

# Auxílio ao Diagnóstico Precoce de Patologias da Mama Usando Imagens Térmicas e Técnicas de Mineração de Dados

Roger Resmini<sup>1</sup>, Tiago B. Borchardt<sup>1</sup>, Aura Conci<sup>1</sup>, Rita C. F. Lima<sup>2</sup>

Instituto de Computação – Universidade Federal Fluminense (UFF)  
24.210-240 – Niterói – RJ – Brazil

Departamento de Engenharia Mecânica – Universidade Federal de Pernambuco (UFPE)  
52.071-530 – Recife – PE – Brasil

{rresmini,tbonini,aconci}@ic.uff.br,ritalima@ufpe.br

**Abstract.** *Cancer is the most common death cause among women in the world. Accord INCA (National Cancer Institute - Brazil), early detection and treatments are the main way of cure. Among developed countries the overlife rate after five years is 85%, while the overlife in developing countries is 60%. For 2012 in Brazil, are expected 52680 new cases. Cancer is a disease in which cells reproduce uncontrollably because they have lost their function in tissue, usually due to mutations. This reproduction increases the local temperature because new blood vessels, neo-angiogenesis, are required by cancer cells. The medical thermography is a way to register the skin temperature of volunteers, and after that, to make some analysis and find thermal patterns. The human body is a symmetric thermal system and some changes in the pattern between right and left breast could identify some pathology. This paper explores the feature extractions of breast thermal images and after that classify the volunteers among with or without pathology. We extract statistic features, fractal dimension and geo-statistic features. The results are very promises with accuracy near of 90% and area under ROC curve near of 0.900%. We test three classifiers: SVM, KNN and Naïve Bayes. And two feature reduction techniques: PCA and Information Gain Ratio.*

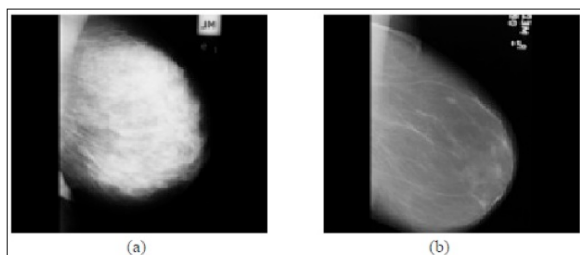
**Resumo.** *O câncer de mama é a maior causa de morte entre mulheres do mundo todo. Segundo o INCA (Instituto Nacional do Câncer), a detecção e tratamento precoce são as principais formas de cura. Dentre os países desenvolvidos a taxa de sobrevivência após cinco anos é de 85%, enquanto que em países em desenvolvimento é de 60%. Para 2012, no Brasil, são esperados 52680 novos casos. O câncer, caracteriza-se por uma proliferação (reprodução) descontrolada entre células que perderam sua função no tecido a que fazem parte, geralmente devido a mutações. O aumento da população celular mutante faz com que seja necessária uma maior nutrição da região, consequentemente uma maior vascularização para conduzir os nutrientes. Essa nova vascularização, neoangiogênese, ocasiona um aumento de temperatura local. A termografia médica é uma maneira de registrar a temperatura da pele das voluntárias para posterior análise e identificação de padrões térmicos. Tendo em vista que o corpo humano é um sistema que busca a simetria térmica, a presença de alteração no padrão térmico entre as mamas esquerda e direita, é um importante indício de presença de patologias. Este trabalho tem por objetivo explorar a extração de características de*

*imagens térmicas das mamas e posterior classificação das voluntárias entre com e sem patologia. Para tal, utilizou-se características estatísticas simples, dimensão fractal e características de geoestatística. Os resultados se mostraram bastante promissores com acurácia próxima de 90% e área abaixo da curva ROC próxima de 0,9%. Foram testados três classificadores, SVM, KNN e Naïve Bayes e duas técnicas de redução de características: PCA e Ganho de Informação.*

## 1. Introdução

Dados recentes do Instituto Nacional do Câncer [Inca 2011] apontam o câncer de mama como o segundo câncer com maior frequência no mundo, sendo este o mais comum entre as mulheres. Para reduzir a taxa de mortalidade por câncer é necessário o desenvolvimento de técnicas para detecção precoce da doença, o que permite que o tratamento seja iniciado ainda em estágio inicial [Inca 2011; Ng 2001]. Existem estudos indicando que no caso do câncer de mama, a detecção no início da doença eleva para 85% a possibilidade de cura, enquanto que quando detectada em estágio avançado esse percentual cai para 10% [Ng 2001].

Atualmente, os principais aliados da detecção do câncer de mama são o autoexame, que tem baixo custo, e a mamografia, que é o principal exame utilizado para o diagnóstico em estágios iniciais do câncer de mama (exame padrão-ouro). No entanto, uma limitação encontrada pela mamografia é a dificuldade de detectar tumores em pacientes mais jovens. Este fato deve-se as mamas de mulheres jovens serem constituídas de tecido glandular, tornando-as mais densas e firmes [Koay 2004]. Gradativamente, devido à diminuição dos níveis hormonais, esse tecido tende a ser substituído por tecido adiposo (menopausa). Esta nova constituição da mama a torna menos densa facilitando a detecção de regiões suspeitas através da mamografia [Bezerra 2007]. Na Figura 1, pode ser vista uma comparação entre duas mamografias, em (a) de uma paciente jovem e em (b) de uma paciente com mais idade. Pode-se perceber que na imagem (b) existe um contraste bem maior, facilitando a identificação de possíveis patologias.



**Figura 1. Mamografias: (a) mama jovem e (b) mama idosa.**

Esta limitação da mamografia, juntamente com o crescente número de casos de câncer em pacientes mais jovens, motiva o desenvolvimento de novas técnicas para a detecção precoce de patologias na mama, dentre as quais, destaca-se a termografia.

A termografia começou a ser utilizada na mastologia em 1959, porém os equipamentos disponíveis na época fizeram com que o método fosse desacreditado e não recomendado para o diagnóstico de doenças na mama. Com o avanço tecnológico das novas gerações de câmeras termográficas, pesquisas envolvendo a termografia voltaram a ser

realizadas, agora utilizando novas técnicas tanto de diagnóstico como de processamento de imagens médicas [Moghbel 2011].

O objetivo deste trabalho é prover a melhoria de trabalho anterior, apresentado a seguir, através de uso de novos classificadores e métodos de seleção de atributos aplicados a três bases de dados diferentes. Estas bases de dados são constituídas de características extraídas a partir de texturas de imagens termográficas utilizando diversos extratores de características.

O trabalho está organizado da seguinte maneira: na Seção 2 é feita uma breve explicação do exame de termografia; na Seção 3 é explicado o trabalho no qual este trabalho foi baseado e expostos resultados da literatura; na Seção 4 a metodologia proposta e implementada é apresentada; na Seção 5 são apresentados os resultados obtidos e por fim, na Conclusão busca-se avaliar os resultados obtidos e discutir a viabilidade de trabalhos futuros. Este trabalho foi realizado com suporte da CAPES.

## **2. Termografia**

A termografia é um exame fisiológico (funcional), diferente de outros exames tais como mamografia, ultrassonografia e ressonância magnética (sem uso de contraste), que são anatômicos (estruturais). As imagens térmicas, termogramas, são adquiridas por uma câmera termográfica que é sensível à radiação infravermelha. Todos os objetos que tenham temperatura acima do zero absoluto ( $-273^{\circ}\text{C}$ ), emitem radiação infravermelha na sua superfície [Ng 2008].

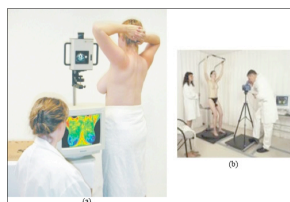
A medida da radiação infravermelha emitida pela pele pode ser convertida diretamente em um valor acurado de temperatura, devido ao alto valor de emissividade do corpo humano.

O sensor térmico de captura da radiação infravermelha varia de acordo com a câmera. Neste trabalho, são utilizadas imagens adquiridas por uma câmera termográfica modelo S45 da FLIR. O sensor presente nesta câmera tem sensibilidade de  $0,08^{\circ}\text{C}$  em um intervalo de captura de  $-40^{\circ}\text{C}$  à  $2000^{\circ}\text{C}$ , gerando imagens com uma resolução de 320 por 240 pixels [Flir Systems 2009]. A sensibilidade térmica desta câmera pode ser considerada dentro da faixa usada em aplicações médicas, conforme é relatado na literatura:  $0,5^{\circ}\text{C}$  [Kuruganti 2002],  $0,03^{\circ}\text{C}$  e  $0,09^{\circ}\text{C}$  [Umadevi 2011]. Esta câmera estima a temperatura do objeto analisado baseando-se em três componentes de radiação: a radiação emitida pelo objeto de interesse, a radiação refletida pelos objetos adjacentes e a radiação emitida pelo meio.

Para que a câmera possa estimar a influência da radiação do meio em que estão sendo obtidas as imagens, é necessário configurar alguns parâmetros antes: a emissividade do objeto de interesse (no caso do corpo humano este valor é 0,98), a temperatura ambiente, a umidade relativa do ar e a distância entre a câmera e o objeto.

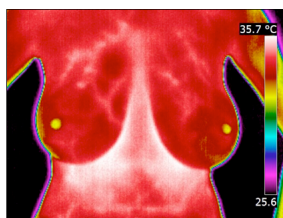
A partir de algoritmos internos, baseados na lei de Stefan-Boltzmann, a câmera converte a irradiação medida em uma imagem representativa de temperaturas pontuais da cena capturada. Esta imagem é apresentada em pseudocores que podem variar conforme a paleta de cores utilizada. As paletas são compostas por cento e vinte (120) cores discretas, com exceção da paleta Medical que apresenta somente dez (10) cores. Cada cor representa um intervalo de temperaturas. Ao lado de cada imagem é mostrada a paleta utilizada e os limites de temperatura representados na imagem.

Na mastologia a termografia é utilizada para a análise das variações térmicas que ocorrem nas mamas. Uma variação anormal de temperatura entre uma mama e outra pode indicar alguma patologia. A termografia é indicada para pacientes de qualquer idade, não tendo contraindicações e podendo ser repetida quantas vezes forem necessárias, sendo considerado um exame complementar para o diagnóstico de patologias. O processo de aquisição das imagens térmicas da mama é simples e totalmente indolor, como pode ser visto na Figura 2.



**Figura 2. Captura realizada: (a) Infraredmed (2009) e (b) Sobraterm (2010).**

Na Figura 3 é apresentado um exemplo de termograma, onde pode ser observada a paleta de cores na lateral da imagem.



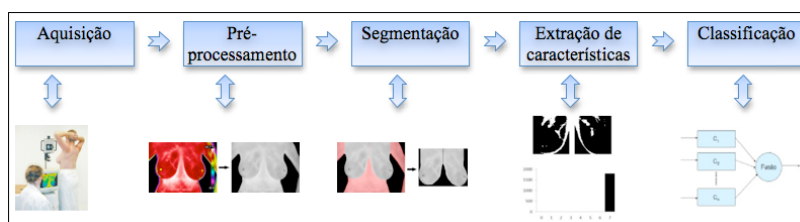
**Figura 3. Exemplo de termograma utilizado no projeto.**

As imagens utilizadas neste trabalho são provenientes do projeto “Processamento e análise de imagens aplicadas à mastologia”, que foi elaborado em uma parceria entre o Instituto de Computação da Universidade Federal Fluminense (UFF) e o Departamento de Engenharia Mecânica da Universidade Federal de Pernambuco (UFPE). Os principais objetivos deste projeto são o desenvolvimento de métodos de simulação e análise dos perfis de temperatura das mamas baseados em imagens térmicas e modelos 3D gerados através delas e a elaboração de métodos automáticos para o auxílio ao diagnóstico (CAD – *Computer Aided Diagnosis*). Atualmente estas imagens estão disponíveis em um banco de dados no site: <http://visual.ic.uff.br/proeng>.

### **3. Trabalhos Relacionados**

Técnicas de auxílio ao diagnóstico baseadas em imagens são muito utilizadas para os mais variados tipos de patologia. As técnicas desenvolvidas para detectar patologias da mama trabalham considerando a comparação de características entre as duas mamas. Em todos os tipos de exames (mamografia, ultrassom, ressonância magnética e termografia) espera-se encontrar certa simetria entre uma mama e a outra. Assim, qualquer assimetria encontrada pode ser indício de alguma alteração ou patologia na mama. No caso das imagens térmicas, uma diferença razoavelmente alta dos padrões de distribuição de temperatura entre as mamas pode ser considerada um sinal de alerta.

Este trabalho tem como base a dissertação de mestrado de Resmini (2011) e a Figura 4 mostra um esquema da metodologia proposta.



**Figura 4. Esquema da metodologia proposta por Resmini (2011).**

Neste trabalho, será discutida a fase de classificação que obteve os resultados mostrados na Tabela 1, para a classificação dos diagnósticos de exames de termografia, onde foram utilizadas vinte e quatro (24) imagens de pacientes com patologia e quatro (4) imagens de pacientes sem patologia. A baixa taxa de especificidade se deve ao fato do método ter apontado apenas um diagnóstico como realmente sem patologia, enquanto os outros 3 foram apontados como possuindo patologia, quando na realidade eram saudáveis. A taxa de sensibilidade alta é um ponto positivo, pois poucos diagnósticos foram apontados como sem patologia, quando na realidade tinham patologia. Mesmo assim essa taxa deve ser melhorada.

**Tabela 1. Resultados de Resmini (2011).**

	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
Resmini	82,14	91,70	25,00

Para alcançar tais resultados, um total de setecentas e doze características (712), atributos, foram extraídas. Para a diminuição da dimensionalidade do espaço de atributos, foi utilizada uma técnica de PCA que reduziu para dezessete (17), a quantidade de características a serem utilizadas. As características extraídas são divididas em três categorias: medidas estatísticas simples (média, desvio padrão, intervalo entre o maior e menor tom de cinza da imagem e a quantização do último tom de uma posterização de 8 bins), dimensão fractal (*Higuchi*) e medidas de geoestatística (índice de *Moran* e coeficiente de *Geary*). A base de dados utilizada, possuía um total de vinte e oito voluntárias, sendo quatro com diagnóstico sem patologia e vinte e quatro com diagnóstico com patologia (tumor maligno ou benigno).

O trabalho de Resmini (2011) descreve vários resultados encontrados na literatura, mas o melhor resultado é de Schaefer *et al.* (2009), que no total considera trinta e oito características dos termogramas. Estas características foram analisadas utilizando lógica *fuzzy* com uma técnica de validação cruzada. Os principais resultados alcançados por Schaefer *et al.* (2009) foram 79,86% de acurácia, 79,86% de sensibilidade e 79,49% de especificidade, usando 14 partições.

Um outro trabalho é de Acharya *et al.* (2010), que descreve a extração de dezesseis (16) características, mas somente quatro (4) características são utilizadas, pois os autores definiram-las como clinicamente significantes comparadas com as outras. Os autores relatam obter 88,10% de acurácia, 85,71 de sensibilidade e 90,48% de especificidade.

Com a intenção de melhorar os resultados encontrados por Resmini (2011), buscou-se uma nova forma de seleção de atributos, proporcionando além da diminuição de processamento, a melhoria dos resultados da classificação.

#### 4. Metodologia

Três abordagens foram adotadas. Na primeira, extraiu-se o mesmo número de características descritas em Resmini (2011), setecentas e doze utilizando uma base de dados de vinte e oito imagens (quatro sem patologia e vinte e quatro com patologia). Esta é a mesma base utilizada no trabalho anterior e aqui será chamada de “antiga”. Na segunda abordagem, utilizou-se a base anterior acrescida de seis imagens de voluntárias sem patologia, totalizando dez imagens de voluntárias sem patologia e vinte e quatro imagens de voluntárias com patologia. Esta base será nomeada por “nova”. Na terceira abordagem, utilizou-se a subtração entre as características extraídas da imagem da mama direita pelas características extraídas da imagem da mama esquerda de cada voluntária. Esta base será chamada de “subtração”. Não utilizou-se registro para que a geometria de uma mama se adequasse a geometria da outra. Com a subtração das características, espera-se trabalhar com informações relacionadas as duas mamas e não características isoladas de cada mama como nas outras bases. Utilizou-se duas classes: com patologia e sem patologia. A Tabela 2, mostra as base utilizadas.

**Tabela 2. Bases de dados utilizadas.**

Base	Tamanho	Com patologia	Sem patologia	Atributos
Antiga	28 imagens	24 imagens	4 imagens	712
Nova	34 imagens	24 imagens	10 imagens	712
Subtração	34 imagens	24 imagens	10 imagens	356

Utilizou-se três classificadores: IbK (KNN), *Naïve Bayes* e SVM. Este último, também foi utilizado no trabalho anterior. Todos os classificadores estão disponíveis na ferramenta WEKA (2011), uma ferramenta que proporciona vários algoritmos de aprendizado de máquina para tarefas de mineração de dados: classificação, agrupamento (*clusterization*) e extração de regras de decisão. Para cada classificador, fez-se diversos experimentos e as melhores configurações obtidas foram:

1. SVM, utilizou-se um *SVMTType = nu-SVC*, *nu = 0,5* e *kernelType = linear*;
2. IbK, utilizou-se um *KNN = 5* e *distanceWeighting = weight by 1 - distance*;
3. *Naïve Bayes*, com configuração padrão do WEKA.

Os demais parâmetros não foram alterados e permaneceram os padrões do WEKA.

Como descrito em Pereira (2011), algumas importante motivações de utilizar seleção de atributos são: melhoria da acurácia do classificados, aceleração da tarefa de classificação, simplificação do modelo de classificação gerado.

Utilizou-se então, com intensão de melhorar a acurácia obtida na literatura e no trabalho anterior de Resmini (2011), a seleção de atributos utilizando uma técnica de PCA (novamente) e uma técnica de “Ganho de Informação” (*InfoGainRatio*), ambos presentes na ferramenta WEKA.

## 5. Resultados

Nesta Seção descreve os experimentos realizados e expõe os resultados obtidos. Para avaliar o desempenho de cada experimento, utilizou-se a medida de acurácia (ACC), que é o total de acertos, e a área abaixo da curva ROC, razão entre a sensibilidade e a especificidade (sensibilidade/1-especificidade).

A Tabela 3, mostra os resultados obtidos para a base de dados “antiga”, onde os resultados obtidos com o classificador SVM, são os mesmos obtido no trabalho anterior.

**Tabela 3. Resultados obtidos com a base “antiga”.**

	SVM	IbK	<i>Naïve Bayes</i>
ACC (%)	82,14	85,71	75,00
ROC (%)	0,580	0,708	0,449

Embora a acurácia do classificador SVM esteja maior que encontrado em Schaefer *et al.* (2009), a área abaixo da curva ROC ficou bastante baixa, isso devido a especificidade do método ter sido baixa. Um dos motivos para isso, acredita-se ser a base desbalanceada, pois, foram utilizados apenas quatro diagnósticos sem patologia, e esta quantidade de diagnóstico pode não ter sido suficiente para o classificador reconhecer um padrão neste tipo de diagnóstico. O classificador IbK mostrou melhor desempenho para esta base.

A Tabela 4, mostra os resultados obtidos para a base “nova”, onde o classificador SVM melhorou sutilmente a acurácia, mas melhorou razoavelmente a área abaixo da curva ROC. O classificador IbK, melhorou ainda mais o desempenho, mostrando-se promissor para esta base.

**Tabela 4. Resultados obtidos com a base “nova”.**

	SVM	IbK	<i>Naïve Bayes</i>
ACC (%)	82,35	88,23	73,52
ROC (%)	0,700	0,842	0,746

A Tabela 5, mostra os resultados obtidos para a base “subtração”, onde o classificador SVM piorou o desempenho e o classificador IbK diminuiu a acurácia mas manteve a área abaixo da curva ROC obtida com a base anterior. O classificador *Naïve Bayes*, teve uma melhora de desempenho sutil se comparado com a base anterior.

**Tabela 5. Resultados obtidos com a base “subtração”.**

	SVM	IbK	<i>Naïve Bayes</i>
ACC (%)	76,47	85,29	79,41
ROC (%)	0,658	0,842	0,750

Em seguida, utilizou-se a seleção de atributos com a intenção de melhorar o desempenho dos classificadores.

A Tabela 6 (linhas 1 e 2), mostram os resultados obtidos com a base “subtração” e uso da seleção de atributos fixando em 10% (trinta e cinco atributos). Os classificadores SVM e *Naïve Bayes*, mostraram melhores valores de acurácia, mas o segundo obteve uma área abaixo da curva ROC superior. O classificador IbK, embora tenha obtido uma área abaixo da curva ROC praticamente igual ao classificador *Naïve Bayes*, obteve uma acurácia razoavelmente inferior aos outros dois classificadores.

A Tabela 6 (linhas 3 e 4), mostram os resultados obtidos com a base “subtração” e uso da seleção de atributos fixando em 20% (setenta atributos). O classificador SVM diminuiu consideravelmente o desempenho, enquanto o classificador IbK melhorou a

acurácia e área abaixo da curva ROC. O classificador *Naïve Bayes* diminuiu o desempenho ficando com a mesma acurácia do classificador IbK.

**Tabela 6. Resultados obtidos com a base “subtração” e 10% dos atributos.**

Linha	Critério de Avaliação	SVM	IbK	Naïve Bayes	Atributos
1	ACC (%)	88,23	82,35	88,23	10%
2	ROC (%)	0,800	0,858	0,857	10%
3	ACC (%)	76,47	85,29	85,29	20%
4	ROC (%)	0,688	0,867	0,827	20%
5	ACC (%)	70,58	85,29	85,29	30%
6	ROC (%)	0,617	0,904	0,876	30%

A Tabela 6 (linhas 5 e 6), mostram os resultados obtidos com a base “subtração” e uso da seleção de atributos fixando em 30% (cento e cinco atributos).

Depois utilizou-se estas mesmas proporções para o Ganho de Informação utilizando a base “nova”, com setecentos e doze atributos. A Tabela 7, mostra o desempenho de todos os classificadores para a variação de Ganho de Informação de 5% (trinta e cinco atributos), 10% (setenta e um atributos) e 15% (cento e seis atributos).

**Tabela 7. Resultados obtidos com a base “nova”.**

	SVM	IbK	Naïve Bayes	InfoGainRatio
ACC (%)	88,23	85,29	85,29	5%
ROC (%)	0,858	0,908	0,845	5%
ACC (%)	85,29	85,29	85,29	10%
ROC (%)	0,808	0,888	0,816	10%
ACC (%)	85,29	82,35	88,23	15%
ROC (%)	0,779	0,863	0,822	15%

Enquanto o classificador SVM obteve melhor acurácia com uma menor quantidade de atributos, o classificador *Naïve Bayes* obteve maior acurácia com uma quantidade de atributos maior. O classificador IbK manteve-se estável nos dois primeiros testes e declinou sutilmente o desempenho no último teste.

**Tabela 8. Resultados obtidos com as bases “nova” e “subtração” e PCA.**

	SVM	IbK	Naive Bayes	Base	Atributos
ACC (%)	88,23	79,41	76,47	Nova	22
ROC (%)	0,800	0,658	0,817	Nova	22
ACC (%)	88,23	67,64	79,41	Subtração	18
ROC (%)	0,800	0,592	0,747	Subtração	18

Por fim, utilizou-se uma técnica de PCA para a seleção de atributos nas bases “nova” e “subtração”, Tabela 8. O classificador SVM, obteve o melhor desempenho com uma acurácia de 88,23% e área abaixo da curva ROC de 0,800%, nas duas bases. O



classificador IbK se saiu pior na base “subtração” e o classificador *Naïve Bayes* melhorou o desempenho na base “subtração”.

## 6. Conclusão

Este trabalho buscou melhorar os resultados obtidos na dissertação de Resmini (2011), para a proposta de um método de auxílio ao diagnóstico de patologias da mama utilizando características extraídas da textura de imagens térmicas. Utilizando-se outros classificadores além do utilizado no trabalho anterior e também outra técnica de seleção de atributos.

Foram analisados três classificadores: SVM, IbK e *Naïve Bayes*. Todos eles presentes na ferramenta WEKA. Estes classificadores foram aplicados a três bases de dados: uma chamada “antiga”, com vinte e oito imagens (quatro sem patologia e vinte e quatro com patologia), outra base chamada “nova”, com as vinte e oito imagens da base antiga, acrescidas de mais seis sem patologia totalizando trinta e quatro imagens. Por fim uma base chamada “subtração”, com trinta e quatro imagens. As duas primeiras bases possuem setecentos e doze atributos, enquanto a última tem trezentos e cinquenta e seis atributos que são resultantes da subtração das características extraídas da mama direita pelas características extraídas da mama esquerda. Com a intenção de melhorar o resultado da classificação, aplicou-se a seleção de atributos utilizando-se de uma técnica de PCA e Ganho de Informação.

O classificador SVM obteve seu melhor desempenho com a base “nova”, setecentos e doze atributos e trinta e quatro imagens, com uma acurácia de 88,23% e área abaixo da curva ROC de 0,858% e utilizando a seleção de atributos por Ganho de Informação de 5%. Estes valores superaram os resultados obtidos no trabalho anterior onde a acurácia foi de 82,14% e a área abaixo da curva ROC de 0,580%. Também superou o trabalho citado da literatura, Schaefer (2009), onde a acurácia foi de 79,53% (este trabalho não usou a área abaixo da curva ROC como medida). Também superou a acurácia de 88,10% relatada por Acharya (2010). Essas comparações de resultados são puramente ilustrativas, pois as bases são diferentes e não faz sentido compará-las.

O classificador *Naïve Bayes* teve seu melhor desempenho com a base “subtração”, trezentos e cinquenta e seis atributos e trinta e quatro imagens. Obteve uma acurácia de 88,23% e área abaixo da curva ROC de 0,857% e seleção de atributo por Ganho de Informação de 10%, trinta e cinco atributos.

O classificador IbK obteve seu melhor desempenho com a base “nova”, setecentos e doze atributos e trinta e quatro imagens. Alcançou uma acurácia de 88,23% e área abaixo da curva ROC de 0,842% sem seleção de atributos.

A seleção de atributos por Ganho de Informação, mostrou ter ajudado na melhoria do desempenho dos dois métodos que tiveram melhores resultados, SVM e *Naïve Bayes*, respectivamente.

## Referências

Acharya, U.R., Ng, E. Y. K., Tan, J. H., e Sree, S. V. (2010) “Thermography based breast cancer detection using texture features and support vector machine”, *Journal of Medical Systems*, pp. 01-08.

- Bezerra, L. (2007) “Uso de imagens termográficas em tumores mamários para validação de simulação computacional”, Dissertação de Mestrado. Departamento de Engenharia Mecânica, Universidade Federal de Pernambuco.
- Flir Systems. (2004) ThermaCAM TM S45 - Manual do operador.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I. H. (2009) “The WEKA Data Mining Software: An Update”, SIGKDD Explorations, vol. 11.
- Inca. (2011) “Estimativa 2012/2013”. Acesso em 1/12/2011, disponível em: <<http://www.inca.gov.br/estimativa/2012/index.asp?ID=5>>.
- Infraredmed. (2009) “Diagnóstico por infravermelho: Como é realizado”. Acesso em 1/12/2011, disponível em: <[http://www.infraredmed.org/exame\\_como.php](http://www.infraredmed.org/exame_como.php)>.
- Koay, J., Herry, C., Frize, M. (2004) “Analysis of Breast Thermography with an Artificial Neural Network”, Proceedings of the 26th Annual International Conference of the IEEE (EMBS), San Francisco, CA, USA, pp. 1159-1162.
- Kuruganti, P. T., e Qi, H. (2002) “Asymmetry analysis in breast cancer detection using thermal infrared images”, Proceedings of Second Joint EMBS/BMES Conference, Houston, TX, USA, Vol. 2, No. 1, pp. 1129-1130.
- Moghbel, M. e Mashohor, S. (2011) “A review of computer assisted detection/diagnosis (CAD) in breast thermography for breast cancer detection”, Artificial Intelligence Review, Springer Netherlands, pp. 1-9. Acesso em 20/11/2011, disponível em: <<http://dx.doi.org/10.1007/s10462-011-9274-2>>.
- Ng, E. (2008) “A review of thermography as promising non-invasive detection modality for breast tumor”, International Journal of Thermal Sciences, vol. 48, pp. 849-859.
- Ng, E., e Sudarshan, N. (2001) “Numerical computation as a tool to aid thermographic”, Journal of Medical Engineering and Technology, vol. 25, ed. 2, pp. 53-60.
- Pereira, R. P., Plastino, A., Zadrozny, B., Merschmann, L. H. C. e Freitas, A. A. (2011) “Lazy attribute selection: Choosing attributes at classification time”, Intelligent Data Analysis Journal, IOS Press, vol. 15, ed. 5, pp. 715-732.
- Resmini, R. (2011) “Análise de Imagens Térmicas da Mama Usando Descritores de Textura”, Dissertação de Mestrado. Instituto de Computação, Universidade Federal Fluminense.
- Schaefer, G., Zavissek, M., Nakashima, T. (2009) “Thermography based breast cancer analysis using statistical features and fuzzy classification”, Pattern Recognition, vol. 42, ed. 6, pp. 1133–1137.
- Sobraterm. (2010) Sociedade Brasileira de Termologia. Disponível em: <<http://www.termologia.org>>.
- Umadevi, V., Raghavan, S. V., e Jaipurkar, S. (2011) “Framework for estimating tumour parameters using thermal imaging”, Indian Journal of Medical Research, Vol. 134, pp. 725-731, November 2011.