

## SEGMENTAÇÃO E CARACTERIZAÇÃO DE NÓDULOS PULMONARES

L. M. Nunes\*, A. C. Patrocínio\*, M. L. N. Franco\*, A. P. P. Froner\*\* e A. M. M. da Silva \*\*

\*Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia-MG, Brasil

\*\* Programa de Pós-graduação em Engenharia Elétrica PUC-RS, Porto Alegre-RS, Brasil

e-mail: laramarins.nunes@gmail.com

**Resumo:** Segundo dados do INCA, a neoplasia pulmonar é o tipo de tumor maligno mais incidente no mundo e apresenta sobrevida curta após o diagnóstico. Na maioria das vezes, o diagnóstico deste tipo de câncer é obtido em estágios avançados da doença, o que diminui a chance de cura. Com isso, estudos estão sendo desenvolvidos em busca de técnicas para detecção precoce de neoplasia pulmonar. Este trabalho apresenta um algoritmo automático para segmentação de nódulos após a seleção de uma região de interesse em imagens obtidas em exames de TC de pulmão. A partir da imagem segmentada, foram calculados os valores de alguns descritores de imagens, os quais foram relacionados com o processo de caracterização de lesões. Além disso, foi feita uma avaliação da significância dos atributos em separar diferentes tipos de lesões. As medidas extraídas foram colocadas em uma tabela e utilizadas na construção de curvas de distribuições Gaussianas das mesmas, as quais permitiram comparação entre valores obtidos para cada lesão. O estudo das curvas obtidas indica que alguns dos descritores utilizados podem proporcionar bons resultados na caracterização de lesões. O poder de separabilidade de cada atributo foi avaliado entre três grupos de lesões: carcinoma, adenocarcinoma e lesões benignas. A caracterização de lesões através de técnicas de processamento de imagens pode trazer grandes benefícios relacionados ao diagnóstico precoce de câncer de pulmão, já que pode representar grande avanço no desenvolvimento de uma técnica não invasiva para definição do diagnóstico da doença.

**Palavras-chave:** Nódulos pulmonares, atributos, caracterização, descritores de imagem, segmentação.

**Abstract:** According to data from INCA, lung cancer is the most frequent type of malignant tumor in the world and presents short survival after diagnosis. Most often, the diagnosis of this cancer is obtained in advanced stages of the disease, which decreases the chance of cure. Thus, studies are being developed in search of techniques for early detection of lung cancer. This paper presents an automatic algorithm for segmentation of nodules after selecting a region of interest in images obtained at CT of the lung. From the segmented image, the values of some descriptors of images, which are related to the characterization of the lesions were determined. In addition, an assessment of the significance of attributes to separate different types of lesions was made. The measures taken were placed on a table and used in the construction of Gaussian's curves distributions of them, which allowed comparison between values obtained for each lesion. The study of curves obtained indicates that some of the descriptors can

provide good results in the characterization of lesions. The power of separability of each attribute was evaluated among three groups of lesions: carcinoma, adenocarcinoma and benign lesions. The characterization of injuries using techniques of image processing can bring significant benefits related to early diagnosis of lung cancer, since it can represent a great advance in the development of a noninvasive technique for diagnosis of the disease.

**Keywords:** lung nodules, attributes, characterization, image descriptors, segmentation.

### Introdução

Segundo o Instituto Nacional de Câncer (INCA), o câncer de pulmão é o tipo de tumor maligno com maior incidência no mundo desde a década de 1920, apresentando aumento de 2% ao ano. No Brasil, o número de mortes devido à neoplasia pulmonar foi de 21.867 no ano de 2010 [1]. Dados do *American Cancer Society* apontam que, devido ao diagnóstico em estágios avançados, a sobrevida de cinco anos após diagnóstico deste tipo de câncer é de aproximadamente 16%. O diagnóstico e tratamento da doença durante fase inicial pode fazer com que essa taxa chegue a 53% [2].

Em 2004, S.J. Swensen, mais da metade dos nódulos submetidos à biópsia ou procedimentos cirúrgicos são identificados como tumores benignos [3]. Vários estudos estão sendo feitos em busca de técnicas para detecção precoce da neoplasia pulmonar maligna [4], [5], [6], e, embora o uso de procedimentos não-invasivos como imagens por Tomografia Computadorizada (TC) tenha aumentado consideravelmente, técnicas invasivas como punção ou biópsia ainda são necessárias para definição do diagnóstico [4]. Em alguns casos, as alterações dos tecidos pulmonares são imperceptíveis aos olhos humanos, dificultando o processo de diferenciação entre nódulos malignos e benignos [5]. Além disso, a interpretação das imagens adquiridas depende da habilidade dos radiologistas [6].

Diante deste cenário, vários sistemas baseados em técnicas de processamento de imagens digitais e outras ferramentas computacionais estão sendo desenvolvidos para auxiliar no diagnóstico de câncer de pulmão. Esses sistemas, chamados computer-aided diagnosis (CAD), realizam medidas relacionadas à textura, forma, crescimento de volume ou área, irrigação e localização dos nódulos detectados e relacionam-nas com malignidade ou benignidade do tumor.

Este trabalho tem como objetivo a caracterização de nódulos de pulmões através de extração de atributos a partir de imagens segmentadas por algoritmo de segmentação automática.

### Materiais e métodos

**Imagens utilizadas** – As imagens utilizadas foram adquiridas no Hospital São Lucas da PUCRS, localizado na cidade de Porto Alegre - RS. O tomógrafo usado é um *multislice* com 16 canais da marca *Siemens*, modelo *Emotion 16*.

No total foram analisados exames de TC de 27 pacientes de ambos os sexos e com idade entre 18 e 81 anos. O que totalizou 498 fatias contendo partes de nódulos.

As imagens apresentam janelamento com nível -600 HU e largura da janela 1200 HU.

Além das imagens dos exames de TC, foram fornecidos laudos anatomopatológicos e radiológicos de alguns dos casos estudados e, ainda, uma tabela contendo medidas como área, perímetro e média de intensidade da imagem. Estas medidas foram extraídas pelo software do tomógrafo utilizado considerando-se imagens de 15 pacientes após segmentação manual das mesmas realizadas por profissional da saúde. A tabela foi utilizada para comparações das mesmas medidas obtidas a partir do processamento desenvolvido neste trabalho.

**Segmentação** – A segmentação das imagens foi feita considerando as seguintes etapas:

- Determinação das fatias de interesse - Inicialmente, foi feita uma análise visual de todas as fatias do exame, determinando quais fatias apresentavam parte da lesão pulmonar (fatias de interesse).
- Seleção da ROI – A região de interesse (ROI) selecionada para cada exame corresponde a uma área quadrada que abrange toda a lesão em todas as fatias de interesse. Sendo assim, são consideradas as fatias contendo a maior área de cada lesão para a seleção manual da ROI e, em seguida, a ROI selecionada foi aplicada ao restante das fatias de interesse pelo algoritmo.
- Thresholding*- Após a definição da região de interesse, foi gerada uma nova sequência de imagens, as quais eram compostas por  $n$  imagens  $m \times m$ , sendo  $n$  o número de fatias de interesse e  $m \times m$  as dimensões das ROI, tanto  $n$  (número de fatias) quanto  $m$  (dimensão da ROI), variam de paciente para paciente. Em seguida, foi aplicado o método de Otsu [7] para a binarização da nova sequência de imagens.
- Operação morfológica de abertura da imagem - Com as imagens binarizadas, foi aplicada operação morfológica de abertura [7], a fim de eliminar objetos menores que o elemento estruturante utilizado, pois assim eliminou-se partes segmentadas na ROI que não pertenciam ao nódulo. Para todas as imagens, foi utilizado elemento estruturante em forma de disco, entretanto a determinação do tamanho desse elemento

estruturante foi feita a partir de testes para cada exame, separadamente.

**Extração de atributos** – A partir da imagem binarizada, foram calculados os 7 momentos invariantes de Hu [8], a área e perímetro das lesões em cada uma das fatias de interesse. As medidas de área e perímetro foram relacionadas através de medidas de compactidade [9] e irregularidade [9] segundo as equações a seguir:

$$C = \frac{P}{4\pi A} \quad (1)$$

$$I = \frac{A}{P^2} \quad (2)$$

Considerando as imagens com fundo preto e objeto (lesão) em tons de cinza, foram calculadas as matrizes de co-ocorrência com ângulos de  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ . Posteriormente, foi calculada a homogeneidade das lesões, a qual foi dada pela média dos valores encontrados utilizando cada uma das matrizes de co-ocorrência geradas [7].

Por fim, foi extraída a intensidade de pixel média das lesões, desconsiderando, para isto, os pixels de fundo das imagens.

**Testes realizados** – As medidas de área, perímetro e volume extraídas foram tabeladas, a fim de facilitar a comparação entre valores obtidos pelo algoritmo desenvolvido e valores obtidos diretamente do tomógrafo, considerando-se a segmentação manual dos nódulos.

Para análise da eficácia dos atributos extraídos na caracterização de lesões, foi gerada uma tabela para cada paciente com as medidas de todos os atributos citados para cada uma das fatias de interesse. A partir desses valores, foram construídas as curvas de distribuições Gaussianas para cada atributo, considerando o valor da fatia com maior área da lesão e a média dos valores de todas as fatias, facilitando a visualização de diferenças significativas entre os vários tipos de lesão.

Com os resultados dos laudos anatomopatológicos dos exames analisados, as imagens foram divididas em três grupos distintos: adenocarcinoma, carcinoma e lesões benignas. Para cada grupo, foram calculadas a média de cada atributo extraído das imagens e a curva de distribuição Gaussiana a fim de visualizar as diferenças entre as lesões a partir dos atributos.

### Resultados

Considerando a fatia com maior área da lesão de cada exame, foi definida a ROI das imagens. A fatia original que contém a maior área do nódulo para um dos pacientes e o resultado do algoritmo de segmentação desenvolvido neste trabalho estão representados nas Figuras 1 e 2, respectivamente.

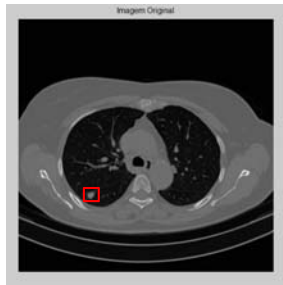


Figura 1: Imagem original da fatia com maior área da lesão de um dos pacientes com região do nódulo destacada em vermelho.

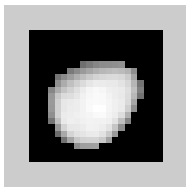


Figura 2: Resultado da segmentação do nódulo representado Figura 1.

A área e o perímetro de cada nódulo na fatia com maior área da lesão foram calculados a partir do resultado do algoritmo de segmentação.

A Tabela 1 mostra os valores de área calculados de forma semiautomática pelo tomógrafo (considerou-se a segmentação manual feita por profissional da saúde) e os valores calculados pelo algoritmo desenvolvido, as medidas no tomógrafo foram feitas para 15 pacientes. As medidas de área utilizadas estão em mm<sup>2</sup> e a diferença percentual dos resultados foi calculada a fim de facilitar a comparação e determinar o erro.

Tabela 1: Comparação entre áreas extraídas pelo tomógrafo e pelo algoritmo desenvolvido neste trabalho

Paciente	Área (Tomógrafo)	Área (Otsu/Matlab)	Diferença entre Áreas (%)
01	-	316,05	-
02	122,90	110,88	9,78
03	618,03	688,44	10,23
04	266,65	236,98	11,13
05	235,68	239,80	1,72
06	-	292,84	-
07	1722,88	1996,27	13,69
08	160,07	142,28	11,11
09	173,53	164,15	5,40
10	89,25	78,61	11,93
11	386,57	357,18	7,60
12	1963,43	2756,93	28,78
13	1765,93	2261,35	21,91
14	2187,17	2559,39	14,54
15	285,234	330,018	13,57

O mesmo foi feito a seguir (Tabela 2) para os valores de perímetro do nódulo, sendo então calculados em mm.

Tabela 2: Comparação entre perímetros extraídos pelo tomógrafo e pelo algoritmo desenvolvido neste trabalho

Paciente	Perímetro (Tomógrafo)	Perímetro (Otsu/Matlab)	Diferença entre Perímetros (%)
01	-	80,99	-
02	53,58	61,64	13,07
03	120,06	113,52	5,45
04	74,71	73,04	2,23
05	70,50	70,08	0,60
06	-	72,59	-
07	216,79	219,3	1,14
08	64,28	54,02	15,96
09	60,47	57,40	5,07
10	46,48	40,12	13,69
11	121,92	114,08	6,43
12	328,29	250,29	23,76
13	204,17	214,9	4,99
14	261,89	264,00	0,80
15	83,07	82,62	0,55

A Figura 3 representa as curvas Gaussianas de alguns dos atributos extraídos. As curvas em azul representam lesões benignas, em vermelho representam lesões do tipo carcinoma e, em verde, lesões do tipo carcinoma.



(a)



(b)



(c)

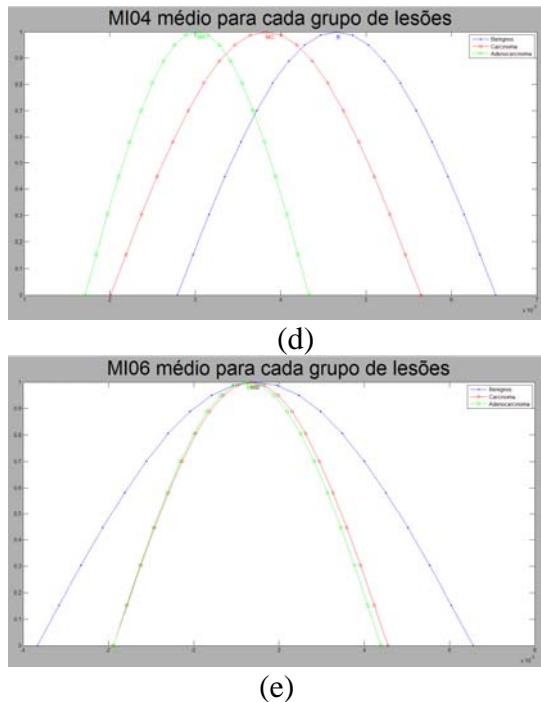


Figura 3: Curvas Gaussianas da (a) Irregularidade, (b) MI01, (c) MI02, (d) MI04, (e) MI06 médios para cada grupo de lesões.

### Discussão

O algoritmo de segmentação desenvolvido neste trabalho apresentou resultados com erros inferiores a 10% para a maioria das lesões analisadas. Entretanto, para lesões de alta densidade adjacentes à parede pulmonar, a segmentação resultante foi menos satisfatória. Isso se deve ao baixo contraste entre tecidos da parede pulmonar e que compõem a lesão.

Os erros médios para medidas de área e perímetro calculadas pelo algoritmo desenvolvido em relação às medidas do tomógrafo utilizado foram, respectivamente, 12,41% e 7,21%, conforme mostrados nas Tabelas 1 e 2. O cálculo da área e do perímetro do tomógrafo considerou a seleção manual de pontos pelo especialista, o que pode influenciar na precisão das medidas. Além disso, o perímetro pode ser medido por pixels externos ou internos à borda do objeto, os quais proporcionam diferentes valores para uma mesma imagem.

Com relação ao poder de *separabilidade* dos grupos, os resultados apresentados indicam que alguns dos atributos extraídos não se diferem entre os tipos de lesões estudados como mostra a Figura 3 (e), onde para o atributo MI06 apresenta as curvas de distribuição bastante sobrepostas. Entretanto, as curvas das distribuições Gaussianas da irregularidade, 1º, 2º e 4º Momentos Invariantes de Hu (Figura 3 a, b, c e d) mostraram bom grau de *separabilidade* entre os grupos analisados com poucas áreas sobrepostas.

### Conclusão

Caracterizar lesões por meio de processamento de imagens pode trazer grandes benefícios relacionados ao diagnóstico do câncer de pulmão, já que ainda hoje a diferenciação entre nódulos pulmonares malignos e benignos necessita de exames invasivos. Este processo requer, primeiramente, uma boa segmentação das lesões, para que não haja perda de informações da lesão, minimizando a chance de resultados errôneos. Outra etapa importante para caracterização de lesões é a seleção dos atributos a serem considerados neste processo, ou seja, a determinação de quais medidas permitem uma diferenciação dos tipos de lesões pulmonares. A extração de atributos fornece informações quantitativas, as quais podem ser de grande utilidade na definição do diagnóstico de pacientes.

### Referências

- [1] Tipos de Câncer: Pulmão at <http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/pulmao>.
- [2] American Cancer Society - Cancer Facts & Figure 2010 at <http://www.cancer.org/>
- [3] Swensen SJ et al. Lung nodule enhancement at CT: multicenter study. *Radiology*. 2000; 214:73–80.
- [4] Netto SMB, Silva AC, Nunes RA, Gattass M. Analysis of directional patterns of lung nodules in computerized tomography using Getis statistics and their accumulated forms as malignancy and benignity indicators. *Pattern Recognition Letters*. 2012; 33:1734-1740.
- [5] Silva, A C, Carvalho, P C P, Gattass, M. Diagnosis of lung nodule using semivariogram and geometric measures in computerized tomography images. *Computer Methods and Programs in Biomedicine*. 2005; 79: 31-38.
- [6] Kamiya H, Murayama S, Kakinohana Y, Miyara T. Pulmonary nodules: a quantitative method of diagnosis by evaluating nodule perimeter difference to approximate oval using three-dimensional CT images. *Clinical Imaging*. 2011; 35:123-126.
- [7] Gonzales RC, Woods RE. *Processamento Digital de Imagens*. 3ª ed. São Paulo: Pearson; 2009
- [8] Hu MK. Visual Pattern Recognition by Moments Invariant. *IRE Trans. Inf. Theory*. 1962; 8:179-187.
- [9] Martí J, Cufí X, Regincós J, Español J, Barceló C. Shape-based feature selection for microcalcification evaluation. *Proc.SPIE: Image Processing*. 1998; 3338:1215-1224.