

# Métricas Morfológicas para a Classificação de Tumores de Mama

Lizianne P. Marques Souto<sup>1,2</sup>, Thiago K. L. dos Santos<sup>1</sup>, Marcelino P. dos Santos Silva<sup>1</sup>

<sup>1</sup>Universidade do Estado do Rio Grande do Norte (UERN)

<sup>2</sup>Universidade Federal Rural do Semiárido (UFERSA)

{lizianne.priscilla, thiagokleyton}@gmail.com,

{marcelinopereira}@uern.br

**Abstract.** *In this paper a methodology and a computer-aided diagnosis system for detection of breast cancer were proposed. The approach included the extraction of metrics from tumors in mammograms and the use of a machine learning algorithm to classify them morphologically. Images from MIAS repository were used for the experiments. The results showed the efficacy of the proposed method.*

**Resumo.** *Neste artigo foi proposta uma metodologia e um sistema de diagnóstico auxiliado por computador para a detecção do câncer de mama. A abordagem contemplou a extração de métricas a partir de tumores em mamografias e o uso de um algoritmo de aprendizagem de máquina para classificá-las morfológicamente. Imagens do repositório MIAS foram utilizadas para os experimentos. Os resultados mostraram a eficácia do método proposto.*

## 1. Introdução

O aumento no surgimento de novos casos de câncer de mama tem repercutido como um dos graves problemas de saúde pública no mundo. Países em desenvolvimento, como o Brasil, são os mais afetados com essa doença devido à falta de informação e precarização de parte dos serviços de saúde enfrentados pela população. O câncer de mama também é uma das principais causas de mortes em mulheres com idade entre 39 e 58 anos. No ano de 2010 cerca de 12.705 mulheres tiveram mortes ocasionadas pelo câncer de mama no Brasil [INCA 2012]. A Organização Mundial da Saúde (OMS) projeta para o ano de 2030 uma média de 27 milhões de novos casos de câncer de mama no mundo, o que resultará em 17 milhões de mortes.

A maneira mais eficaz de se combater essa doença ainda é a prevenção secundária, onde é indicada a realização periódica do auto-exame das mamas, do exame clínico realizado pelo profissional de saúde e de exames por imagens, como a mamografia, pois com a descoberta precoce existe a possibilidade da realização de tratamentos que aumentam as chances de sobrevivência do paciente. A mamografia, comparada aos demais exames apresentados, é a forma mais eficaz para a detecção de tumores em fase inicial, a partir de 1 mm, enquanto no auto-exame das mamas só é possível perceber tumores a partir de 1,5 cm [Leite et al. 2011]. Portanto, a mamografia apresenta até 88% de sensibilidade e até 99% de especificidade [INCA 2002]. A sensibilidade é a capacidade de detectar

verdadeiro positivo, ou seja, em casos em que existe a doença, ela é detectada. Já a especificidade está relacionada à detecção de verdadeiro negativo, em outras palavras significa a detecção correta de casos livres de anomalias.

A leitura de mamografias normalmente é realizada por radiologistas. Essa tarefa demanda formação e muita experiência do profissional, que devido a fatores como baixa qualidade da imagem, tamanho e variação morfológica das lesões nem sempre propiciam uma avaliação precisa e uniforme. Sendo assim, estima-se que a sensibilidade dos radiologistas no rastreamento do cancro de mama varia entre 65% e 75% [Skaane et al. 1997].

Uma alternativa viável para auxiliar os radiologistas na leitura de imagens e no diagnóstico precoce do câncer de mama é a utilização de esquemas de diagnóstico auxiliado por computador (*Computer-aided diagnosis - CAD*). Os sistemas CAD diminuem a incerteza do especialista no diagnóstico fornecendo uma segunda opinião do caso. São úteis para evitar distorções na interpretação de lesões, tratamentos incorretos e biópsias cirúrgicas desnecessárias. De modo geral, os sistemas CAD melhoram a qualidade da imagem (e conseqüentemente) a visualização e localização de lesões suspeitas, extraem características a partir de imagens e classificam os achados de acordo com sua probabilidade de malignidade.

Neste contexto, é proposto neste artigo o desenvolvimento de um CAD mamográfico para auxiliar radiologistas no rastreamento do cancro de mama. A metodologia proposta utiliza processamento digital de imagens, extração de atributos para representar a assinatura morfológica de tumores e um método de classificação que permite distingui-los entre benignos e malignos.

## 2. Computer-Aided Diagnosis - CAD

Apesar da mamografia ser considerada o melhor método de detecção precoce do câncer de mama, por motivos como limitações próprias do especialista, cerca de 10% a 30% das lesões mamárias não são identificadas [Calas et al. 2012]. Uma solução para esse problema é a utilização de esquemas de diagnóstico auxiliado por computador (CAD) que provêm uma dupla leitura de imagens, aumentando em até 15% a sensibilidade no rastreamento de mamografia [Thurfjell et al. 1994].

A literatura abrange várias metodologias de esquemas CAD voltados ao diagnóstico do câncer de mama. Dentre elas podemos citar o trabalho de [Todd and Naghdy 2004] que teve como foco a utilização de métodos para a extração de características relacionadas à malignidade de achados em mamografias e seu processo de classificação. Para a identificação das lesões e a segmentação das regiões de interesse, utilizou-se um algoritmo de detecção de margens e o algoritmo *fuzzy region growing*. 43 imagens do repositório DDSM (*Digital Database for Screening Mammography*) foram utilizadas para extrair um subconjunto de 13 atributos baseados nas propriedades (forma e margem) dos achados que foram classificados por uma árvore de decisão binária.

Iniciativa correlata nesta área foi o trabalho realizado por [Surendiran and Vadivel 2012] que teve como objetivo a classificação de mamografias utilizando métricas geométricas (morfológicas) e atributos extraídos da margem das regiões de interesse para a detecção do câncer de mama. Foram extraídos 17 atributos de 940 imagens da base de dados DDSM. Para a identificação das regiões de interesse nas

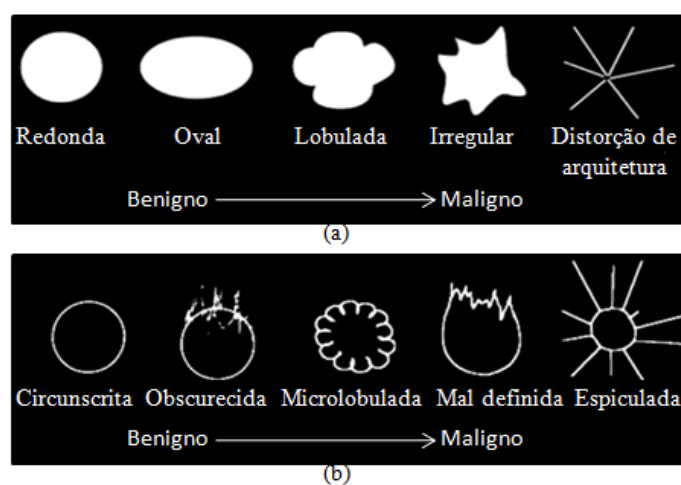
mamografias foi utilizado o algoritmo de Limiarização (*Thresholding*). Na classificação foi utilizado o CART, algoritmo baseado em árvore de regressão.

O trabalho de [Pérez et al. 2012] consistiu na avaliação de métodos de seleção de características para classificação do câncer de mama. Para isso, foram utilizadas 220 imagens do banco de dados BCDR (*Breast Cancer Digital Repository*), das quais foram extraídas características de forma e textura das regiões de interesse. Esses recursos foram combinados com dados clínicos dos pacientes, totalizando 29 atributos divididos em subconjuntos com diferentes combinações de dados. Os subconjuntos foram classificados utilizando *Feed Forward Backpropagation Neural Network* (FFBP), *Support Vector Machine* (SVM) e *Árvore de Decisão* (DTJ48). Os autores obtiveram um subconjunto de características com capacidade de discriminação para classificação do câncer de mama.

Em uma outra proposta, [Junior et al. 2006] desenvolveu uma metodologia para auxiliar no desempenho de esquemas CAD em detectar automaticamente massas em imagens de mamografia. O algoritmo K-means foi utilizado para segmentar 56 imagens obtidas a partir do banco de dados MIAS. Foram extraídas 5 características de textura da região de interesse e 3 métricas geométricas para formar a assinatura das lesões. No total de 33 imagens segmentadas corretamente, 22 imagens foram utilizadas para o treinamento do algoritmo e 11 para a classificação, utilizando o algoritmo de aprendizagem supervisionada, K-NN (*K-Nearest Neighbor*).

### 3. Morfologia de Tumores

Os tumores (benignos ou malignos) são resultantes do crescimento e multiplicação de células de maneira desordenada. Quando maligno, o tumor é considerado cancerígeno, tende a invadir tecidos adjacentes e, por isso, possuem um formato mais irregular e margem indefinida [Alvarenga et al. 2003]. Por outro lado, tumores benignos não cancerígenos, tendem a possuir um formato mais arredondado ou ovalado e margem bem definida (Figura 1). Sendo assim, geralmente os tumores são visualmente distinguíveis pela sua forma e margem na mamografia.



**Figura 1. (a) Morfologia de tumores e (b) Classificação de suas margens**

Fonte: [Kopans 2000]

Quanto ao formato dos tumores, é possível fazer uma análise baseando-se nos

seguintes parâmetros:

- Tamanho: depende do tamanho do tumor, ele pode ser detectado por meio de exames clínicos (palpável). Em casos de tumores muito pequenos eles só podem ser identificados pela mamografia (não palpável);
- Forma: os tumores podem apresentar formato redondo, oval, lobulado, irregular e espiculado. Ao visualizar a Figura 1(a), nota-se que a malignidade está relacionada ao formato irregular do tumor;
- Margem: tumores malignos geralmente possuem contorno irregular ou indefinido;
- Densidade: tumores malignos tendem a apresentar alta densidade.

#### 4. Segmentação de Imagem

Com o objetivo de delimitar a região de interesse (*Region of Interest* - ROI) com a lesão destacada do fundo da imagem, utilizamos o método de segmentação de imagens baseado no crescimento de regiões (*region growing*) [Bins et al. 1996]. O método baseia-se no fundamento de que uma imagem é composta por um conjunto de regiões, que por sua vez são formadas por um conjunto de pixels. Assim, o crescimento é resultado do agrupamento de regiões/pixels vizinhos de acordo com a similaridade (distância euclidiana) de tons de cinza ou textura. A definição dos parâmetros de crescimento está relacionada à natureza do problema e do tipo da imagem a ser trabalhada. No caso das imagens de mamografia deste trabalho, a segmentação foi realizada com base nos níveis de cinza.

Na Figura 2(a) é possível observar o resultado da segmentação da mamografia da Figura 2(b) com o algoritmo *region growing*. Na Figura 2(c) a região de interesse é selecionada, desprezando-se o restante da imagem.

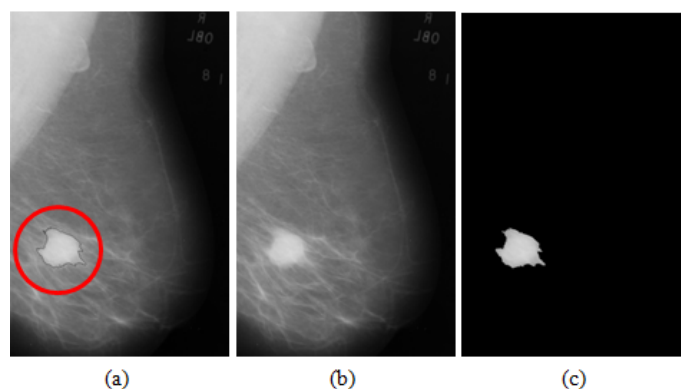


Figura 2. Imagem: (a) segmentada (b) original (c) ROI

#### 5. Extração de Métricas

A extração de métricas consiste na obtenção de características suficientemente relevantes para representar as lesões (ROI's) encontradas nas imagens. Os atributos encontram-se distribuídos nas principais categorias: textura, gradiente e geométrica (morfológica).

De forma intuitiva, textura descreve características da imagem como aspereza, uniformidade e regularidade. De acordo com [Ferreira and Borges 2005], o reconhecimento de características de textura são um desafio pelo fato de não apresentarem um padrão regular e por serem dependentes de escala. Em [Duarte et al. 2013] é citado que

características como textura, pelo fato de não serem morfológicas, são consideradas problemáticas pois dificultam a interpretação das variações de informações internas (pixels) da região pelo olho humano.

O gradiente é ideal para encontrar mudanças nos níveis de cinza e a direção da borda das regiões nas imagens [Gonzalez and Woods 2008]. Os atributos geométricos descrevem as propriedades morfológicas das regiões de interesse como área, perímetro, circularidade, entre outros [Al-Shamlan and El-Zaart 2010]. Baseando-se na alta representatividade das métricas geométricas no diagnóstico médico para identificar qualquer objeto de interesse na imagem [Al-Shamlan and El-Zaart 2010], neste trabalho optou-se por extrair as métricas descritas na Tabela 1.

**Tabela 1. Métricas morfológicas. Fonte: [McGarigal and Marks 1995]**

<b>Métrica</b>	<b>Equação</b>	<b>Descrição</b>	<b>Aplicação</b>
<i>Area (A)</i>		Retorna a área da região. Quando medida em pixels, é igual ao número destes.	Tumores malignos são maiores que os benignos.
<i>Circle</i>	$1 - \frac{A}{\pi(\text{raio}^2)}$	Retorna 0 para regiões circulares e próximo de 1 para regiões lineares.	Tumores benignos possuem formato circular/oval.
<i>Perimeter (P)</i>		Retorna o perímetro da região. Quando medida em pixels, é igual ao número destes na borda da região.	Utilizado para cálculo de métricas que refletem complexidade e compactação de regiões.
<i>Fractal</i>	$2 \frac{\log(0.25 * P)}{\log(A)}$	Índice que mede a complexidade de forma da região.	A forma de tumores benignos é menos complexa que a forma de tumores malignos.
<i>Shape</i>	$\frac{P}{4\sqrt{A}}$	Índice de forma da região. Retorna 1 para regiões mais compactas e aumenta de acordo com a irregularidade.	Tumores benignos apresentam forma mais compacta que os malignos.
<i>Perimeter-Area</i>	$\frac{P}{A}$	Razão entre o perímetro e a área de uma região, que é um indicador da complexidade do formato da região.	Tumores malignos possuem forma irregular.

## 6. Classificação e Árvore de Decisão

O objetivo da classificação, no contexto de processamento de imagens, é identificar características ou padrões das estruturas encontradas em uma imagem e atribuí-las a determinadas classes [Solomon and Breckon 2011]. Em casos que se tem conhecimento prévio sobre os atributos e suas respectivas classes, a classificação é supervisionada, caso contrário, a classificação é não supervisionada e os atributos são agrupados por critérios de similaridades formando assim os *clusters*. Neste contexto, um método de aprendizagem supervisionada para realizar inferência indutiva muito utilizado no diagnóstico de

casos médicos é a árvore de decisão [Mitchell 1997]. Seu funcionamento consiste em aplicar uma estratégia para dividir recursivamente um problema complexo em subproblemas e, em seguida, combinar as soluções encontradas em formato de árvore para produzir uma solução final do problema [Faceli et al. 2011].

Formalmente, uma árvore de decisão pode ser definida como um grafo acíclico direcionado composto pelo nó raiz (inicial), pelos nós de divisão e nós folha. Os nós de divisão contêm um teste para os possíveis valores que o atributo possa assumir. Estes nós possuem arestas para outros nós de divisão ou para os nós folha, os quais retornam uma resposta (rótulo). Os atributos dos nós de divisão são escolhidos de maneira que eles sejam suficientemente representativos para aumentar o ganho de informação ao término do processo de aprendizagem. Com isso, faz-se necessária a utilização de uma medida da Teoria da Informação chamada Entropia, para medir quão informativo um nó é [Mitchell 1997]. Dessa forma, os atributos mais representativos são escolhidos pelo algoritmo com o objetivo de minimizar a entropia (dificuldade de prever) dos dados e o tamanho da árvore.

Existem inúmeros algoritmos de árvore de decisão, porém um dos mais bem-sucedidos e representativos é C4.5 [Faceli et al. 2011]. Proposto por [Quinlan 1993], o C4.5 é uma extensão do algoritmo ID3 (*Iterative Dichotomiser 3*) e possui uma versão Java o J48. Esse algoritmo foi utilizado em uma outra abordagem por [Silva et al. 2005], que conseguiu distinguir diferentes padrões espaciais presentes em imagens de sensoramento remoto.

## 7. Descrição da Metodologia e do CAD

A metodologia desenvolvida neste trabalho, implementada em um sistema CAD mamográfico, é dividida em duas etapas: treinamento e classificação. Na etapa de treinamento são realizadas as seguintes tarefas (Figura 3):

- Seleção: a mamografia que será submetida ao processo é selecionada no banco de imagens. Na etapa de treinamento as imagens selecionadas já possuem um diagnóstico;
- Segmentação: o algoritmo de processamento digital de imagem, *region growing*, é aplicado para segmentar a região de interesse na imagem;
- Extração de métricas: da ROI são calculadas métricas como área, perímetro, perímetro-área, shape, fractal e circularidade;
- Criação do modelo: métricas extraídas de um conjunto de imagens e seus respectivos diagnósticos são utilizados como entrada pelo algoritmo J48;
- Modelo: é a árvore de decisão gerada na etapa de treinamento pelo J48, a qual caracteriza os tipos de lesões em mamografias que passaram pela etapa de treinamento. O modelo é utilizado na etapa de classificação.

A etapa de classificação é semelhante à etapa de treinamento (Figura 3):

- Seleção: a mamografia que passará pelo processo de classificação é selecionada. Nesta etapa a mamografia não possui diagnóstico;
- Segmentação: o objetivo desta tarefa é a identificação da ROI e segmentação da mesma;
- Extração de métricas: são obtidos atributos (shape, circle, fractal, etc.) da ROI;
- Classificação: as métricas extraídas são submetidas ao modelo (árvore de decisão) gerado na etapa anterior (treinamento) para que este processo de classificação indique o diagnóstico da lesão;

- ROI classificada: tem-se como resultado dessa etapa a região de interesse classificada como benigna ou maligna. Esta informação pode auxiliar o médico fornecendo uma segunda opinião sobre o caso ou aumentando a certeza no diagnóstico.

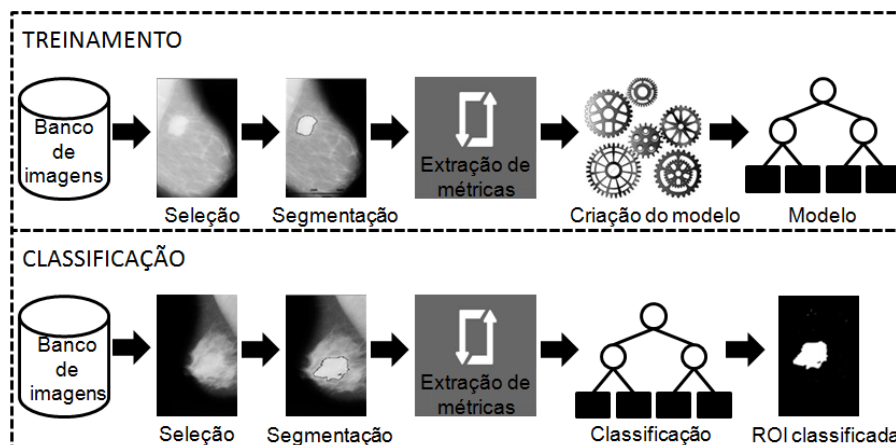


Figura 3. Visão geral da metodologia

## 8. Resultados e Discussão

Os experimentos desse trabalho foram realizados com imagens obtidas a partir do banco de dados MIAS (*Mammographic Image Analysis Society*) no site: <http://peipa.essex.ac.uk/info/mias.html>. As 322 mamografias disponibilizadas encontram-se subdivididas em categorias de acordo com o diagnóstico apresentado: normal, benigno ou maligno. Para as imagens que apresentam tumor, é fornecida uma referência de localização (raio do tumor) e a posição da mama (esquerda/direita). Também é descrito o tipo de tecido da mama e a classe de anomalia da lesão (massa bem definida ou mal definida, calcificação, massa espiculada, distorção de arquitetura, etc).

Sendo assim, para os testes do CAD foram selecionadas 50 mamografias apresentando massas (tumores). Dentre elas, foram utilizadas para o treinamento 10 mamografias com tumores benignos e 10 com tumores malignos. Para a etapa de classificação foram 10 mamografias com tumores benignos e 20 com tumores malignos. O número maior de mamografias com tumores malignos na classificação deve-se à necessidade de uma amostra maior para adequada verificação da sensibilidade do método.

A identificação da ROI se deu de maneira semi-automática, ou seja, a região suspeita é indicada na mamografia e o algoritmo realiza a segmentação. Com isso, pôde-se observar que o *region growing* apresentou desempenho inferior na segmentação de mamografias que apresentam tumores sobrepostos em tecido glandular denso, visto que esse tecido mamário tende a obscurecer a visibilidade e a mascarar as lesões. Tumores malignos com formato extremamente irregular e limite de massa indistinto também contribuíram para a diminuição do desempenho do algoritmo, pelo fato dos tons de cinza dos pixels da ROI serem muito similares aos tons de cinza dos pixels das regiões vizinhas. Nesses casos trabalhou-se diminuindo o parâmetro de variação dos tons de cinza (distância euclidiana entre 2 e 7). Em casos de tumores benignos, que apresentam formato bem definido e normalmente se destacam do restante da imagem, não houve dificuldade na segmentação e trabalhou-se com distância euclidiana maior, entre 7 e 12.

Após a segmentação, as métricas extraídas serviram como entrada do algoritmo J48 para criar o modelo de treinamento e, em seguida, realizar a classificação das regiões. Com base na árvore de decisão gerada (Figura 4), as métricas *circle* e *perimeter-area* foram fundamentais na separação dos atributos em classes, para a representação das lesões e, conseqüentemente, no diagnóstico dos tumores de mama.

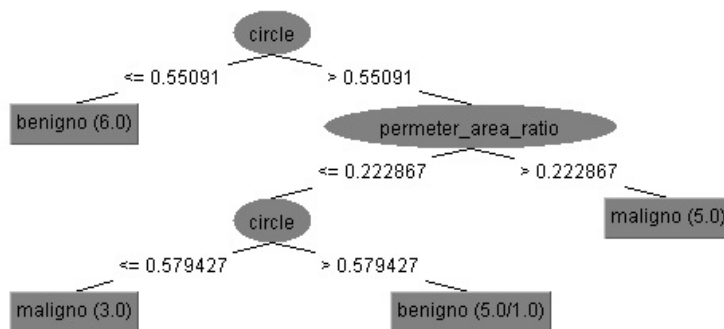
Utilizamos as seguintes medidas estatísticas para quantificar o desempenho da nossa abordagem (o termo benigno foi considerado como “negativo” e o maligno como “positivo”): verdadeiro positivo (VP); falso positivo (FP); verdadeiro negativo (VN); falso negativo (FN). O termo VP refere-se a casos em que o tumor é maligno e a mamografia é classificada corretamente. No caso de FP o tumor é benigno, porém é classificado erroneamente como maligno. Em VN o tumor é benigno e foi classificado corretamente, e em FN o tumor maligno é classificado como benigno. Os resultados foram calculados utilizando as seguintes equações:

- Sensibilidade =  $\frac{VP}{(VP+FN)}$
- Especificidade =  $\frac{VN}{(VN+FP)}$
- Acurácia =  $\frac{(VP+VN)}{(VP+VN+FP+FN)}$

Os resultados obtidos na classificação encontram-se na Tabela 2, onde é possível observar que o CAD apresentou 75% de sensibilidade, 80% de especificidade e 76% de acurácia. Com base nestes resultados, acreditamos que o principal fator que impede uma melhor taxa de sensibilidade é a dificuldade na segmentação da ROI pela complexidade do caso e baixa qualidade da imagem.

**Tabela 2. Resultados da classificação.**

Tipo do tumor	Classificado como maligno	Classificado como benigno	Total	Taxa de Classificação
<b>Maligno</b>	15 (VP)	5 (FN)	20	75% (sensibilidade)
<b>Benigno</b>	2 (FP)	8 (VN)	10	80% (especificidade)
<b>Total</b>	-	-	30	76% (acurácia)



**Figura 4. Árvore de decisão (modelo)**



## 9. Conclusão

A mamografia é considerada um dos métodos mais eficazes para a descoberta precoce do câncer de mama. No entanto, sua leitura e interpretação é um processo muito complexo que demanda muita atenção e ampla experiência do profissional. Neste artigo foi proposto um sistema de diagnóstico auxiliado por computador para auxiliar radiologistas, fornecendo uma segunda opinião sobre o diagnóstico de lesões em mamografias. Por meio do método proposto, 6 métricas morfológicas foram extraídas das ROI's, possibilitando a classificação correta de 76% das mamografias, comprovando a representatividade das métricas utilizadas e a eficácia do CAD.

Com o intuito de melhorar os resultados apresentados, pretende-se incluir como trabalho futuro a aplicação de filtros para melhorar o contraste das imagens, pois sua qualidade influencia no desempenho do algoritmo de segmentação e consequentemente na classificação. Também pretende-se aumentar o número de imagens para a realização dos testes e ampliar o conjunto de métricas geométricas para melhor representar a morfologia dos tumores, objetivando o aumento da acurácia do sistema e a diminuição da taxa de falso negativo.

## Agradecimentos

Os autores agradecem à CAPES, ao CNPq e à FAPERN pelo financiamento ao trabalho, bem como ao Centro de Oncologia e Hematologia de Mossoró pelo apoio técnico-científico à pesquisa.

## Referências

- Al-Shamlan, H. and El-Zaart, A. (2010). Feature extraction values for breast cancer mammography images. *International Conference on Bioinformatics and Biomedical Technology*, pages 335–340.
- Alvarenga, A. V., Infantosi, A. F. C., de Azevedo, C. M., and de Albuquerque Pereira, W. C. (2003). Aplicação de operadores morfológicos na segmentação e determinação do contorno de tumores de mama em imagens por ultra-som. *Revista Brasileira de Engenharia Biomédica*, 19(2):91–101.
- Bins, L., Fonseca, L., and Erthal, G. S. (1996). Satellite imagery segmentation: a region growing approach. *In: VIII Brazilian Symposium on Remote Sensing*, pages 677–680.
- Calas, M. J. G., Gutflen, B., and de A. Pereira, W. C. (2012). Cad and mammography: why use this tool? *Colégio Brasileiro de Radiologia e Diagnóstico por Imagem*, 1:46–52.
- Duarte, Y. A. S., do Nascimento, M. Z., and Oliveiras, D. L. L. (2013). Algoritmo de extração de textura baseado em wavelet e clbp para classificação de lesões em mamogramas. *Workshop de Informática Médica*.
- Faceli, K., Lorena, A. C., Gama, J., and de Carvalho, A. C. P. L. F. (2011). *Inteligência artificial: uma abordagem de aprendizado de máquina*, volume 1. LTC, Rio de Janeiro, 1 edition.

- Ferreira, C. B. R. and Borges, D. L. (2005). Uma estratégia de seleção de um subconjunto mínimo de características wavelet em uma abordagem multirresolução para classificação de tumores em mamogramas. *V Workshop de Informática Médica*, 1.
- Gonzalez, R. C. and Woods, R. E. (2008). *Digital Image Processing*. Prentice Hall, 5 edition.
- INCA (2002). Prevenção e controle de câncer. *Revista Brasileira de Cancerologia*, pages 317–332.
- INCA (2012). Câncer de mama.
- Junior, G. B., Silva, A. C., de Paiva, A. C., and de Oliveira, A. C. M. (2006). Identificação de massas em mamografias usando textura, geometria e algoritmos de agrupamento e classificação. *VI Workshop de Informática Médica - WIM*, pages 94–104.
- Kopans, D. (2000). *Imagem da mama*. Rio de Janeiro: MEDSI Editora Médica e Científica Ltda, 2 edition.
- Leite, G. C., Leite, J. S. S., Meneses, F. G. A., Santos, D. A., and Silva, J. S. (2011). A utilização de técnicas de limiarização para auxílio no diagnóstico de câncer de mama. *III Encontro Nacional De Engenharia Biomecânica*.
- McGarigal, K. and Marks, B. (1995). Fragstats: spatial pattern analysis program for quantifying landscape structure. *USDA Forestry Service Technical Report PNW*, page 134.
- Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill, New York. United States of America.
- Pérez, N., Guevara, M. A., and Silva, A. (2012). Evaluation of feature selection methods for breast cancer classifications. *15th International Conference on Experimental Mechanics*, page 10.
- Quinlan, R. (1993). *C4.5: programs for machine learning*. CA: Morgan Kaufmann Publishers.
- Silva, M. P. S., Câmara, G., Souza, R. C. M., Valeriano, D. M., and Escada, M. I. S. (2005). Mining patterns of change in remote sensing image databases. *Fifth IEEE International Conference on Data Mining*, pages 362–369.
- Skaane, P., Engedal, K., and Skjennald, A. (1997). Interobserver variation in the interpretation of breast imaging. *Acta Radiol*, pages 497–502.
- Solomon, C. and Breckon, T. (2011). *Fundamentals of digital image processing: a practical approach with examples in matlab*. Wiley-Blackwell.
- Surendiran, B. and Vadivel, A. (2012). Mammogram mass classification using various geometric shape and margin features for early detection of breast cancer. *Int. J. Medical Engineering and Informatics*, 4(1):36–54.
- Thurfjell, E., Lernevall, K., and Taube, A. (1994). Benefit of independent double reading in a populationbased mammography screening program. *Radiology*, pages 241–244.
- Todd, A. C. and Naghdy, G. (2004). Method for breast cancer classification based solely on morphological descriptors. *Medical Imaging*, pages 857–867.