

# Análise de Imagens Termográficas para a Classificação de Alterações nas Mamas

Francisco Evangelista N. Filho, Brena Kelly S. Lima, Maria Raquel L. de Couto e Iális Cavalcante de P. Júnior

**Resumo**— A Termografia Infravermelha (TI) é uma técnica de diagnóstico antiga, que está conquistando espaço na Medicina moderna pelos seus diversos benefícios. É amplamente utilizada no diagnóstico de alterações nas mamas e tireóide ou até mesmo diabetes. Este artigo apresenta técnicas de processamento de imagens termográficas e uma análise de desempenho dos classificadores LDA (Linear Discriminant Analysis), QDA (Quadratic Discriminant Analysis) e RBF (Radial Basis Function). Para consolidar a análise, foram propostos dois conjuntos de características distintos, para os quais os classificadores se mostraram eficientes e satisfatórios.

**Palavras-Chave**— Termografia Infravermelha, Inteligência Artificial, Alterações nas Mamas, Câncer

**Abstract**— Infrared Thermography (IT) is an ancient diagnostic technique that is gaining space in modern Medicine because it has many benefits. It is widely used in the diagnosis of changes in the breasts and thyroid or even diabetes. This paper presents techniques of thermographic imaging processing and a analysis of performance of Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Radial Basis Function (RBF). To consolidate the analysis, it were proposed two sets of distinct characteristics, for which the classifiers were efficient and satisfactory.

**Keywords**— Infrared Thermography, Artificial Intelligence, Breasts Changes, Cancer

## I. INTRODUÇÃO

Segundo o Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA), câncer é o nome dado ao conjunto de doenças cujo processo de divisão celular ocorre de forma desordenada. Essas células invadem tecidos e órgãos, podendo ainda espalhar-se para outras regiões do corpo (metástase) [4]. É uma patologia com causas variadas – podem ser externas ou internas –, que pode acometer qualquer ser humano.

O GLOBOCAN, projeto criado em 2012 pela *International Agency for Research on Cancer*, uma divisão da Organização Mundial da Saúde (OMS), mostrou que ocorreram 14,1 milhões de novos casos de câncer (exceto câncer de pele não melanoma) e 8,2 milhões de óbitos, que correspondem a 80% do índice mundial de mortalidade [5].

Para o Brasil, são estimados 1,2 milhões de novos casos de câncer para os anos de 2018/2019. Com exceção do câncer de pele não melanoma, são previstos 60 mil casos de câncer de mama (29,5% dos registros das mulheres) [12].

O câncer de mama é o mais comum entre as mulheres. Em 2012, ele foi responsável por 14,7% dos óbitos do mundo todo

[5]. Já a nível nacional, o último dado estatístico fornecido em 2013 pelo Sistema de Informações sobre Mortalidade (SIM) mostra que 14.388 pessoas faleceram por causa do câncer de mama. Dessas, 14.206 eram mulheres. As elevadas taxas de mortalidade evidenciam uma população pouco preocupada em realizar exames de prevenção, uma etapa importante para diagnóstico precoce e rastreamento de indivíduos falsos assintomáticos. Uma pesquisa realizada pelo INCA nos anos de 2008/2009 mostra que o diagnóstico em estágio inicial aumenta em 88,3% a sobrevida de pacientes com câncer de mama.

Por alguns anos, os principais trabalhos da Literatura relacionados à alterações nas mamas utilizavam a Mamografia, pois é o exame de diagnóstico por imagem padrão [17]. No entanto, com a popularização da Termografia Infravermelha (TI), estudiosos dos campos de Medicina e Visão Computacional passaram a explorá-la. O estudo de outros métodos de diagnóstico por imagem permitiu mensurar e comparar a eficiência de cada um deles (assunto que será abordado na Seção II).

Atualmente, pode-se citar como algumas utilizações da TI: diagnóstico do diabetes [9], verificação de presença do vírus *Aphthovirus* (causador da Febre Aftose ou FMDV) [15], reconhecimento de alterações nas mamas [17], [21], na tireóide [10] e no sistema cardio-vascular de humanos e bovinos [2], [19].

Em se tratando dos métodos de classificação utilizados em problemas que envolvem a Termografia Infravermelha, a Literatura é bastante diversificada, pois os autores costumam utilizar, várias técnicas por trabalho, sobretudo supervisionadas [22]. Isso se deve ao fato de que a análise de classificadores é um campo importante da Visão Computacional. São citados como exemplos de classificadores encontradas na Literatura associada à TI: o *Support Vector Machine* (SVM) utilizado por (GOGOI, U. R., 2017) e (RESMINI, R., 2011) e *Fuzzy* empregado por (G., Schaefer, 2009).

O artigo será organizado da seguinte forma: dará-se-á uma fundamentação teórica sobre Termografia Infravermelha na Seção II. As técnicas utilizadas serão exploradas na Seção III. Já a Seção IV exhibe uma análise de desempenho das simulações computacionais realizadas no *software* MATLAB. Finalmente, as conclusões sobre este trabalho serão apresentadas na Seção V.

## II. TERMOGRAFIA INFRAVERMELHA

A temperatura do corpo é uma medida de integridade física e patológica, além de ser um dos cinco sinais vitais, que

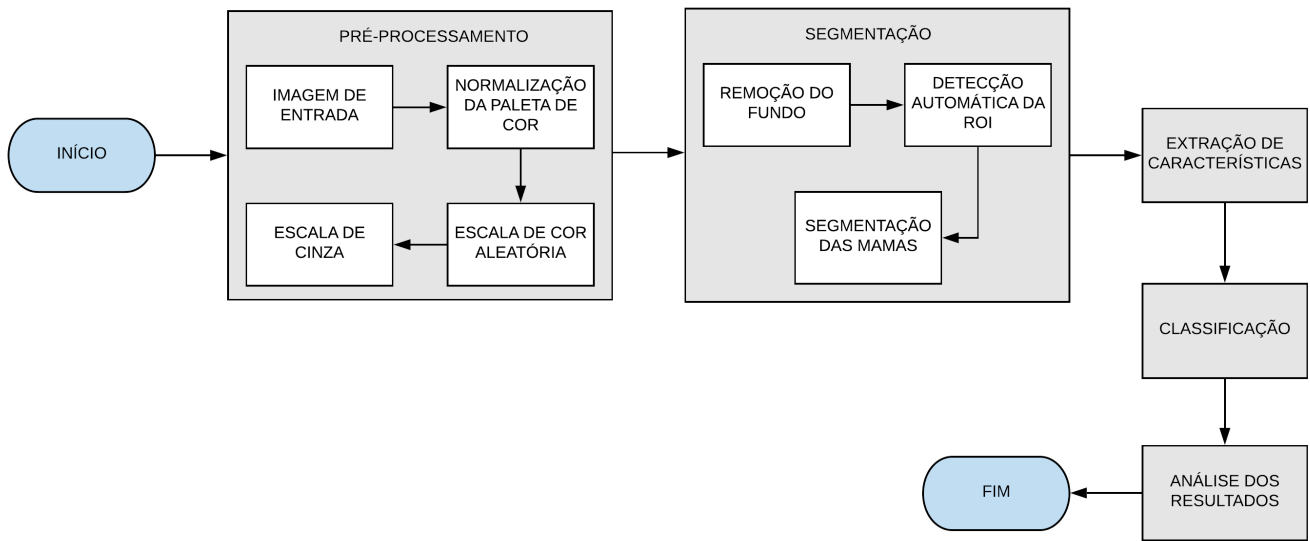


Fig. 1. Passos do algoritmo para detecção de pacientes com anomalias mamárias.

podem ser definidos como estatísticas fisiológicas utilizadas para avaliar as funções corporais básicas (temperatura, pulso, pressão arterial, frequência respiratória e dor) [3], [17].

A termorregulação é um processo fisiológico, estrutural e comportamental capaz de regular a troca de calor do corpo com o meio externo. Isso mantém a homeostase e, como consequência, faz com que a temperatura dos órgãos varie muito pouco. A pele é a estrutura responsável por realizar a troca do calor interno, que irradia até sua superfície, com o do ambiente externo. Uma vez em sua superfície, a temperatura pode ser medida – o instrumento mais utilizado é o termômetro, que fornece uma estimativa global.

Entretanto, a estimativa global fornecida pelo termômetro pode não ser suficiente para uma análise mais detalhada. Para isso, a Medicina passou a utilizar a Termografia Infravermelha, que fornece a matriz de temperatura de todos os pontos de uma cena. A troca de calor do corpo com o meio externo é radiação eletromagnética que se propaga na faixa do infravermelho e pode ser descrita pela Equação 1, a Lei de Stefan-Boltzmann.

$$W = A \cdot \sigma \cdot \epsilon \cdot T^4 \quad (1)$$

A Lei de Stefan-Boltzmann é válida para objetos cuja temperatura é superior à  $-273K$  (zero absoluto) e afirma que a radiação total emitida por um objeto é diretamente proporcional à sua área, emissividade e quarta potência de sua temperatura absoluta, onde  $A$  é a área em  $m^2$ ,  $\sigma$  a constante de proporcionalidade de Stefan-Boltzmann,  $\epsilon$  a emissividade e  $T$  a temperatura absoluta em *Kelvin*. Como a emissividade da pele humana é extremamente alta ( $\epsilon \approx 0,98$ ), a radiação infravermelha emitida pode ser convertida diretamente para um valor de temperatura [21]. Esse processo é realizado por câmeras que possuem sensores infravermelhos que captam a radiação emitida por todos os pontos, os transforma em sinais elétricos e originam os termogramas ou matriz de temperaturas.

As primeiras utilizações da TI para diagnóstico do câncer de mama foram nos Estados Unidos, entre os anos de 1973 e 1982 [17]. É um método considerado recente, sobretudo a nível de Brasil, porém ele conquista cada vez mais espaço na Medicina pelos seus benefícios, como a não utilização de radiação ionizante e acesso venoso ou outro processo invasivo. Esses benefícios reduzem, consideravelmente, danos e riscos aos pacientes. Uma pequena comparação entre os principais exames de diagnóstico por imagem – Termografia Infravermelha (TI), Mamografia (MMG), Ultrassonografia (USG) e Ressonância Magnética (RM) – utilizados para detectar o câncer de mama é mostrada na Tabela I.

TABELA I  
COMPARAÇÃO ENTRE EXAMES DE DIAGNÓSTICO POR IMAGEM [17],  
ONDE  $y$  INDICA UM CUSTO QUALQUER.

Exame	Invasivo	Contraindicação	Custo
TI	Não	Não	$y$
MMG	Muito	Sim	100 $y$
USG	Muito Pouco	Sim	100 $y$
RM	Pouco	Sim	800 $y$

Segundo (NG, 2009), "além de a termografia ter potencial para detecção do câncer de mama de forma precoce, ela é capaz de identificar o câncer dez anos antes quando comparado a outras formas de diagnóstico, como a mamografia".

Cabe ressaltar que o risco de câncer aumenta 2% a cada exposição à radiação ionizante – utilizada na mamografia. Essa porcentagem se eleva em mulheres que não atingiram a menopausa, pois as mamas são mais sensíveis à radiação [21].

### III. METODOLOGIA

Um fluxograma de todo o processo que será descrito pode ser visto na Figura 1.

### A. Aquisição de Imagens

As imagens foram obtidas e disponibilizadas *on-line* pelo projeto VisualLab - Universidade Federal Fluminense (UFF) -, coordenado pela Professora Dra. Aura Conci e pelo grupo do Hospital das Clínicas - Universidade Federal de Pernambuco (UFPE) -, coordenado pela Professora Dra. Rita de Cássia Fernandes de Lima. A aquisição de imagens foi realizada em ambiente controlado através do protocolo estático [21].

No total, são 233 imagens (13 da UFPE) que se dividem em dois grupos: pacientes com mamas normais e anormais (neoplasias diversas). O primeiro grupo possui 179 imagens e o segundo 54. Um exemplo de Termografia Infravermelha estática pode ser visto na Figura 2.

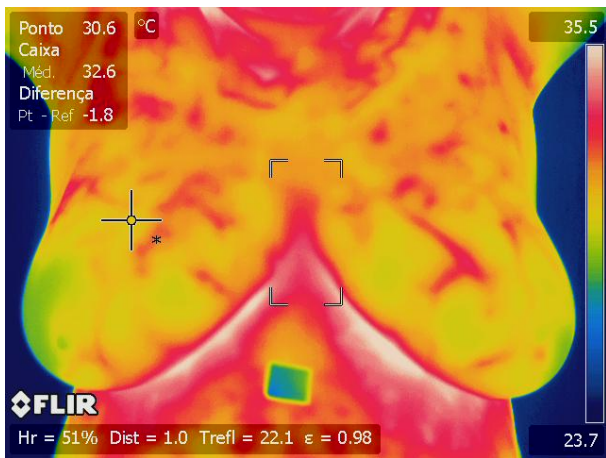


Fig. 2. Termografia Infravermelha das mamas de uma paciente [20].

Na TI os *pixels* que se aproximam do valor 255, equivalente à cor branca, identificam as temperaturas mais altas (vermelho, laranja, amarelo, etc.), enquanto os *pixels* que se aproximam do valor 0, equivalente à cor preta, simbolizam as baixas temperaturas (verde, azul, violeta, etc.).

Algumas imagens do banco não estavam de acordo com a paleta de cores da Termografia Infravermelha. Por esse motivo, foi realizada uma normalização, com auxílio do *software* FLIR Tools<sup>1</sup>, onde todas as 233 imagens foram ajustadas para a escala de cinza, como pode ser visto na Figura 3. Em outras palavras, as imagens foram inicialmente salvas em um dos seguintes padrões: da imagem, *artic*, *iron*, *lava*, *rainbow* ou *rainbow HC* – Figura 4 – e posteriormente foram convertidas para *gray* ou escala de cinza no português. Na Figura 1, essas escalas, exceto a escala de cinza, são exibidas no bloco "pré-processamento" como "escala de cor aleatória". Cabe ressaltar que, na escala de cinza, as temperaturas seguem a mesma lógica da paleta de cor da termografia.

### B. Pré-processamento

A etapa de pré-processamento consiste em aplicar transformações nas imagens de modo a melhorá-las, como por exemplo suavização, rotação, ajuste de escala, correção de brilho, contraste e saturação. É um processo importante, pois aumenta a probabilidade das etapas futuras serem bem

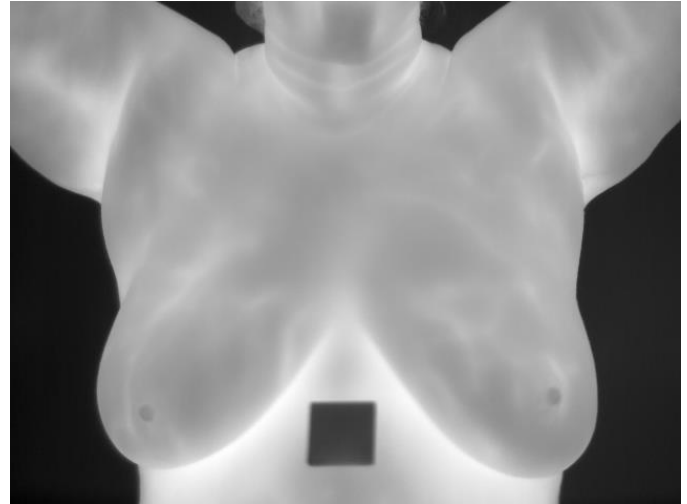


Fig. 3. Termografia Infravermelha das mamas de uma paciente em escala *gray*.

sucedidas. Como falado anteriormente na Seção III-A, o pré-processamento deste trabalho consistiu na conversão dos tons das imagens para a escala de cinza. Um exemplo desse processo pode ser visto na Figura 4, onde na esquerda da imagem se encontram as escalas disponibilizadas pelo FLIR Tools.

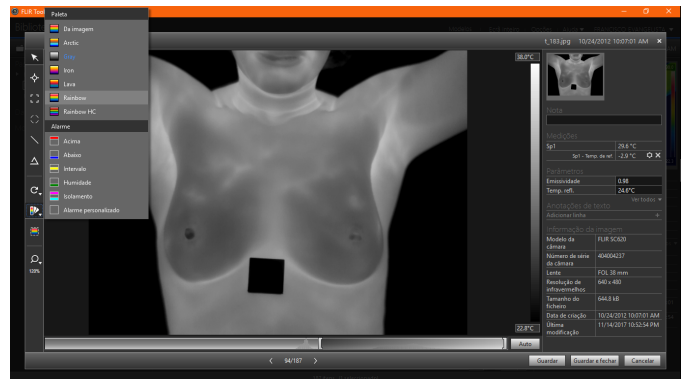


Fig. 4. Exemplo de utilização do *software* FLIR Tools (normalização da escala de cor).

### C. Segmentação

A segmentação subdivide uma imagem em suas regiões constituintes [7]. A *Region of Interest* (ROI) ou Região de Interesse – região que será utilizada, enquanto as demais serão descartadas, como o próprio nome sugere – é particular de cada problema. Cabe ressaltar que a acurácia da classificação está diretamente ligada ao sucesso dessa etapa.

Para a etapa de segmentação foi utilizado do algoritmo automático apresentado em [8]. Inicialmente, o fundo da imagem foi removido através de uma processo *thresholding*; em seguida, foi utilizado o operador de detecção de borda de Canny e um operador de gradiente. Eles são responsáveis por auxiliar na identificação de bordas e das curvas que caracterizam o corpo e os seios, respectivamente. Após a

<sup>1</sup><http://www.flir.com.br/home/>

detecção do corpo e das pregas inframamárias, toda porção não relevante é removida e, diferentemente de [8], as mamas não são separadas. Essa decisão é decorrente da não utilização de muitas características geométricas. O resultado do processo anteriormente descrito é exibido na Figura 5.

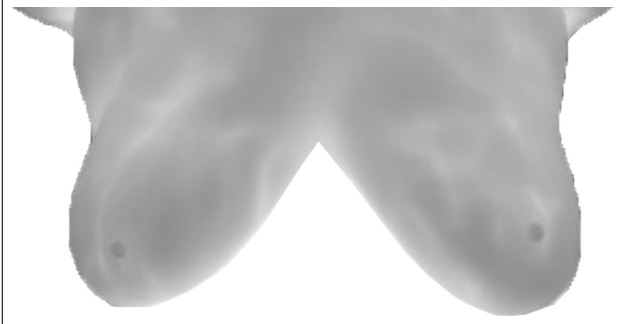


Fig. 5. Exemplo de imagem segmentada. Adaptado de [8].

#### D. Extração de características

Como o próprio nome sugere, essa etapa consiste na extração de informações para propósitos de segmentação ou classificação [1]. De acordo com (GOGOI, U. R. *et al.* 2017.) "um conjunto de 24 características, composto por 7 características estatísticas e 17 características de textura, foi extraído de ambas as mamas de um termograma que quantifica os padrões térmicos simétricos e assimétricos". (GOGOI, U. R. *et al.* 2017.) realiza uma análise das 24 características para dois bancos de imagens distintos (sendo o *DMR Database* utilizado neste trabalho), concluindo quais são pertinentes para cada um dos bancos. Como nem todas mostraram-se relevantes, foram utilizadas, no total, 13 dessas 24 características – as de maior relevância –, sendo 6 estatísticas e 7 de textura e cor. A partir daí, formaram-se dois conjuntos não disjuntos, como segue:

- 1) **Estatísticas:** média, entropia, curtose e variância; **textura e cor:** correlação, variância e soma da variância.
- 2) **Estatísticas:** média e assimetria; **textura e cor:** correlação, dissimilaridade, energia, entropia, soma da entropia e variância.

As definições matemáticas dessas características podem ser encontradas em [1], [7], [22].

#### E. Classificação

Após os vetores de características das imagens serem obtidos, ocorre a classificação. Essa etapa consiste em processar informações (características) para obter um resultado, como por exemplo "normal" e "anormal".

As 233 imagens foram separadas em duas classes: pacientes com e sem alterações na mama e para realizar a classificação foi criada uma matriz de características  $M_{i,j}$  – para cada um dos grupos –, onde cada linha  $i$  indica o vetor de características  $\vec{V}_i$  da amostra ou imagem  $i$ .

Três classificadores foram propostos e utilizados para realização dessa etapa: *Linear Discriminant Analysis* (LDA) [14], [16], *Quadratic Discriminant Analysis* (QDA) [14] e *Radial Basis Function* (RBF) [11].

Para a validação de dados dos três classificadores foi utilizada a estratégia *hold-out*, que consiste em separar o conjunto de dados em dois grupos disjuntos, denominados conjunto de treinamento e conjunto de teste [1]. Foi definido, através de tentativas, a utilização de [75%] do conjunto de características para treinamento e os outros [25%] para teste (tanto do grupo 1 quanto do grupo 2).

Realizaram-se 40 iterações – alguns testes com diferentes números de iterações foram realizados e esse retornou resultados satisfatórios – utilizando o método acima descrito e o resultado de cada uma foi armazenado em um vetor definido na Equação 2, onde  $x_i$  corresponde ao valor da  $i$ -ésima iteração.

$$\vec{\Psi} = [x_1 \quad x_2 \quad x_3 \quad \dots \quad x_k] \quad \text{em que} \quad 1 \leq k \leq 40 \quad (2)$$

Por fim, foi calculada a média do vetor  $\vec{\Psi}$ .

## IV. RESULTADOS

Os resultados obtidos pela metodologia proposta são mostrados na Tabela II, onde  $\mu(\vec{\Psi})$  equivale ao desempenho (em porcentagem) do classificador, ou seja, a quantidade de acertos dadas as classes definidas na Seção III-E.

TABELA II

COMPARAÇÃO DE DESEMPENHO DOS CLASSIFICADORES UTILIZADOS.

Classificador	Características	Desempenho (%)
LDA	1	77
	2	82
QDA	1	88
	2	91
RBF	1	96
	2	100

É possível notar que todas as estratégias de classificação se mostraram eficientes, sendo a rede neural RBF a melhor delas. Comparando os resultados obtidos neste trabalho com [22]: a utilização do classificador RBF mostrou-se superior e o LDA apresentou o pior desempenho; já realizando uma comparação com [21]: o QDA e o LDA mostraram desempenhos não tão bons e, novamente, o RBF se sobressaiu.

## V. CONCLUSÕES

Neste trabalho foi analisado o desempenho de 3 classificadores para realizar o diagnóstico de alterações mamárias. As características de cor e textura sobressairam-se perante as estatísticas, fato que condiz com a descrição do problema. Ademais, futuramente pretendemos estender o banco de imagens e a quantidade de características de maneira que possamos implementar técnicas de classificação mais robustas, além de testar outros *setups*.

## REFERÊNCIAS

- [1] BACKERS, J. J. d. M. S. J. A. R. *Introdução à Visão Computacional usando Matlab*. [S.l.]: Alta Books, 2016. v. 1.
- [2] BAGAVATHIAPPAN, S. *et al.* Infrared thermal imaging for detection of peripheral vascular disorders. *Journal of Medical Physics*. Association of Medical Physicists of India (AMPI), 2009. vol. 34, p. 43-47.
- [3] BRIOSCHI, M. L. *et al.* *Princípios e Indicações da Termografia Médica*. [S.l.]: Andreoli, 2010.

- [4] CÂNCER Instituto Nacional de. *O que é o câncer?* 2018. Disponível em: <[http://www.inca.gov.br/conteudo\\_view.asp?id=322](http://www.inca.gov.br/conteudo_view.asp?id=322)>. Acesso em: 28 mar. 2018.
- [5] FERLAY, J. et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in globocan 2012. *International Journal of Cancer*, v. 136, n. 5, p. 359–386, 2015.
- [6] GOGOI, U. R. et al. Discriminative feature selection for breast abnormality detection and accurate classification of thermograms. *International Conference on Innovations in Electronics, Signal Processing and Communication*, 2017.
- [7] GONZALEZ, R. E. W. R. C. *Digital Image Processing*. [S.l.]: Pearson, 2007. v. 3.
- [8] KAPOOR S. V. A. V. PRASAD, S. P. P. Image segmentation and asymmetry analysis of breast thermograms for tumor detection. *International Journal of Computer Applications*, v. 50, n. 9, 2012.
- [9] LAHIRIS, B. B.; BAGAVATHIAPPAN, S.; RAJ, Baldev.; PHILIP, John. Infrared Thermography for Detection of Diabetic Neuropathy and Vascular Disorder. Em: NG, Eddie YK.; ETEHADTAVAKOL, Mahnaz. *Application of Infrared to Biomedical Sciences*. 1.ed. Singapore: Springer, 2017. cap. 13, p. 217-247.
- [10] MAHAJAN Pallavi; MADHE Swati. Hypo and hyperthyroid disorder detection from thermal images using Bayesian Classifier. *International Conference on Advances in Communication and Computing Technologies (ICACACT)*, 2014.
- [11] MCCORMICK, C. *Radial Basis Function Network*. 2013. Disponível em: <<http://mccormickml.com/2013/08/15/radial-basis-function-network-rbfn-tutorial/>>. Acesso em: 15 abril. 2018.
- [12] MINISTÉRIO DA SAÚDE. Instituto Nacional de Câncer José Alencar Gomes da Silva. *Estimativa 2018: incidência de câncer no brasil*. Rio de Janeiro, 2017. 130 p.
- [13] NG, E. Y. K. A review of thermography as promising non-invasive detection modality for breast tumor. *International Journal of Thermal Sciences*, v. 48, n. 5, p. 849–859, Mai. 2009.
- [14] PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.
- [15] RAINWATER-LOVETTA, Kaitlin; PACHECOA, Juan M.; PACKER, Craig; RODRIGUEZ, Luis L. Detection of foot-and-mouth disease virus infected cattle using infrared thermography. *The Veterinary Journal*. Elsevier, 2009. vol. 180, p. 317-324.
- [16] RASCHKA, S. *Linear Discriminant Analysis*. 2014. Disponível em: <[http://sebastianraschka.com/Articles/2014-python\\_lda.html](http://sebastianraschka.com/Articles/2014-python_lda.html)>. Acesso em: 15 abril. 2018.
- [17] RESMINI, R. *Análise de imagens térmicas da mama usando descritores de textura*. Dissertação (Programa de Pós-Graduação em Computação) — Universidade Federal Fluminense, Instituto de Computação, Niterói, 2011.
- [18] G., Schaefer; ZAVISEK, M.; NAKASHIMA, T. Thermography Based Breast Cancer Analysis Using Statistical Features and Fuzzy Classification. *Elsevier-Pattern Recognition*, 47 (2009) 1133-1137.
- [19] SCHAEFER, A. L. et al. The non-invasive and automated detection of bovine respiratory disease onset in receiver calves using infrared thermography. *Research in Veterinary Science*. Elsevier, 2012. vol. 92, p. 928-935.
- [20] SILVA D. C. M. SAADE, G. O. S. O. A. S. A. P. R. B. A. C. L. A new database for breast research with infrared image. *Journal of Medical Imaging and Health Informatics*, v. 4, n. 1, p. 92–100, 2014.
- [21] SILVA, L. F. da. *Uma Análise Híbrida para Detecção de Anomalias da Mama usando Séries Temporais de Temperatura*. Tese (Programa de Pós-Graduação em Computação) — Universidade Federal Fluminense, Instituto de Computação, Niterói, 2015.
- [22] YOUSSEF Ben Y. et al. Statistical Features and Classification of Normal and Abnormal Mammograms. *International Conference on Multimedia Computing and Systems (ICMCS)*, 2014.