

DIAGNÓSTICO PRECOCE DE DOENÇAS MAMÁRIAS USANDO IMAGENS TÉRMICAS E APRENDIZADO DE MÁQUINA

Roger Resmini¹, Aura Conci¹, Tiago Bonini Borchart¹, Rita de Cássia Fernandes de Lima², Anselmo Antunes Montenegro¹, Cristina Asvolinsque Pantaleão¹

¹Universidade Federal Fluminense - UFF

²Universidade Federal de Pernambuco - UFPE

rresmini@ic.uff.br, aconci@ic.uff.br, tbonini@ic.uff.br, ritalima@ufpe.br,
anselmo@ic.uff.br, cristinaasvolinsque@gmail.com.

Resumo

O câncer é uma doença que se origina de células mutantes, sem causas bem conhecidas ainda, que se reproduzem descontroladamente, aumentando a perfusão sanguínea e, conseqüentemente, ocasionando um aumento da temperatura da região tumoral. Essa temperatura é irradiada para a pele e pode ser medida por câmeras térmicas. Tendo em vista que o corpo humano é um sistema praticamente simétrico em relação ao plano sagital (i.e. ao plano que divide o corpo em parte direita e esquerda), a presença de uma grande alteração no padrão térmico entre as mamas esquerda e direita, é um importante indício de presença de patologias. Este trabalho tem por objetivo verificar a viabilidade do uso de técnicas de reconhecimento de padrões na classificação das imagens disponíveis no projeto ProENG com pacientes saudáveis ou com portadoras de alguma patologia da mama. Para tanto, destas imagens são extraídas características que permitirão a sua classificação através de técnicas de Inteligência Artificial. Utilizou-se características de três grupos distintos: estatísticas simples, baseadas na geometria fractal e características de fundamentação geoestatística. Foram testados três classificadores, SVM, KNN e Naïve Bayes e duas técnicas de redução de características: PCA e Ganho de Informação. Os resultados se mostraram bastante promissores com uma acurácia próxima de 90% e área abaixo da curva ROC próxima de 0,9.

Palavras-chave: Termografia. Câncer de mama. Auxílio ao diagnóstico.

Abstract

Cancer is a class of diseases characterized by out-of-control cell growth, they have lost their function in tissue and do not die. This reproduction increases the local temperature because new blood vessels, neo-angiogenesis, are promoted by cancer cells. The medical thermography is a way to acquire the skin temperature and analyze these patterns. The human body is almost symmetric considering the sagittal plane that is the plane that divides the body in right and left parts, when there are great changes in the temperature pattern between right and left breast possible pathology must be investigated. This work aims to explore the possibilities of pattern recognition techniques on the classification of the images from the ProENG project as from healthy or pathological mamma. Three different groups of feature are extracted from the thermal images of this project: statistic features, fractal geometry based features and geo-statistic features. Three classifiers have been tested: SVM, KNN and Naïve Bayes. Additionally two feature reduction techniques have been used: PCA and Information Gain Ratio. The results are promising: 90% of accuracy and 0.9 for the area under ROC.

Keywords: Thermography. Breast cancer. Computed Aided Diagnosis.

1. Introdução

Dados recentes do Instituto Nacional do

Câncer (INCA, 2012) apontam o câncer de mama como o segundo câncer com maior frequência no Brasil, sendo este o mais comum entre as mulheres. Este tipo de

câncer é responsável por 22% dos novos casos de câncer a cada ano. No entanto se diagnosticado e tratado oportunamente, o prognóstico é relativamente bom. Segundo a Organização Mundial da Saúde (OMS, 2012) nos anos de 60 e 70 registrou-se um aumento de 10 vezes nas taxas de incidência ajustadas por idade nos Registros de Câncer de Base Populacional de diversos continentes (INCA, 2012). No Brasil, as taxas de mortalidade por câncer de mama continuam elevadas, muito provavelmente porque a doença ainda é diagnosticada em estágios avançados. Na população mundial, a sobrevida média após cinco anos é de 61%. O câncer de mama é relativamente raro antes dos 35 anos, mas acima desta faixa etária sua incidência cresce rápida e progressivamente (INCA, 2012). 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 (NG, 2001) (INCA2, 2012). 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 detectado em estágio avançado, esse percentual cai para 10% (NG, 2001).

O fator hereditariedade, mulheres com histórico familiar de câncer de mama, é responsável por apenas 10% do total de casos, porém estas mulheres apresentam maior risco de desenvolver a doença (INCA3, 2012). Esse grupo deve ser acompanhado por médico a partir dos 35 anos. É o profissional de saúde quem vai decidir quais exames a paciente deverá fazer. Primeira menstruação precoce, menopausa tardia (após os 50 anos), primeira gravidez após os 30 anos ou nunca ter engravidado também constituem fatores de risco para o câncer de mama (INCA3, 2012).

Atualmente, os dois principais aliados da detecção do câncer de mama são o exame clínico, 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 (INCA3, 2012). O Exame Clínico das Mamas se realizado por um médico ou enfermeira treinados, pode detectar tumor de até 1 (um) centímetro. Deve ser feito uma vez por ano pelas mulheres entre 40 e 49 anos (INCA, 2012). A mamografia permite a detecção precoce do câncer, ao mostrar lesões em fase inicial, muito pequenas (medindo milímetros), deve ser realizada a cada dois anos por mulheres entre 50 e 69 anos ou segundo recomendação médica. É realizada em um aparelho de Raios X apropriado, chamado mamógrafo. Nele, a mama é comprimida de forma a fornecer melhores imagens, e, portanto, melhor capacidade de diagnóstico. O desconforto provocado é suportável. Segundo a Lei nº 11.664, (LEI 11664, 2008) (INCA3, 2012), todas as brasileiras têm direito à mamografia a partir dos 40 anos pelo Sistema Único de Saúde. Essa lei, embora tenha suscitado interpretações divergentes, não altera as recomendações de faixa etária para rastreamento de mulheres saudáveis: dos 50 aos 69 anos (INCA3, 2012).

Assim sendo, observa-se que o Exame Clínico das Mamas só detecta problemas em estágio já bastante avançado (INCA3, 2012). Quanto a mamografia, esta se mostra muito adequada para a detecção de microcalcificações em pacientes pós menopausa, quando a estrutura das mamas é formada por muito tecido adiposo, mas representa grande dificuldade da detecção precoce de tumores em pacientes mais jovens (INCA3, 2012). Isto deve-se ao fato de mamas de mulheres jovens serem constituídas principalmente de tecido glandular, tornando-as mais densas, firmes e com dificuldades de identificação diferencial entre tecidos e achados relevantes pela mamografia (KOAY, 2004). Com a idade, gradativamente, devido à diminuição dos níveis hormonais e necessidades reprodutivas, esse tecido tende a ser substituído por tecido adiposo. Esta nova constituição da mama a torna menos densa facilitando a detecção de regiões suspeitas através da mamografia (BEZERRA, 2007).

Além disso, a mamografia requer maior exposição aos Raios X que o uso destes para radiografias convencionais. Segundo estudos recentes cada vez que a mamaé exposta aos Raios X, há um aumento do riscodo câncerde2%, se a mama for de uma paciente na pré-menopausa, e portanto mais sensível à radiação (ARABI, MUTTAN e SUJI, 2010). Um cálculo simples mostra que após 35 exposições, o risco de câncer de uma pessoa será dobrado, apenas devido as mamografias que realizar. Esta característica da mamografia, juntamente com a constatação do aumento da longevidade e a geralmente maior agressividade dos cânceres em pacientes mais jovens, motiva a busca de novas técnicas para a detecção precoce de patologias das mamas.

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 das mamas. Com o avanço tecnológico das novas 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 identificar as possibilidades do uso de técnicas de reconhecimento de padrões na classificação das imagens térmicas. Pretende-se prover a melhoria de trabalhos anteriores no tema através de uso de novos classificadores e métodos de seleção de atributos aplicados a três bases de dados diferentes.

O trabalho está organizado da seguinte maneira: na Seção 2, é feita uma breve explicação do exame de termografia; na Seção 3, considera-se trabalhos anteriores e 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.

2. Termografia

Todos os objetos que têm temperatura acima

do zero absoluto (-273°C), emitem radiação infravermelha na sua superfície (NG, 2008). A radiação infravermelha emitida pela pele pode ser convertida em um valor de temperatura, devido ao alto valor de emissividade do corpo humano (BEZERRA, 2007). As imagens térmicas, termogramas, são adquiridas por uma câmera termográfica que é sensível à radiação infravermelha. O sensor térmico de captura da radiação infravermelha varia de acordo com a câmera. Neste trabalho, são utilizadas imagens adquiridas por uma câmera termográfica ThermaCAMFLIRS45no Hospital dasClínicasda UniversidadeFederaldePernambuco (HC-UFPE). O projeto de aquisição das imagens foiregistrado noMinistério da Saúde brasileiro após ter sidopreviamente aprovado peloComitê de Ética daUFPE (CEP / CCS /UFPEN °279/05).As imagens adquiridasestão disponíveis emum banco de dadospúblico(<http://150.161.110.168/termo>). O sensor presente nesta câmera tem sensibilidade de 0,08°C em um intervalo de captura de -40°C à 2000°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, que, de acordo com a literatura, dependendo dos autores seria de 0,5°C (KURUGANTI, 2002), ou na faixa de 0,03°C a 0,09°C (UMADEVI, 2011). Esta câmera converte o valor adquirido na temperatura da superfície 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 humidade 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 matriz de pontos representativos da temperatura de pontos 2D da cena capturada. Se esta matriz for representada em níveis de cinza dentro do intervalo de temperaturas da cena ou se apresentada em pseudocores (que podem variar conforme a paleta de cores utilizada), poderá ser vista como uma imagem pelo olhos humanos. Cada tom representa um intervalo de temperaturas. Ao lado de cada imagem adquirida por essa câmera, é mostrada a paleta utilizada e os limites de temperatura representados na imagem. A Figura 1, mostra algumas destas possibilidades. Assim, a representação digital de um termograma pode ser definida como uma função tridimensional $M(i,j,t)$, onde i e j denotam a posição de cada elemento na matriz e t é o valor da temperatura naquela posição (i,j) ou ponto. Esta matriz de temperatura pode ser obtida usando o software proprietário da FLIR.

Essa representação matricial pode se transformar também em uma imagem em pseudocores se o valor térmico de cada ponto (i,j,t) , for relacionado (indexado como um ponteiro) ao mapa de cores. Portanto, a cor de cada ponto na imagem depende da cor para a qual é mapeada (paleta) e dos limites definidos pela escala de temperatura para cada imagem. Assim, as imagens de temperatura ou infravermelhas (IR), são originadas da representação como tom das temperaturas de uma cena. A Figura 1 mostra alguns termogramas visualizados por três diferentes paletas. Todas as imagens na Figura 1 são da mesma cena e portanto representam a mesma matriz de temperatura. As paletas usuais, disponíveis no software fornecido com o equipamento, são compostas por cento e vinte (120) cores discretas, com exceção da paleta Medical que apresenta dez (10) tons de cor.

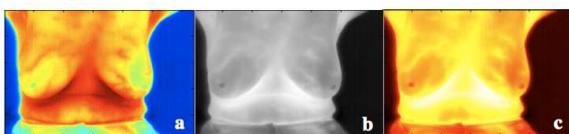


Figura 1 – Três diferentes representações da mesma cena a) paleta Rainbow; b) paleta Grayscale; e c) paleta Iron; (OLIVEIRA, 2012).

Na mastologia, a termografia é utilizada para a análise das variações térmicas que ocorrem nas mamas. O uso da termografia das mamas como exame de rotina da mama é indicado para pacientes de qualquer idade, não tendo contra-indicações e podendo ser repetida quantas vezes forem necessárias, podendo ser considerado um exame complementar ao exame clínico no auxílio ao diagnóstico de doenças das mamas. O processo de aquisição das imagens térmicas das mamas é simples, totalmente indolor e feito preferencialmente com uso de um aparato auxiliar no posicionamento dos braços das pacientes sobre a cabeça. A Figura 2, mostra o aparato desenvolvido pela UFPE para captura destas imagens.

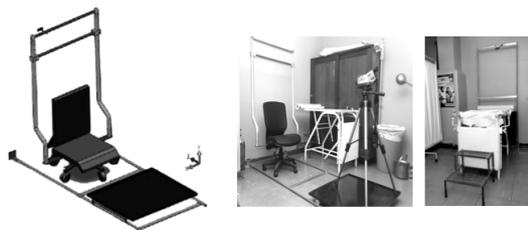


Figura 2 – Exemplo do aparato usado para aquisição das imagens desenvolvido pela UFPE (OLIVEIRA, 2012).

Na Figura 3, são apresentados exemplos de termogramas, em diversas poses de captura: (a) T1 – frontal com as mãos na cintura, (b) T2 – frontal com as mãos na cabeça, (c) T2 – com grade metálica, (d) LIMD – Lateral Interna da Mama Direita, (e) LIME – Lateral Interna da Mama Esquerda, (f) LEMD – Lateral Externa da Mama Direita e (g) LEME – Lateral Externa da Mama Esquerda.

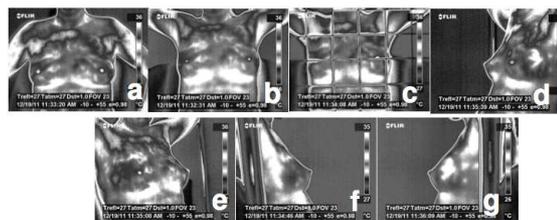


Figura 3 – Posições de capturas, (OLIVEIRA, 2012).

As imagens adquiridas obedecem a um protocolo padrão. Sempre se registra a temperatura ambiente durante o exame e espera-se que cada paciente tenha a estabilização da sua temperatura no ambiente da sala de teste. Todos os funcionários que participam no processo de aquisição de imagens termográficas, também estão em equilíbrio térmico com a sala (encontram-se na sala muito antes da aquisição de imagem e entrada da paciente, sendo que esta espera por 10 minutos antes da aquisição das imagens). Esta padronização é usada para reduzir ao mínimo quaisquer oscilações de temperatura no interior da sala durante a aquisição da imagem.

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, grandes assimetrias podem ser indício de alguma alteração ou doença da mama. No caso das imagens térmicas, uma diferença razoavelmente alta dos padrões de distribuição de temperatura entre as mamas pode ser considerado um sinal de alerta.

Entre os resultados encontrados na literatura, o trabalho de Schaefer, Zavisek e Nakashima (2009) consideram trinta e oito características obtida de termogramas. Estas características foram analisadas utilizando lógica fuzzy com uma técnica de validação cruzada. Os melhores resultados alcançados por Schaefer, Zavisek e Nakashima (2009) foram 79,53% de acurácia, 79,86% de sensibilidade e 79,49% de especificidade, usando 14 partições.

Wiecek *et al.*(2008), relatam vários possíveis extratores de características, mas mostram resultados somente com a classificação utilizando a média aritmética

das temperaturas: 80,95% de acurácia.

Arora *et al.*(2008) analisam um software de CAD térmico com três modos de operação: triagem, clínico e rede neural artificial. Os autores não detalham as características extraídas e nem como o software trabalha, mas expõe as acurácias obtidas nos três modos: triagem, 66,70%; clínico, 71,40%; e rede neural artificial, 81,80%.

Outro trabalho interessante é o de Acharya *et al.* (2010), que faz a extração de dezesseis (16) características inicialmente, embora que no decorrer do método somente quatro (4) características sejam empregadas, pois foram consideradas como clinicamente significantes comparadas com as outras. Neste trabalho, os autores chegam a obter 88,10% de acurácia, 85,71% de sensibilidade e 90,48% de especificidade.

4. Metodologia

Este trabalho tem como base o trabalho desenvolvido anteriormente (RESMINI, 2011), onde nos testes realizados foram utilizadas vinte e quatro (24) imagens de pacientes com patologia e quatro (4) imagens de pacientes sem patologia. O diagnóstico foi determinado por laudo médico, após cirurgia e biópsia. Após a aquisição das imagens, elas foram convertidas para tons de cinza e o fundo foi removido (pré-processamento). Então as imagens foram segmentadas para utilizarem-se nas fases seguintes, apenas a região de interesse (do inglês, *Region Of Interest* – ROI) que neste caso eram as regiões das mamas. Na fase de extração de características, os algoritmos desenvolvidos varreram as ROIs extraindo informações segundo conceitos de estatística simples, geometria fractal e geoprocessamento. Por fim, na fase de classificação, utilizou-se o classificador SVM (do inglês, *Support Vector Machine* ou máquina de vetores de suporte). A Figura 4 mostra metodologia utilizada.

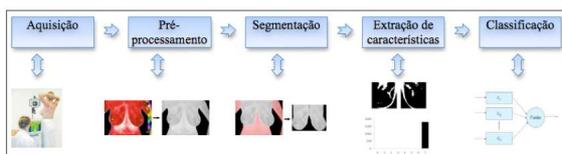


Figura 4 – Passos da metodologia proposta por Resmini (2011).

O trabalho anterior obteve 82,14% de acurácia (RESMINI, 2011). Para alcançar tais resultados, um total de setecentas e doze características (712), atributos, foram extraídas de cada paciente. Para a diminuição da dimensionalidade do espaço de atributos, foi utilizada a técnica de Análise dos Componentes Principais (do inglês, *Principal Component Analysis - PCA*) que reduziu para dezessete (17) a quantidade de características a serem utilizadas.

As características extraídas foram divididas em três grupos distintos. No primeiro grupo, medidas estatísticas simples foram utilizadas: a média dos tons de cinza; o desvio padrão destes tons; o intervalo de tons, ou seja, a diferença entre o maior e menor tom de cinza da ROI; e o número de *pixels* com intensidade acima de 87,5% do valor mais intenso da ROI, ou seja, a partir do histograma da ROI quantizado em 8 tons (*bins* ou níveis), computou-se o número de *pixels* pertencentes ao último tom quantizado (que contém os *pixels* mais claros e simbolizam regiões mais quentes da ROI).

No segundo grupo, foi utilizada a dimensão fractal de Higuchi (PARAMANATHAN e UTHAYAKUMAR, 2008). Como medidas de geoestatística, último grupo, foram usados o índice de Moran e o coeficiente de Geary (SHEKAR e XIONG, 2008).

No trabalho atual, discute-se principalmente as fases de extração de características e classificação, onde três abordagens foram adotadas. Na primeira, extraiu-se o mesmo número de características descritas no trabalho anterior (RESMINI, 2011), setecentas e doze (712), utilizando um total de vinte e oito (28) imagens (quatro (4) sem patologia e vinte e

quatro (24) com patologia). Esta base denominou-se “base de dados antiga”.

Na segunda abordagem, utilizou-se os dados anteriores acrescidos de seis (6) imagens de voluntárias sem patologia, totalizando dez (10) imagens de voluntárias sem patologia e vinte e quatro (24) imagens de voluntárias com patologia. Esta base foi chamada de “base 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. Estes dados formaram o que chamou-se de “base de subtração”. Com a subtração das características, esperava-se trabalhar com informações relacionadas as duas mamas e não apenas características isoladas de cada mama, como nas outras bases.

Base	Antiga	Nova	Subtração
Total pacientes	28	34	34
Doentes	24	24	24
Saudáveis	4	10	10
Atributos	712	712	356

Tabela 1. Número de imagens e atributos extraídos das bases de dados utilizadas.

Para a classificação utilizou-se duas classes: Doentes ou “com patologia” e Saudáveis ou “sem patologia”. A Tabela 2 mostra o número de pacientes total que constitui essas bases, número de casos com ou sem patologias e o número de características extraídas de cada uma delas.

Utilizou-se três classificadores: IbK (kNN), Naïve Bayes e SVM. Este último, também foi utilizado no trabalho anterior. O algoritmo kNN (do inglês, *k-nearest neighbor*), ou *k* vizinhos mais próximos (KEOGH, 2011), é um algoritmo de aprendizado de máquina supervisionado do tipo *lazy* (AHA, 1991). Algoritmos de classificação ditos supervisionados são aqueles que utilizam a informação do atributo classe (RUSSELL e NORVIG, 2004). Na tarefa de classificação, a mostra, neste caso todas as características extraídas das pacientes, é dividida em duas (2)

fases: a fase de treino e a fase de teste (RUSSELL e NORVIG, 2004). O fato de um algoritmo ser do tipo *lazy* significa que a discretização dos atributos é feita durante a fase de teste enquanto que um algoritmo do tipo *eager* realiza a discretização na fase de treino (HSU, HUANG e WONG, 2011). Em uma dada amostra “A”, o vizinho mais próximo a um dado objeto “o” é o objeto A_i que minimiza a distância (o, A_i), onde a distância é medida seguindo alguma métrica definida (KEOGH, 2011).

O classificador Naïve Bayes é um algoritmo de aprendizado simples, que utiliza Regras de Bayes juntamente com uma forte suposição de que os atributos são condicionalmente independentes, dada a classe. Embora esta suposição de independência seja muitas vezes violada, na prática, muitas vezes o algoritmo fornece a acurácia da classificação competitiva, juntamente com a sua eficiência computacional e muitas outras características desejáveis. Isso leva o Naïve Bayes a ser amplamente utilizado na prática. Naïve Bayes fornece um mecanismo para a utilização da informação em dados de amostra para estimar a probabilidade posterior “P (y | x)” de cada classe de “y”, dado um objeto “x”. Uma vez que temos essas estimativas, pode-se usá-los para classificação ou outras aplicações de suporte à decisão (WEBB, 2011).

O algoritmo linear SVM é apto a executar várias tarefas dentro do contexto de mineração de dados. Em classificação de duas classes, o algoritmo determina um hiper-plano que separa as duas classes de dados com uma margem tão ampla quanto possível. Isso leva a boa generalização da acurácia em dados invisíveis e suporta métodos de otimização especializados através de ajuste dos parâmetros do algoritmo.

Todos esses 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 até determinar-se as melhores configurações, que foram:

1. SVM, utilizou-se um $SVMType = nu-SVC$, $nu = 0,5$ e $kernelType = linear$;

2. IbK, utilizou-se um parâmetro $kNN = 5$ e $distanceWeighting = weight\ by\ 1 - distance$;

3. Naïve Bayes, com configuração padrão do WEKA.

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

Segundo Pereira *et al.* (2011), utilizar a seleção de atributos tem finalidades muito relevantes como: melhoria da acurácia do classificador, aceleração da tarefa de classificação, simplificação do modelo de classificação gerado. Por isso, utilizou-se a seleção de atributos por uma técnica de PCA e uma técnica de “Ganho de Informação” (InfoGainRatio), ambos presentes na ferramenta WEKA. Isto foi feito com intenção de melhorar a acurácia obtida na literatura e em trabalho anterior (RESMINI, 2011).

5. Análise dos Resultados

Para avaliar o desempenho dos experimentos, utilizou-se a área abaixo da curva ROC (do inglês, *Receiver Operating Characteristic*) e a acurácia (ACC). A análise da curva ROC é uma ferramenta útil para avaliar o desempenho de testes diagnósticos em geral. Provê uma visualização gráfica da relação entre sensibilidade e especificidade: **sensibilidade / 1 – especificidade** (ZOU, O’MALLEY, e MAURI, 2007). A Sensibilidade (SENS) ou *recall*, mede a proporção de diagnósticos (no contexto deste trabalho) positivos que são corretamente identificados como tais:

$$SENS = \frac{VP}{VP + FN}$$

A especificidade (ESPEC), mede a proporção de diagnósticos negativos que são corretamente classificados como tais:

$$ESPEC = \frac{VN}{VN + FP}$$

A área sob a curva ROC, AUC (do inglês, *Area Under Curve*), é uma medida importante para demonstrar a relação entre a sensibilidade e especificidade. Uma curva ROC, Figura 5, é um gráfico da sensibilidade no eixo y contra (1-especificidade) no eixo x para diferentes valores de um limiar *l*. A linha diagonal de 45° ligando o ponto (0,0) ao ponto (1,1) é uma curva ROC aleatória (C). A curva ROC ótima é a linha de ligação (0,0) para (0,1) e (0,1) para (1,1) (A). Geralmente, as curvas ROC situam-se entre esses 2 extremos (B). A área sob a curva é uma medida de síntese que essencialmente mede a acurácia médias de diagnóstico em todo o espectro de valores de teste. Analisando a área sob a curva ROC, percebe-se que valores próximos à 1 são melhores do que o caso oposto (ZOU, O'MALLEY, e MAURI, 2007).

Resultados verdadeiros positivos são diagnósticos classificados como com patologia que são realmente de pacientes portadores de alguma doença. Resultados verdadeiros negativos são diagnósticos classificados como sem patologia que realmente são de pacientes saudáveis. Acurácia (ACC) é a proporção de resultados verdadeiros, ou seja:

$$ACC = \frac{VP + VN}{VP + VN + FP + FN}$$

Embora a acurácia do classificador SVM esteja maior que encontrado no trabalho de Schaefer, Zavissek e Nakashima (2009), a área abaixo da curva ROC ficou bastante baixa. Isso se deve à especificidade do método que foi baixa. Um dos motivos para isso acredita-se ser a base não balanceada, 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.

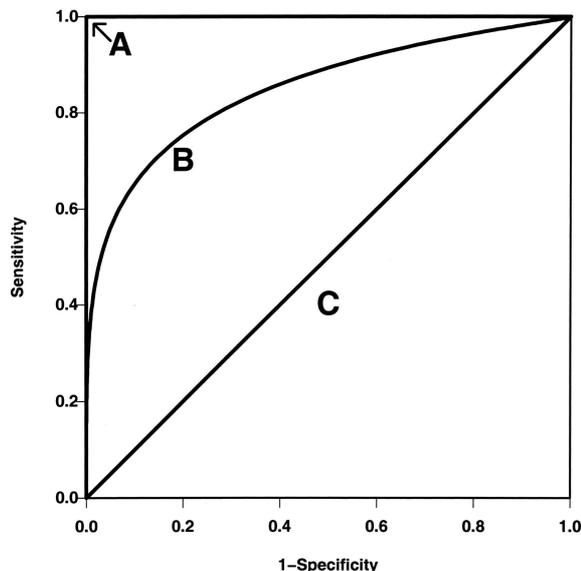


Figura 5. Três curvas ROC hipotéticas: (A) AUC = 1, precisão do diagnóstico ótimo; (B) AUC = 0,85, uma típica curva ROC; e (C) AUC = 0,5, uma linha diagonal que ilustra caso aleatório (ZOU, O'MALLEY e MAURI, 2007).

A Tabela 2, 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.

Base	SVM	IbK	Naïve B.
ACC (%)	82,14	85,71	75,00
ROC (%)	0,580	0,708	0,449

Tabela 2 – Resultados obtidos com a base “antiga”.

A Tabela 3, 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.

Base	SVM	IbK	Naïve B.
ACC (%)	82,35	88,23	73,52
ROC (%)	0,700	0,842	0,750

Tabela 3 – Resultados obtidos com a base “nova”.

A Tabela 4, 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.

Base	SVM	IbK	Naïve B.
ACC (%)	76,47	85,29	79,41
ROC (%)	0,658	0,842	0,750

Tabela 4 – Resultados obtidos com a base “subtração”.

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

A Tabela 5 (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), ou seja, do total de atributos da base, apenas os 10% mais significativos foram considerados pelo classificador.

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 5 (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.

Critério	SVM	IbK	NB	Atrib.
ACC %	88,23	82,35	88,23	10%
ROC %	0,800	0,588	0,857	10%
ACC %	76,47	85,29	85,29	20%
ROC %	0,688	0,867	0,827	20%
ACC %	70,58	85,29	85,29	30%
ROC %	0,617	0,904	0,876	30%

Tabela 5 – Resultados obtidos com a base

“subtração” e 10% dos atributos.

A Tabela 5 (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).

Utilizou-se estas mesmas proporções para calcular o “Ganho de Informação” utilizando a base “nova”, com setecentos e doze atributos. A Tabela 6, 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).

Critério	SVM	IbK	NB	IGR
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

Tabela 6 – Resultados obtidos com a base “nova”.

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.

Critério	ACC (%)	ROC (%)	ACC (%)	ROC (%)
SVM	88,23	0,800	88,23	0,800
IbK	79,41	0,658	67,64	0,592
Naïve Bayes	76,47	0,817	79,41	0,747
Base	Nova	Nova	Sub	Sub
Atrib.	22	22	18	18

Tabela 7 – Resultados obtidos com a base “nova” e “subtração” e PCA.

Por fim, utilizou-se uma técnica de PCA para a seleção de atributos nas bases “nova” e “subtração”. A Tabela 7 mostra esses resultados, onde observa-se que 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 e Trabalhos Futuros

Este trabalho apresenta uma proposta de um método de auxílio ao diagnóstico de patologias das mamas utilizando características extraídas da textura de imagens térmicas. Utilizando-se outros classificadores, além do utilizado em trabalho anterior (RESMINI, 2011), e também outra técnica de seleção de atributos. Para constar, no trabalho anterior citado, a base de dados tinha um total de vinte e oito pacientes, sendo quatro com diagnóstico sem patologia e vinte e quatro com diagnóstico com patologia (tumor maligno ou benigno).

Foram analisados três (3) classificadores: SVM, IbK (kNN) e Naïve Bayes. Todos eles presentes na ferramenta WEKA. Estes classificadores foram aplicados a três (3) bases de dados: uma chamada “antiga”, com vinte e oito (28) imagens (quatro (4) sem patologia e vinte e quatro (24) com patologia), outra base chamada “nova”, com as vinte e oito (28) imagens da base antiga, acrescidas de mais seis (6) sem patologia totalizando trinta e quatro (34) imagens. Uma terceira base de dados com trinta e quatro casos ainda foi considerada.

As duas primeiras bases possuem setecentos e doze (712) atributos, enquanto a última tem trezentos e cinquenta e seis (356) 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 duas abordagens de seleção de atributos: uma usando uma técnica de PCA e outra usando uma técnica de 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% (RESMINI, 2011). Este trabalho foi melhor que os resultados de Wiecek *et al.* (2008) que obteve uma acurácia de 80,95%. Superou trabalho Arora *et al.* (2008), que obteve uma acurácia de 81,80% usando o modo rede neural artificial. Foi também ainda melhor que o trabalho de Schaefer, Zavisek e Nakashima (2009), onde a acurácia foi de 79,53%. Por fim, superou ainda, a acurácia de 88,10% relatada por Acharya *et al.* (2010). Os trabalhos citados não utilizaram a área abaixo da curva ROC (AUC) como medida de avaliação dos métodos propostos. Vale lembrar que a comparação entre métodos diferentes que usam bases também diferentes não é justa para nenhuma das partes, portanto a comparação usada neste trabalho é de caráter ilustrativo.

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.

O mais importante a concluir deste trabalho não é que o classificador SVM obteve um desempenho sutilmente superior aos outros classificadores, mas que com o aumento da base de dados, tornando mais balanceada, mas ainda não ideal (50% para cada diagnóstico), mostrou ser a principal contribuição.

Como limitação do método, pode-se citar

que ele ainda é bastante custoso, pois tem-se que segmentar as imagens, extrair as características, reduzir o número de características e classificar, onde cada fase é feita em um software diferente, com exceção das duas últimas tarefas que são feitas no WEKA, ou seja, o método ainda está bastante manual. O método deve funcionar com imagens de mamografia também, pois já há publicações que utilizam as mesmas técnicas demonstradas neste trabalho.

Na continuação deste trabalho pretende-se aumentar a base (o grupo de pesquisa teve um projeto para aquisição de imagens térmicas aprovado pelo comitê de ética e irá adquirir imagens de pacientes do Hospital Universitário Antonio Pedro - HUAP da UFF em Niterói). Também a melhoria da qualidade das termografia deve contribuir para a melhoria de desempenho do método. Um novo protocolo de aquisição e armazenamento está em estudo para ser usado no HUAP. O aumento de casos com diagnósticos comprovado na base de dados é lento pois é preciso uma continuidade no acompanhamento médico e mesmo alguns diagnósticos apenas podem ser estabelecidos depois de cirurgia e biópsia.

Outro ponto interessante a investigar é quanto as características extraídas. Embora exista muitas publicações relatando o uso de técnicas empregadas com sucesso na mamografia e sendo migradas para a termografia, seria de muito mais valia utilizar características específicas da termografia como a vascularização que em testes prévios do novo protocolo de aquisição demonstrou ser promissor.

7. Agradecimentos

Este trabalho foi desenvolvido com suporte financeiro da CAPES.

Referências

ACHARYA, U.R.; NG, E. Y. K.; TAN, J. H. e SREE, S. V. **Thermography based breast cancer detection using texture features and support vector machine.** Journal of Medical Systems, pp. 01-08, 2010.

ARABI, P. M.; MUTTAN, S. e SUJI, R. J. **Image enhancement for detection of early breast carcinoma by external irradiation.** International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 01-09, 2010.

ARORA, N., MARTINS, D., RUGGERIO, D., TOUSIMIS, E., SWISTEL, A., OSBORNE, M. P. **Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer.** The American Journal of Surgery Vol. 196, pp. 523-526, 2008.

BEZERRA, L. **Uso de imagens termográficas em tumores mamários para validação de simulação computacional.** Dissertação de Mestrado. Departamento de Engenharia Mecânica, Universidade Federal de Pernambuco, 2007.

FLIR SYSTEMS **ThermaCAM TM S45:** Manual do operador. 2004.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P. E WITTEN, I. H. **The WEKA Data Mining Software: An Update.** SIGKDD Explorations, vol. 11, 2009.

HSU, C. N., HUANG, H. J. e WONG, T. T. (2000) *apud* YANG, Y. **Encyclopedia of Machine Learning.** Springer, editores: SAMMUT, C. e WEBB, G. I., p. 288, 2011.

INCA. **Câncer de mama.** Disponível em: <http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama>. Acesso em: 5 jul. 2012.

INCA2. **Estimativa.** Disponível em: <http://www.inca.gov.br/estimativa/2012/index.asp?ID=5>. Acesso em: 3 maio 2012.

INCA3. **Câncer de mama: detecção precoce.** Disponível em: http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama/deteccao_precoce. Acesso em: 5 jul. 2012.

INFRARED MED. Diagnóstico por infravermelho: Como é realizado. Disponível em: http://www.infraredmed.org/exame_como.php. Acesso em: 1 dez. 2011.

KEOGH, E. **Encyclopedia of Machine Learning.** Springer, editors: SAMMUT, C. e WEBB, G. I., p. 714, 2011.

KOAY, J.; HERRY, C. e FRIZE, M. **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, 2004.

KURUGANTI, P. T. e QI, H. **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, 2002.

LEI 11664. **Dispõe sobre a efetivação de ações de saúde que assegurem a prevenção, a detecção, o tratamento e o seguimento dos cânceres do colo uterino e de mama, no âmbito do Sistema Único de Saúde – SUS.** Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/111664.htm. Acesso: 6 jul. 2012.

MOGHBEL, M. e MASHOHOR, S. **A review of computer assisted detection/diagnosis (CAD) in breast thermography for breast cancer detection.** Artificial Intelligence Review, Springer Netherlands, pp. 1-9. Disponível em: <http://dx.doi.org/10.1007/s10462-011-9274-2>. Acesso: 20 nov. 2011.

NG, E. **A review of thermography as promising non-invasive detection modality for breast tumor.** International Journal of Thermal Sciences, vol. 48, pp. 849-859, 2008.

NG, E. e SUDARSHAN, N. **Numerical computation as a tool to aid**

thermographic. Journal of Medical Engineering and Technology, vol. 25, ed. 2, pp. 53-60, 2001.

OLIVEIRA, M. M. **Desenvolvimento de um protocolo e construção de um aparato mecânico para padronizar a aquisição de imagens termográficas da mama.** Dissertação de Mestrado, Departamento de Engenharia Mecânica, Universidade Federal de Pernambuco, 2012.

OMS. **Organização Mundial de Saúde.** Disponível em: http://www.who.int/cancer/publications/world_cancer_report2008/en/index.html. Acesso em: 06 jul. 2012.

PARAMANATHAN, P. e UTHAYAKUMAR, R. **An algorithm for computing the fractal dimension of waveforms.** Applied Mathematics and Computation vol. 195, ed. 2, pp. 598-603, 2008.

PEREIRA, R. P.; PLASTINO, A.; ZADROZNY, B.; MERSCHMANN, L. H. C. e FREITAS, A. A. **Lazy attribute selection: Choosing attributes at classification time.** Intelligent Data Analysis Journal, IOS Press, vol. 15, ed. 5, pp. 715-732, 2011.

PROENG. **Processamento e Análise de Imagens Aplicada à Mastologia.** Disponível em: <http://visual.ic.uff.br/proeng>. Acesso em: 10 dez. 2011.

RESMINI, R. **Análise de Imagens Térmicas da Mama Usando Descritores de Textura.** Dissertação de Mestrado. Instituto de Computação, Universidade Federal Fluminense, 2011.

RUSSEL, S. e NORVIG, P. **Inteligência Artificial.** Elsevier, tradução da 2ª ed., pp. 630 e 640, 2004.

SCHAEFER, G.; ZAVISEK, M. e NAKASHIMA, T. **Thermography based breast cancer analysis using statistical**

features and fuzzy classification. Pattern Recognition, vol. 42, ed. 6, pp. 1133–1137, 2009.

SHEKAR, S. e XIONG, H., Eds. **Encyclopedia of GIS.** Springer, 2008. ISBN 978-0-387-35973-1.

SOBRATERM. **Sociedade Brasileira de Termologia.** Disponível : <http://www.termologia.org>. Acesso em: 10 dez. 2011.

UMADEVI, V.; RAGHAVAN, S. V. e JAIPURKAR, S. **Framework for estimating tumour parameters using thermal imaging.** Indian Journal of Medical Research, Vol. 134, pp. 725-731, 2011.

WEBB, G. I. **Encyclopedia of Machine Learning.** Springer, editors: SAMMUT, C. e WEBB, G. I., p. 713, 2011.

WIECEK, B., STRZELECKI, M., JAKUBOWSKA, T., WYSOCKI, M., PESZYNSKI, C. D. **Medical Infrared Imaging.** CRC Press, editors: DIAKIDES, N. A., BRONZINO, J. D., Cap. 12, pp. 12.1-12.13, 2008.

ZOU, K. H., O'MALLEY, A. J, e MAURI, L. **Statistical Primer for Cardiovascular Research: Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models.** Circulation, Vol. 115, pp. 654-657, 2007.