificação se fundamentam.

Para ser feita a classificação é necessário a utilização de alguma técnica de decisão. Existem diversas técnicas para isso, desde as que usam critérios bem simples, como funções discriminantes e distância mínima, ao uso de técnicas de aprendizado automático e Inteligência artificial (clusterização, redes neurais artificiais, lógica nebulosa, algoritmos genéticos).

Nesta dissertação serão utilizadas técnicas de aprendizado de máquina. Foram usadas cento e trinta e três características extraídas (noventa e sete através de Lacunaridade e trinta e seis por Coeficiente de Hurst) para identificar o grupo a que a paciente pertence (saudável ou com patologia). Com este objetivo foi utilizado o software WEKA (<http://www.cs.waikato.ac.nz/ml/weka/>). Todas as técnicas de aprendizado de máquina contidas neste software foram utilizadas. São elas: Bayes Logistic Regression (BLR), Bayes Net (BN), Complement Naive Bayes (CNB), Discriminative Multinomial Naive Bayes Text (DMNBT), Naive Bayes (NB), Naive Bayes Multinomial (NBM), Naive Bayes Multinomial Updateable (NBMU), Naive Bayes Simple (NBS), Naive Bayes Updateable (NBU), Logistic (L), Multilayer Perceptron (MP), Radial Basis Function Network (RBFN), Simple Logistic (SL), Sequential Minimal Optimization (SMO), Voted Perceptron (VP), Instance Based on Nearest Neighbour Classifier (IB), Instance Based on K-nearest Neighbours Classifier (IBK), Instance Based Classifier (KStar), Locally Weighted Learning (LWL), Adaboost (AB), Attribute Selected Classifier (ASC), Bagging (B), Classification Via Clustering (CVC), Classification Via Regression (CVR), Parameter Selection by Cross-Validation (CVPS), Dagging (D), Decorate (Dec), Ensemble of Nested Dichotomies (END), Ensemble Selection (ES), Filtered Classifier (FC), Grading (Grad), Logit Boost (LB), Multi Boost (MB), Multi Class Classifier (MCC), Multi Scheme (MS), Ordinal Class Classifier (OCC), Raced Incremental Logit Boost (RILB), Random Committee (RC), Random Sub Space (RSS), Rotation Forest (RF), Stacking (Stac), Stacking C (StacC), Threshold Selection (TS), Vote (V), Class Balanced Nested Dichotomies (CBND), Data Near Balanced Nested Dichotomies (DNBND), Nested Dichotomies (ND), Field Limiting Ring (FLR), Hyper Pipes (HP), Voting Feature Intervals (VFI), Alternating Decision Tree (ADTree), Best First Tree

(BFTree), Decision Stump (DS), Functional Trees (FT), Class for generating a pruned or unpruned decision tree (J48), Class for generating a grafted decision tree (J48graft), Logitboost Alternating Decision Tree (LADTree), Logistic Model Trees (LMT), Naive Bayes Tree (NBTree), Random Forest (RFor), Random Tree (RTree), Reduced Error Pruning Tree (REPTree), Simple Cart (SC), Conjunctive Rule (CR), Decision Table (DT), Decision Table Naive Bayes (DTNB), Repeated Incremental Pruning (JRip), Non Nested Generalized Exemplars (NNge), 1-R classifier (OneR), Partial Decision Tree (PART), Ripple Down Rule (Ridor), 0-R classifier (ZeroR) (Witten e Eibe, 2005).

Essas técnicas são responsáveis pela criação de modelos computacionais capazes de identificar quais são as características relacionadas a cada grupo (paciente saudável e paciente com patologia). Após isso, o modelo construído pode classificar novas imagens através de suas características.

4. Características utilizadas

Na geometria euclidiana são apresentadas definições para dimensão baseada no número de elementos necessários para descrever os pontos de um corpo. Porém, esta geometria não consegue caracterizar de modo preciso formas da natureza (uma montanha ou fragmentos de rochas, contorno da costa de um país ou de uma folha) devido ao seu alto grau de irregularidade. Por outro lado, a geometria fractal usa outras definições que levam a dimensões fracionárias, por exemplo: 5/7 ou 3/11. Essas dimensões são capazes de quantificar mais precisamente a complexidade e a irregularidade de formas naturais que as dimensões da geometria euclidiana. Nesta dissertação são usadas duas medidas fractais (Coeficiente de Hurst e Lacunaridade) para a caracterização das texturas (Conci et al., 2008) das imagens térmicas.

4.1 Coeficiente de Hurst

O Coeficiente de Hurst é uma medida que avalia a densidade de imagens. Com isso, através desse coeficiente, é possível associar índices numéricos a imagens para posteriormente ser realizada a classificação. Essa medida vem sendo utilizada em diversas aplicações, por exemplo: análise da complexidade da textura de imagens da vegetação e áreas de floresta (Freitas, 2007), na elaboração de projetos urbanos e arquitetônicos (Martins e Librantz, 2006), obtenção e análise de imagens de transformação de textura (Hott et al., 2005). Como pôde se observar, este coeficiente tem sido utilizado para a caracterização de texturas. No entanto, o uso deste coeficiente para a caracterização de imagens térmicas é desconhecido. Nessa dissertação, o Coeficiente de Hurst será usado como atributo de comparação entre imagens térmicas de pacientes saudáveis e pacientes com patologias (Serrano et al., 2010).

Com o intuito de esclarecer como o cálculo do coeficiente de Hurst é efetuado, considere a Figura 4.1 que representa uma janela de 5 x 5 pixels com parte de uma imagem em níveis de cinza.

95	102	90	103	97
101	105	80	93	95
94	99	92	90	107
108	103	75	67	85
112	97	60	72	82

Figura 4.1 Janela de 5 x 5 pixels com parte de uma imagem em níveis de cinza

Para efetuar esse cálculo é necessário escolher uma métrica para fornecer a distância do pixel central da imagem (x_c, y_c) aos demais pixels (x_i, y_i). Como exemplos de métricas têm-se:

Distância euclidiana:

$$D((x_c, y_c), (x_i, y_i)) = \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2}$$

Distância “city block”, quarteirão ou Manhattan:

$$D((x_c, y_c), (x_i, y_i)) = |x_c - x_i| + |y_c - y_i|$$

Distância “sup”:

$$D((x_c, y_c), (x_i, y_i)) = \text{Máx}(|x_c - x_i|, |y_c - y_i|)$$

As Figuras 4.2, 4.3, e 4.4 representam as distâncias euclidiana, “city block” e distância “sup” de cada pixel (x_i, y_i) em relação ao pixel central (x_c, y_c) da Figura 4.1.



Figura 4.2 Distância euclidiana

4	3	2	3	4
3	2	1	2	3
2	1	0	1	2
3	2	1	2	3
4	3	2	3	4

Figura 4.3 Distância “city block”

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

Figura 4.4 Distância “sup”

Qualquer uma das três métricas pode ser utilizada no cálculo do coeficiente de Hurst (H). Nesta dissertação, a distância euclidiana foi a métrica escolhida.

A seguir calcula-se a diferença entre o maior e o menor nível de cinza dentre os pixels que estão em uma mesma região. Entenda-se, neste caso, por região, os pixels que estão a uma distância “d” do pixel central, e também, os pixels que possuem uma distância “r” do pixel central, onde $r < d$.

Para a primeira região ($d = 1$) da imagem (Figura 4.1), são identificados o maior (99) e o menor (75) nível de cinza. Fazendo a diferença (Δg) entre esses tons tem-se 24. Para a segunda região ($d = \sqrt{2}$), o maior nível de cinza é 105 e o menor nível é 67, portanto $\Delta g = 38$. Para a terceira região ($d = 2$), o maior nível de cinza é 107 e o menor é 60, logo $\Delta g = 47$. Esse processo é feito sucessivamente para as demais regiões da imagem (neste caso $d = \sqrt{5}$ e $d = \sqrt{8}$). Em seguida é computado o logaritmo natural das distâncias euclidianas e, também, das diferenças de níveis de cinza Δg . A Tabela 4.1 mostra toda computação feita utilizando a Figura 4.1.

Tabela 4.1 Elementos para o cálculo de H

Distância	$\ln(\text{Distância})$	Δg	$\ln(\Delta g)$
1	0	24	3,17805383
1,414213562	0,34657359	38	3,63758616
2	0,693147181	47	3,850147602
2,236067977	0,804718956	48	3,871201011
2,828427125	1,039720771	52	3,951243719

Após é construído um gráfico com os pontos formados pelo logaritmo natural das distâncias ($\ln d$) e pelo logaritmo natural das diferenças de nível de cinza ($\ln \Delta g$). A Figura 4.5 mostra o gráfico resultante da utilização dos dados da Tabela 4.1.

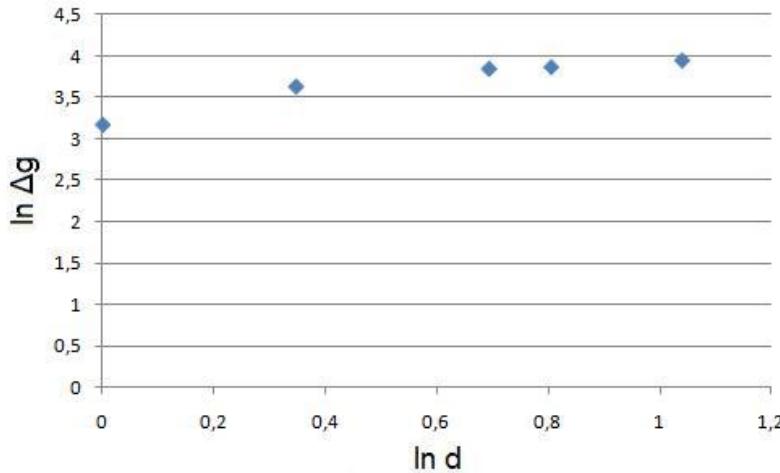


Figura 4.5 Gráfico de $\ln \Delta g$ x $\ln d$

Em seguida, através do método dos mínimos quadrados, é obtida a reta $y = ax + b$ que melhor se ajusta ao conjunto de pontos formados por $\ln \Delta g$ x $\ln d$. Os coeficientes a e b são obtidos da seguinte forma:

$$a = \frac{n \sum (\ln d \ln \Delta g) - \sum (\ln d) \sum (\ln \Delta g)}{n \sum (\ln d)^2 - \sum (\ln d)^2}$$

$$b = \frac{\sum (\ln \Delta g)}{n} - a \frac{\sum (\ln d)}{n}$$

A Figura 4.6 mostra o gráfico da Figura 4.5 com a reta que melhor se ajusta ao seu conjunto de pontos.

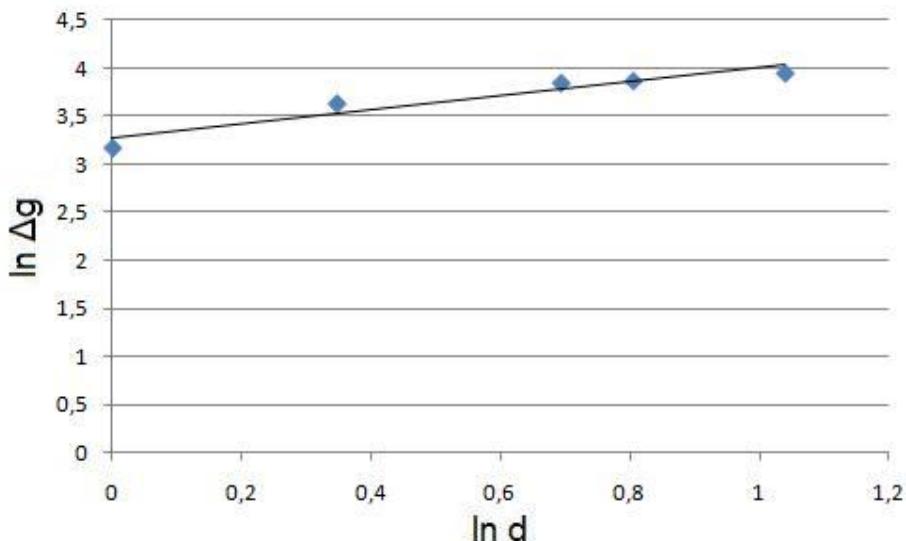


Figura 4.6 Reta obtida através do método dos mínimos quadrados

O resultado obtido foi $y = 0,7336x + 3,2745$. O coeficiente angular desta reta ($a = 0,7336$) é definido como coeficiente de Hurst da imagem.

4.1.1 Proposta do uso do Coeficiente de Hurst

As imagens “mama direita” e “mama esquerda” são percorridas por uma janela móvel (deslocada pixel a pixel para cada linha e coluna da imagem) que calcula o coeficiente de Hurst. Foram utilizadas janelas quadradas de lado $w = 5, 7, 9, 11, 13$ e 15 . Após isso, avalia-se a média e o desvio padrão dos coeficientes de Hurst calculados para cada janela w , dando origem às quatro primeiras características associadas ao H . A Figura 4.7 mostra um esquema que resumem quais foram os passos dados para a obtenção das primeiras quatro características para cada janela de lado w .

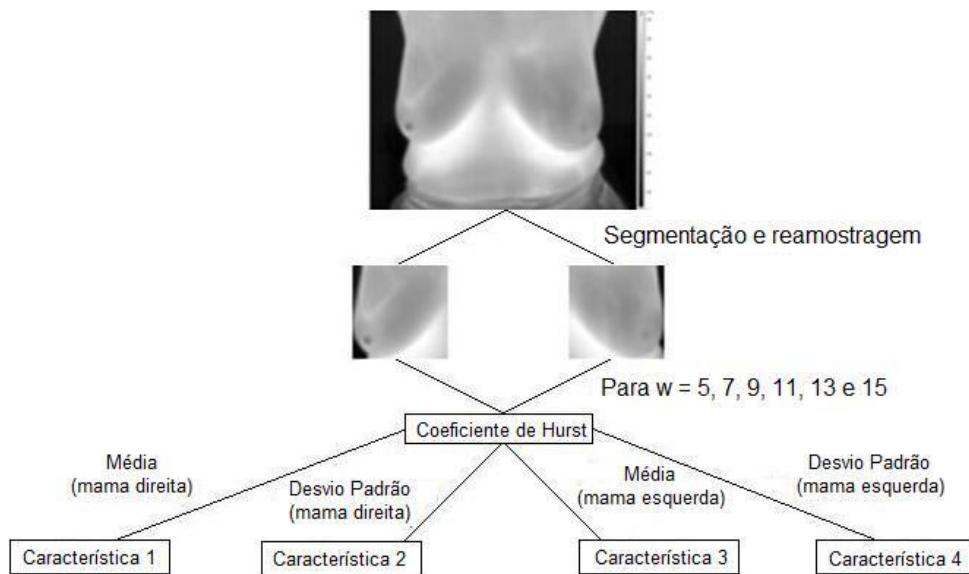


Figura 4.7 Esquema da primeira análise

Como foram utilizadas seis janelas de lado w ($w = 5, 7, 9, 11, 13$ e 15) e para cada janela extraiu-se quatro características, então esta primeira análise nos fornece 24 características. Através de uma segunda análise, usando a imagem Subtração, é possível extrair outras doze características. Estas foram extraídas calculando a média (característica 5) e o desvio padrão (característica 6) dos coeficientes de Hurst para cada uma das janelas de lado

w. A Figura 4.8 resume os passos que foram dados para a geração das características 5 e 6.

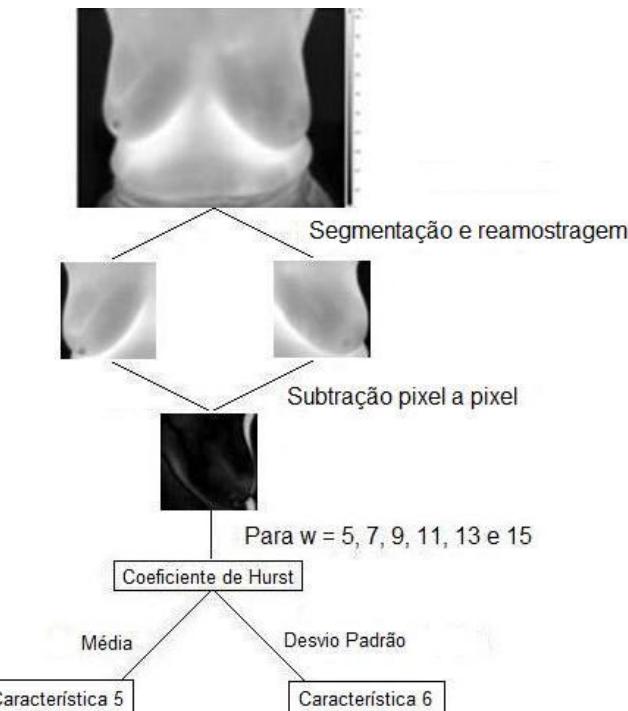


Figura 4.8 Esquema da segunda análise

Portanto, para cada imagem térmica, através das abordagens propostas (primeira e segunda análise) usando o coeficiente de Hurst foi realizada a extração de 36 características.

Para trabalhos futuros sugere-se que as janelas de tamanho w sejam deslocadas $3 \times 3, 5 \times 5, 7 \times 7, \dots, W \times W$ pixels ao invés de ser deslocada pixel a pixel para cada linha e coluna da imagem conforme feito nesta dissertação. Além disso, sugere-se também que para cada janela de tamanho w não sejam utilizadas todas as distâncias “ d ” para o cálculo de H . Por exemplo, para uma janela w , sejam utilizadas somente as distâncias “ d ” tais que $d < w/2$. Com isso, H é calculado para apenas parte da imagem contida na janela w . Ambas as sugestões de trabalhos futuros visam à análise de texturas de áreas e, no caso das imagens térmicas, essas texturas têm variação ao longo da imagem, por isso a importância de caracterizar as texturas de partes da imagem.

4.2 Lacunaridade

A lacunaridade é outra medida fractal. Ela avalia a quantidade ou a distribuição de lacunas de um objeto (Mandelbrot, 1983). Assim como o Coeficiente de Hurst, o uso da lacunaridade tem o objetivo de caracterizar texturas. Uma alta lacunaridade é atribuída a objetos que apresentam lacunas largas ou buracos. Quando o objeto em questão for uma imagem, quanto maior a variação na distribuição dos pixels da imagem, maior será o valor atribuído à lacunaridade. Sendo assim, um alto valor para a lacunaridade será associado à imagem que apresentar uma distribuição heterogênea da textura. Naturalmente, um baixo valor para a lacunaridade será associado a imagens cuja variação da distribuição de pixels se dá de forma homogênea. Aplicações da lacunaridade são encontradas em diversas áreas, entre elas: ecologia, medicina, processamento de imagem, entre outras (Rauch, 2007).

Um algoritmo de lacunaridade implementado anteriormente (Melo et al., 2006; Melo et al., 2006-a) pelo grupo de pesquisas em geometria fractal da UFF será utilizado para caracterizar as imagens térmicas. Neste algoritmo, a terceira coordenada (tom de cinza) é usada e a imagem é vista como uma coleção de voxels. A Figura 4.9 mostra uma imagem de 3x3 pixels em tons de cinza que será utilizada nas etapas seguintes desta abordagem para o entendimento deste algoritmo.

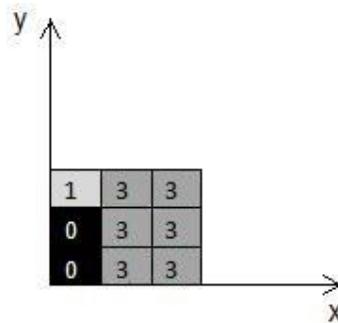


Figura 4.9 Imagem de 3 x 3 pixels em tons de cinza

A Figura 4.10 mostra a imagem da Figura 4.9 como uma coleção de voxels.

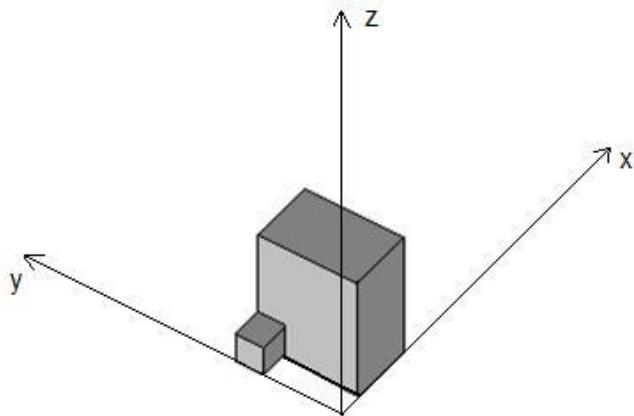


Figura 4.10 Imagem da Figura 4.9 em 3D

Os passos deste algoritmo são:

1- A imagem (vista como uma coleção de voxels) é ajustada por uma caixa delimitadora AABB (axial aligned bounding box). O tamanho S da AABB depende do tamanho do objeto (coleção voxels). A Figura 4.11 mostra a imagem da Figura 4.10 delimitada por uma AABB.

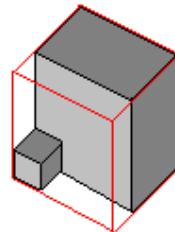


Figura 4.11 Exemplo de AABB

2- Após, o método “Gliding box” (Allain e Cloitre, 1991) é utilizado. Então, o primeiro voxel de uma caixa “deslizante” ($s \times s \times s$) é posicionado no canto do primeiro voxel da AABB e, em seguida, é computada a quantidade de voxels não-vazios no interior da caixa deslizante. Depois disso o primeiro voxel da caixa deslizante é posicionado no segundo voxel da AABB e novamente é computada a quantidade de voxels não-vazios no interior da caixa deslizante. Isso será feito até a última posição (N) da AABB. Com isso tem-se um grupo de elementos n_i , $i \in (1, 2, \dots, N)$, onde cada elemento corresponde à quantidade de voxels ocupados dentro da caixa deslizante na posição i . A notação $n_i (M_i, s)$ será usada para representar a freqüência de caixas de lado s com massa M_i . O

número total de caixas de lado s , $N(s)$, que percorrem a AABB é dado em função de seu tamanho (S), $N(s) = (S - s + 1)^3$. Uma função probabilidade que será utilizada no cálculo da lacunaridade é definida da seguinte maneira: $Q(M_i, s) = n_i (M_i, s) / N(s)$ (Papoulis, 1965). Esta função expressa a probabilidade de uma caixa deslizante de lado s conter M_i voxels não vazios. Então, define-se Lacunaridade para uma caixa deslizante de lado s , $\Lambda(s)$, do seguinte modo:

$$\Lambda(s) = \frac{\sum_{i=1}^N M_i^2 Q(M_i, s)}{(\sum_{i=1}^N M_i Q(M_i, s))^2}$$

Note que o cálculo de $\Lambda(1)$ não precisa ser realizado, pois este valor é apenas uma função da porcentagem de voxels ocupados. Além disso, $\Lambda(1)$ não se relaciona com os detalhes da distribuição de pixels. Sendo assim, deve-se calcular $\Lambda(s)$ para $s \geq 2$.

Computando a Lacunaridade para a Figura 4.9 com $s = 2$, tem-se: $n_i = \{8, 8, 8, 8, 5, 4, 4, 4\}$. A Tabela 4.2 mostra os resultados preliminares necessários para o cálculo de $\Lambda(2)$.

Tabela 4.2 Cálculo de $\Lambda(2)$

Massa i	$n_i (M_i, s)$	$Q(M_i, s)$	$M_i Q(M_i, s)$	$M_i^2 Q(M_i, s)$
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	3	0,375	1,5	6
5	1	0,125	0,625	3,125
6	0	0	0	0
7	0	0	0	0
8	4	0,5	4	32
Σ	8	1	6,125	41,125

Portanto $\Lambda(2) \approx 1,096$.

4.2.1 Proposta do uso da Lacunaridade

Anteriormente foi utilizado o Coeficiente de Hurst para a extração de características. Agora será feita a extração de características das imagens (MD, ME e S) usando outra medida fractal: a Lacunaridade. Inicialmente, para as imagens MD e ME (inteiras) foram calculados os valores de Lacunaridade $\Lambda(s)$ para $2 \leq s \leq 25$. Com isso, quarenta e oito características são extraídas. A Figura 4.12 apresenta um esquema com os passos seguidos para a obtenção dessas quarenta e oito características.

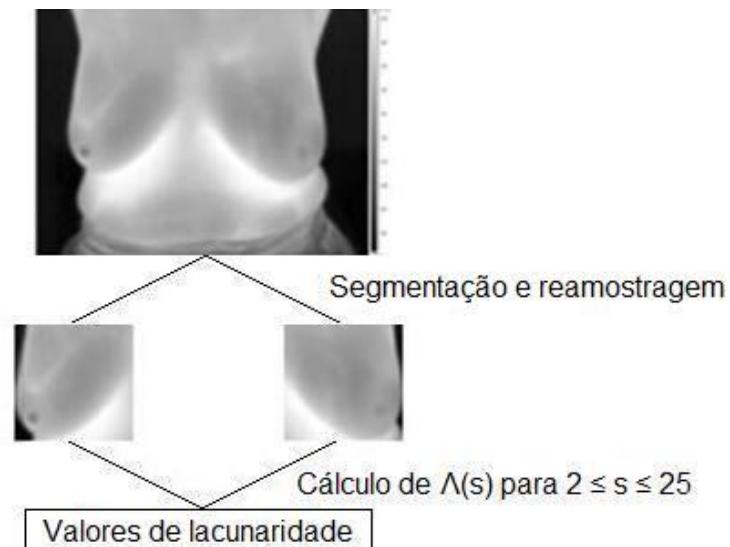


Figura 4.12 Passos para a extração de 48 características

Em seguida foram calculados os valores de lacunaridade para a imagem S. Com isso, computando $\Lambda(s)$ para $2 \leq s \leq 25$ são geradas vinte e quatro características (chamadas de “24 características da imagem S”). A Figura 4.13 mostra um esquema com um resumo dos passos utilizados para a extração dessas características.

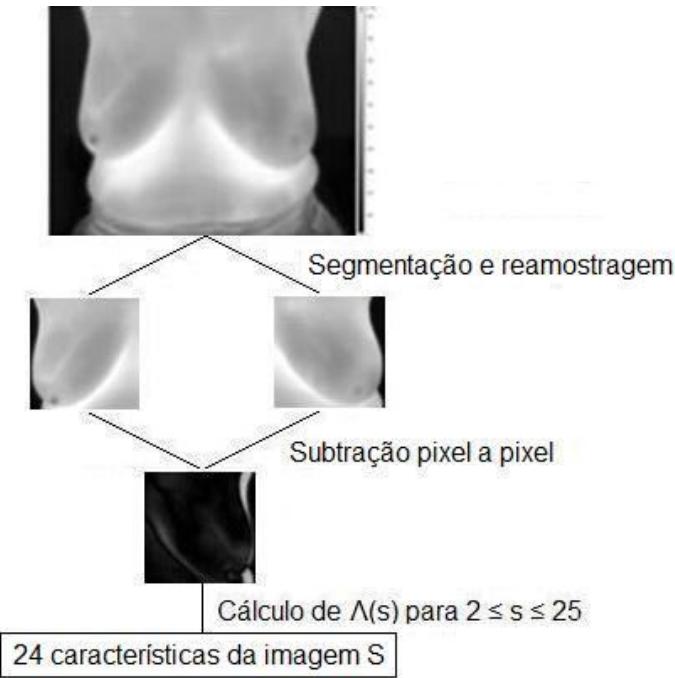


Figura 4.13 Passos para a extração de 24 características da imagem Subtração

Através da análise dos gráficos do \ln (Lacunaridade) x \ln (tamanho da caixa deslizante s) é possível obter outras vinte e cinco características a serem descritas mais adiante. A Figura 4.14 e a Figura 4.15 apresentam os gráficos do \ln (Lacunaridade) x \ln (tamanho da caixa deslizante s) de uma paciente saudável e de uma paciente com patologia respectivamente. Nestes gráficos, cada curva corresponde a uma das mamas da paciente.

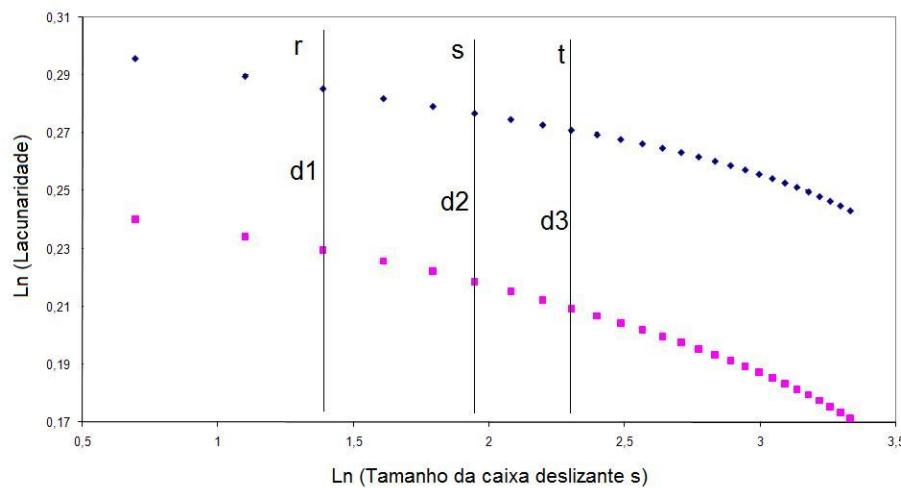


Figura 4.14 Gráfico de Lacunaridade típico de uma paciente saudável

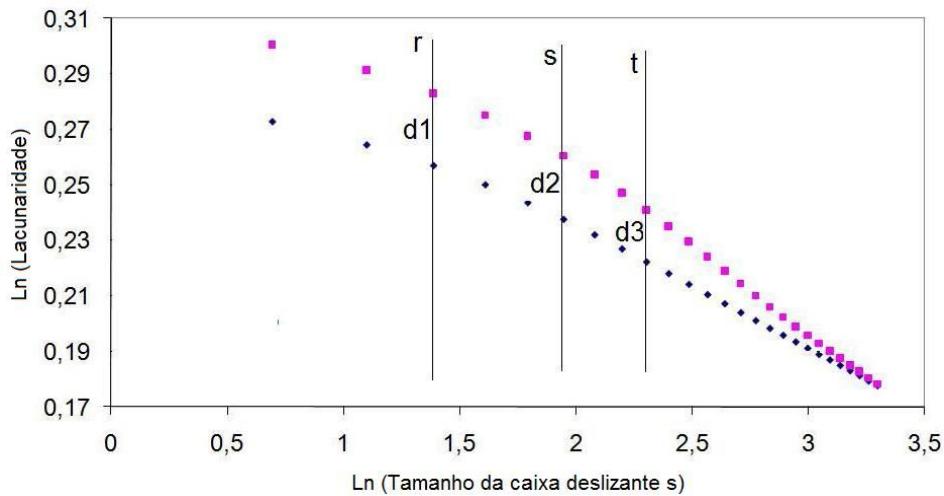


Figura 4.15 Gráfico de Lacunaridade típico de uma paciente com patologia

Observando os gráficos foi percebido que pacientes com mamas saudáveis apresentavam curvas de lacunaridade similares, porém transladas em relação ao eixo vertical. Pacientes com patologias na mama apresentavam curvas de lacunaridade que tendiam a se interceptar em um ponto.

Após ter percebido isso é necessário “transformar” essa idéia em números para que possam ser interpretados pelo computador no estágio de classificação. Para isso, considere as retas r, s e t paralelas ao eixo \ln (Lacunaridade) como nas Figuras 4.14 e 4.15. Cada reta intercepta cada curva em um ponto gerando segmentos de tamanhos d_1 , d_2 e d_3 respectivamente. Quando se trata do gráfico de uma paciente saudável, d_1 , d_2 e d_3 tendem a ter o mesmo tamanho. Enquanto que nos gráficos de pacientes com alguma patologia, $d_1 > d_2 > d_3$. Naturalmente, o desvio padrão de d_1 , d_2 e d_3 de pacientes saudáveis será menor do que o desvio padrão de d_1 , d_2 e d_3 de pacientes com patologia, pois essa medida estatística avalia a regularidade de dados. A partir desse momento, que se possui uma interpretação matemática do comportamento das curvas, pode ser feita a descrição da maneira como as vinte e cinco características serão extraídas. Para isso considere as seguintes notações: $\Delta D(s)$ e $\Delta E(s)$ para nos referirmos aos valores de lacunaridade da mama direita e da mama esquerda respectivamente. Fazendo $|\Delta D(s) - \Delta E(s)|$, para $2 \leq s \leq 25$ tem-se vinte e quatro características. Essas serão chamadas de “24 características dif” (correspondentes à d_1 , d_2 , d_3 , d_4 , d_5 , ..., d_{24}). A Figura

4.16 mostra um esquema com as etapas para a extração dessas características.

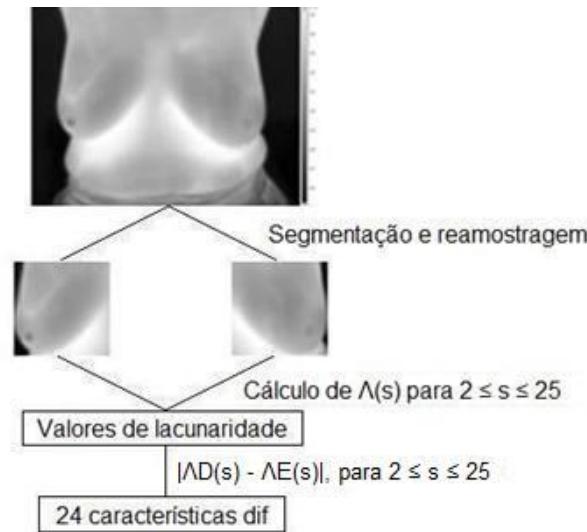


Figura 4.16 Passos para a extração de 24 características

Calculando o desvio padrão destas vinte e quatro características é obtida uma nova característica (característica vinte e cinco). Esta (que será chamada de DP) descreve o comportamento das curvas nos gráficos do \ln (Lacunaridade) x \ln (tamanho da caixa deslizante s). A Figura 4.17 mostra um esquema que resume os passos que foram dados para a extração desta característica.

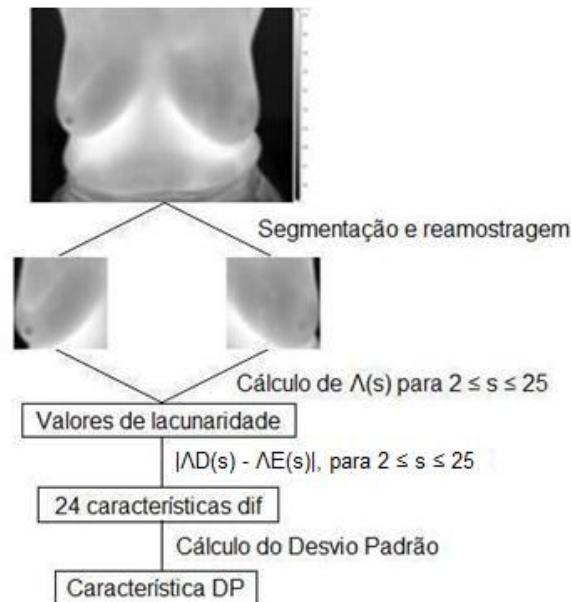


Figura 4.17 Passos para a extração da característica DP

Com isso, através da lacunaridade foram extraídas as seguintes características: vinte e quatro características de lacunaridade da mama direita ($\Lambda D(s)$ para $2 \leq s \leq 25$), vinte e quatro características de lacunaridade da mama esquerda ($\Lambda E(s)$ para $2 \leq s \leq 25$), vinte e quatro características de lacunaridade da imagem S ($\Lambda S(s)$ para $2 \leq s \leq 25$), vinte e quatro características de lacunaridade através de $|\Lambda D(s) - \Lambda E(s)|$ para $2 \leq s \leq 25$ (24 características dif), uma característica (característica DP) obtida através do desvio padrão de $|\Lambda D(s) - \Lambda E(s)|$ para $2 \leq s \leq 25$. Sendo assim, usando a medida fractal Lacunaridade foi possível extrair noventa e sete características de cada imagem térmica. A Figura 4.18 mostra um esquema com as características obtidas utilizando a Lacunaridade.

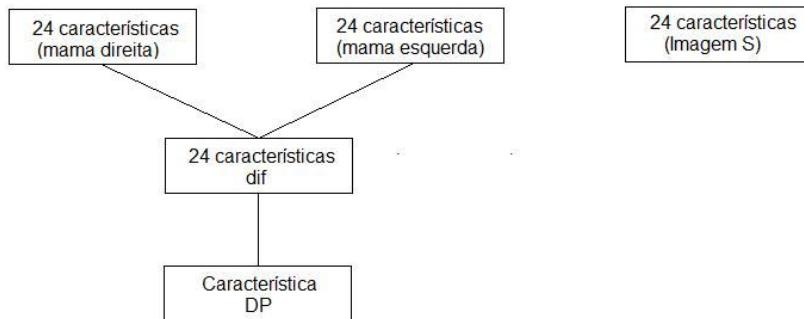


Figura 4.18 Todas as características obtidas através da Lacunaridade

Neste trabalho usaram-se as imagens (MD, ME e S) inteiras para o cálculo da Lacunaridade. Sugere-se, como trabalho futuro, a verificação do cálculo da lacunaridade utilizando janelas (como foi feito com o Coeficiente de Hurst). Além disso, cada imagem (MD, ME e S) pode ser dividida em um número qualquer de partes iguais e, em seguida, ser calculada a Lacunaridade de cada uma das partes. As sugestões de trabalhos futuros apresentadas têm como objetivo a análise de texturas de áreas e, no caso das imagens térmicas, essas texturas têm variação ao longo da imagem, por isso a importância de characterizar as texturas de partes da imagem.

Assim como as características extraídas através do coeficiente de Hurst, as características obtidas por Lacunaridade (pelo modo que foram extraídas) buscam avaliar a simetria da distribuição de temperaturas nas mamas das pacientes.

5. Avaliação dos resultados obtidos

A análise da curva ROC proporciona uma forma de selecionar abordagens ótimas e descartar outras “sub-ótimas”. Ela será usada, neste capítulo, para “ranquear” as metodologias propostas do uso do coeficiente de Hurst e da Lacunaridade.

Com o objetivo de analisar o grau de viabilidade das características extraídas para descrever a simetria da distribuição de temperaturas das mamas nas imagens térmicas das pacientes, dividiram-se as características em 14 grupos. A seguir o conjunto de características de cada grupo é descrito.

Grupo 1: Composto por todas as cento e trinta e três características.

Grupo 2: Composto por todas as características extraídas através do Coeficiente de Hurst (trinta e seis características).

Grupo 3: Composto pelas características $C = 1, 2, 3$ e 4 para $w= 5, 7, 9, 11, 13$ e 15 (vinte e quatro características).

Grupo 4: Composto pelas características $C = 2$ e 4 para $w= 5, 7, 9, 11, 13$ e 15 (doze características).

Grupo 5: Composto pelas características $C = 1$ e 3 para $w= 5, 7, 9, 11, 13$ e 15 (doze características).

Grupo 6: Composto pelas características $C = 5$ e 6 para $w= 5, 7, 9, 11, 13$ e 15 (doze características).

Grupo 7: Composto pela característica $C = 5$ para $w= 5, 7, 9, 11, 13$ e 15 (seis características).

Grupo 8: Composto pela característica $C = 6$ para $w= 5, 7, 9, 11, 13$ e 15 (seis características).

Grupo 9: Composto por todas as características extraídas através de Lacunaridade (noventa e sete características).

Grupo 10: Composto por $\Delta D(s)$ e $\Delta E(s)$ com $2 \leq s \leq 25$ (quarenta e oito características).

Grupo 11: Composto por $|\Delta D(s) - \Delta E(s)|$ com $2 \leq s \leq 25$ (vinte e quatro características).

Grupo 12: Composto pelo desvio padrão de $|\Delta D(s) - \Delta E(s)|$ com $2 \leq s \leq 25$ (uma característica).

Grupo 13: Composto por $\Lambda S(s)$ com $2 \leq s \leq 25$ (vinte e quatro características).

Grupo 14: Composto pelas características dos grupos 12 e 13 (vinte e cinco características).

A Figura 5.1 mostra um esquema com a organização dos grupos de características.

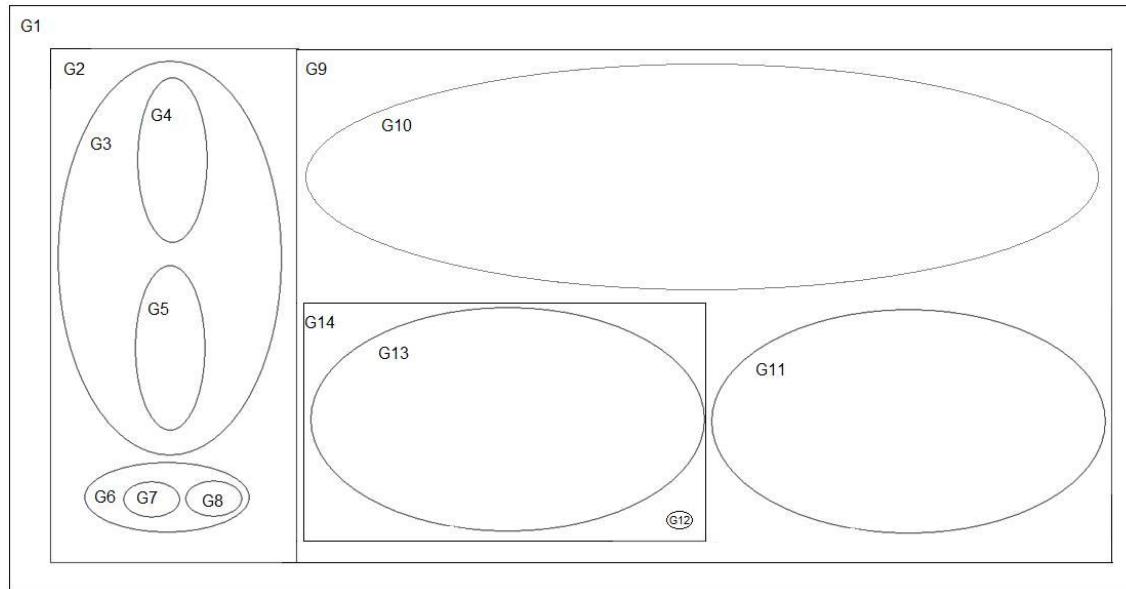


Figura 5.1 Grupos de características

Através da comparação dos resultados obtidos num determinado teste com os resultados fornecidos por avaliações de especialistas (médicos) determinam-se a sensibilidade e a especificidade. Define-se como “sensibilidade” a probabilidade de um indivíduo doente apresentar um teste positivo. A “especificidade” é definida como a probabilidade de um indivíduo saudável apresentar um teste negativo.

A curva ROC (*Receiver Operating Characteristic*) é uma forma de representação da relação entre a sensibilidade e a especificidade. A construção desta curva é dada pelo gráfico que representa a sensibilidade versus $(1 - \text{especificidade})$ para um conjunto de limiares (também conhecidos como valores de ponto de corte). Quando o resultado da aplicação de um teste diagnóstico quantitativo trata-se de uma variável contínua e se quer transformá-la numa variável que forneça apenas dois resultados (saudável ou doente)

então é necessária a utilização de algum valor na escala contínua que diferencie as duas classes. Esse valor é denominado “ponto de corte”.

A área abaixo da curva ROC é proporcional à exatidão de um teste diagnóstico. Com isso, quanto maior o valor da área abaixo da curva ROC maior será a exatidão do teste. O resultado máximo que pode ser atribuído à área abaixo da curva ROC é 1.

Com o uso do software WEKA foi possível obter o resultado da área abaixo da curva ROC utilizando os classificadores citados no Capítulo 3 com as características extraídas no Capítulo 4. Neste software foi selecionada a opção “validação cruzada”. Com isso, um classificador é projetado usando os dados extraídos de 27 pacientes (total de pacientes - 1) e é testado usando os dados da paciente restante. Isso é repetido 28 vezes. Assim, os dados de cada paciente são usados como um caso de teste e os demais são usados para projetar o classificador. Isso é feito para evitar que se construa um modelo tendencioso.

As Tabelas 5.1 e 5.2 mostram os resultados das áreas abaixo das curvas ROC dos quatorze grupos.

Tabela 5.1 Resultados das áreas abaixo das curvas ROC para os grupos 1 ao 7

Técnica	G1	G2	G3	G4	G5	G6	G7
BLR	0,5	0,5	0,5	0,5	0,5	0,5	0,5
BN	0	0	0	0	0	0	0
CNB	0,375	0,5	0,5	0,458	0,458	0,5	0,5
DMNBT	0,125	0,125	0,125	0,125	0,125	0,125	0,125
NB	0,708	0,875	0,927	0,958	0,667	0,49	0,594
NBM	0,396	0	0	0	0	0	0
NBMU	0,5	0	0,01	0,042	0,01	0	0
NBS	0,792	0,854	0,885	0,927	0,667	0,484	0,625
NBU	0,708	0,875	0,927	0,958	0,667	0,49	0,594
L	0,503	0,469	0,38	0,26	0,635	0,589	0,146
MP	0,667	0,729	0,75	0,719	0,583	0,146	0,385
RBFN	0,438	0,552	0,833	0,948	0,604	0,583	0,438
SL	0,615	0,354	0,438	0,417	0,698	0,469	0,323
SMO	0,417	0,479	0,5	0,5	0,5	0,5	0,5
VP	0,135	0,083	0,031	0,031	0,047	0,125	0,125
IB1	0,521	0,813	0,667	0,667	0,708	0,563	0,417
IBk	0,521	0,813	0,667	0,667	0,708	0,563	0,417
KS	0,656	0,698	0,771	0,844	0,599	0,688	0,344

LWL	0,417	0,469	0,469	0,417	0,49	0,438	0,51
ABM1	0,604	0,49	0,365	0,542	0,521	0,344	0,448
ASC	0	0	0	0	0	0	0
B	0,25	0,276	0,323	0,276	0,552	0,323	0,313
CVC	0,771	0,375	0,771	0,792	0,646	0,5	0,708
CVR	0,167	0,167	0,188	0,375	0,688	0	0
CVPS	0	0	0	0	0	0	0
D	0,323	0,333	0,438	0,438	0,479	0,281	0,333
Dec	0,635	0,438	0,521	0,604	0,615	0,458	0,26
END	0	0,396	0,375	0,495	0,109	0,729	0
ES	0,521	0,521	0,531	0,563	0,481	0,479	0,469
FC	0	0	0	0	0	0	0
Grad	0,5	0,5	0,5	0,5	0,5	0,5	0,5
LB	0,51	0,354	0,542	0,563	0,573	0,396	0,333
MBAB	0,396	0,417	0,24	0,396	0,479	0,469	0,531
MCC	0,503	0,469	0,38	0,26	0,635	0,589	0,146
MS	0	0	0	0	0	0	0
OCC	0	0,396	0,375	0,495	0,109	0,729	0
RILB	0	0	0	0	0	0	0
RC	0,516	0,865	0,667	0,833	0,734	0,411	0,391
RSS	0,083	0,083	0,083	0,167	0,25	0,125	0,125
RF	0,75	0,521	0,422	0,927	0,448	0,365	0,094
Stac	0	0	0	0	0	0	0
StacC	0	0	0	0	0	0	0
TS	0,51	0,37	0,38	0,359	0,594	0,542	0,115
V	0	0	0	0	0	0	0
CBND	0	0,396	0,375	0,495	0,109	0,729	0
DNBND	0	0,396	0,375	0,495	0,109	0,729	0
ND	0	0,396	0,375	0,495	0,109	0,729	0
FLR	0,5	0,5	0,5	0,5	0,5	0,479	0,479
HP	0,583	0,708	0,688	0,724	0,474	0,672	0,599
VFI	0,427	0,823	0,792	0,813	0,552	0,719	0,667
AD Tree	0,51	0,448	0,656	0,708	0,5	0,521	0,667
BF Tree	0,109	0,073	0,182	0,313	0,542	0,365	0,563
DS	0,167	0,417	0,417	0,5	0,625	0,547	0,51
FT	0,5	0,5	0,5	0,5	0,5	0,5	0,5
J48	0	0,396	0,375	0,495	0,109	0,729	0
J48graft	0,042	0,438	0,396	0,396	0,109	0,766	0
LAD Tree	0,667	0,042	0,031	0,063	0,51	0,24	0,271
LMT	0,615	0,396	0,479	0,417	0,698	0,469	0,323
NB Tree	0	0	0	0	0	0	0
RF	0,531	0,714	0,698	0,734	0,625	0,5	0,443
Random Tree	0,583	0,708	0,479	0,438	0,563	0,521	0,458
REP Tree	0,125	0,125	0,125	0,125	0,167	0,125	0,208
SC	0	0	0	0	0,286	0	0

CR	0,458	0,417	0,417	0,438	0,417	0,438	0,438
DT	0,078	0,104	0,104	0,104	0,104	0,104	0,104
DTNB	0,078	0,104	0,104	0,104	0,104	0,104	0,104
Jrip	0	0	0,109	0,292	0,656	0	0,292
Nnge	0,5	0,5	0,479	0,458	0,75	0,5	0,458
One R	0,583	0,417	0,438	0,5	0,438	0,458	0,5
Part	0	0,396	0,375	0,495	0,109	0,729	0
Ridor	0,5	0,5	0,5	0,479	0,458	0,458	0,458
Zeror	0	0	0	0	0	0	0

Para o Grupo 1, a maioria dos classificadores não apresentou resultado bom (entre 0,8 e 0,9) ou excelente (entre 0,9 e 1,0). Isso ocorreu devido ao fato de algumas características não serem tão precisas na descrição das imagens térmicas. Com isso, essas características “poluem” a construção do modelo de classificação. Por isso a importância de separar todas as características em diversos grupos para avaliar quais características descrevem de forma precisa as imagens térmicas. Conforme destacado em negrito, na tabela dos resultados para o Grupo 1, os classificadores que apresentaram os melhores resultados foram: Naive Bayes Simple (0,792), Classification Via Clustering (0,771), Rotation Forest (0,75), Naive Bayes (0,708) e Naive Bayes Updateable (0,708).

Como pode se perceber através da tabela com os resultados para o Grupo 2, foi possível obter melhores resultados do que o Grupo 1. Para este Grupo 2 foram obtidos resultados bons para os classificadores Naive Bayes e Naive Bayes Updateable (0,875), Random Committee (0,865), Naive Bayes Simple (0,854), Voting Feature Intervals (0,823) e Instance Based on K-nearest Neighbours Classifier (0,813). Outros resultados regulares (entre 0,7 e 0,8) foram alcançados, são eles: Multilayer Perceptron (0,729), Random Forest (0,714), Hyper Pipes (0,708) e Random Tree (0,708).

Para o Grupo 3, com os classificadores Naive Bayes e Naive Bayes Updateable foram obtidos resultados excelentes (área abaixo da curva ROC = 0,927). Note que com apenas vinte e quatro características foi possível obter um valor elevadíssimo para a área abaixo da curva ROC. Os classificadores Naive Bayes Simple (0,885) e Radial Basis Function Network (0,833) apresentaram resultados bons. Outros classificadores apresentaram resultados regulares para esse Grupo 3, são eles: Voting Feature Intervals (0,792),

Instance Based Classifier (0,771), Classification Via Clustering (0,771) e Multilayer Perceptron (0,75).

Os melhores resultados dessa dissertação foram obtidos utilizando as características do Grupo 4. Este grupo, com apenas doze características, obteve diversos resultados excelentes, são eles: Naive Bayes (0,958), Naive Bayes Updateable (0,958), Radial Basis Function Network (0,948), Naive Bayes Simple (0,927) e Rotation Forest (0,927).

Devido aos resultados não tão elevados, quando comparados com grupos 1, 2 e 3 pode-se concluir que o Grupo 5 é um dos que “polui” a descrição das imagens térmicas destes grupos. Non Nested Generalized Exemplars (0,75), Random Committee (0,734), Instance Based on Nearest Neighbour Classifier (0,708) e Instance Based on K-nearest Neighbours Classifier (0,708) apresentaram os melhores resultados para este grupo.

Assim como o Grupo 5, o Grupo 6 não apresentou resultados bons ou excelentes e, por isso, também contribuiu para a “poluição” da descrição das imagens térmicas do Grupo 1 e do Grupo 2. Os melhores resultados para este grupo foram alcançados com os classificadores Class for generating a grafted decision tree (0,766), Ensemble of Nested Dichotomies (0,729), Ordinal Class Classifier (0,729), Class Balanced Nested Dichotomies (0,729), Data Near Balanced Nested Dichotomies (0,729), Nested Dichotomies (0,729), Class for generating a pruned or unpruned decision tree (0,729), Partial Decision Tree (0,729) e Voting Feature Intervals (0,719).

Da mesma forma que os Grupos 5 e 6, o Grupo 7 não alcançou resultados bons ou excelentes. Deve ser observado o que os três grupos têm em comum: o fato de utilizarem a média. Classification Via Clustering apresentou o melhor resultado (0,708) para este grupo.

Tabela 5.2 Resultados das áreas ROC para os grupos 8 ao 14

Técnica	G8	G9	G10	G11	G12	G13	G14
BLR	0,5	0,5	0,5	0,5	0,5	0,5	0,5
BN	0	0	0	0	0	0	0
CNB	0,5	0,479	0,396	0,438	0,5	0,667	0,458
DMNBT	0,125	0,125	0,125	0,125	0,125	0,125	0,125
NB	0,438	0,49	0,488	0,313	0,625	0,443	0,453
NBM	0	0,365	0	0	0	0,292	0,208

NBMU	0	0,5	0,135	0,24	0	0,073	0,073
NBS	0,458	0,557	0,542	0,281	0,698	0,417	0,422
NBU	0,438	0,49	0,488	0,313	0,625	0,443	0,453
L	0,365	0,552	0,063	0,531	0,771	0,38	0,615
MP	0,396	0,469	0,479	0,542	0,688	0,5	0,719
RBFN	0,542	0,302	0,427	0,099	0,531	0,5	0,542
SL	0,542	0,823	0,646	0,25	0,792	0	0,771
SMO	0,5	0,5	0,5	0,5	0,5	0,5	0,5
VP	0,104	0,135	0,016	0,135	0,021	0,146	0,146
IB1	0,729	0,479	0,438	0,479	0,563	0,438	0,563
IBk	0,729	0,479	0,438	0,479	0,563	0,438	0,563
KS	0,771	0,406	0,406	0,38	0,625	0,344	0,49
LWL	0,547	0,635	0,531	0,073	0,677	0,417	0,594
ABM1	0,625	0,677	0,406	0,063	0,615	0,073	0,677
ASC	0	0	0	0	0,688	0	0,688
B	0,375	0,292	0,359	0,172	0,323	0,25	0,328
CVC	0,563	0,688	0,625	0,479	0,813	0,417	0,438
CVR	0	0,688	0,49	0	0,833	0	0,792
CVPS	0	0	0	0	0	0	0
D	0,172	0,328	0,5	0	0,542	0,25	0,25
Dec	0,292	0,604	0,51	0,208	0,708	0,24	0,667
END	0,833	0,688	0	0	0,688	0	0,688
ES	0,479	0,417	0,479	0,479	0,406	0,479	0,396
FC	0	0	0	0	0	0	0
Grad	0,5	0,5	0,5	0,5	0,5	0,5	0,5
LB	0,625	0,656	0,365	0,073	0,615	0,406	0,698
MBAB	0,583	0,594	0,323	0,198	0,698	0,229	0,604
MCC	0,365	0,552	0,063	0,531	0,771	0,38	0,615
MS	0	0	0	0	0	0	0
OCC	0,833	0,688	0	0	0,688	0	0,688
RILB	0	0	0	0	0	0	0
RC	0,745	0,568	0,391	0,271	0,563	0,542	0,557
RSS	0,125	0,25	0,125	0,125	0,333	0,125	0,333
RF	0,49	0,552	0,302	0,542	0,688	0,271	0,25
Stac	0	0	0	0	0	0	0
StacC	0	0	0	0	0	0	0
TS	0,344	0,506	0,063	0,526	0,417	0,453	0,536
V	0	0	0	0	0	0	0
CBND	0,833	0,688	0	0	0,688	0	0,688
DNBND	0,833	0,688	0	0	0,688	0	0,688
ND	0,833	0,688	0	0	0,688	0	0,688
FLR	0,458	0,5	0,438	0,458	0,563	0,458	0,479
HP	0,615	0,536	0,531	0,354	0,563	0,5	0,536
VFI	0,656	0,344	0,479	0,271	0,583	0,417	0,469
AD Tree	0,781	0,656	0,406	0,208	0,615	0,219	0,615

BF Tree	0,219	0,635	0	0	0,667	0	0,667
DS	0,625	0,458	0,51	0	0,458	0	0,458
FT	0,5	0,5	0,5	0,5	0,5	0,5	0,5
J48	0,833	0,688	0	0	0,688	0	0,688
J48graft	0,875	0,229	0,042	0	0,688	0	0,458
LAD Tree	0,615	0,646	0,135	0,302	0,781	0,563	0,854
LMT	0,583	0,823	0,646	0,25	0,792	0	0,771
NB Tree	0	0	0	0	0	0	0
RF	0,594	0,448	0,292	0,297	0,745	0,464	0,63
Random Tree	0,688	0,583	0,396	0,438	0,563	0,438	0,604
REP Tree	0,125	0,25	0,125	0,125	0,333	0,125	0,333
SC	0	0,333	0	0	0,375	0	0,375
CR	0,438	0,479	0,438	0,458	0,521	0,458	0,521
DT	0,104	0,078	0,104	0,104	0,078	0,104	0,078
DTNB	0,104	0,078	0,104	0,104	0,078	0,104	0,078
Jrip	0,458	0,656	0	0	0,479	0	0,688
Nnge	0,458	0,438	0,5	0,458	0,563	0,5	0,667
One R	0,458	0,583	0,5	0,5	0,708	0,5	0,708
Part	0,833	0,688	0	0	0,688	0	0,688
Ridor	0,458	0,708	0,458	0,5	0,458	0,5	0,458
Zeror	0	0	0	0	0	0	0

Para o Grupo 8, com apenas seis características, muitos classificadores conseguiram obter resultados bons para a área abaixo da curva ROC. São eles: Class for generating a grafted decision tree (0,875), Ensemble of Nested Dichotomies (0,833), Ordinal Class Classifier (0,833), Class Balanced Nested Dichotomies (0,833), Data Near Balanced Nested Dichotomies (0,833), Nested Dichotomies (0,833), Class for generating a pruned or unpruned decision tree (0,833) e Partial Decision Tree (0,833).

Embora alguns classificadores tenham apresentado resultados bons (Simple Logistic (0,823) e Logistic Model Trees (0,823)) para o conjunto das características de lacunaridade (Grupo 9), é provável que utilizando todos os classificadores para alguns subgrupos de características do Grupo 9 se consigam melhores resultados. A seguir, as análises dos resultados desses subgrupos.

Como se pode perceber, o Grupo 10 e o Grupo 11 não apresentaram resultados bons.

Conforme esperado, um dos subgrupos (Grupo 12) do Grupo 9 apresentou melhor resultado de classificação (Classification Via Regression

(0,833)) do que este. Além disso, como foi visto anteriormente, o Grupo 10 e o Grupo 11 não tiveram resultados satisfatórios, no entanto para o Grupo 12 (combinação das características dos Grupos 10 e 11) foram alcançados resultados bons (Classification Via Regression (0,833) e Classification Via Clustering (0,813)). Sendo assim, além de se fazer uma seleção adequada de métodos que descrevam as imagens térmicas no processo de extração de características, é importante a análise e a interpretação dos resultados obtidos nesse processo para que se possam fazer as operações devidas e, com isso, conseguir melhores resultados.

O Grupo 13 não apresentou resultados bons.

Note que, o melhor resultado para o Grupo 12 foi 0,833 (Classification Via Regression) e que o Grupo 13 não obteve resultados bons. No entanto, unindo estes dois grupos de características (Grupo 14) foi possível conseguir um resultado melhor: Logitboost Alternating Decision Tree (0,854). Isso ocorreu pois ambos os grupos (Grupo 12 e Grupo 13) analisam a simetria na distribuição de temperaturas das mamas e quando unidos acabam descrevendo de forma mais precisa as imagens. Novamente, isso mostra que não basta apenas extrair características das imagens, neste problema tão complexo é necessário também saber como mesclar e utilizá-las para a obtenção de melhores resultados.

6. Conclusões e trabalhos futuros

Neste capítulo serão descritas as conclusões relativas ao uso do Coeficiente de Hurst e da Lacunaridade para a caracterização de imagens térmicas e, também, direções para trabalhos futuros nesta linha.

Através das Tabelas 5.1 e 5.2 pode se perceber que para as abordagens propostas houve grupos que conseguiram apresentar resultados excelentes (G3 e G4), grupos com resultados bons (G2, G8, G9, G12, G14) e grupos com resultados regulares (G1, G5, G6, G7). Por outro lado, alguns grupos apresentaram somente resultados ruins ou insignificantes (G10, G11, G13). Isso mostra que não é simplesmente o uso do Coeficiente de Hurst ou da Lacunaridade que geram resultados excelentes ou bons, mas sim a forma como estes são abordados.

Os Grupos 10, 11 e 13 apresentaram somente resultados ruins ou insignificantes. Deve ser observado que os Grupos 10 e 13 são compostos pelos resultados dos cálculos de lacunaridade para as imagens MD, ME e S. Além disso, o Grupo 11 é composto por características resultantes de subtrações das características do Grupo 10. Isso reforça a idéia de que não é somente o uso da Lacunaridade que proporciona bons resultados.

Observe que o Grupo 12, composto por uma única característica (resultado da “transformação” dos valores obtidos pelo algoritmo de lacunaridade) apresentou resultados bons. Além disso, o grupo 12 “impulsionou” os resultados (bons) apresentados pelos Grupos 9 e 14.

O Grupo 14 apresentou resultados melhores do que cada um dos grupos que o compõem (Grupos 12 e 13).

Dentre os grupos que tiveram no máximo resultados regulares tem-se: G1, G5, G6 e G7. Três desses grupos (G5, G6 e G7) tem a medida estatística média em comum.

É importante notar que o Grupo 4 com apenas doze características (bem selecionadas) obteve resultados excelentes. O Grupo 3 com vinte e quatro características também obteve tais resultados.

O Grupo 2 apresentou resultados bons utilizando trinta e seis características. Enquanto que o Grupo 8 conseguiu obter estes resultados com apenas seis.

É notável que os grupos que utilizaram a medida estatística desvio padrão alcançaram os melhores resultados. Enquanto que os piores foram obtidos pelos grupos que utilizaram a medida estatística média.

Diversas aplicações já utilizaram o coeficiente de Hurst (H) para a análise de imagens. Esta dissertação apresentou uma nova abordagem para o uso deste coeficiente. Diferentemente de H , a Lacunaridade até então não vinha sendo muito utilizada para a caracterização de imagens.

Após a análise dos resultados obtidos, pode-se concluir que não é somente o uso do coeficiente de Hurst e da Lacunaridade os responsáveis pelos resultados excelentes ou bons que foram alcançados. Estes são frutos das abordagens propostas para o uso dessas medidas fractais aqui apresentadas. Sendo assim estas medidas como foram exploradas nesta dissertação podem ser usadas para descrever adequadamente imagens térmicas de mamas.

6.1 Trabalhos Futuros

Para esta linha são sugeridos os seguintes trabalhos futuros:

- 1) No cálculo do coeficiente de Hurst, avaliar a influência de alterar o critério de parada (a distância “d” usada em cada janela). Ou seja, sugere-se que para cada janela de tamanho w não sejam utilizadas todas as distâncias “d” para o cálculo de H . Por exemplo, para uma janela w , sejam utilizadas somente as distâncias “d” tais que $d < w/2$. Com isso, H é calculado para apenas parte da imagem contida na janela w .
- 2) Sugere-se, ainda no cálculo de H , investigar se a metodologia ainda continua eficiente se as janelas de tamanho w forem deslocadas 3x3, 5x5, 7x7, ..., $W \times W$ pixels ao invés de serem deslocadas pixel a pixel para cada linha e coluna da imagem conforme feito nesta dissertação.
- 3) O uso de outras métricas, ao invés da euclidiana.
- 4) Combinação das sugestões de 1 a 3.

- 5) Utilização de outras medidas estatísticas. Por exemplo: curtose, “skewness” e momentos estatísticos de ordem superior (Papoulis, 1965).
- 6) Extração automática da região de interesse, como vem sendo testado (Serrano et al., 2009).
- 7) Teste de outras formas de combinar as imagens MD e ME. Por exemplo, ao invés do módulo da subtração poderia ter sido feito a soma pixel a pixel, ou mesmo operações lógicas entre as imagens.
- 8) Extrair características das regiões de interesse sem reamostrá-las. Nesta dissertação isso feito devido ao alto custo computacional do algoritmo de lacunaridade.
- 9) Verificar a possibilidade de usar técnicas de decisão mais simples como, por exemplo, o uso de funções discriminantes para a classificação das imagens. Isso havia sido tentado em trabalho anterior para um número menor de imagens e talvez possa ser reconsiderado à medida que os casos de imagens com diagnóstico fechado do banco evoluam (Silveira et al., 2009).
- 10) Testar a metodologia para um número maior de imagens térmicas.
- 11) Como o banco de imagens apresenta imagens da mesma paciente de lado e mesmo de frente em outras posições, pode-se tentar unir os dados das diversas posições de captura para aprimorar a análise.
- 12) Combinar as imagens em outras posições com as imagens utilizadas.
- 13) Combinação de exames de mamografia, ultrassom, ressonância magnética e imagens térmicas.
- 14) Fazer a modelagem 3D a partir das imagens térmicas (Castro et al., 2009).
- 15) Otimização do algoritmo de Lacunaridade.
- 16) Extração de características utilizando a medida fractal Sucessividade.
- 17) Extração de características utilizando a dimensão fractal pelos seus diferentes métodos de cálculo (“Box- Counting”, “Differential Box-Counting”, “Differential Box-Counting” Modificado e método da contagem de D-Cubos).
- 18) Testar técnicas “não fractais” para a extração e análise das características de texturas das imagens térmicas usadas. Por exemplo: matrizes de co-ocorrência, funções de autocorrelação e histograma.
- 19) Outras combinações dos grupos de características extraídas com o coeficiente de Hurst e com a Lacunaridade.

Pelo comentado pode ser observado que essa dissertação apenas inicia a investigação do uso de imagens térmicas e técnicas de reconhecimento de padrões no auxílio ao diagnóstico de patologias da mama, havendo muito a ser feito em todos os níveis nesta linha de pesquisa até poder chegar-se a uma identificação automática mais precisa de ser ou não saudável a paciente representada pela imagem em análise, ou mesmo de poder-se identificar o tipo de problema presente na imagem adquirida.

Referências bibliográficas

Allain, M. Cloitre, M., Characterizing the lacunarity of random and deterministic fractal sets, *Physical Review A*, 44, 3552-3558, 1991.

Araújo, M.C., Desenvolvimento de um banco de dados (e de imagens) para gerenciamento do uso da termografia no ambulatório de mastologia do Hospital das clínicas da UFPE, (Dissertação de Mestrado – Departamento de Engenharia Mecânica/Universidade Federal de Pernambuco), 2009.

Araújo, MC, Lima, RCF, Magnani, FS, Silva, RTN, Santos, FGS.,The use of a database as an auxiliar tool in thermographic diagnosis for early detection of breast diseases, ENCIT, Belo Horizonte, MG, Brazil, 2008.

Araújo, M. C., Lima, R. C. F., Magnani, F. S., Santos, F.G.C., Nunes, R., Uma investigação preliminar da possibilidade de utilização de câmera por infravermelho para detecção de diferentes patologias em clima tropical, XX Congresso Brasileiro de Engenharia Biomédica, SP,pp. 105-108, 2006.

Barros Filho, M., Sobreira, F., "Accuracy of lacunarity algorithms in texture classification of high spatial resolution images from urban areas". In: XXI Congress of International Society of Photogrammetry and Remote Sensing, 2008, Beijing, China. XXI Congress of Internation Society of Photogrammetry and Remote Sensing. Beijing, China, 2008.

Barros Filho, M., Sobreira, F., Assessing texture pattern in slums across scales: an unsupervised approach. Casa Working Paper, Centre for advanced Spatial Analysis, University College London, 2005.

Barros Filho, M., Sobreira, F., analysing spatial patterns in slums: a multiscale approach. In: Congresso Internacional de Planejamento Urbano Regional Integrado e Sustentável, São Carlos (SP). São Carlos: Pluris, 2005-a.

Batty, M., Longley, P. Fractal Cities: geometry of Form and Function. 1 ed. London: Academic Press, 1994.

Bezerra, L. A., Uso de imagens termográficas em tumores mamários para validação de simulação computacional, (Dissertação de Mestrado – Departamento de Engenharia Mecânica/Universidade Federal de Pernambuco), 2007.

Block, V., Schellnhuber, H., Efficient box-counting determination of generalized fractal dimensions, *Physical Review A*, 42,4 pp. 1869-1874, 1990.

Bronzino, J.D., "Infrared imaging of the breast – an overview", The Biomedical Engineering handbook – Medical Systems and Devices, 3o ed., CRC, 2006.

Bui, A.A.T., Taira, R. K., Medical Imaging Informatics, Ed. Springer, 2010.

Castro, F. J., Vasconcelos, S., **Serrano, R.C.**, Motta, L., Menezes, P.M., Bezerra, L. A., Lima, R.C.F., Conci, A., Um sistema para pré processamento de imagens térmicas e modelagem tridimensional aplicadas à Mastologia. – Encontro Nacional de Engenharia Biomecânica, 2009.

Conci, A., Azevedo, E., Leta, F.R., Computação Gráfica: Teoria e prática, vol. 2 Editora Campus (Elsevier), Rio de Janeiro, Brazil, 2008.

Conci A , Nunes, E. O., Multi-bands image analysis using local fractal dimension, Proceedings of SIBGRAPI-Brazilian Symp. on Comp. Graphics, Image Proc. and Vision, 91-99, 2001.

Conci, A., Proença, C. B., A Fractal Image Analysis System for Fabric Inspection Based on a Box-Counting Method, Computer Networks and ISDN Systems, 30, 1887-1895, 1998.

Dong, P., Test of a new lacunarity estimation method for image texture analysis, International Journal of Remote Sensing, v. 21, Issue 17, pp. 3369-3373, 2000.

Einstein, J. Hai-Shan Wu, Gil, J.: Self-affinity and lacunarity of chromatin textures in benign and malign breast epithelial cell nuclei, Physical Review Letters, 80. 397-400, 1998.

Fernandes, J., Conci, A., Uso da lacunaridade no controle de qualidade visual de padrões industriais. III Congresso Nacional de Engenharia Mecânica, 2004.

Frazer, G.W., Wulder, M.A., Niemann, K.O.: Simulation and quantification of the fine-scale spatial pattern and heterogeneity of forest canopy structure: A lacunarity-based method designed for analysis of continuous canopy heights, Forest Ecology and Management 214, 65-90, 2005.

Freitas, R. M., Técnicas de análises de séries temporais aplicadas à detecção de desflorestamento em tempo real, INPE, 2007.

Gautherie, M. Atlas of breast thermography with specific guidelines for examination and interpretation (Milan, Italy:PAPUSA), 1989.

Guimarães, M. S., Abordagens difuso-neurais para análise de mamogramas, (Dissertação de Mestrado – Instituto de Computação/Universidade Federal Fluminense), 1999.

Hott, M. C., Soares, V.P., Ribeiro, C.A.A.S., Griffith, J.J., Análise fractal de textura usando um operador de Hurst em uma imagem TM/Landsat-5, Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, 2005.

Hu, L., Gupta, A., Gore, J.P., Xu, L.X., Effect of Forced Convection on the Skin Thermal Expression of Breast Cancer, Journal of Biomechanical Engineering, 126 (2), pp. 204-211, 2004.

- Keller, J., R. Chen, S.: Texture description and segmentation through fractal geometry. Computer Vision Graphics and Image Processing 45, 150-160, 1989.
- Keyserlingk, J. R., Ahlgren, P. D., Yu, E., Belliveau, N., Yassa, M., Functional infrared imaging of the breast. IEEE Engineering in Medicine and Biology, pages 30 to 41, May/June 2000.
- Love, T.J., Analysis and application of thermography in medical diagnosis, in Shitzer, A.; Eberhart, R. C.: Heat transfer in medicine and biology, New York: Plenum Publishing Corporation, 1985.
- Mandelbrot, B., The fractal geometry of nature, Freeman, New York, 1983.
- Martins, A.M.S.M., Librantz, A.F.H., A geometria fractal e suas aplicações em arquitetura e urbanismo, Exacta, vol.4, UNINOVE, 2006.
- Menezes, P.M., **Serrano, R.C.**, Zamith, M., Knechtel, M., Montenegro, A., Clua, E.W.G, Conci, A., Bezerra, L.A., Lima, R.C.F., Reconstruindo as imagens termográficas a partir dos arquivos JPEG em false color para auxilio no diagnóstico mastológico - trab 83- II Encontro Nacional de Engenharia Biomecânica - Florianópolis- SC, 2009.
- Melo, R. H. C., Using fractal characteristics such as fractal dimension, lacunarity and succolarity to characterize texture patterns on images, (Dissertação de Mestrado – Instituto de Computação/Universidade Federal Fluminense), 2007.
- Melo, R.H.C., Vieira, E.A., Conci, A., Comparing two approaches to compute lacunarity of mammograms, Proceedings of 13th International Conference on Systems, Signals and Image IWSSIP, Budapest, Hungary, pp. 299-302, 2006.
- Melo, R.H.C., Vieira, E. A., Conci, A., Characterizing the lacunarity of objects and image sets and its use as a technique for the analysis of textural patterns, Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science, Publisher Springer Berlim/Heidelberg, ACIVS - IEEE Benelux Signal Processing Chapter, v. 4179, pp. 208 – 219, Antwerp, Belgium, September 2006-a.
- Melo, R.H.C., Conci, A., Identificação de tumores dos seios pela análise de suas imagens. I Encontro Nacional de Engenharia Biomecânica - ENEBI, 2007.
- Müssigmann, U.: Texture Analysis by Fractal Dimension. In: J.L. Encarnação, H. O. Peitgen. G. Sakas, G. Englert (eds.): Fractal geometry and Computer Graphics. Springer-Verlag, Berlin, 217-230, 1992.
- Ng, E.Y.K., Sudharsan, N.M., Effect of blood flow, tumor and cold stress in a female breast: A novel time-accurate computer simulation, Proceedings of the Institution of Mechanical Engineering in Medicine, 215 (4), pp.392-404, 2001- a.

Ng, E.Y.K., Sudarshan, N. M., Numerical computation as a tool to aid thermo-graphic interpretation. Journal of Medical Engineering and Technology, March/April 2001-b.

Ng, E. Y. K., Ung, L. N., Ng, F. C. and Sim, L. S. J. Statistical analysis of healthy and malignant breast thermography. Journal of Medical Engineering and Technology, November/December 2001.

Nunes, E. O., Emprego da dimensão fractal para análise de texturas em imagens multiespectrais, (Dissertação de Mestrado – Instituto de Computação/Universidade Federal Fluminense), 2001.

Olsen, O. and Gotzsche, P. C, Cochrane review on screening for breast cancer with mammography. The Lancet, 9290:1340 to1342, October 2001.

Papoulis, A.: Probability, Random Variables, and Stochastic Processes. In: McGraw-Hill Series in Systems Science, McGraw-Hill, New York, 1965.

Plotnick, R.E., Gardner, R.H., O'Neill, R.V: Lacunarity indices as measures of landscape texture. Landscape Ecology 8,201-211, 1993.

Rauch, E., Introduction to Lacunarity. Link:
<http://www.swiss.ai.mit.edu/~rauch/lacunarity/lacunarity.html>.

Santos, L.C., Desenvolvimento de ferramenta computacional para análise paramétrica da influência da posição e do tamanho de um tumor de mama em perfis de temperatura, (Dissertação de Mestrado – Departamento de Engenharia Mecânica/Universidade Federal de Pernambuco), 2009.

Santos, S. K. S. L., Bezerra, L. A., Bezerra, K. M., Lyra, P. R. M., Lima, R. de C. F. de, Data Acquisition from Infrared Images for the Bidimensional Numerical Calculations of Temperatures, August, 18-22, Salvador, Bahia, Brazil, 2008.

Serrano, R. C., Conci, A., Zamith, M., Lima, R.C.F., About the feasibility of Hurst coefficient in thermal images for ealy diagnosis of breast diseases, 11th Pan-American Congress of Applied Mechanics, Foz do Iguaçu, Paraná - BRAZIL 6pp, January 04 - 08, 2010.

Serrano, R.C., Motta, L., Batista, M., Conci, A., Using a new method in thermal images to diagnose early breast diseases, XXIInd BrazilianSymposium on Computer Graphics and Image Processing - SIBGRAPI, Rio de Janeiro, Brazil, October 2009.

Silveira, O.T.F., Conci, A., **Serrano, R.C.**, Mello, R., Lima, R.C.F., Paper ID: 549374 On Using Lacunarity for Diagnosis of Breast Diseases Considering Thermal Images, presented in Session: Image processing for medical applications, 4pp. IWSSIP 2009:16th International Workshop on Systems, Signals and Image Processing for the year 2009.

Soares, L. M., Auxílio ao diagnóstico de mamografias por computação visual, (Dissertação de Mestrado – Instituto de Computação/Universidade Federal Fluminense), 1998.

Sonka, M., Hlavac, V., Boyle, R., *Image Processing Analysis and Machine Vision*, 3rd Ed. Thomson, 2008.

Valous, N. A., Mendoza, F., Sun, Da-Wen, Allen, P., "Texture appearance characterization of pre-sliced pork ham images using fractal metrics: Fourier analysis dimension and lacunarity", *Food Research International*, vol. 42, Issue 3, pp. 353-362, ISSN 0963-9969, 2009.

Witten, I.H., Eibe, F., 2005. "Data mining: Practical machine learning tools and techniques", 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

Yahara, T., Koga, T., Yoshida, S. Nakagawa, S., Deguchi, H., Shirouzo,K., Relationship between micro vessel density and thermo graphic hot areas in breast cancer, *Surgery Today*, 33 (4), pp. 243-248, 2003.

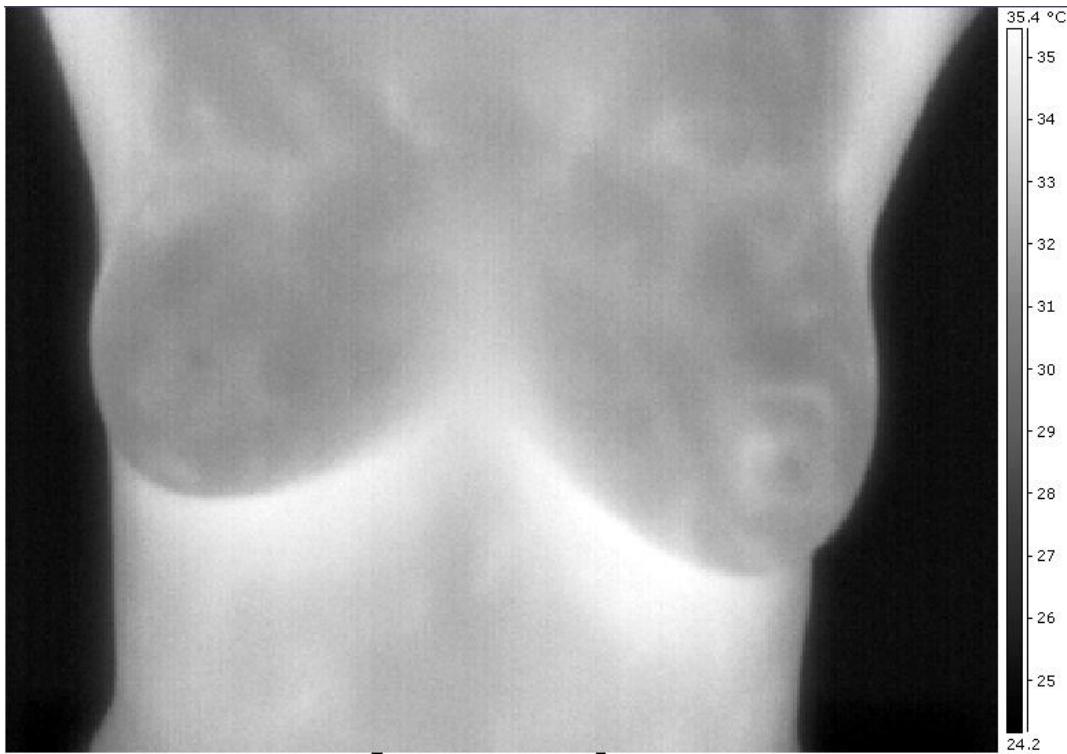
Apêndice

A seguir, apresentam-se todas as vinte e oito imagens térmicas utilizadas nesta dissertação. Essas imagens foram divididas nos seguintes grupos de pacientes: Saudáveis (4), cisto (8), fibroadenoma (11) e carcinoma (5). Cada imagem é seguida do seu número de identificação no site <http://200.20.11.171/proeng/>.

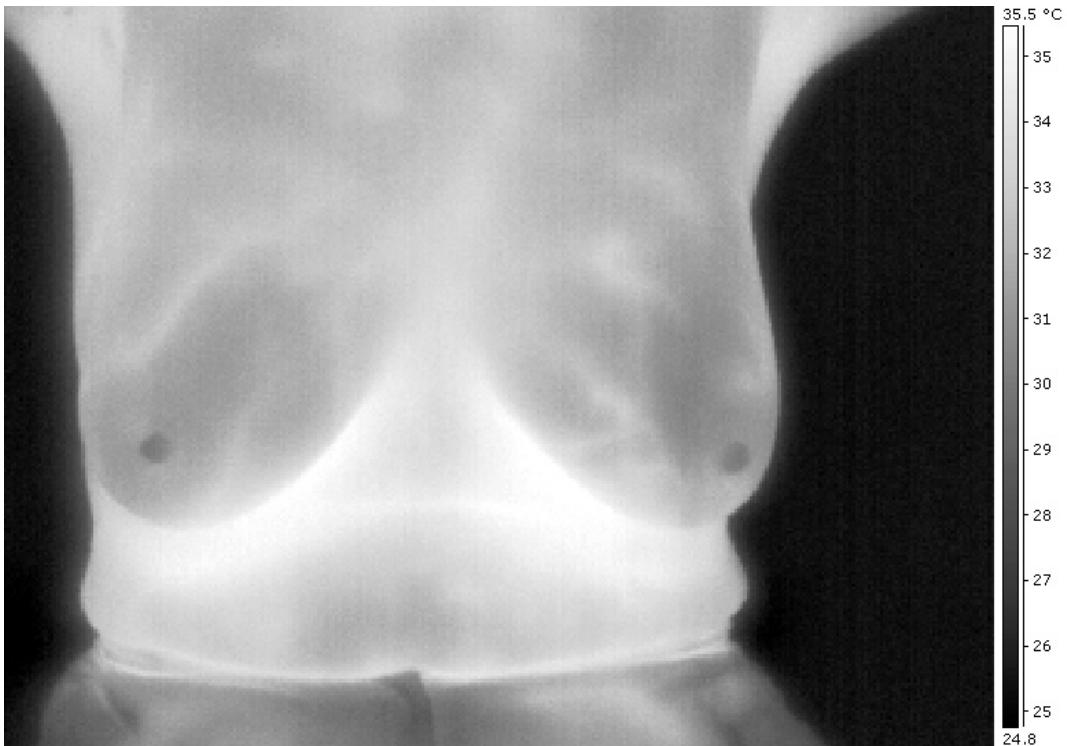
Saudáveis



0149



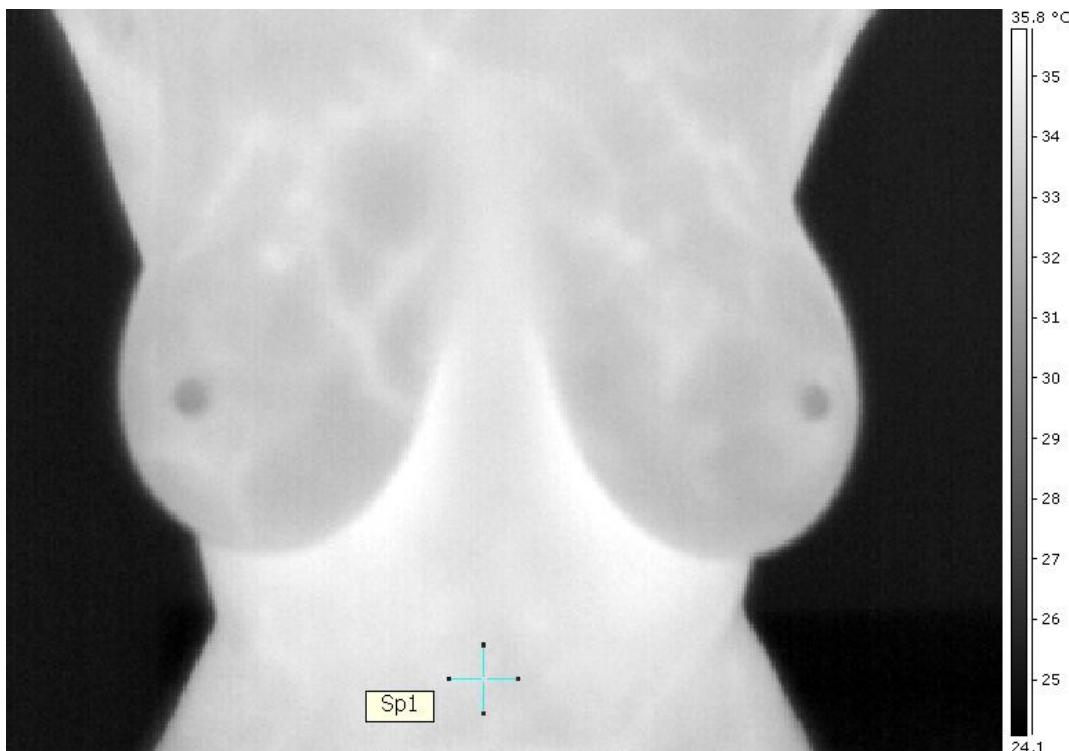
3777



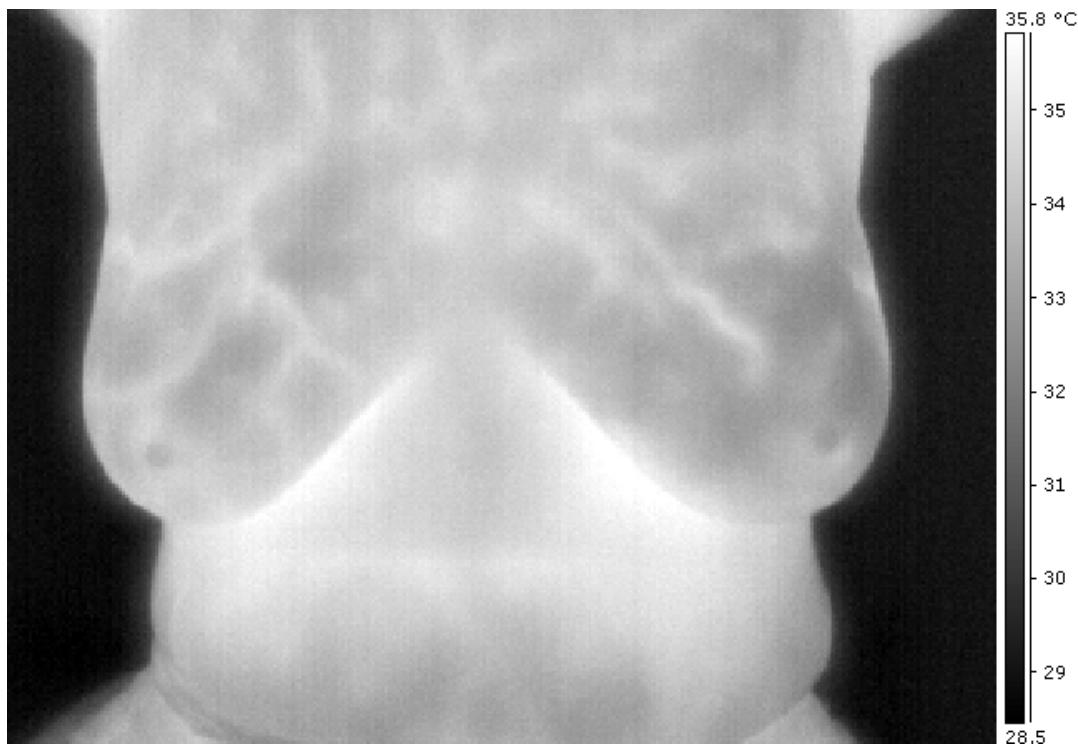
3849



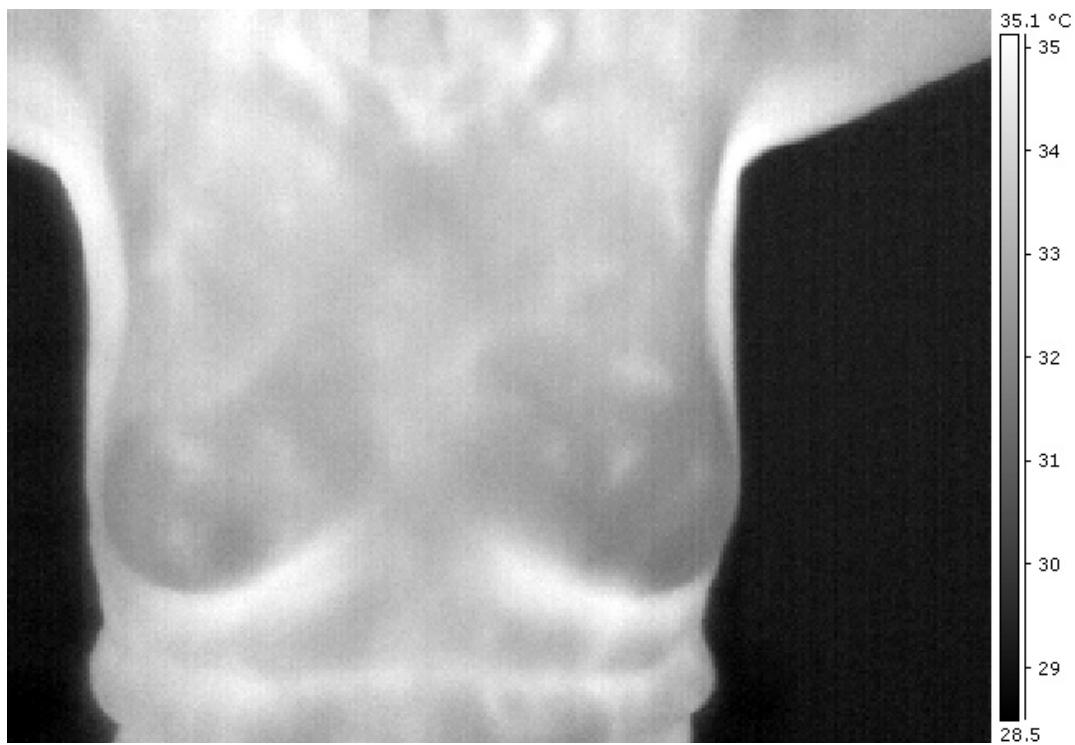
5531

Cisto

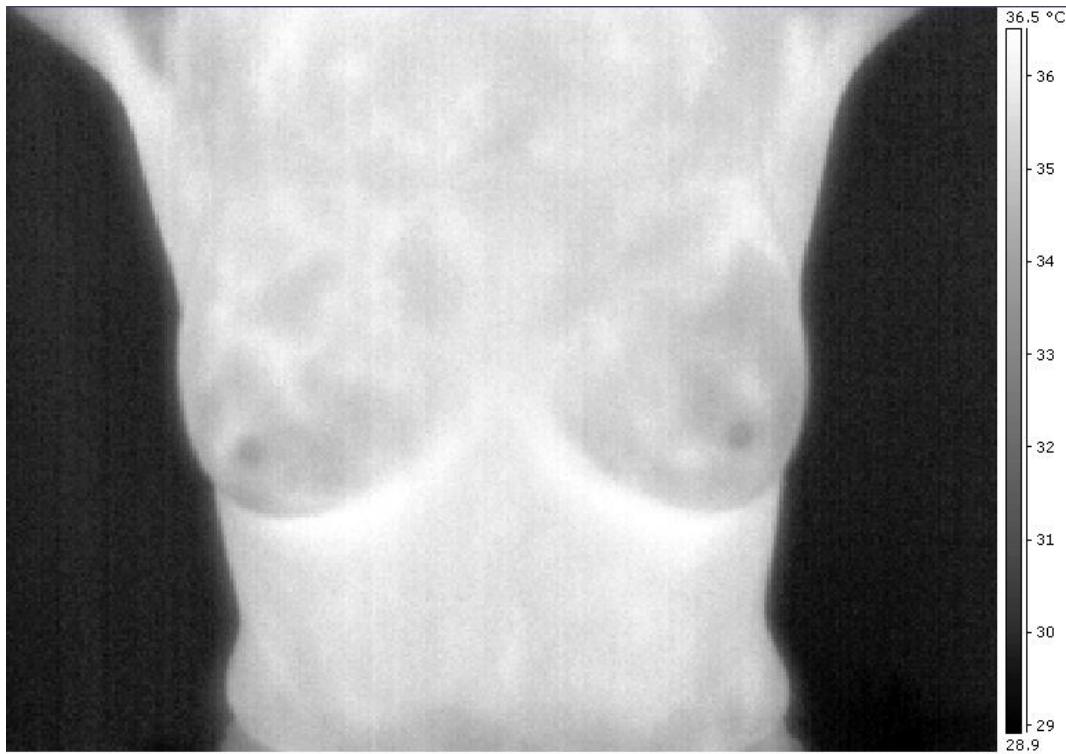
0100



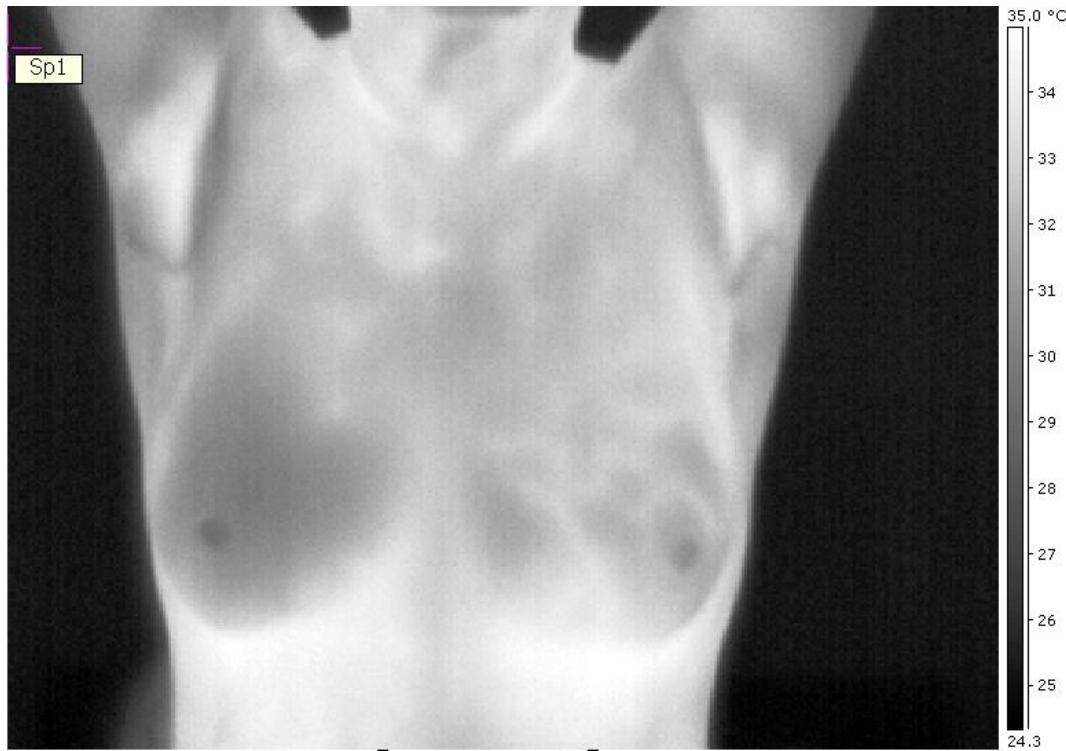
1030



1041



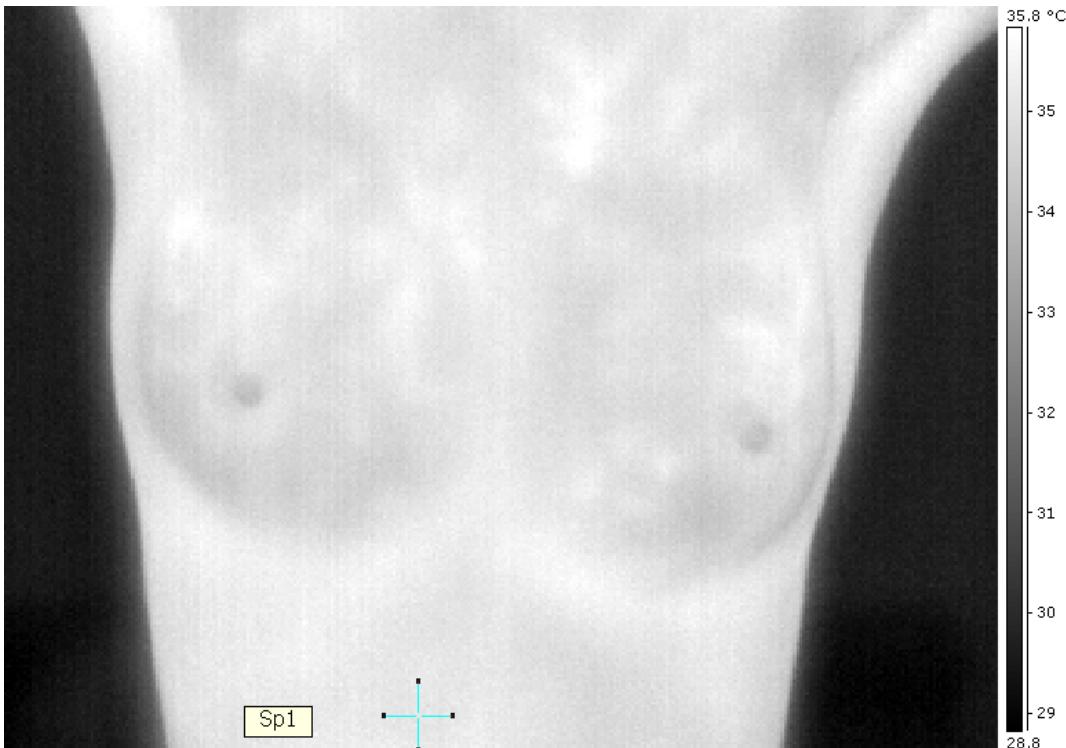
3437



3727



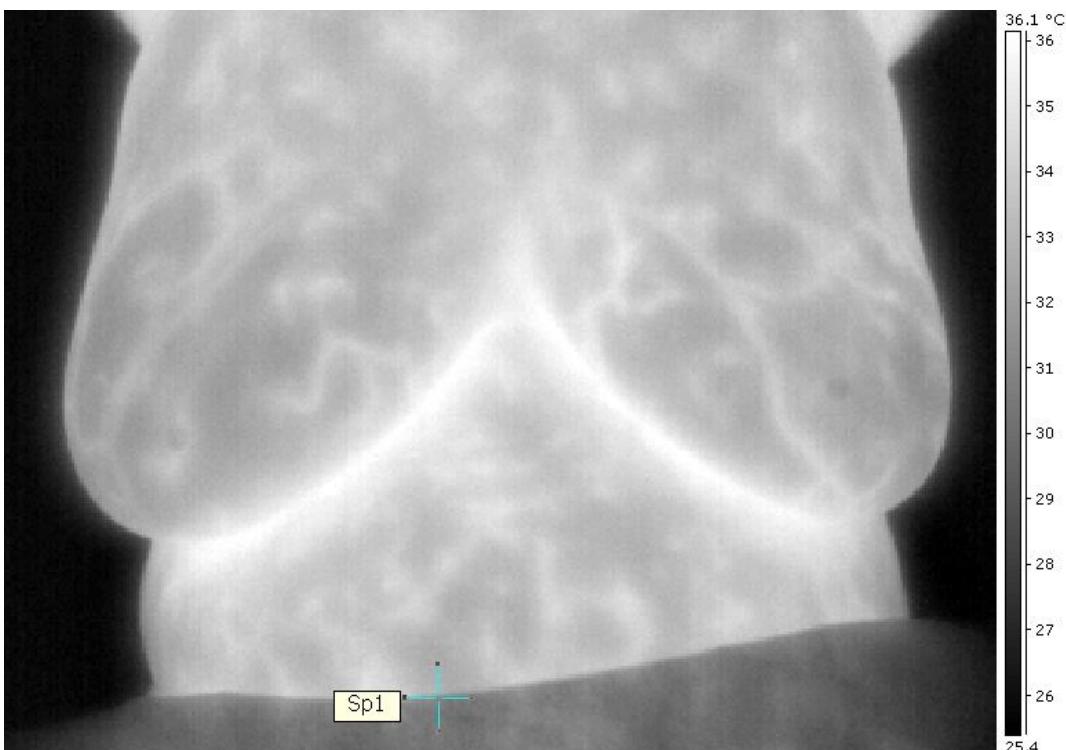
3927



4875



5317

Fibroadenoma

0089



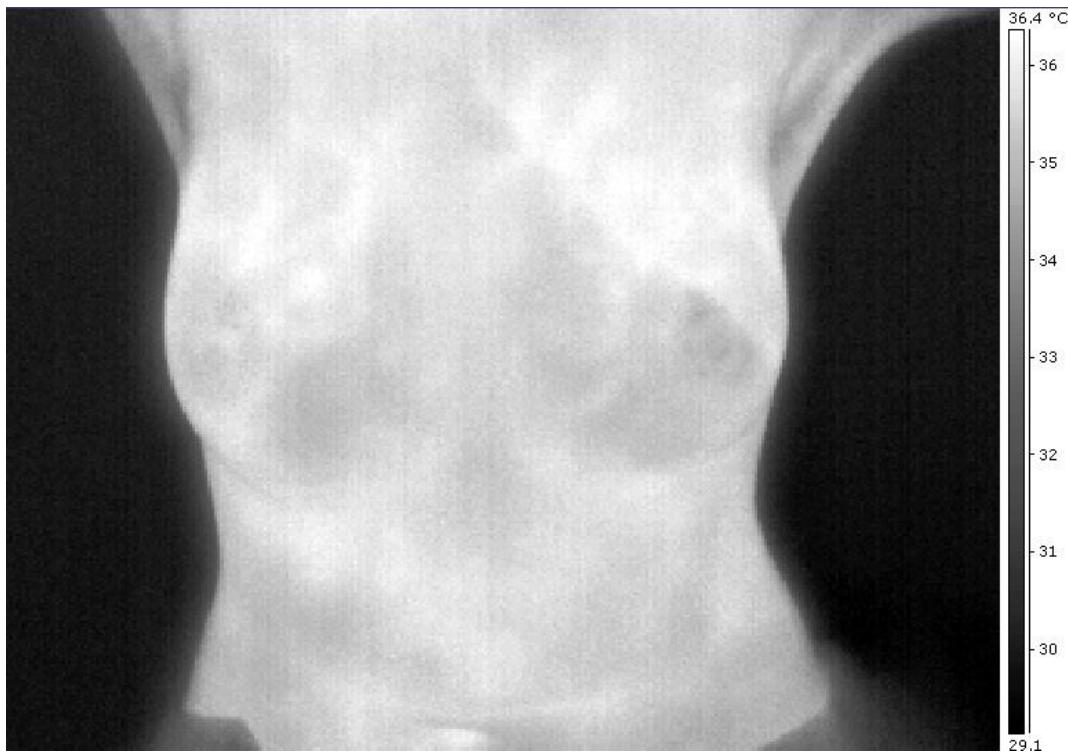
0756



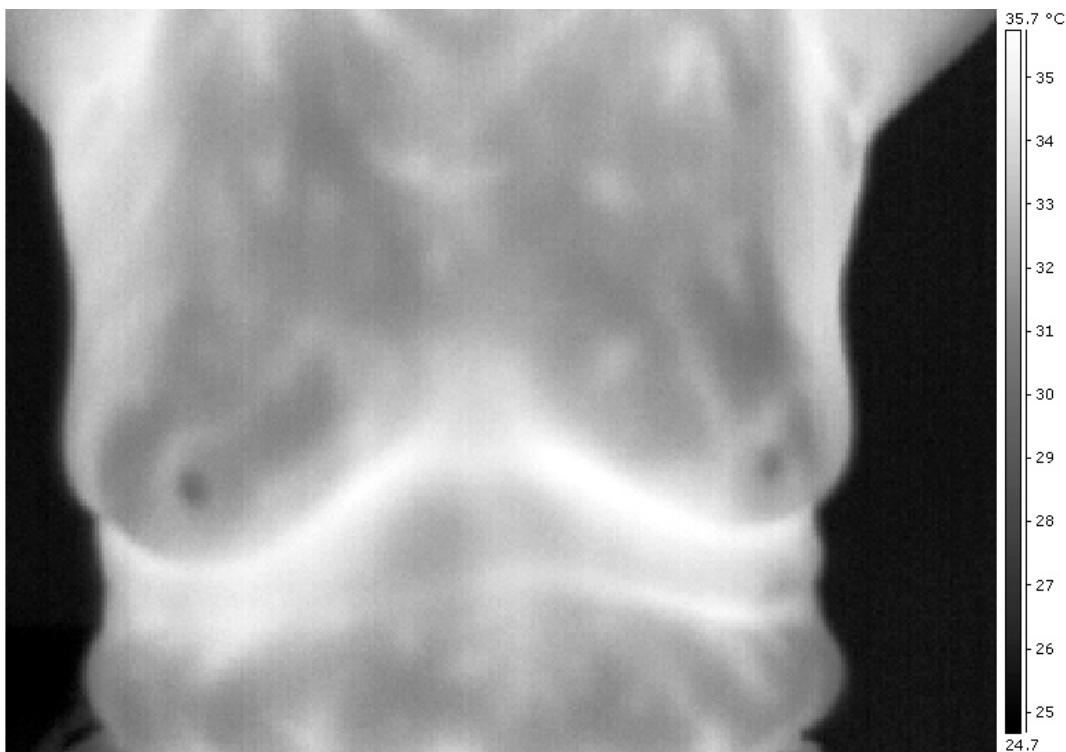
1026



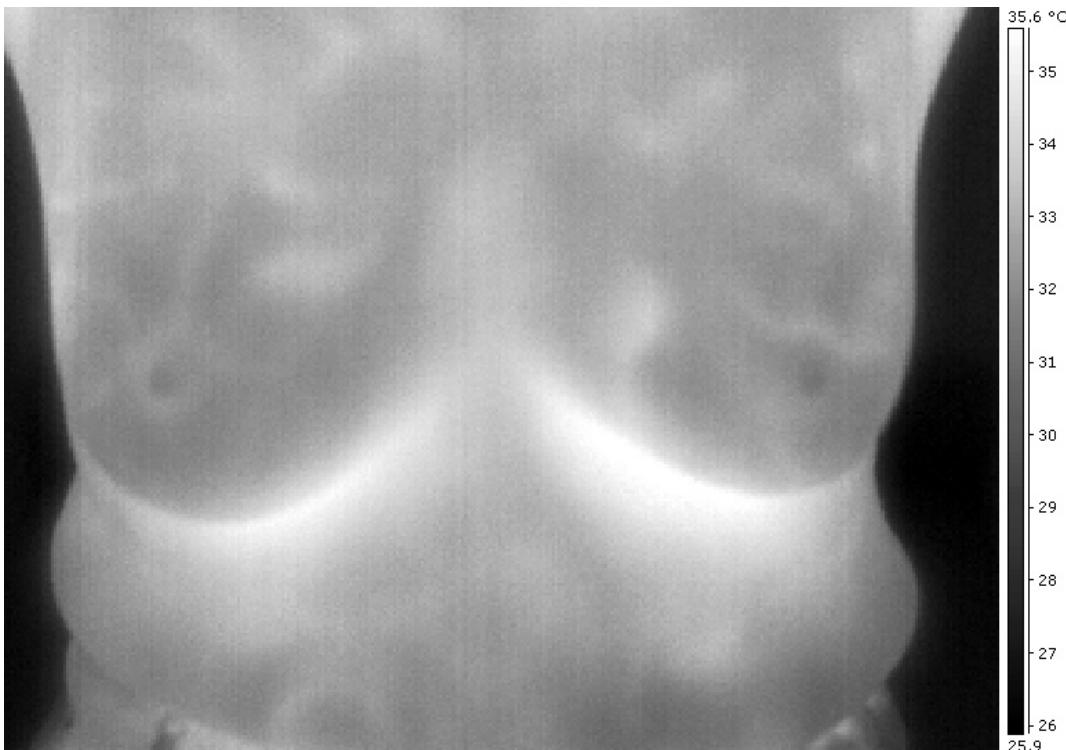
2889



3441



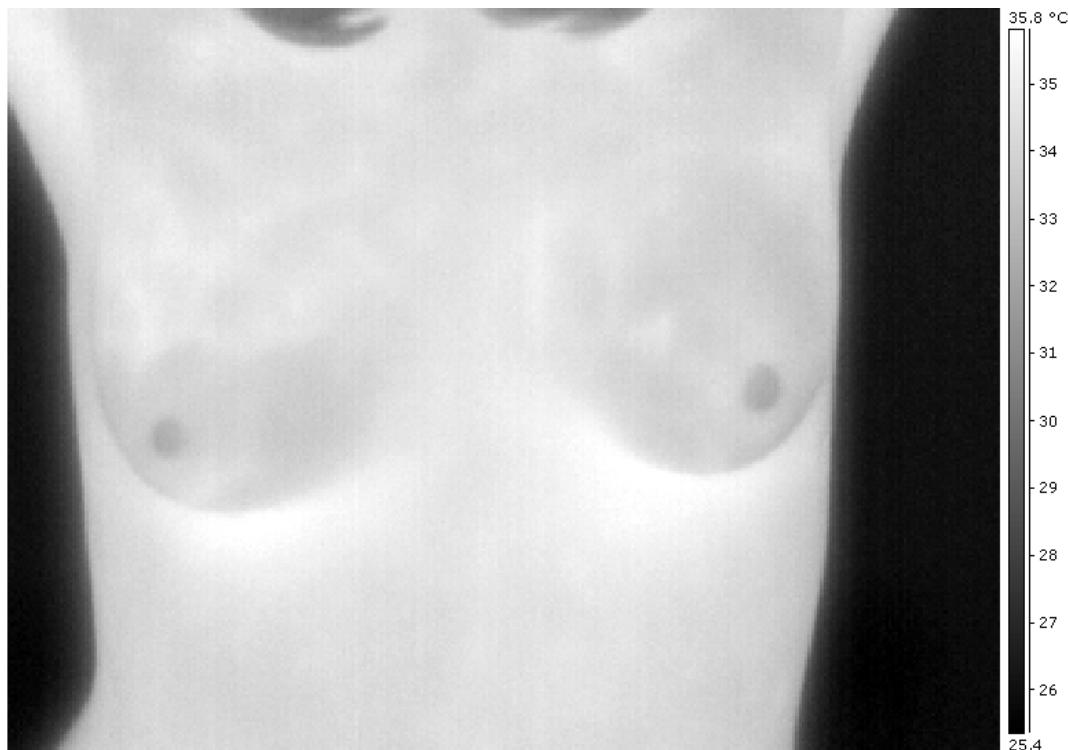
3839



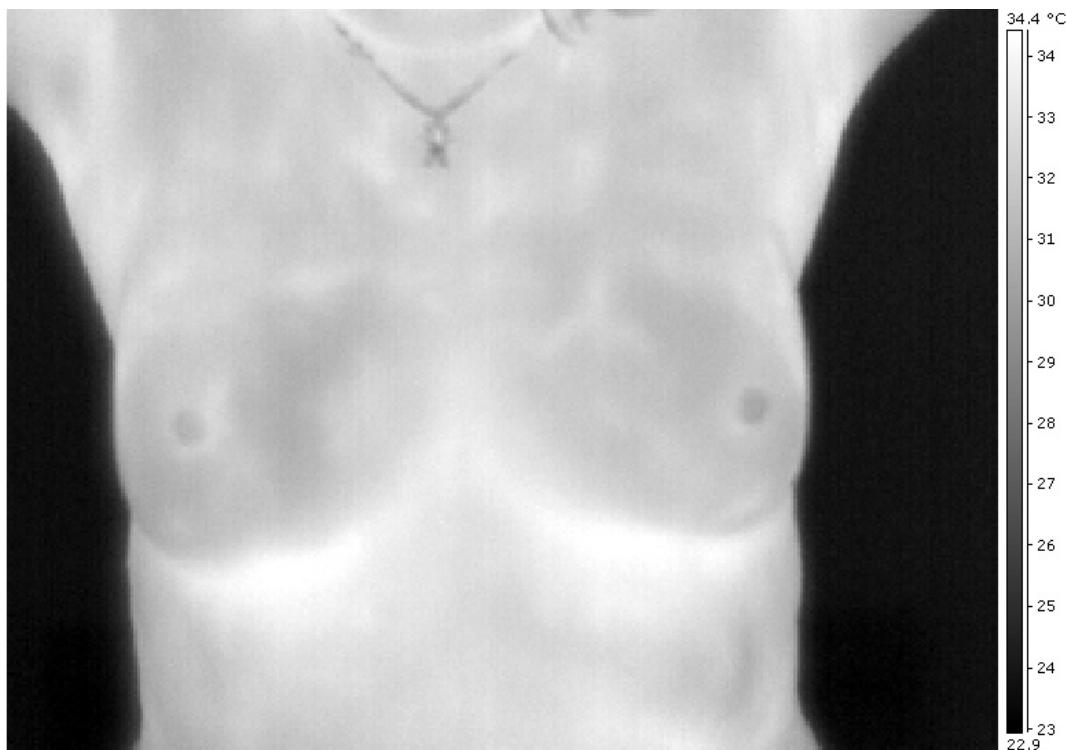
4006



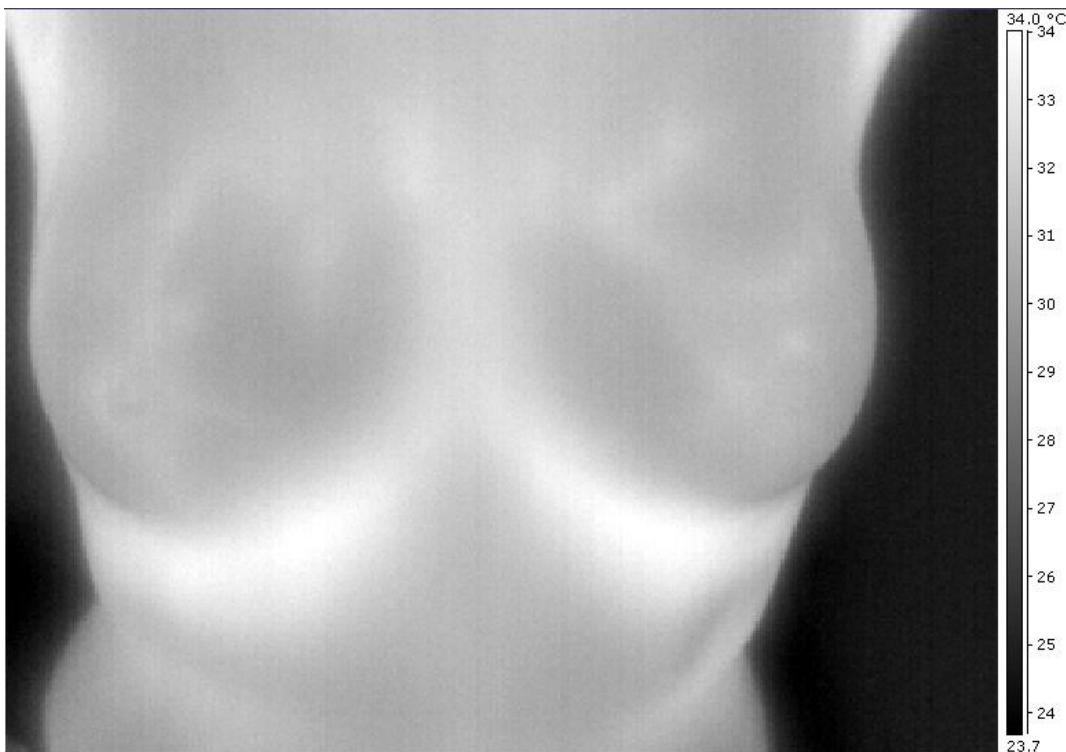
5358



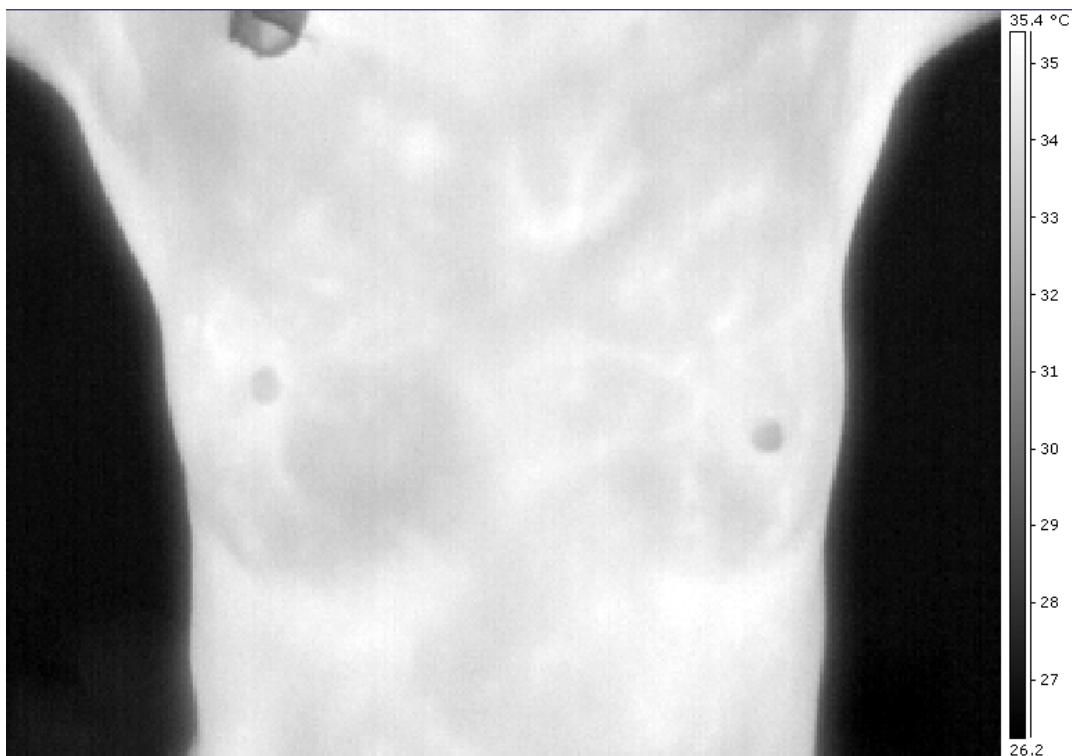
5563



5670



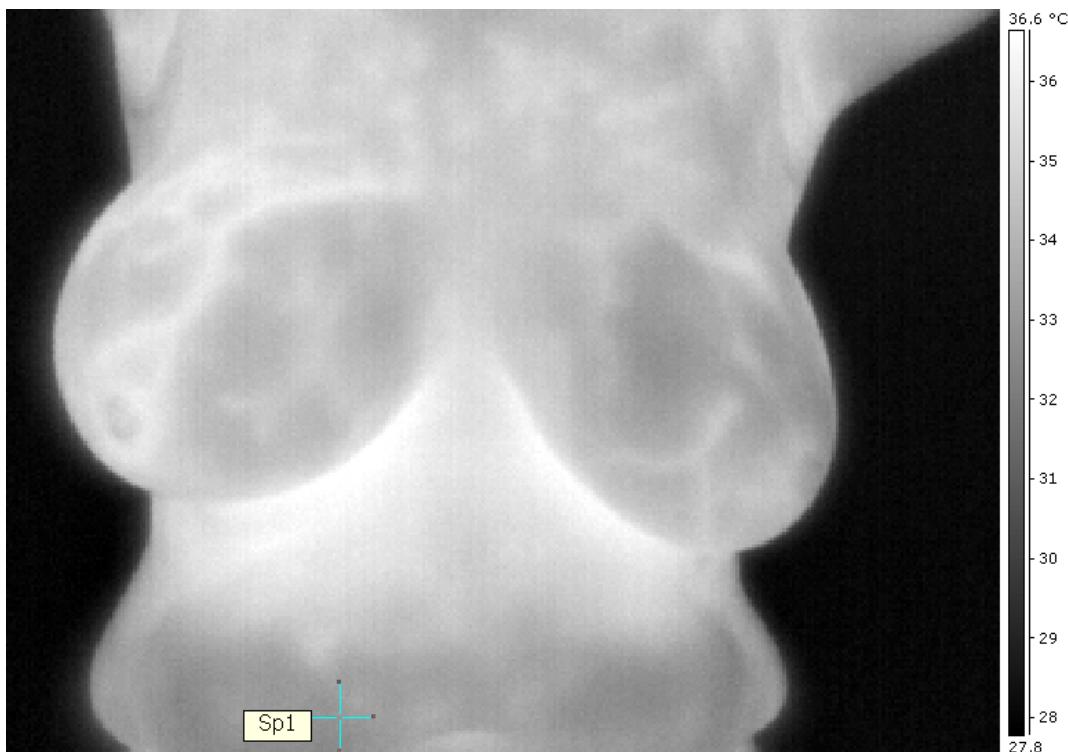
5756

Carcinoma

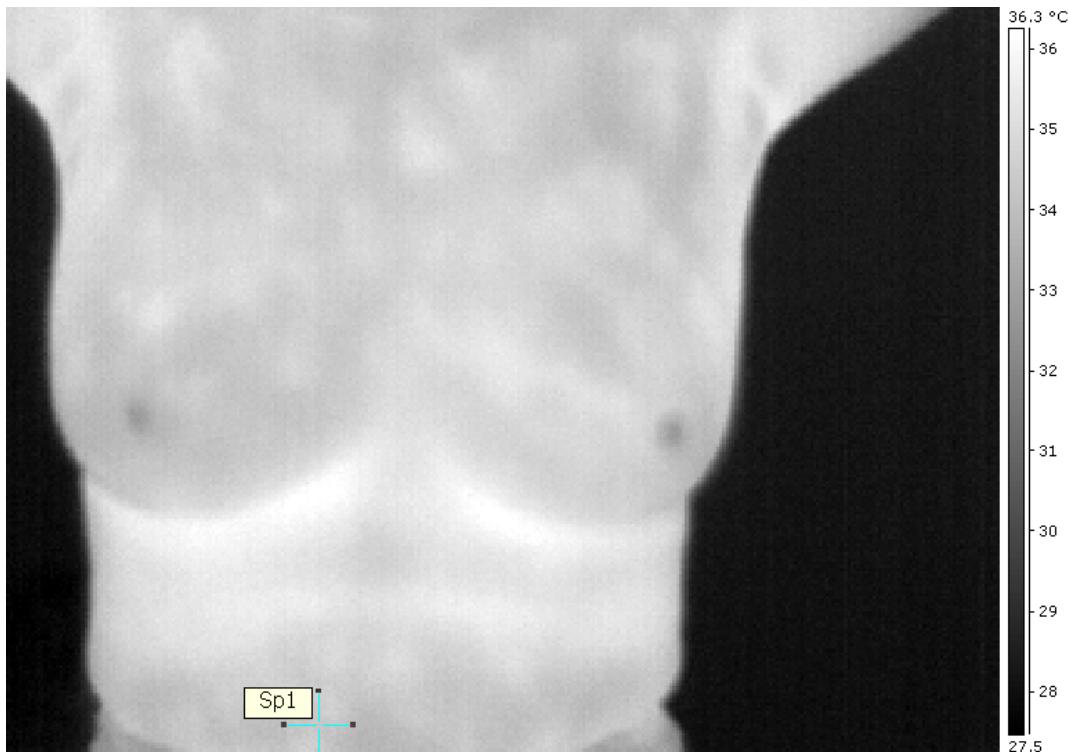
0716



3604



5902



5929



7464