

Classificação de Tumores e Massas de Mama Utilizando um Comitê de Perceptrons de Múltiplas Camadas

Leandro Augusto da Silva, Emilio del Moral Hernandez e Rangaraj M. Rangayyan

Abstract- This paper addresses a new approach using the committee machine to classify masses in mammograms as benign or malignant. Three shape factors and three measures of edge sharpness were used for the classification of 37 regions of interest (ROIs) related to benign masses and 20 ROIs of malignant tumors. The committee machine is a group of classifiers used to resolve a difficult task. In this work, we used a group of multi-layer perceptrons (MLP) as a committee machine, and the classification results were realized by combining all of the classifiers' responses. The classifier's performance was evaluated by the area under the receiver operating characteristics (ROC) curve (A_z). The A_z result for the committee machine was compared with the A_z results of the MLP and single-layer perceptrons (SLP) neural networks. With the shape features plus edge-sharpness features, the A_z result for the committee machine (0.88) was significantly better than that for the MLP (0.76) and SLP (0.54).

I. INTRODUÇÃO

O câncer de mama é o tipo de câncer que se manifesta com maior frequência entre as mulheres brasileiras [1]. Desde a origem da doença, ainda desconhecida, a melhor maneira para a cura é a rápida detecção, conseguida através do exame de mamografia, o qual é capaz de detectar o tumor antes que ele se torne palpável. A avaliação do exame de mamografia e o diagnóstico são feitos por radiologistas e requerem bastante habilidades.

Para auxiliar os radiologistas na avaliação e no diagnóstico vêm se realizando estudos para o desenvolvimento de sistemas computadorizados. Estes sistemas são conhecidos como CAD (*Computer-Aided Diagnosis*). Os sistemas CAD são desenvolvidos utilizando duas técnicas que se complementam, que são as técnicas de processamento de imagens e as técnicas de inteligência computacional.

As técnicas de processamento de imagens são aplicadas diretamente nos exames de mamografia ou nas regiões suspeitas de câncer, ou regiões de interesse (ROIs, *Regions of Interest*), da mamografia para detectar e extrair características de uma eventual lesão na mama. A lesão pode ser calcificações ou massas (também chamadas de nódulos) que surgem na mama. Com o objetivo de detectar e extrair características das lesões, diferen-

tes técnicas de processamento de imagens vêm sendo pesquisadas. Alguns pesquisadores estudam técnicas para detectar e extrair automaticamente, da mamografia, características de massas na mama [2],[3],[4]. Existem também, pesquisadores que estudam técnicas apenas para detectar automaticamente as microcalcificações [5],[6],[7] ou a massa [6],[8] ou alguma outra anormalidade, como a diferença entre a orientação da textura [9]. Outros pesquisadores desenvolvem técnicas de processamento de imagens para detectar, nas ROIs, marcações manuscritas feitas por radiologistas e então extrair características da massa como forma, nitidez da borda e textura [10],[11],[12]. Ao extrair as características da lesão é necessário o uso de outra técnica para classificar se o caso em estudo trata-se de um tumor maligno ou de uma lesão, por exemplo, uma massa benigna.

Esta outra técnica é a inteligência computacional. As Redes Neurais Artificiais (RNAs) são um exemplo de técnica de inteligência computacional que vem sendo utilizada para a classificação de características das microcalcificações [7],[13] e das características de massas na mama [14],[15],[16],[17],[18]. Diferentes arquiteturas de RNAs como Perceptrons de Múltiplas Camadas (MLP, de *Multi-Layer Perceptrons*) [7],[13],[14], Perceptrons de Camada Única (SLP de *Single-Layer Perceptrons*) [14], Comitê de MLPs [15],[16],[17] e Redes de Base Radial [18], vêm sendo empregado na tentativa de minimizar erros de diagnóstico.

Para este trabalho, as características de forma e de borda extraídas do contorno da massa, previamente marcado manuscritamente pelo radiologista, em uma ROI do mamograma, serão utilizadas [10],[11],[12]. A classificação será feita por um comitê de redes neurais do tipo MLP, inicialmente investigada por Silva et. al. em [15],[16],[17]. Neste trabalho, novos experimentos com o número de iterações do comitê de MLPs e também do MLP e do SLP, esses dois últimos classificadores utilizados para comparação, serão conduzidos. Os resultados dos experimentos de classificação serão apresentados pelo cálculo da área sobre a curva ROC (*Receiving Operating Characteristic*) denominado por (A_z) [19],[20],[21].

Nas próximas seções serão descritas as informações sobre a base de dados, quais características extraídas de massas na mama e como são extraídas, assim como serão abordados os detalhes sobre o classificador comitê de MLPs e brevemente será apresentado o método de avaliação do desempenho do classificador. Por fim, descrevem-se os experimentos realizados e os resultados obtidos.

L. A. Silva - Departamento de Sistemas Eletrônicos, Universidade de São Paulo, São Paulo, Brasil. e-mail: leandro.augusto@ieee.org

E. Del-Moral-Hernandez - Departamento de Sistemas Eletrônicos, Universidade de São Paulo, São Paulo, Brasil. e-mail: emilio_del_moral@ieee.org

R.M. Rangayyan - Departments of Electrical & Computer Engineering and Radiology, University of Calgary, Calgary, AB, Canada T2N1N4 e-mail: ranga@ucalgary.ca

II. EXTRAÇÃO DE CARACTERÍSTICAS DE MASSAS NA MAMA

Foram utilizadas neste trabalho um conjunto com 57 ROIs de mamografias digitalizadas com uma resolução de $50 \mu\text{m}$ com 12 bits por pixel, sendo 37 ROIs diagnosticadas como massas benignas e 20 como tumores malignos [22]. As ROIs aqui descritas consistem de mamografias do programa da província canadense de Alberta, para a detecção precoce do câncer de mama (*Alberta Program for the Early Detection of Breast Cancer* [23]).

O diagnóstico a partir de uma mamografia é realizado por radiologistas que fazem um rastreamento na imagem de mamografia em busca de algum indicador de doença como o câncer. Algumas anormalidades são facilmente perceptíveis na mama, entre outras, destacam-se assimetria entre as mamas, distorção no tecido da mama, aumento na densidade do tecido da mama e a presença significativa de massas (nódulos) e calcificações [6].

Ao perceber uma anormalidade, por exemplo a massa, o radiologista faz uma marcação manuscrita na mesma. A partir dessa ROI que envolve a marcação do radiologista são realizadas investigações de algumas características do contorno feito na mama como borda, densidade e tamanho. Em caso de indicadores malignos, observa-se que a forma da massa é mais irregular do que em massas benignas. A maioria das massas benignas possui o contorno bem circunscrito, compacto, e aproximadamente elíptico (possivelmente lobulado). Por outro lado, as lesões malignas geralmente apresentam bordas borradas, sua aparência é irregular e são cercadas por espículos radiantes e lineares (forma de estrela) [6],[11],[10]. Porém, algumas massas benignas podem ter bordas espiculadas ou bordas borradas e algumas lesões malignas podem ter forma arredondada e bem definida [6],[11],[10]. Com base nestas informações, diversos autores vêm pesquisando métodos computacionais e técnicas de processamento de imagens para a extração destas características [2],[3],[8],[9],[10],[11],[12],[5]. Para este trabalho, serão utilizadas as características pesquisadas por Rangayyan et al. [10],[11] que são as características de forma da borda e da região em torno da borda da marcação manuscrita.

As características de forma extraídas da borda da massa de cada ROI foram denominadas conforme [11]. Essas características são: compactness (C), fractional concavity (F_{cc}) e spiculation index (SI) [11]. Através destas características é possível medir a complexidade da borda da massa. A característica de forma C é uma medida realizada através de uma expressão normalizada, que indica 0 para bordas circulares e aumentada até o valor máximo 1, conforme a borda se torna mais irregular [11]. A característica F_{cc} mede à relação entre o comprimento cumulativo da concavidade da borda e o comprimento total da borda. Bordas arredondas ou ovais, os valores apresentados são baixos [11], por outro lado, as bordas com o contorno espiculado ou lobulado deverão apresentar valores maiores. A característica SI é o cálculo do grau de “espicularidade” da borda, realizado a partir do método proposto por Rangayyan et al. [11], que se trata de um modelo poligonal de borda junto com combinações de comprimento do segmento, larguras de base e ângulos das espicularidade. Estas características são chamadas nas próximas seções de características de forma.

O segundo conjunto de características que será utilizado é o

da nitidez da borda. Este é extraído das regiões ao redor da borda da massa [10],[12]. As características são: Acutance (A), contraste (*contrast*) e coeficiente de variação (CV). Para a extração dessas características foi selecionada uma faixa da região ao redor da borda e a partir desta faixa calculam-se as características da nitidez da borda. O cálculo da característica A é realizado pela computação da variação (descontinuidade) dos tons de cinza da faixa extraída e a partir da identificação dessa descontinuidade são calculados o contraste da borda e a média do CV . Detalhes do cálculo destas medidas podem ser encontrados em Mudigonda et al. [12]. Nas próximas seções essas características são chamadas de características de borda.

As características de forma (C , F_{cc} e SI) e borda da massa (A , *contrast* e CV) extraídas das 57 ROIs de mamografia (37 benignas e 20 malignas) [22] compõem o conjunto de características utilizados no treinamento e teste do comitê de MLPs, do MLP e do SLP. Na seção seguinte será discutido sobre o classificador comitê de MLPs.

III. CLASSIFICADOR COMITÊ DE MLPs

A idéia básica é treinar um comitê de MLPs para resolver uma tarefa computacionalmente complexa. Em uma máquina de comitê, como comumente chamado este tipo de classificador, os resultados individuais de cada especialista, neste trabalho o MLP, são combinados para se conseguir um melhor desempenho de generalização. Este processo é conhecido também como “princípio de dividir e conquistar” [24],[25],[26].

A simplicidade computacional é conseguida por meio da distribuição da tarefa do aprendizado, entre um número de especialistas. A combinação dos resultados determinado por cada especialista, compõe a máquina de comitê. Uma maneira de implementar a máquina de comitê é através do algoritmo AdaBoost, proposto por Schapire [26] (veja também Haykin [24]) e brevemente explicado a seguir.

Dada uma arquitetura de rede neural (RN) usada para implementar os especialistas do comitê, nomeados $RN_1, RN_2, \dots, RN_{ne}, \dots, RN_{NTE}$, (sendo NTE o número máximo de especialistas permitido) e a partir de um dado conjunto de padrões de treinamento $Tr = \{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots, (x_N, y_N)\}$ onde, x_i pertence ao mesmo domínio ou espaço instanciado pelo conjunto de padrões de treinamento X , y_i pertence ao conjunto de classes de saída Y e N é o número de padrões do conjunto de treinamento, o algoritmo AdaBoost apresenta uma hipótese final com base na combinação da hipótese de cada especialista.

O algoritmo é baseado na definição de um conjunto adaptativo de pesos probabilísticos associados com os elementos do conjunto de padrões de treinamento. O particular peso probabilístico distribuído sobre o conjunto de padrões de treinamento em cada iteração ne é chamado $D_{ne}(i)$, $i = 1, 2, \dots, N$.

Inicialmente, para $ne = 1$, todos os pesos probabilísticos são de mesmos valores. Os pesos são atualizados em cada iteração ne com base no sucesso ou falha da hipótese h_{ne-1} determinada pelo especialista RN_{ne-1} .

Quando o especialista RN_{ne} recebe um conjunto de padrões de treinamento ($Tr_{ne \sim D_{ne}}$) com os pesos probabilísticos D_{ne} , este apresenta uma hipótese $h_{ne} : X \rightarrow \{-1, 1\}$. A amostra do conjunto de padrões de treinamento classificada erroneamente

por h_{ne} terá o seu peso probabilístico aumentado enquanto que a amostra classificada corretamente terá o seu peso probabilístico diminuído quando considerado durante o treinamento do novo especialista. Na próxima iteração, o especialista RN_{ne+1} é forçado a focar nas amostras mais “difíceis” (classificações incorretas) do conjunto de padrões de treinamento. Este comportamento adaptativo pode ser observado na Figura 1.

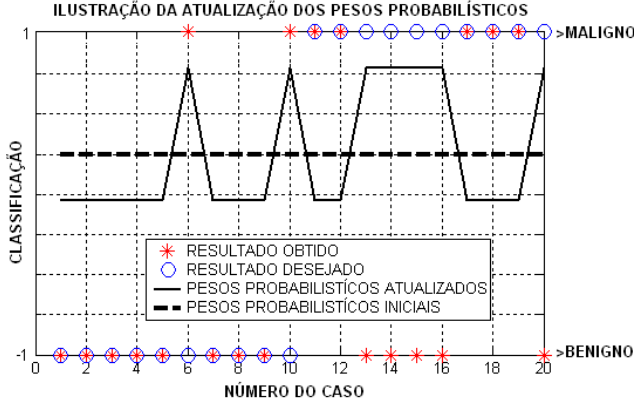


Figura 1. Ilustração dos pesos probabilísticos. Atualização do peso probabilístico do conjunto de padrões de treinamento. Perceba que o peso probabilístico das amostras classificadas incorretamente é aumentado, enquanto que o peso probabilístico das amostras classificadas corretamente é diminuído. Neste exemplo foram usados 20 casos (amostras) (1 a 10 benigno e 11 a 20 maligno). A saída desejada é -1 para caso benigno e 1 para caso maligno. Os valores dos pesos probabilísticos, não pertencem à escala vertical do gráfico, eles foram assim colocados para facilitar a explicação do conceito, pois os valores reais dos pesos probabilísticos são distribuídos de forma que a soma de todos os pesos probabilísticos resulta no valor 1 assim, se colocados na escala vertical do gráfico (-1 a 1) a visualização da atualização dos pesos probabilísticos seria prejudicada.

A qualidade do resultado da hipótese (h_{ne}) é calculada pelo seu erro:

$$\epsilon_{ne} = \sum_{i=1}^N D_{ne}(i) |h_{ne}(x(i)) - y(i)|. \quad (1)$$

onde N é a cardinalidade do conjunto de treinamento.

Uma vez recebida a hipótese h_{ne} do especialista é necessário calcular um parâmetro α_{ne} para mensurar a importância com que h_{ne} foi determinada:

$$\alpha_{ne} = \frac{1}{2} \ln \left[\frac{1 - \epsilon_{ne}}{\epsilon_{ne}} \right]. \quad (2)$$

A próxima distribuição D_{ne+1} é atualizada conforme mostrado na equação 3:

$$D_{ne+1}(i) = \frac{D_{ne}(i)}{Z_{ne}} \begin{cases} e^{-\alpha_{ne}} & \text{se (classificação correta)} \\ e^{\alpha_{ne}} & \text{se (classificação incorreta)} \end{cases} \quad (3)$$

onde Z_{ne} é um fator de normalização (escolhido de modo que D_{ne+1} seja normalizado).

Dessa maneira, o algoritmo AdaBoost aumenta o peso do padrão de treinamento classificado incorretamente por h_{ne} e diminui o peso do padrão de treinamento classificado corretamente (veja a Figura 1). Assim, os pesos tendem a se concentrar nos padrões do conjunto de treinamento mais difíceis.

Esse processo é repetido até que o $\epsilon_{ne} \leq erro_{max}$ ou enquanto ne não atinja o número máximo permitido ($ne \leq NTE$). No final deste processo iterativo, temos um comitê de redes neurais $RN_1, RN_2, \dots, RN_{ne}, \dots, RN_{NTE}$, (veja a Figura 2). A hipótese final $H(x)$ será dada por:

$$H(x) = \text{sign} \left[\left(\sum_{ne=1}^{NTE} \alpha_{ne} h_{ne}(x) \right) \right]. \quad (4)$$

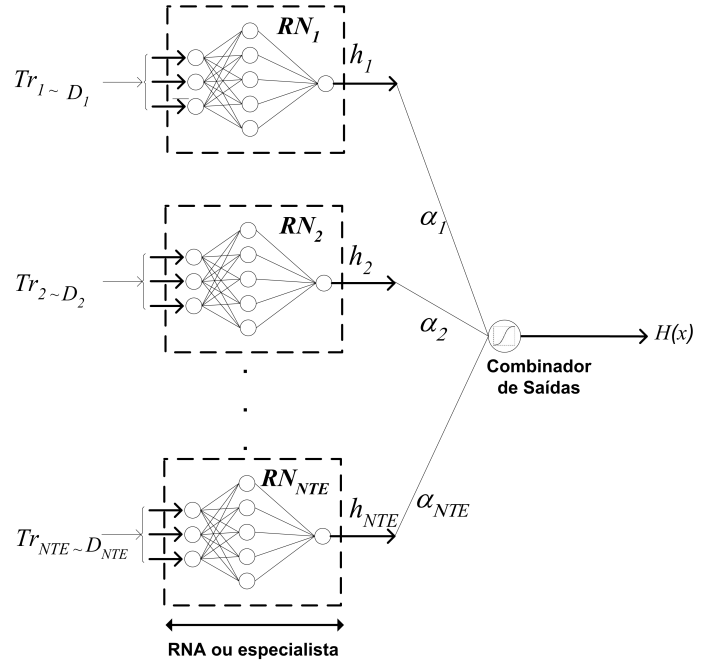


Figura 2. Ilustração do comitê de RN.

A decisão final de classificação será obtida pela discriminação dos pesos que foram obtidos através da resposta de cada RN . O especialista RN aqui usado foi o bem conhecido MLP (veja [24]). Para a análise do desempenho do classificador, foi utilizado o método da área sobre a curva ROC, A_z , descrito na seção seguinte.

IV. ÁREA SOBRE A CURVA ROC (A_z)

Um método de especificar o desempenho de um classificador é pelo cálculo da **taxa verdadeiro-positivo** (TVP), porcentagem de pacientes doentes para o qual o teste resulta positivo e através da **taxa falso-positivo** (TFP), porcentagem de pacientes sem a doença para a qual o teste resulta positivo [19]. Uma vez tendo o classificador treinado, quando um conjunto de pacientes é submetido a um teste de classificação, o resultado é um valor de TVP e um valor de TFP .

Porém, para aplicações de auxílio ao diagnóstico é requerido que o classificador treinado apresente valores diferentes de TVP e TFP [19]. Para isso é necessário que o classificador tenha a possibilidade de variar algum parâmetro. A Figura 3, ilustra possíveis resultados de exames (eixo x), para diversos pacientes examinados (eixo y), diagnosticados em tumor maligno ou massa benigna. Percebe-se neste exemplo (Figura 3) que ao variar o parâmetro *valor de corte*, têm-se diferentes valo-

res de TVP e TFP , ou como serão chamados a seguir, pontos de operação.

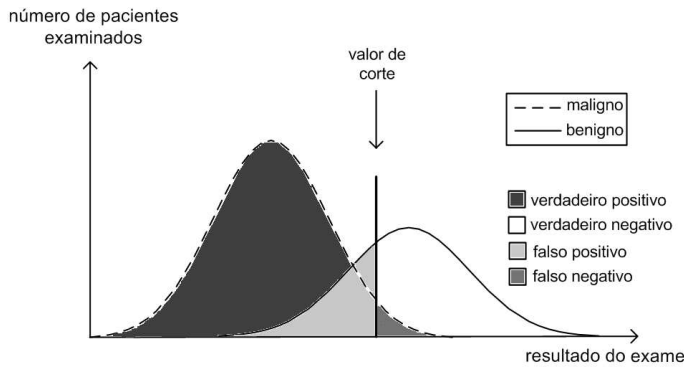


Figura 3. Ilustração de um resultado de pacientes diagnosticados como maligno ou benigno. Se a resposta estiver distribuição normal, como nesta ilustração, o erro na classificações é inevitável, seja manual ou automática. Assim, procura-se encontrar o *valor de corte* que resulte no erro é menor. A cada escolha de *valor de corte*, os quatro casos possíveis de classificação (VP, VN, FP e FN) são medidos.

É possível notar que erros são inevitáveis, e que para cada *valor de corte* escolhido temos quatro possíveis casos de classificação:

Verdadeiro Positivo (VP) o tumor é classificado como sendo câncer (maligno) e comprova ser câncer;

Verdadeiro Negativo (VN) a massa é classificada como não sendo câncer (benigna) e comprova não ser câncer;

Falso Negativo (FN) o tumor é classificado como não sendo câncer (benigno), mas comprova ser câncer (maligno).

Falso Positivo (FP) a massa é classificada como sendo câncer (maligna) e comprova não ser câncer (benigna).

Um método para contrastar a relação entre os resultados da variação do parâmetro *valor de corte* é a curva ROC. A curva ROC é um gráfico de pontos de operação onde é mostrado a relação entre TVP (sensibilidade) versus TFP (1-especificidade). O cálculo da *sensibilidade* e da *especificidade* são mostrados, a seguir, nas equações 5 e 6:

$$sensibilidade = \frac{VP}{VP + FN} \quad (5)$$

$$especificidade = \frac{VN}{VN + FP} \quad (6)$$

Na Figura 4 são apresentados exemplos de curva ROC .

A capacidade de discriminação de um classificador em diagnosticar se um paciente tem lesão maligna ou massa benigna é medida pelo cálculo da área sob a curva ROC, geralmente chamada de A_z [19],[20],[21].

A seguir serão apresentados os resultados de A_z para o comitê de MLPs, assim como serão mostrados os resultados para o MLP e também para o classificador linear SLP.

V. EXPERIMENTOS DE CLASSIFICAÇÃO

Para a realização dos experimentos de classificação foram usados 57 ROIs como descrito na seção anterior. Para garantir que as características extraídas de cada ROI sejam realmente

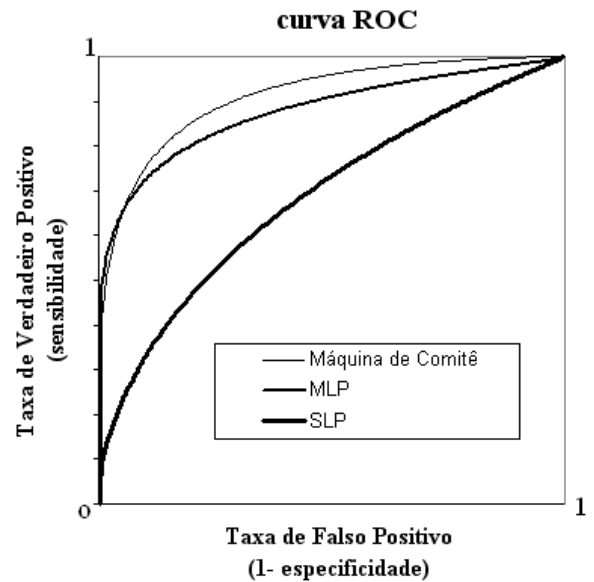


Figura 4. Curva ROC. Ilustração do estudo comparativo dos classificadores: Máquina de Comitê, MLP e SLP. Resultado obtido através das características de borda (A , *contraste* e CV).

pertencentes a tumores malignos e massas benignas, as massas referentes as ROI analisadas foram submetidas ao exame de biópsia. Por esta razão, não se encontra disponível um grande número de bases de dados que seriam úteis para ampliar o número de casos utilizados neste trabalho. Foi encontrado um banco de dados de imagens de mamogramas disponibilizadas por Suckling [27], mas por considerar alguns casos peculiares e futuras comparações com outros resultados obtidos a partir de André e Rangayyan [14] e Alto et al. [22], os experimentos foram realizados unicamente sobre o banco de dados usado no trabalho de Alto et al. [22].

A distribuição do conjunto de padrões (características de forma e de borda) pode ser observada nas Figuras 5 e 6, respectivamente.

AMOSTRAS A SEREM CLASSIFICADAS UTILIZANDO AS CARACTERÍSTICAS DE FORMA

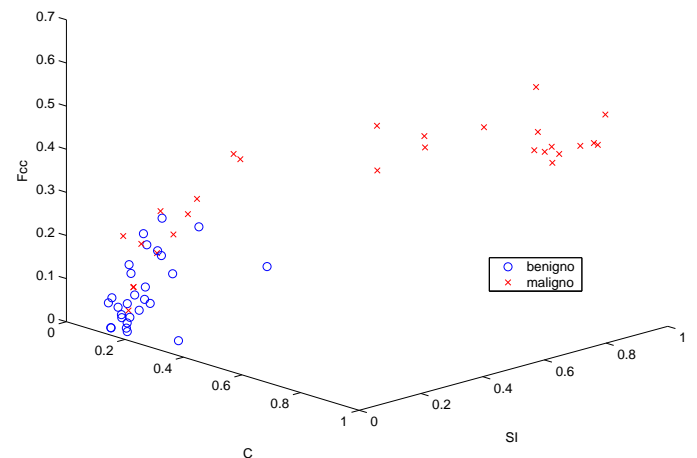


Figura 5. Distribuição no espaço das características de forma (F_{cc} , SI e C).

Os experimentos realizados envolvem diferentes agrupamentos das características e foram submetidos a treinamentos e testes dos classificadores e os valores de A_z obtidos foram contrastados entre si. Com esse objetivo, foram feitos alguns experimentos usando o conjunto de características: separados em forma (F_{cc} , SI , C) e borda (*contraste*, A , CV); somados (F_{cc} , SI , C + *contraste*, A , CV); individuais, apenas de forma (F_{cc} ou SI) ou de borda (*contraste* ou A) e também características individuais de forma somada com de borda (F_{cc} + A ou SI + A). As combinações feitas podem ser observadas na Tabela I.

AMOSTRAS A SEREM CLASSIFICADAS UTILIZANDO AS CARACTERÍSTICAS DE BORDA

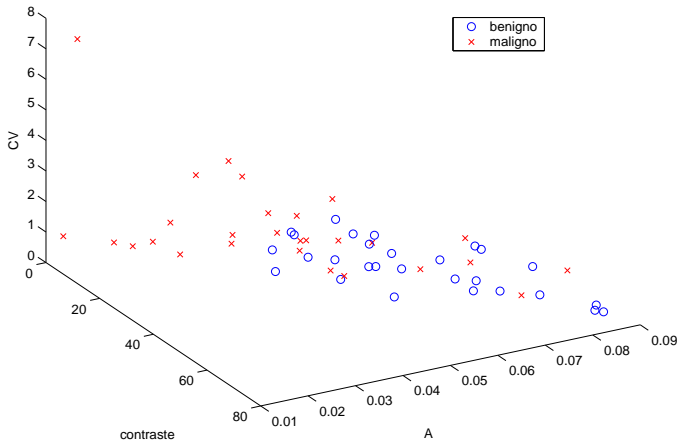


Figura 6. Distribuição no espaço das características de borda (CV , A e *contraste*).

Para cada experimento realizado, foram escolhidos aleatoriamente dois grupos de conjunto de características, onde um grupo foi usado para compor o conjunto de padrões de treinamento e o outro para o conjunto de padrões de teste. Dessa maneira, 20 diferentes grupos de teste foram escolhidos aleatoriamente e o cálculo de A_z apresentado pela média e desvio padrão das 20 escolhas aleatórias. Os resultados de classificação são mostrados na Tabela I.

Estes experimentos já foram tentados por Silva et. al. em [15],[16]. Os experimentos aqui conduzidos diferem na configuração dos classificadores, como o número de iterações possíveis de cada classificador (comitê de MLPs, MLP e SLP), e o número de neurônios da camada escondida do MLP.

Para os experimentos com máquina de comitê, o especialista MLP [15],[16] foi configurado com três neurônios na camada de entrada, cinco neurônios na camada escondida e um neurônio na camada de saída (3:5:1). O treinamento de cada rede em geral, foi finalizado com um erro médio quadrático (RMS) entre a saída desejada e a saída apresentada pelo MLP de aproximadamente 0.3 e nos casos onde não conseguíamos um RMS menor ou igual ao desejado (0,3), o treinamento era parado com 5.000 iterações (critério de parada). O número de MLPs utilizados no comitê quando utilizados as características de borda, foi de 200, e quando usado às características de forma foram utilizados 100 MLPs, mesmo quando conjugado com as características de borda.

Foi repetida a mesma configuração dos especialistas de co-

Características Usadas	Comitê de MLPs	MLP	SLP
	A_z	A_z	A_z
F_{cc} , SI , C	$0,99 \pm 0,01$	$0,99 \pm 0,02$	$0,99 \pm 0,04$
<i>contraste</i> , A , CV	$0,87 \pm 0,06$	$0,84 \pm 0,11$	$0,63 \pm 0,11$
F_{cc} , SI , C + <i>contraste</i> , A , CV	$0,88 \pm 0,07$	$0,76 \pm 0,07$	$0,54 \pm 0,11$
F_{cc}	$0,99 \pm 0,02$	$0,96 \pm 0,04$	$0,95 \pm 0,01$
SI	$0,93 \pm 0,05$	$0,84 \pm 0,07$	$0,75 \pm 0,02$
A	$0,78 \pm 0,07$	$0,70 \pm 0,07$	$0,56 \pm 0,10$
<i>contraste</i>	$0,68 \pm 0,09$	$0,66 \pm 0,10$	$0,58 \pm 0,16$
F_{cc} + A	$0,98 \pm 0,02$	$0,97 \pm 0,04$	$0,99 \pm 0,01$
SI + A	$0,93 \pm 0,04$	$0,90 \pm 0,05$	$0,86 \pm 0,02$

TABELA I

TABELA QUE MOSTRA OS RESULTADOS DE A_z PARA OS DIFERENTES CLASSIFICADORES UTILIZADOS E PARA OS DIFERENTES AGRUPAMENTOS DAS CARACTERÍSTICAS DE FORMA E DE BORDA.

mitê para o MLP utilizado no contraste dos resultados, (3:5:1), porém com um RMS desejado de 0,1 e número máximo de iterações de 10.000. Para a arquitetura SLP, foi feito um treinamento para que um erro mínimo entre a saída desejada e a apresentada pelo SLP fosse menor que 0,3 ou até que a máxima iteração (10.000) fosse atingida.

Quanto ao tempo de treinamento, o comitê de MLPs consumia na ordem de minutos de treinamento, como o MLP. Por outro lado, o SLP gastou segundos no seu treinamento.

Como descrito na seção anterior, para a construção da curva ROC é necessário variar um valor de corte para obter os pontos de operação. No comitê de MLPs, o parâmetro variado de modo a representar uma mudança no valor de corte foi o limiar de ativação da função discriminante do combinador dos MLPs. Para a arquitetura MLP, variou-se o limiar de ativação da função sigmóide do neurônio da camada de saída, e no SLP, variou-se o limiar da função discriminante de saída. Desta maneira, foram obtidos os pontos de operação para a construção das curvas ROC e conseqüentemente foi calculado o A_z de cada classificador.

VI. RESULTADOS E DISCUSSÕES

Com as alterações realizadas no número de iterações e no número de neurônios da camada escondida do MLP usado na comparação, os resultados de A_z são ligeiramente melhores que os apresentados em nossos trabalhos anteriores [15],[16]. A diferença mais significativa aqui apresentada está nos resultados apresentados pelo SLP, pois com um número maior de iterações os resultados das características que deixavam dúvidas de separabilidade linear puderam ser evidenciados.

Verifica-se neste trabalho que para os casos onde o conjunto de padrões aparenta não ter separabilidade linear, identificada pela Figura 6 e pelo resultado de A_z do SLP, o comitê de MLPs apresentou considerável resultado de A_z diante dos outros classificadores contrastados, MLP e SLP.

Com o propósito de melhorar o valor de A_z para as características não separáveis linearmente, pretende-se em trabalhos

futuros, identificar melhores escolhas do valor de corte e também testar outros classificadores como especialistas do comitê.

Uma completa análise estatística dos resultados comparativos apresentados na Tabela I, com teste de hipótese utilizando a distribuição de Student-t, é apresentado por Silva [17].

VII. AGRADECIMENTOS

Nós gostaríamos de agradecer ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pelo apoio financeiro a este projeto, a Escola Politécnica da Universidade de São Paulo. Gostaríamos de agradecer Hilary Alto, do Departamento de Engenharia Elétrica e Computadores da University of Calgary, Calgary, AB, Canada, por nos fornecer a base de dados de características.

REFERÊNCIAS

- [1] INCA, "Instituto nacional do câncer - estimativa 2005: Incidência de câncer no Brasil," Ministério da Saúde. Secretaria Nacional de Assistência à Saúde. 2004.
- [2] B.S. Sahiner, H.P. Chan, N. Petrick, M.A. Helvie, and M.M. Goodsitt, "Computerized characterization of masses on mammograms: The rubber band straightening transform and texture analysis," *Medical Physics*, vol. 25, pp. 516–526. 1998.
- [3] B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and L. M. Hadjiiski, "Improvement of mammographic mass characterization using spiculation measures on morphological features," *Medical Physics*, vol. 28, pp. 1455–1465. July 2001.
- [4] N. R. Mudigonda, R. M. Rangayyan, and J. E. L. Desautels, "Detection of breast masses in mammograms by density slicing and texture flow-field analysis," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 1215–1227. December 2001.
- [5] L. Shen, R. M. Rangayyan, and J. E. L. Desautels, "Application of shape analysis to mammographic calcifications," *IEEE Transactions on Medical Imaging*, vol. 13, pp. 263–274. June 1994.
- [6] E. Claridge and J. H. Richter, "Characterization of mammographic lesions," *Proc. 2nd Int. Workshop on Digital Mammography*, pp. 241–250. July 1994.
- [7] S. Yu and L. Guan, "A cad system for the automatic detection of clustered microcalcifications in digitized mammogram films," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 115–126. February 2000.
- [8] V. T. Santos, *Segmentação de imagens mamográficas para detecção de nódulos em mamas densas*, tese de mestrado apresentada à Universidade de São Paulo, São Carlos, SP, 2002.
- [9] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, and A. F. Frere, "Analysis of asymmetry in mammograms via directional filtering with gabor wavelets," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 953–964. September 2001.
- [10] R. M. Rangayyan, N. M. El-Faramawy, J. E. L. Desautels, and O. A. Alim, "Measures of acutance and shape for classification of breast tumors," *IEEE Transactions on Medical Imaging*, vol. 16, pp. 799–810. 1997.
- [11] R. M. Rangayyan, N. R. Mudigonda, and J. E. L. Desautels, "Boundary modelling and shape analysis methods for classification of mammographic masses," *Medical and Biological Engineering and Computing*, vol. 38, pp. 487–495. May 2000.
- [12] N. R. Mudigonda, R. M. Rangayyan, and J. E. L. Desautels, "Gradient and texture analysis for the classification of mamographic masses," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 1032–1043. October 2000.
- [13] S. Halkiotis, J. Mantas, and T. Botsis, "Computer-aided detection of clustered microcalcifications in digital mammograms," *Special Issue: Proceedings of the fifth European Systems Science Congress*, vol. 2. October 2002.
- [14] T. Andre and R. M. Rangayyan, "Classification of tumors and masses in mammograms using neural networks with shape and texture features," *25th Annual International Conference of the IEEE Transactions on Engineering in Medicine and Biology Society*, pp. 17–21. September 2003.
- [15] L. A. Silva, E. Del-Moral-Hernandez, and R. M. Rangayyan, "Classificação de tumores e massas de mama utilizando máquinas de comitê," *Proceedings of the International Federation for Medical and Biological Engineering: Proceedings of the III Latin American Conference on Biomedical Engineering*, João Pessoa, Paraíba, Brazil, vol. 5, pp. 935–938. Setembro 2004.
- [16] L. A. Silva, E. Del-Moral-Hernandez, and R. M. Rangayyan, "Classification of breast masses using a committee machine of neural networks," *Proceedings of the IASTED International Conference on Biomedical Engineering*, Innsbruck, Austria, pp. 168–173. 16-18 February 2005.
- [17] L. A. Silva, *Classificação de Nódulos Mamográficos utilizando um Comitê de Redes Neurais Artificiais*, Dissertação de Mestrado apresentada à Escola Politécnica da Universidade de São Paulo, São Paulo, SP, 2005.
- [18] R. Espírito-Santo, R. D. Lopes, and R. M. Rangayyan, "Classification of mammographic masses using radial basis functions and simulated annealing with shape, acutance, and texture features," *Proceedings of the IASTED International Conference on Biomedical Engineering*, Innsbruck, Austria, pp. 164–167. 16-18 February 2005.
- [19] K. Woods and K. Bowyer, "Generating roc curves for artificial neural networks," *IEEE Transactions on Medical Imaging*, vol. 13, pp. 329–337. June 1997.
- [20] J. M. DeLeo and G. Campbell, "Receiver operating characteristic (roc) methodology in artificial neural network with biomedical applications," *Proc. World Congress on Neural Networks*, vol. 2, pp. 735–740. 1995.
- [21] J. M. DeLeo, "Measuring classifier intelligence," *Proceedings of the 2002 PerMIS Workshop*, vol. 990, pp. 26–31. September 2002.
- [22] H. Alto, R.M. Rangayyan, and J.E.L. Desautels, "Content-based retrieval and analysis of mammographic masses," to appear in *Journal of Electronic Imaging*. 2005.
- [23] Alberta-Cancer-Board, "Screen test: Alberta program for the early detection of breast cancer," 2099-01 Biennial Report, vol. Edmonton, Alberta, Canada. 2001.
- [24] S. Haykin, *Neural networks: A comprehensive foundation*, Prentice Hall, Upper Saddle River, NJ, 1999.
- [25] R. Duda, P. Hart, and D.G. Stork, *Pattern Classification and Scene Analysis*, John Wiley Professio, Wiley, NY, 2000.
- [26] R. E. Schapire, "The boosting approach to machine learning: An overview," *MSRI Workshop on Nonlinear Estimation and Classification*, pp. 1–23. December 2002.
- [27] J. Suckling, J. Parker, D. R. Dance, J. Hutt S. Astley, C. R. M. Doggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. L. Kok, P. Taylor, D. Betal, , and J. Savage, "The mammographic image analysis society digital mammogram data-base," *Proc. 2nd Int. Workshop Digital Mammography*, pp. 375–378. 1994.