

1 Introdução

Lorem ipsum dolor sit amet, soleat oblique inciderint mei at. Mea harum legere delicata ut. Zril maiorum ea eum, affert adipiscing ad per, veniam nostro sit in. Nec ea falli oratio expetendis. Alii impetus prodesset mea ad, eum ei erroribus appellantur.

Eum etiam interpretaris ei, ad probo labores nam. Possit indoctum est no, nulla persius elaboraret sed ne. Vide summo pertinacia ea nec, et ius vocibus feugait. Dicat fuisset ex vix, dicta dolores pri et. Ex tempor petentium consulatu cum, dico eius dicat no est, ea cum liber virtute. Convenire posidonium mea in, pertinax scribentur eloquentiam eam ei, usu et labore dictas complectitur.

Ad sit eruditi dolorum, purto omittantur vix ut. Mel ad fugit adolescens, et nisl oblique molestiae pri, in incorrupte elaboraret vim. Vis erat praesent in, pri ad eros pertinax. Ex adhuc meliore delicatissimi cum, ut ius justo nulla, in ferri ceteros nam.

At mei quas admodum appellantur, cum ex dolores intellegam, sale soluta probatus ea eum. Rebum bonorum in pri, ad nam graeco insolens, ex esse facete has. Qui ei elit referrentur, eum alii scripta ne. Ei ius tantas mucius latine, justo percipitur cu ius, eum dicam mollis iisque eu. Nam dico tota splendide cu, pri an omittam dignissim voluptatum. Aequae tacimates consulatu eu usu, an pro pertinax electram.

Iudicabit corrumpit scribentur ad est, appareat vituperatoribus cu vix, case movet at quo. Dolorum maiestatis disputando no sit. At etiam dolorem usu, lorem summo quo an, pro dolore nusquam quaerendum eu. Eos suas mazim invenire et, accusam fabellas id nam, te eum doctus consequuntur. Conceptam elaboraret ad nec, est in etiam minimum fabella.

2 Objetivos

2.1 Objetivos Gerais

Desenvolver uma metodologia computacional que analise um exame termográfico dinâmico das mamas, a fim de encontrar padrões que diferenciem mamas saudáveis de mamas doentes.

A principal motivação do trabalho é ajudar médicos especialistas no diagnóstico precoce do câncer de mama, portanto qualquer informação que possa discriminar mamas saudáveis de mamas doentes é relevante.

2.2 Objetivos Específicos

Alguns objetivos extras são necessários para completar os objetivos gerais, sendo eles:

- Desenvolver técnicas que permitam a correção de postura em exames termográficos dinâmicos. Tais técnicas são conhecidas como registro de imagens, onde o resultado final pode variar muito de uma técnica utilizada para outra. Portanto, deve-se utilizar a que melhor se encaixe no problema.
- Analisar a variação de temperatura entre os diferentes tempos do exame, gerando características que sejam capazes de diferenciar os exames de pacientes saudáveis de exames de pacientes doentes.
- Utilizar as características extraídas para gerar um classificador automático, que seja capaz de classificar corretamente novos exames.

3 Fundamentação Teórica

Nesta seção, serão abordados os conceitos necessários para o entendimento da metodologia proposta.

3.1 A Termografia

A termografia é uma técnica que permite a visualização dos raios do espectro infravermelho de forma a mapear a temperatura de um objeto. A termografia infravermelha da mama é um tipo de exame que detecta a radiação infravermelha emitida pela superfície da mama produzindo um mapa de temperatura conhecido como termograma. A grande vantagem deste tipo de exame em relação aos exames mais conhecidos, como a mamografia, é que este não utiliza radiação no processo de obtenção das imagens e também não causa incômodo físico ao paciente, pois não é necessário a compressão das mamas, como ocorre na mamografia. Outro fator a se levar em consideração é o custo extremamente baixo e o fácil manuseio do equipamento necessário para obtenção do exame (??). A Figura 1 apresenta uma termografia em pseudocor, onde a temperatura varia de acordo com a paleta de cores à direita da imagem.

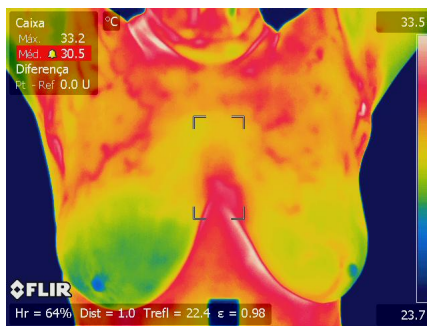


Figura 1: Termografia mamária.

3.1.1 Termografia Estática

A termografia estática (TI), é o tipo de termografia onde a medição de temperatura é feita uma única vez. No caso de TI das mamas são procedimentos (??) de preparação do ambiente e cuidados extras ao paciente para que o exame tenha resultado satisfatório, pois vários fatores podem influenciar o resultado final (??).

3.1.2 Termografia Dinâmica

A termografia dinâmica (TID), é um tipo de termografia onde a medição de temperatura é realizada através de várias etapas de tempo, ou seja, TID mede as mudanças de temperatura sobre um determinado período de tempo. Originalmente conceituada por (??), onde notou-se que mudanças bruscas na temperatura da pele produziam informações valiosas, que não podem ser obtidos por meio de uma TI.

A TID apresenta uma característica importante na detecção de lesões mamárias, pois áreas saudáveis apresentam um comportamento distinto de áreas com algum tipo de patologia. Neoplásias são associadas com a angiogênese que causa um aumento de vascularização na região, sendo que, os vasos recém formados

apresentam a características de possuir poucas terminações nervosas, o que causa um comportamento irregular a estímulos externos, que é detectável através de um exame de termografia dinâmico (??).

Na prática a TID monitora as mudanças na temperatura da pele, onde o estímulo térmico através de uma corrente de ar produz um contraste entre tecidos saudáveis e doentes. Regiões saudáveis da mama tendem a apresentar uma diminuição de temperatura com o estímulo de ar, enquanto as regiões doentes tendem a permanecer estáticas (??).

3.2 Registro de Imagens

Quando imagens que foram extraídas de diferentes tempos, pontos de vista ou aparelhos precisam ser comparadas, ocorre um problema de alinhamento das coordenadas dessas imagens. É necessário efetuar alguma forma de processamento para que essas imagens possam ser comparadas adequadamente. O *Registro de Imagens* é uma técnica que utiliza uma transformação T para mapear a posição e o valor de intensidade de um pixel p da imagem A para q na imagem B , como pode ser visto pela Equação 3.1. Figura 2 mostra o efeito desse mapeamento.

$$T : p \rightarrow q \Leftrightarrow T(p) = q \quad (3.1)$$

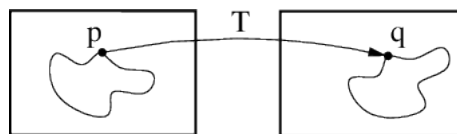


Figura 2: Registro de imagens é a tarefa de achar uma transformação espacial de uma imagem em outra. (??)

Na literatura, é possível encontrar diversas definições para o problema de registro de imagens. De acordo com (??), o processo de registro de imagens é transformação de conjuntos distintos de dados para um mesmo sistema de coordenadas. Em (??) é definido como um processo que determina correspondências entre as características de imagens extraídas em diferentes momentos, pontos de vista ou aparelhos. Tais correspondências podem ser utilizadas para aplicar transformações (rotação, translação, alongamento, etc.) em uma imagem de forma que seja possível a comparação entre as duas imagens. A forma mais intuitiva de utilização do registro é para corrigir diferenças na posição entre scans. Registro de imagens adiciona valor à imagens, permitindo que imagens estruturais (CT, MR, ultrassom) e funcionais (PET, SPECT, funcional MRI) sejam vistas e analisadas no mesmo sistema de coordenadas, e facilita o uso de novas imagens, como para monitorar e quantificar a evolução de uma doença à medida que o tempo passa (??).

A Figura 3 mostra o fluxograma básico de um processo de registro de imagens. O conjunto de entrada básico de um processo de registro de imagens é composto por:

- Imagem fixa: imagem estática cujo espaço de coordenadas é o objetivo.
- Imagem móvel: imagem que será transformada para o espaço de coordenadas da imagem fixa.
- Transformação: função que irá ser responsável por mapear os pixels da imagem móvel na imagem fixa. Geralmente é responsável por dar nome ao registro.
- Métrica: uma medida que indica o quanto duas imagens são equivalentes.

- Interpolador: uma técnica para interpolar os valores da imagem móvel quando são remostrados através da transformação.
- Otimizador: o método utilizado para achar os melhores parâmetros da transformação que otimizam a métrica entre as duas imagens.

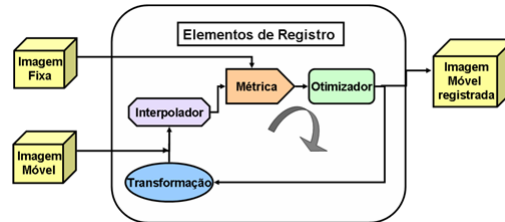


Figura 3: Fluxograma do processo de registro de imagens.

(??)

3.2.1 Métricas de similaridade

As métricas de similaridade são provavelmente o elemento mais crítico no problema do registro de imagens, pois define o objetivo final do processo que é medir o quanto a imagem móvel é equivalente à imagem fixa após a transformação ser aplicada.

3.2.1.1 Correlação Cruzada Normalizada

A Correlação Cruzada Normalizada é uma métrica insensível à fatores multiplicativos entre as imagens. Produz uma função de custo com picos afiados e mínimos bem definidos. Por outro lado tem um raio de captura relativamente pequeno. Sua aplicação é limitada a imagens de mesma modalidade (adquiridas com o mesmo tipo de aparelho) (??). A correlação cruzada normalizada é dada pela Equação 3.2:

$$C(f, m) = -1 \cdot \frac{\sum_{i=1}^N (f_i \cdot m_i)}{\sqrt{\sum_{i=1}^N f_i^2 \cdot \sum_{i=1}^N m_i^2}}, \quad (3.2)$$

onde f e m são os valores de pixels em forma de vetor das imagens fixa e móvel, respectivamente, i indica a posição do pixel em f e m , N é a quantidade pixels a ser considerado. Note que a equação é multiplicada por -1 . Esse fator é responsável por fazer o otimizador procurar os valores que mais se aproximem de 0, que é quando duas imagens são ditas equivalentes.

3.2.2 Transformação

A transformação é responsável por mapear os pixels da imagem móvel para a imagem fixa. Podemos dividir as transformações em dois conjuntos distintos, rígidas e deformáveis.

As transformações rígidas aplicam sobre a imagem móvel apenas transformações simples como operações de translação e rotação, dessa forma a imagem móvel não sofre deformações, sofrendo apenas o alinhamento de suas coordenadas. Registros que utilizam transformações rígidas são chamados de registro rígidos e na prática não são utilizados individualmente, pois não são capazes de corrigir erros locais que necessitam de transformações mais complexas para serem resolvidas. Apesar disso, são bastante utilizados como pré-processamento para os chamados registros deformáveis.

As transformações deformáveis são capazes de efetuar deformações nas imagens, dessa forma é possível corrigir erros que uma transformação rígida não é capaz de corrigir. Registro de imagens que utilizam esse tipo de transformação são geralmente chamados de registros não-rígidos ou deformáveis.

3.2.2.1 Translação

A translação é uma transformação rígida bastante simples. Ela mapeia todos os pontos da imagem móvel adicionando um vetor à eles. Dado um ponto 2D $p = (x, y)$, e o vetor de translação $\vec{t} = (t_x, t_y)$, podemos dizer que o ponto transladado q é dado por

$$q = p + \vec{t} = (x + t_x, y + t_y). \quad (3.3)$$

Utilizando a notação de matrizes em coordenadas homogêneas temos

$$q = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \quad (3.4)$$

A Figura 4 demonstra o resultado da transformação de translação. Vemos que apenas as posições dos pixels foram alteradas.

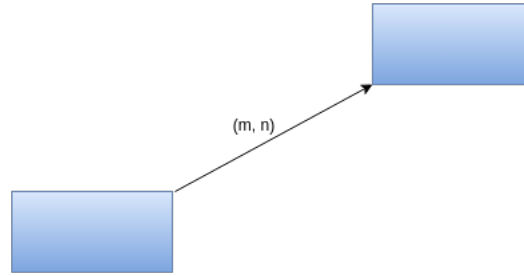


Figura 4: Efeito da translação.

Elaborada pelo autor

3.2.2.2 Rotação

A transformação de rotação é uma transformação rígida capaz de rotacionar os objetos a partir de um centro de rotação. Dado um ponto $p = (x, y)$ e um ângulo de rotação θ em radianos, o ponto q é o resultado da rotação de p em θ radianos dado por

$$q = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3.5)$$

A Figura 5 representa uma transformação de 90° em torno da origem dos eixos.

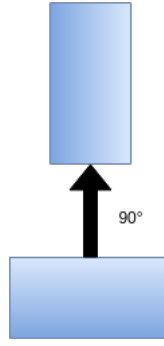


Figura 5: Rotação de $\theta = \frac{\pi}{2} = 90^\circ$.

Geralmente em registro de imagens, a translação e rotação são aplicadas em conjunto, para que a rotação não altere a posição do objeto. Dessa forma, aplicamos uma operação de translação para trazer o objeto para a origem das coordenadas, aplicamos a rotação, então aplicamos uma translação inversa para levar o objeto de volta para sua posição inicial.

3.2.2.3 Escala

A transformação de escala é uma transformação não-rígida capaz alterar a forma de um objeto. Dado um ponto $p = (x, y)$ e o fator de escala $s = (s_x, s_y)$, podemos dizer que o ponto q é dado por

$$q = p \cdot s = (x \cdot s_x, y \cdot s_y). \quad (3.6)$$

Em notação de matrizes homogêneas temos

$$q = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \quad (3.7)$$

Na Figura 6 vemos o efeito da transformação de escala. Caso os parâmetros sejam iguais $s_x = s_y$, não mudamos a forma do objeto. Porém, caso $s_x \neq s_y$, temos uma transformação que altera a forma do objeto transformado.

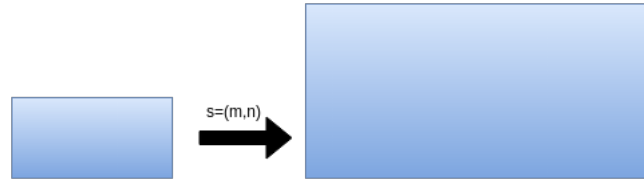


Figura 6: Transformação de escala.

Elaborada pelo autor

3.2.2.4 Transformações FFD baseadas em *B-Splines* cúbicas

Um dos tipos de transformações não-rígidas mais comuns atualmente são as chamadas transformações FFD (deformação de forma livre) baseadas em *b-splines* cúbicas (??). A ideia básica desse tipo de transformação é manipular a grade de pixels a partir dos chamados pontos de controle. Essa manipulação permite criar um campo de deslocamento, que irá mapear os pixels da imagem móvel.

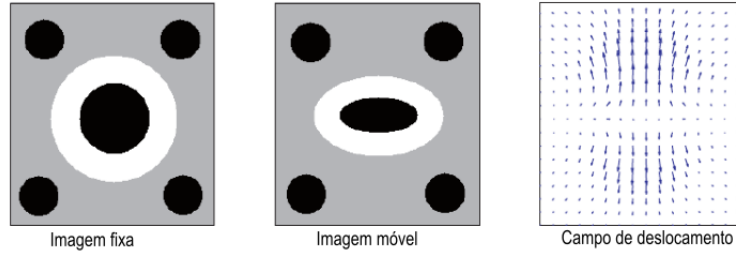


Figura 7: O conceito de campo de deslocamento. Para cada posição de pixel na imagem móvel, existe uma direção e distância que ele precisa se mover para se assemelhar à imagem fixa.

(??)

Seja Φ a grade uniforme de dimensões $n_x \times n_y \times n_z$ com $\Phi_{i,j,k}$ representando o deslocamento do ijk -ésimo ponto de controle. O espaço entre as grandes controle nas direções x , y e z são denotadas por δ_x , δ_y e δ_z , respectivamente. A transformação $T(x : \Phi)$ é definida por

$$T(x : \Phi) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 \beta_l(u) \beta_m(v) \beta_n(w) \Phi_{i+l, j+m, k+n}, \quad (3.8)$$

onde os parâmetros são dados por

$$i = \left\lfloor \frac{x}{\delta_x} \right\rfloor - 1, \quad (3.9)$$

$$j = \left\lfloor \frac{y}{\delta_y} \right\rfloor - 1, \quad (3.10)$$

$$z = \left\lfloor \frac{z}{\delta_z} \right\rfloor - 1, \quad (3.11)$$

$$u = \frac{x}{\delta_x} - (i + 1), \quad (3.12)$$

$$v = \frac{y}{\delta_y} - (j + 1), \quad (3.13)$$

$$w = \frac{z}{\delta_z} - (z + 1). \quad (3.14)$$

As funções β são *B-splines* cúbicas e definidas por

$$\beta_0(t) = (-t^3 + 3t^2 - 3t + 1)/6, \quad (3.15)$$

$$\beta_1(t) = (3t^3 - 6t^2 + 4)/6, \quad (3.16)$$

$$\beta_2(t) = (-3t^3 + 3t^2 + 3t + 1)/6, \quad (3.17)$$

$$\beta_3(t) = t^3/6, \quad (3.18)$$

onde $0 \leq t \leq 1$.

Um parâmetro muito importante desse tipo de transformação é a resolução da grade de controle, pois a partir dela serão gerados os campos de deslocamentos. Uma grade muito espaçosa permite a modelagem de transformações deformáveis globais, enquanto uma grade mais fina modela deformações altamente locais (??).

3.2.3 Interpolador

No processo do registro, a métrica geralmente compara os valores de intensidade dos pixels da imagem fixa com os pixels correspondentes na imagem móvel transformada. Quando transformamos um ponto de um espaço para o outro através de uma transformação, este geralmente irá ser mapeado para uma posição fora da grade de pixels da imagem (Figura 8) . A função do interpolador é calcular o valor de intensidade em uma dada posição de forma correta (??).

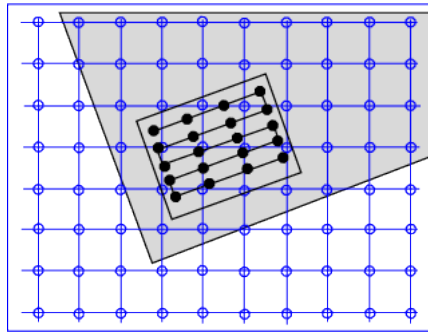


Figura 8: Posições de grade da imagem fixa mapeadas para posições fora da grade na imagem móvel. (??)

3.2.3.1 Interpolador Linear

O interpolador linear assume que os valores de intensidade dos pixels variam de forma linear entre as posições de grade. Dessa forma, os valores interpolados serão contínuos espacialmente, porém o gradiente de intensidade será descontínuo na grade. Se dois pontos conhecidos são dados por (x_0, y_0) e (x_1, y_1) a interpolação linear é dada pela Equação 3.19.

$$y = y_0 + (y_1 - y_0) \frac{x - x_0}{x_1 - x_0} \quad (3.19)$$

3.2.4 Otimizador

Como dito anteriormente, o papel do otimizador é encontrar os melhores parâmetros para a transformação escolhida no processo de registro de imagens. Dessa forma, um amplo conjunto de otimizadores podem ser utilizados, como os *Algoritmos Genéticos* (??) e o *Gradiente Descendente* (??).

3.2.4.1 LBFGS

O *LBFGS* (Limited-memory Broyden-Fletcher-Goldfarb-Shanno) é um método de otimização bastante comum em problemas de registro de imagens. De acordo com (??), é um método quasi-Newton que constrói informações sobre as segundas derivadas durante a otimização e utiliza essa informação para avançar em direção ao mínimo harmônico previsto. Mais especificamente, a inversa da matriz hessiana H^{-1} é construída

iterativamente, começando a partir da matriz diagonal. O método pode ser utilizado de duas formas. Na primeira, uma direção de busca,

$$d_j = F_j H_j^{-1}, \quad (3.20)$$

é identificada em cada iteração, e a minimizador de linha é utilizado para avançar na direção,

$$R_{j+1} = R_j + \lambda d_j. \quad (3.21)$$

A segunda forma é utilizar H^{-1} diretamente para calcular o avanço,

$$R_{j+1} = R_j + F_j H^{-1}. \quad (3.22)$$

O LBFGS utiliza a memória da iteração anterior para construir H^{-1} . O número de iterações é um parâmetro variável e é definido pelo usuário.

3.3 Matrizes de Coocorrência de Níveis de Cinza

A análise de textura em imagens é uma técnica importante na identificação de características em imagens digitais. Uma das primeiras técnicas utilizadas para a extração dessas características foram as matrizes de coocorrências de níveis de cinza também chamadas de *GLCMs* (Gray Level Co-occurrence Matrix) originalmente propostas em (??). Desde então elas vem sendo bastante utilizadas em várias aplicações de análise de textura e permanecem sendo uma ferramenta importante no domínio de análise de texturas (??).

A GLCM é uma técnica que utilizada para extrair características estatísticas de textura de segunda ordem (??), ou seja, a medição considera a relação entre pares de pixels, geralmente vizinhos (??).

3.3.1 Construção de uma GLCM

Seja I uma imagem em níveis de cinza e N a quantidade de níveis de cinza, uma matriz de coocorrência G é uma matriz quadrada de ordem N (??). O elemento da matriz $P(i, j | \Delta x, \Delta y)$ é a frequência relativa em que dois pixels de intensidade (i, j) , separados por uma distância $(\Delta x, \Delta y)$, são vizinhos. Também é possível dizer que o elemento $P(i, j | d, \theta)$ contém os valores da probabilidade de segunda ordem para mudanças entre os níveis de cinza i e j em uma distância d e um ângulo θ (??).

Seja I uma imagem de dimensões $W \times H$, então uma GLCM de ir pode ser definida pela Equação 3.23.

$$G(i, j | \Delta x, \Delta y) = \sum_{p=1}^W \sum_{q=1}^H A \quad (3.23)$$

$$A = \begin{cases} 1, & \text{se } I(p, q) = i \text{ e } I(p + \Delta x, q + \Delta y) = j \\ 0, & \text{caso contrário} \end{cases} \quad (3.24)$$

Como exemplo temos uma imagem I de dimensões 5×5 , com 4 níveis de cinza. Então as matrizes de co-ocorrência $G(i, j | \Delta x = 1, \Delta y = 0)$ e $G(i, j | \Delta x = 1, \Delta y = 1)$ são representadas abaixo.

$$I = \begin{bmatrix} 2 & 0 & 2 & 1 & 1 \\ 2 & 1 & 2 & 3 & 1 \\ 3 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 2 & 2 \\ 3 & 3 & 3 & 1 & 1 \end{bmatrix}$$

$$G(i, j | \Delta x = 1, \Delta y = 0) = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 2 & 2 & 1 \\ 2 & 2 & 1 & 1 \\ 0 & 3 & 1 & 2 \end{bmatrix}$$

$$G(i, j | \Delta x = 1, \Delta y = 1) = \begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 3 & 0 & 1 \\ 0 & 3 & 2 & 1 \\ 2 & 0 & 1 & 0 \end{bmatrix}$$

3.3.2 GLCM Normalizada

Considere $N = \sum_i \sum_j G_d(i, j)$, que é a quantidade coocorrências de pares em G_d . Seja $GN_d(i, j) = \frac{1}{N} G_d(i, j) \cdot GN_d$, chamada de matriz de coocorrência normalizada, onde as entradas (i, j) de $GN_d(i, j)$ são as probabilidades de coocorrência de um determinado pixel com intensidade i ser separado de um pixel de intensidade j por uma distância k em uma determinada direção d (??).

3.3.3 Características de Textura

A partir de uma GLCM, é possível extrair um conjunto de características que descrevem a textura de determinada imagem (??), disponíveis em (??????). De acordo com (??) temos as seguintes características que podem ser extraídas a partir de uma GLCM:

$$Contraste = \sum_{i,j=0}^{N-1} P_{i,j} (i - j)^2 \quad (3.25)$$

$$Dissimilaridade = \sum_{i,j=0}^{N-1} P_{i,j} |i - j| \quad (3.26)$$

$$Homogeneidade = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + (i - j)^2} \quad (3.27)$$

$$ASM = \sum_{i,j=0}^{N-1} P_{i,j}^2 \quad (3.28)$$

$$Energia = \sqrt{ASM} \quad (3.29)$$

$$Correlação = \sum_{i,j=0}^{N-1} P_{i,j} \left[\frac{(i - \mu_i) \cdot (j - \mu_j)}{\sqrt{(\sigma_i^2) \cdot (\sigma_j^2)}} \right] \quad (3.30)$$

$$\mu_i = \sum_{j=0}^{N-1} i \cdot (P_{i,j}) \quad (3.31)$$

$$\mu_j = \sum_{i=0}^{N-1} j \cdot (P_{i,j}) \quad (3.32)$$

$$\sigma_i^2 = \sum_{j=0}^{N-1} P_{i,j} (i - \mu_i)^2 \quad (3.33)$$

$$\sigma_j^2 = \sum_{i=0}^{N-1} P_{i,j} (j - \mu_j)^2 \quad (3.34)$$

onde N é quantidade de níveis de cinza $P_{i,j}$ é a frequência de um par de pixels i e j serem vizinhos.

3.4 Redução de Dimensionalidade

Em *Aprendizado de Máquina*, geralmente possuímos um conjunto de dados previamente disponíveis para utilizar em um classificador qualquer. Nesse conjunto de dados, cada indivíduo é representado por um conjunto de características extraídas do dado original. Por exemplo, em imagens de níveis de cinza é possível extrair as características representadas pelas Equações 3.25, 3.26, 3.27, 3.28, 3.29 e 3.30 a partir de um GLCM qualquer. A quantidade de características que um indivíduo possui é o que chamamos de *dimensão*. Para um conjunto de dados de dimensão $D = 6$, então temos 6 características que descrevem cada indivíduo.

A fácil pensar que quanto maior o número de características, melhor para discriminar os indivíduos. Porém, um conjunto de dados com uma dimensão muito alta implica maior complexidade no problema de classificação. Outro problema é que, em muitos casos, nem todas as características são consideradas importantes para o entendimento de um certo problema(?).

Em termos matemáticos, o problema de *Redução de Dimensionalidade* pode ser descrito da seguinte forma: Dado um conjunto dados com p dimensões $x = (x_1, x_2, \dots, x_p)^T$, encontrar uma representação de menor dimensão $s = (s_1, s_2, \dots, s_k)^T$, onde $k \leq p$, que seja capaz de capturar o conteúdo dos dados originais, de acordo com um critério preestabelecido (?).

Algumas técnicas de redução de dimensionalidade são descritas abaixo.

3.4.1 Seleção de características univariada

A seleção de características univariada seleciona as melhores características baseada em testes estatísticos univariados (?). Por exemplo, podemos selecionar as K melhores características através de um teste estatístico que classifique as características que melhor separam as classes de indivíduos de um conjunto de dados.

3.4.2 Principal Component Analysis

O *Principal Component Analysis* (PCA) é uma técnica de redução linear (?). Por ser baseada na matriz de covariância dos dados, é um método de segunda ordem. O PCA reduz a dimensionalidade encontrando os *principais componentes* (PCs) do conjunto de dados com maior variância, que são vetores formados pelas

combinações das características originais. Por exemplo, o primeiro PC é o que possui maior variância, Temos $s_1 = x^T w_1$, onde o vetor de coeficientes de p dimensões é $w_1 = (w_{1,1}, \dots, w_{1,p})^T$ que resolve:

$$w_1 = \arg \max_{\|w\|=1} \text{Var}\{x^T w\} \quad (3.35)$$

O segundo PC é a combinação linear com a segunda maior variância e ortogonal ao primeiro PC, e assim sucessivamente. Existem tantos PCs quanto o número de características originais (??).

Geralmente, normaliza-se os dados antes da aplicação do PCA, pois o método é dependente da escala. Um dos métodos possíveis para a normalização é escalar cada características entre 0 e 1. Então, assumindo que os dados estão normalizados, podemos calcular a matriz de covariância

$$C = \frac{1}{p-1} \sum_{i=1}^p (x_i - \bar{x}) \cdot (x_i - \bar{x})^T, \quad (3.36)$$

onde \bar{x} é o vetor médio dado por

$$\bar{x} = \frac{1}{p} \sum_{i=1}^p x_i \quad (3.37)$$

Utilizando a relação dos autovetores e autovalores

$$Cv = \lambda v, \quad (3.38)$$

onde v é um autovetor de C e λ o seu respectivo autovalor. Cada autovetor corresponde a um PC, e os respectivos autovalores indicam o grau de variância de cada PC. Assim, escolhendo os k PCs com maior variância absoluta, podemos mapear nossos dados originais em um novo conjunto

$$S = W^T X, \quad (3.39)$$

onde X é o conjunto de dados originais e W uma matriz de transformação

$$W = (v_1, \dots, v_k), \quad (3.40)$$

onde v_1 é o PC de maior variância e v_k o k -ésimo PC de maior variância.

É uma técnica muito útil quando queremos transformar nossas características originais, onde alguma delas não oferecem muita variação nos dados, em um espaço novo com menor dimensão onde as novas características possuem uma melhor dispersão nos dados. A Figura 9 mostra um conjunto de indivíduos gerados por distribuições normais multivariadas com 3 características. Utilizando o PCA para reduzir o espaço dimensional para apenas 2, temos a Figura 10, que exemplifica o efeito da redução.

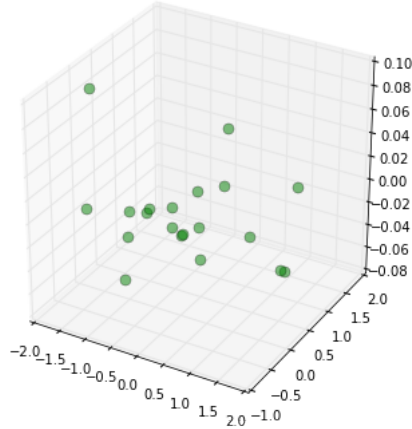


Figura 9: Amostras geradas utilizando um distribuição normal multivariada.
Elaborada pelo autor

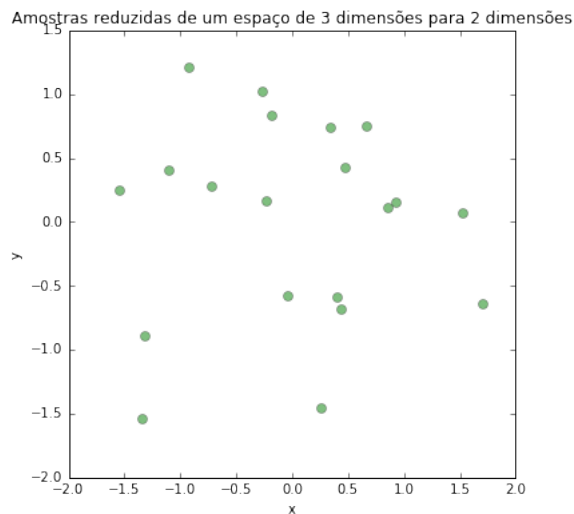


Figura 10: Amostras reduzidas para o espaço bidimensional utilizando os dois principais componentes.
Elaborada pelo autor

3.4.3 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) é uma técnica similar ao PCA, onde buscamos combinações lineares das características para formar um novo espaço de dimensões reduzidas. A diferença é que ao contrário do PCA que busca combinações com maior variância entre os dados, o LDA procura combinações que melhor separam as classes de um determinado problema, portanto podemos dizer que o LDA é uma técnica supervisionada, pois precisamos conhecer as classes que cada indivíduo pertence.

Normalmente é utilizado para redução de dimensionalidade, como um pré-processamento de técnicas de aprendizado de máquina e reconhecimento de padrões(??). Também é possível utilizar o LDA como um classificador. Originalmente proposto em (??) como um problema de *2-class*, que foi generalizado para um problema *multi-class* em (??).

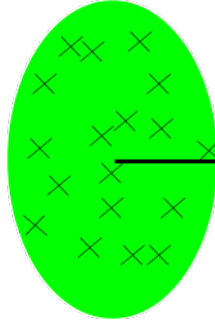


Figura 11: Distância intra-classe.

Elaborada pelo autor

Considerando um conjunto de observações X de dimensões $k \times p$, em que k corresponde à quantidade de indivíduos e p à quantidade de características de cada indivíduo. Para cada indivíduo em X , existe um elemento em y que indica a classe ao qual pertence. O objetivo é encontrar uma transformação W , que maximize a distância entre classes (Figura 12) e minimize a distância intra-classe (Figura 11).

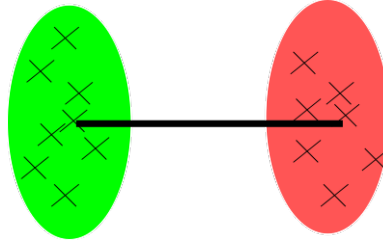


Figura 12: Distância entre-classes.

Elaborada pelo autor

Dessa forma W mapeia X em \bar{X} através da Equação 3.41

$$\bar{X} = W^T \cdot X \quad (3.41)$$

onde W é formado pelos m autovetores com maiores autovalores absolutos da matriz S dada pela Equação 3.42. Uma observação importante é que $m < c - 1$, onde c é a quantidade de classes que existem em y (?).

$$S = S_W^{-1} \cdot S_B \quad (3.42)$$

A matriz S_W é o que chamamos de matriz de dispersão intra-classe e é definida por:

$$S_W = \sum_{i=1}^c S_i \quad (3.43)$$

$$S_i = \sum_{x \in c_i}^n (x - \mu_i) \cdot (x - \mu_i)^T \quad (3.44)$$

$$\mu_i = \frac{1}{n_i} \sum_{x \in c_i}^n x_k \quad (3.45)$$

onde, S_i é matriz de dispersão da classe c_i e μ_i é o vetor que representa os valores médios de cada característica de indivíduos que pertencem a classe c_i .

A matriz S_B é chamada de matriz de dispersão entre-classes e é definida por:

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu) \cdot (\mu_i - \mu)^T \quad (3.46)$$

onde, N_i é a quantidade de indivíduos que pertencem a classe c_i e μ é o vetor que representa a média das características de todos os indivíduos em X .

A vantagem de utilizar o LDA é que simplificamos o problema de classificação criando novas características, a partir das originais, que melhor separam as classes do conjunto de dados. A grande desvantagem é que não é possível saber quais características originais tem maior peso, pois os discriminantes são formados a partir de combinações lineares de todas elas. A Figura 13 apresenta um conjunto de indivíduos que pertencem a três classes distintas. Cada classe foi gerada a partir de uma distribuição normal multivariada, onde os parâmetros diferem entre elas. A Figura 14 apresenta o resultado da aplicação do LDA para efetuar a redução do espaço tridimensional para bidimensional.

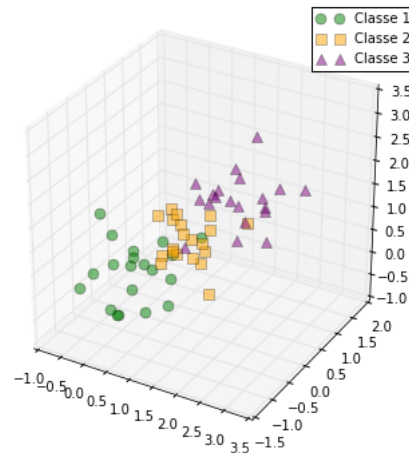


Figura 13: Amostras geradas de distribuições normais multivariadas.

Elaborada pelo autor

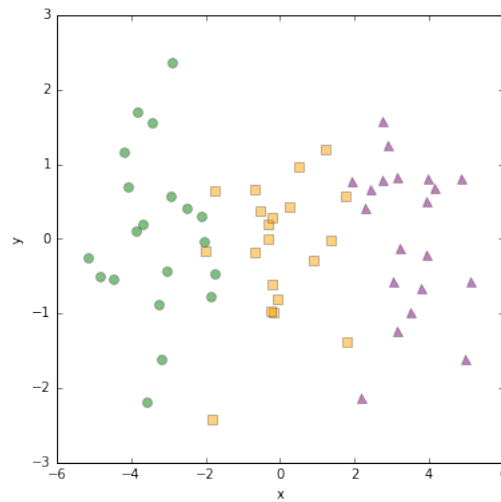


Figura 14: Resultado da aplicação do LDA para reduzir as amostras da Figura 13.

Elaborada pelo autor

3.5 *Support Vector Machines*

Problemas de classificação geralmente envolvem classificar se um certo indivíduo pertence à classe A, B ou C . Atualmente existem diversas técnicas de aprendizado de máquina capazes de realizar essa tarefa com êxito e umas das principais técnicas utilizadas hoje em dia são as *Support Vector Machines* (SVM). Foram originalmente propostas por (??) como um classificador binário, capaz de dizer se um determinado indivíduo pertence à classe A ou B . São parte do grupo de técnicas de aprendizado supervisionado, onde é necessário conhecer previamente um conjunto de indivíduos para que o algoritmo gere um modelo que seja capaz de prever qual a classe uma nova entrada pertence.

Uma SVM constrói hiperplanos em um espaço de alta dimensão, que pode ser utilizado para a tarefa de classificação. Uma boa separação é obtida através do hiperplano que possui maior distância entre os pontos mais próximos de cada classe no conjunto de dados de treino (??).

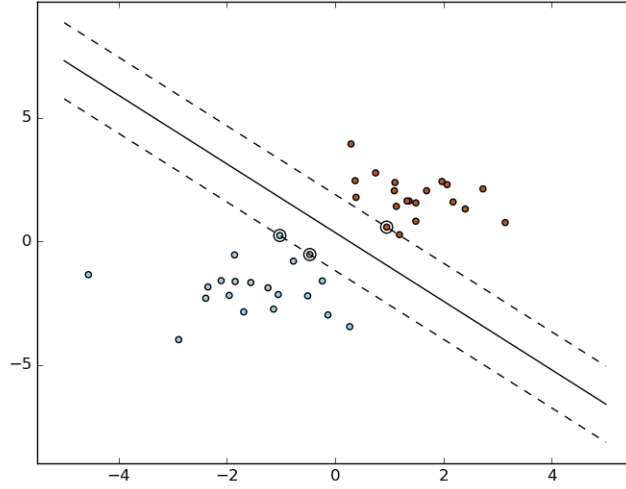


Figura 15: Construção de hiperplanos.

(??)

De acordo com (??), dado (x_k, y_k) o conjunto de amostras para treinamento, sendo que $x_i \in \mathbb{R}^n$ correspondente ao vetor de características do indivíduo i , $y_i \in \{1, -1\}$ corresponde à classe do indivíduo i , sendo $i = 1, 2, \dots, k$. O objetivo do problema de classificação é encontrar uma função $f(x) : \mathbb{R}^n \rightarrow \{1, -1\}$ que seja capaz de estimar corretamente a classe do indivíduo x .

Na etapa de treinamento a função $f(x) = (w \cdot x) + b$ é estimada, de forma que a seguinte relação seja satisfeita:

$$y_i((w \cdot x_i) + b) \geq 1 \quad (3.47)$$

onde w é o vetor normal ao hiperplano e b a distância da função f em relação à origem. Os valores ótimos de w e b são encontrados de acordo com a restrição dada pela Equação 3.47 ao minimizar a equação:

$$\phi(w) = \frac{w^2}{2} \quad (3.48)$$

O SVM possibilita encontrar um hiperplano que minimize a ocorrência de erros nos casos em que a separação ótima entre as classes não seja possível. Com o advento de variáveis de folga, é possível relaxar a restrição da Equação 3.47. Dessa forma o SVM resolve problema de otimização:

$$\min \phi(w, \zeta) = \frac{w^2}{2} C \sum_{i=1}^N \zeta_i \quad (3.49)$$

sujeito à:

$$y_i((w \cdot x_i) + b) + \zeta_i \geq 1 \quad (3.50)$$

onde C é um parâmetro de treinamento que define o equilíbrio entre a complexidade do modelo e o erro de treinamento.

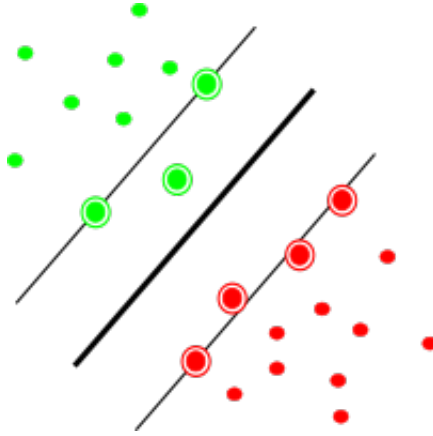


Figura 16: Vetores de suporte.

Elaborada pelo autor

Utilizando a teoria do multiplicadores de Lagrange é possível obter:

$$L(a) = \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i,j=1}^N a_i a_j y_i y_j (x_i, x_j) \quad (3.51)$$

Assim, o objetivo passa a ser encontrar os multiplicadores de Lagrange a_i ótimos que satisfaçam a Equação 3.52 (??).

$$\sum_{i=1}^N a_i y_i = 0, \quad 0 \leq a_i \leq C \quad (3.52)$$

Apenas os pontos onde a restrição imposta pela Equação 3.47 é igual 1, tem correspondentes $a_i \neq 0$. Esses pontos são chamados de vetores de suporte, e estão geometricamente sobre as margens, possuindo grande importância na definição do hiperplano ótimo, pois delimitam as margens do conjunto de treinamento. Na Figura 16, os vetores de suporte são representados por círculos circunscritos.

Pontos além da margem não tem influência suficiente para determinar o hiperplano, porém esses vetores de suporte são essenciais nessa tarefa.

Agora considere o conjunto de amostras da Figura 17. Vemos que não é possível traçar uma reta capaz de separar as classes, tornando o problema um caso não-linear. De acordo com (??), necessitamos de uma transformação não-linear capaz de mapear o conjunto original (espaço dados) para um novo espaço (espaço de características). Esse novo espaço deve apresentar dimensões suficientes para que seja possível realizar a separação linear do conjunto de dados. Dessa forma, o hiperplano de separação é definido como uma função linear de vetores retirados do espaço de características e não do espaço de dados. A construção desse conjunto depende de uma função K , chamada de *kernel* (??). A Equação 3.53 apresenta a forma modificada da 3.52 utilizando uma função K .

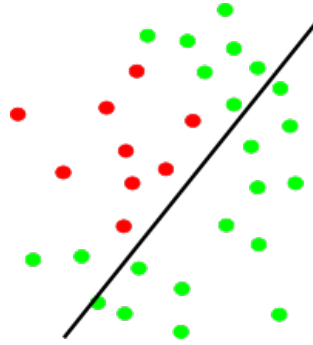


Figura 17: Conjunto de indivíduos não separáveis linearmente.

Elaborada pelo autor

$$L(a) = \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i,j=1}^N a_i a_j y_i y_j K(x_i, x_j) \quad (3.53)$$

Um conjunto de funções amplamente utilizadas em conjunto com o SVM para classificação de dados não linearmente separáveis são as funções de base radial (RBF). Uma função RBF é definida por:

$$K(x_i, y_i) = \exp(-\gamma \|x_i - y_i\|^2) \quad (3.54)$$

onde $\gamma = 1/\sigma^2$, onde σ é a variância.

4 Metodologia

Nesta seção, são apresentados os procedimentos propostos para a realização dos objetivos descritos na Seção 2. A Figura 18 apresenta o fluxo das etapas executadas.

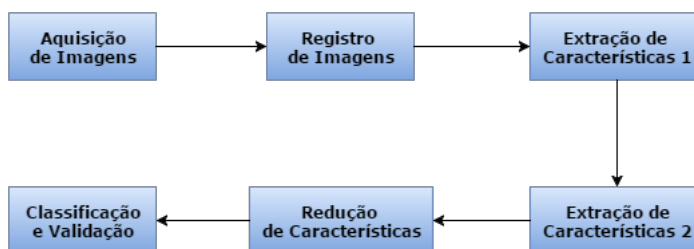


Figura 18: Fluxograma da metodologia.

Elaborada pelo autor

4.1 Aquisição de Imagens

A aquisição de imagens é uma etapa crucial na metodologia proposta, pois sem imagens para casos de teste não é possível validar a metodologia. Atualmente existem poucos bancos de imagem termografias mamárias disponíveis para o público geral, sendo a maioria deles de propósito privado. As imagens utilizadas neste trabalho são provenientes do *Database for Mastology Research with Infrared Image* - DMR-IR, acessível através da interface on-line <http://visual.ic.uff.br/dmi>. Mais informações sobre a base de imagens utilizada podem ser encontradas em (??) e (??).

De acordo com (??), no protocolo de termografias dinâmicas as pacientes são submetidas à um estresse térmico causado pela refrigeração por um ventilador elétrico. Quando a média de temperatura entre as mamas é de 30.5°C (Figura 19), ou 5 minutos de estresse foram aplicados, a refrigeração é interrompida e a aquisição sequencial das imagens é iniciada, extraíndo um total de 20 imagens sequenciais com intervalos de tempo fixos.

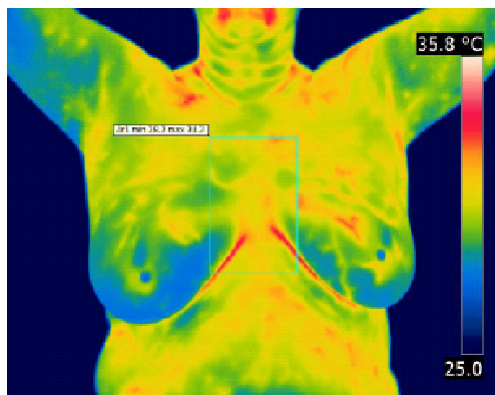
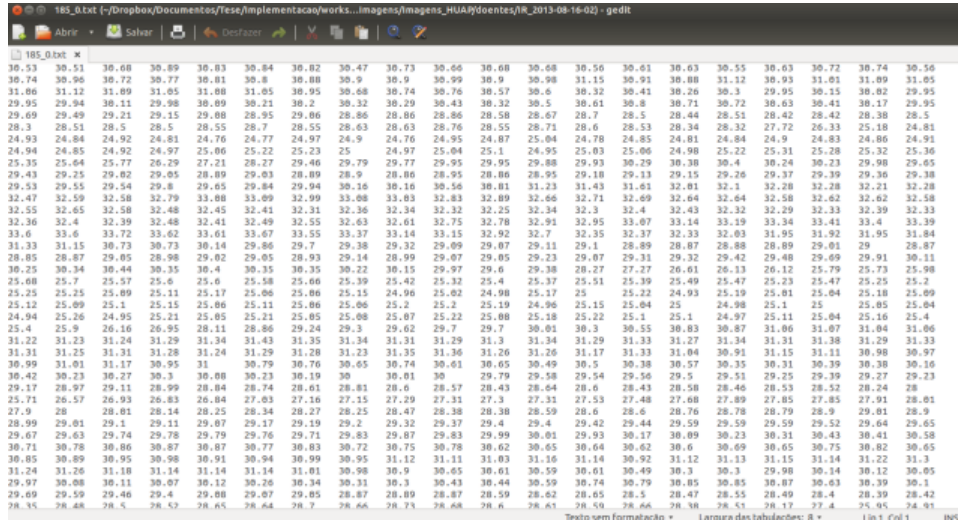


Figura 19: Monitoramento da temperatura média para começar a aquisição sequencial.

(??)

As imagens utilizadas possuem dimensão de 640 pixels de largura e 480 pixels de altura e foram utilizados os valores de temperatura (em graus celsius) obtidos diretamente da câmera termográfica utilizada na aquisição. A Figura 20 apresenta essas matrizes são armazenadas em formato *.txt*, onde cada exame é composto por 20 arquivos separados por tempo de aquisição.



30.53	30.51	30.68	30.89	30.83	30.84	30.82	30.47	30.73	30.66	30.68	30.68	30.56	30.61	30.63	30.55	30.63	30.72	30.74	30.56
30.74	30.90	30.72	30.77	30.81	30.8	30.88	30.9	30.9	30.99	30.9	30.98	31.15	30.91	30.88	31.12	30.93	31.01	31.09	31.05
31.06	31.12	31.09	31.05	31.08	31.05	30.95	30.68	30.74	30.76	30.57	30.6	30.32	30.41	30.26	30.3	29.95	30.15	30.82	29.95
29.95	29.94	30.11	29.98	30.09	30.21	30.2	30.32	30.29	30.43	30.32	30.5	30.61	30.8	30.71	30.72	30.63	30.41	30.17	29.95
29.69	29.49	29.21	29.15	29.08	28.95	29.06	28.86	28.86	28.86	28.58	28.67	28.7	28.5	28.44	28.51	28.42	28.42	28.38	28.5
28.3	28.51	28.5	28.5	28.55	28.7	28.55	28.63	28.63	28.76	28.55	28.71	28.6	28.53	28.34	28.32	27.72	26.33	25.18	24.81
24.93	24.84	24.92	24.81	24.76	24.77	24.97	24.9	24.76	24.95	24.87	25.04	24.78	24.85	24.81	24.84	24.9	24.83	24.86	24.91
24.94	24.85	24.92	24.97	25.06	25.22	25.23	25	24.97	25.04	25.1	24.95	25.03	25.06	24.98	25.22	25.31	25.28	25.32	25.36
25.35	25.64	25.77	26.29	27.21	28.27	29.46	29.79	29.77	29.95	29.95	29.88	29.93	30.29	30.38	30.4	30.24	30.23	29.98	29.65
29.43	29.25	29.02	29.05	28.89	29.03	28.89	28.9	28.86	28.95	29.18	29.13	29.15	29.26	29.37	29.39	29.36	29.36	29.38	
29.53	29.55	29.54	29.8	29.05	29.84	29.94	30.16	30.16	30.56	30.81	31.23	31.43	31.61	32.01	32.1	32.28	32.28	32.21	32.28
32.47	32.59	32.58	32.79	33.08	33.09	32.99	33.08	33.03	32.83	32.89	32.66	32.71	32.69	32.64	32.64	32.58	32.62	32.62	32.58
32.55	32.65	32.58	32.48	32.45	32.41	32.31	32.36	32.34	32.32	32.25	32.34	32.3	32.4	32.43	32.32	32.29	32.33	32.39	32.33
32.36	32.4	32.39	32.48	32.41	32.49	32.55	32.63	32.61	32.75	32.78	32.91	32.95	33.07	33.14	33.19	33.34	33.41	33.4	33.39
33.6	33.6	33.72	33.62	33.61	33.67	33.55	33.37	33.14	33.15	32.92	32.7	32.35	32.37	32.33	32.03	31.95	31.92	31.95	31.84
31.33	31.15	30.73	30.73	30.14	29.66	29.7	29.38	29.32	29.09	29.07	29.11	29.1	28.89	28.87	28.88	28.89	29.01	29	28.87
28.85	28.87	29.05	28.98	29.02	29.05	28.93	29.14	28.99	29.07	29.05	29.23	29.07	29.31	29.32	29.42	29.48	29.69	29.91	30.11
30.25	30.34	30.44	30.35	30.4	30.35	30.22	30.15	29.97	29.6	29.38	28.27	27.27	26.61	26.13	26.12	25.47	25.73	25.73	25.98
25.68	25.7	25.57	25.6	25.6	25.58	25.66	25.39	25.42	25.32	25.4	25.37	25.51	25.39	25.49	25.47	25.23	25.47	25.25	25.2
25.25	25.25	25.09	25.11	25.17	25.06	25.06	25.15	24.96	25.02	24.98	25.17	25	25.22	24.93	25.19	25.01	25.04	25.18	25.09
25.12	25.09	25.1	25.15	25.06	25.11	25.06	25.06	25.2	25.2	25.19	24.96	25.15	25.04	25	24.98	25.1	25	25.05	25.04
24.94	25.26	24.95	25.21	25.05	25.21	25.05	25.08	25.07	25.22	25.08	25.18	25.22	25.1	25.1	24.97	25.11	25.04	25.16	25.4
25.4	25.9	26.16	26.95	28.11	28.86	29.24	29.3	29.62	29.7	29.7	30.01	30.3	30.55	30.83	30.87	31.06	31.07	31.04	31.06
31.22	31.23	31.24	31.29	31.34	31.43	31.35	31.34	31.31	31.29	31.3	31.34	31.29	31.33	31.27	31.34	31.31	31.38	31.29	31.33
31.31	31.25	31.31	31.28	31.24	31.29	31.28	31.23	31.35	31.36	31.26	31.26	31.17	31.33	31.04	30.91	31.15	31.11	30.98	30.97
30.99	31.01	31.17	30.95	31	30.79	30.76	30.65	30.74	30.61	30.65	30.49	30.5	30.38	30.57	30.35	30.31	30.39	30.38	30.16
30.42	30.23	30.27	30.3	30.86	30.23	30.19	30	30.81	30	29.79	29.58	29.54	29.56	29.5	29.51	29.25	29.39	29.27	29.23
29.17	28.97	29.11	28.99	28.84	28.74	28.65	28.83	28.6	28.57	28.43	28.64	28.6	28.43	28.58	28.46	28.53	28.52	28.24	28
25.71	26.57	26.93	26.83	26.84	27.03	27.16	27.15	27.29	27.31	27.3	27.31	27.53	27.48	27.68	27.89	27.85	27.85	27.91	28.01
27.9	28	28.01	28.14	28.25	28.34	28.27	28.25	28.47	28.38	28.38	28.59	28.6	28.6	28.76	28.78	28.79	28.9	29.01	28.9
28.99	29.01	29.1	29.11	29.07	29.17	29.19	29.2	29.32	29.37	29.4	29.4	29.42	29.44	29.59	29.59	29.59	29.52	29.64	29.65
29.67	29.63	29.74	29.78	29.79	29.76	29.71	29.83	29.87	29.83	29.99	30.01	29.93	30.17	30.09	30.23	30.31	30.43	30.41	30.58
30.71	30.78	30.86	30.87	30.87	30.77	30.83	30.72	30.75	30.78	30.62	30.65	30.64	30.62	30.6	30.69	30.65	30.75	30.82	30.65
30.85	30.89	30.95	30.98	30.91	30.94	30.99	30.95	31.12	31.11	31.03	31.16	31.14	30.92	31.12	31.13	31.15	31.14	31.22	31.3
31.24	31.26	31.18	31.14	31.14	31.14	31.01	30.98	30.9	30.65	30.61	30.59	30.61	30.49	30.3	30.3	29.96	30.14	30.12	30.05
29.97	30.08	30.11	30.07	30.12	30.26	30.34	30.31	30.3	30.43	30.44	30.59	30.74	30.79	30.85	30.85	30.87	30.63	30.39	30.1
29.69	29.59	29.46	29.4	29.08	29.07	29.05	28.87	28.89	28.87	28.59	28.62	28.65	28.5	28.47	28.55	28.49	28.4	28.39	28.42
28.35	28.48	28.5	28.52	28.65	28.64	28.7	28.66	28.73	28.68	28.6	28.61	28.59	28.66	28.38	28.51	28.7	27.4	25.95	24.91

Figura 20: Matriz de temperatura.

(??)

No total foram utilizados 70 exames previamente diagnosticados para aplicação desta metodologia, sendo que 35 são de pacientes saudáveis e os 35 restantes de pacientes que apresentam algum tipo de anomalia nas mamas.

4.2 Extração da Região de Interesse

Essa etapa consiste em extrair a região de interesse (ROI) da imagem. Como a metodologia tem o objetivo de detectar anomalias no padrão de vascularização das mamas, é de interesse que as demais regiões de uma TID sejam excluídas do processo. Alguns métodos de segmentação automática são encontrados na literatura (??), porém existem limitações que não tornam possíveis a utilização desse tipo de técnica. Dessa forma, a região de interesse é extraída de forma manual, através da utilização de um programa de visualização e extração de imagens.

Apesar de cada TID possuir um total de 20 sequências de termogramas, utilizamos apenas a sequência no tempo inicial para gerar uma máscara responsável por demarcar a ROI segmentada manualmente através do software ImageJ (????). Utilizando a máscara da sequência inicial é possível extrair as ROIs das demais sequências após a aplicação do registro dos termogramas (Seção 4.3). A Figura 21a demonstra a etapa de segmentação manual utilizando o mouse para desenhar o polígono que definirá a ROI. Como resultado uma máscara binária (Figura 21b) é gerada, para ser utilizada como ROI para todos as 20 sequências de termogramas em um exame.

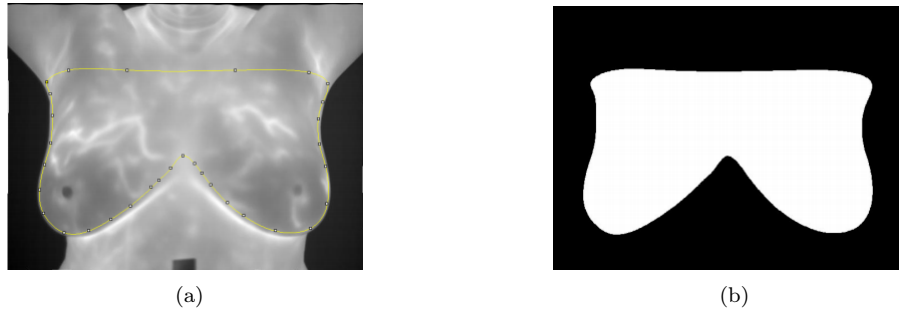


Figura 21: Extração da região das mamas através do software ImageJ.

(??)

4.3 Registro das Imagens

Durante o protocolo de aquisição de imagens, é natural que a paciente execute movimentos involuntários causados pela respiração e ajuste de postura. Esses movimentos causam diferenças espaciais de uma sequência para a outra. As Figuras 22a e 22b são termogramas sequenciais de tempos $t = 1$ e $t = 2$, respectivamente. A Figura 22c apresenta a diferença quadrática $d = (P_t(i, j) - P_t(i, j))$, onde (i, j) é a posição do pixels, sendo $i = 0, \dots, 639$ e $j = 0, \dots, 479$, e $P_t(i, j)$ é o valor de pixel na posição (i, j) no tempo t . É possível notar que existe uma diferença causada pela movimentação involutária da paciente.

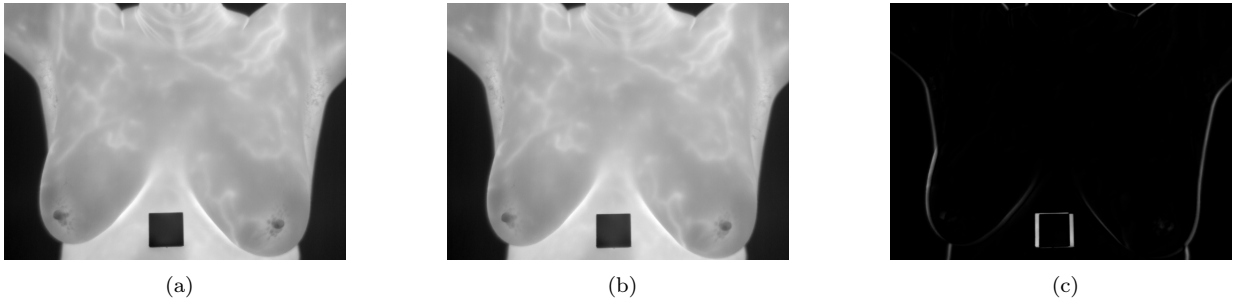


Figura 22: Termogramas dos sequenciais e a diferença de postura entre eles

Elaborada pelo autor.

Para analisar as sequências de termogramas de um determinado exame é necessário corrigir essas diferenças, dessa forma o registro *B-Spline* apresentado na Seção ?? é utilizado. Como dito anteriormente, o registro é uma técnica de projeção de uma imagem no espaço de coordenadas de outra, ou seja, ideal para o problema de correção de postura. Utilizando o primeiro termograma da sequência de um exame como imagem fixa e os termogramas restantes como imagem móvel, executamos o registro para um determinado exame 19 vezes, uma para cada imagem móvel. O registro utiliza a Correlação (Equação X) como métrica de similaridade. O