Artigo Original

Revista Brasileira de Física Médica. 2019; 13(3):38-42.

Lung nodules classification in CT images using texture descriptors

Classificação de nódulos pulmonares em imagens de TC utilizando descritores de textura

Luís V. de Moura[1], Caroline M. Dartora[1,2], Ana M. Marques da Silva[1,2,3]

[1]PUCRS, School of Technology, Medical Image Computing Laboratory, Porto Alegre, Brazil
[2]PUCRS, School of Medicine, Graduate Program in Biomedical Gerontology, Porto Alegre, Brazil
[3]PUCRS, Brain Institute – INSCER, Porto Alegre, Brazil

Abstract

In lung cancer, early diagnosis can improve potentially the prognosis. Accurate interpretation of computed tomography (CT) scans demands significant efforts by radiologists due to the extensive number of slices analyzed in each examination, for each patient. Computer-aided diagnosis (CAD) systems have been applied in several medical fields, but mostly in lung nodules detection and classification. CAD systems for lung lesions classification usually extract different types of features from lesions, such as texture feature, shape and intensity. This exploratory study aims to investigate the performance of lung nodules classification in 2D and 3D CT lesions images using Haralick texture features analysis and binary logistic regression. Expert radiologists manually segmented from a CT dataset of 17 benign and 20 malignant nodules, which have their anatomopathological results. Haralick features were extracted from 2D lesions images, using the largest cross-section nodule area, and from all nodule volume (3D). Principal Component Analysis (PCA) was applied to reduce texture features dimensionality, showing two and three principal components (PC) can explain 85.8% and 96.25% of data variance for 2D lesions, and 72.4% and 91.7% for 3D lesions, respectively. Binary logistic regression using leave-one-out cross-validation for training and test datasets showed no differences in accuracy (63% - 68%), using two or three PC. The higher sensitivity (75%) was acquired using 2D images with two or three PC, while the higher specificity (65%) was obtained using 3D images with two or three PC. Binary logistic regression using a small number of Haralick texture features showed better accuracy in lung nodules classification than visual evaluation by radiologists, although the limited dataset. Further studies are needed to generalize and improve these results.

Palavras-chave: lung nodules; CT; CAD; Haralick; texture.

Resumo

No câncer de pulmão, o diagnóstico precoce pode melhorar potencialmente o prognóstico. A interpretação precisa de imagens de tomografia computadorizada (TC) exige esforços significativos dos radiologistas devido ao extenso conjunto de cortes em cada exame, analisado todos os dias. Os sistemas de diagnóstico assistido por computador (CAD) têm sido aplicados em diversas áreas médicas, mas principalmente na detecção e classificação de nódulos pulmonares. Os sistemas CAD para classificação de lesões pulmonares geralmente extraem diferentes tipos de características de lesões, tais como textura, forma e intensidade. Este estudo exploratório tem como objetivo investigar o desempenho da classificação de nódulos pulmonares em imagens 2D e 3D de lesões em TC usando a análise de características de textura de Haralick e regressão logística binária. Radiologistas especialistas segmentaram manualmente, a partir de um conjunto de dados de TC, 17 nódulos benignos e 20 malignos, com seus respectivos resultados anatomopatológicos. As características de Haralick foram extraídas de imagens de lesões 2D, utilizando a maior área da imagem do corte transversal dos nódulos e de todo o volume de nódulos (3D). A Análise de Componentes Principais (PCA) foi aplicada para reduzir a dimensionalidade das características da textura, mostrando que dois e três componentes principais (PC) podem explicar 85,8% e 96,25% da variação de dados para lesões 2D e 72,4% e 91,7% para lesões 3D, respectivamente. A regressão logística binária usando validação cruzada de exclusão única para conjuntos de dados de treinamento e teste não mostrou diferenças na precisão (63% - 68%), usando dois ou três PC. A maior sensibilidade (75%) foi adquirida com imagens 2D com dois ou três PCs, enquanto a maior especificidade (65%) foi obtida com imagens 3D com dois ou três PCs. A regressão logística binária usando um pequeno número de características da textura Haralick mostrou melhor precisão na classificação dos nódulos pulmonares do que na avaliação visual por radiologistas, embora o conjunto de dados seja limitado. Maiores estudos são necessários para generalizar e melhorar estes resultados.

Keywords*: nódulos pulmonares; CT; CAD; Haralick; textura.*

1. Introduction

Computer-aided diagnosis (CAD) systems have been one of the most used tools in the diagnosis of lung cancer in the last decade, using both clinical information and quantitative analysis from features extracted from images by software. Their purpose is to improve diagnostic accuracy and reproducibility and to act as a second opinion for the radiologist interpretation[1].

CAD systems can be divided into two subgroups: CADe and CADx. CADe are computer-aided detection systems, designed to indicate the location or to produce the segmentation of lesions or abnormalities. CADx are systems designed to help diagnosis, characterizing a region or lesion, based on image descriptors classification, previously learned by the computer using artificial intelligence methods[2].

CAD systems are applied in several diagnostic tasks, but mostly in cancer detection and diagnosis, using different imaging modalities, such as mammography, colonoscopy, and computed tomography (CT) [2].

Regarding lung nodules, an early diagnosis can improve potentially the prognosis[3]. In CT scans, the term "nodule" is used for opacities with a diameter between 3 and 30 mm. Some nodule characteristics like location in the lung, neighboring structures, size, shape, texture, and internal density, as solid, subsolid, or non-solid, can indicate its benignancy or malignancy (cancerous)[4,5]. The likelihood that a nodule can be malignant is about 40%, but the risk with age varies. When a nodule is detected in CT scan, the radiologist usually compares the images with an initial dataset. If there is a change in the nodule characteristics or a new one arises, a bronchoscopy or tissue biopsy is recommended to determine its malignancy[6]. The accurate CT scan interpretation is challenging and demands big efforts by the radiologists due to a large number of slices analyzed for each patient. Therefore, the necessity of invasive pathological tests, like a biopsy, remains necessary[4], and CAD systems can potentially improve the accuracy of the lesion characterization.

Since CAD was developed at the beginning of the 21st century[1], its use for lung abnormalities has been widely studied, implemented, and improved. One of the most CADx methods used for lung abnormalities is feature extraction, a process to acquire a higher-level of information from the image[7]. The texture is one of the key components of human visual perception[7], and spectral and statistical approaches have been developed to extract texture descriptors from digital images mathematically. Haralick (1973) introduced the textural features for image classification using a statistical approach[8], where a gray level co-occurrence matrix (GLCM) is calculated, and 14 textures descriptors can determine local features or global features, from the distribution of local features[7,8].

Texture descriptors were developed mainly for two-dimensional (2D) images, as chest X-ray or CT slices. However, three-dimensional (3D) methods for texture analysis were developed in the last five years. In 2014, Han et al[9] investigated three types of 2D texture descriptors and extended their study to the impact of expansion on the 3D space. They have used a CT database to compare 2D texture features extraction, such as Haralick, Gabor, and Local Binary Patterns. They observed better results when calculating 2D features on all image slices, compared to a single slice. Haralick features were identified as the best choice to differentiate malignant and benign lung nodules, and 3D extension reveals a potential gain in the results when an optimal number of directions to build the GLCM is found.

In 2018, Wei et al.[10] implemented a content-based image retrieval scheme with a two steps similarity metric approach to classify lung nodules based on Mahalanobis and European distances. Three groups of texture descriptors (local binary pattern feature, Gabor features, and Haralick features) were calculated in 366 images of 2D benign and malignant nodules of the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI). The classification accuracy using only Haralick texture descriptors was better than a combination of other methods.

In the last years, CADx systems have been developed using Artificial Intelligence (AI) methods[11–14]. Machine Learning, radiomics, and neural networks are the most used algorithms for lung nodules classification. They show excellent accuracy (more than 80%) using big databases, 2D and 3D images, and a combination of features, as texture, shape, and demographic data. However, health institutions in developing or low-income countries, have no access to this type of technology. A more straightforward approach is required to allow CAD second-opinion systems to this type of institution.

This exploratory study aims to investigate the performance of lung nodules classification in 2D and 3D CT images using Haralick texture analysis and binary logistic regression.

2. Materials and Methods

2.1. Data of the study

Retrospective CT lung images, data from clinical evaluation, and anatomopathological tests of nodules used in this project were obtained in a large hospital, and their use has ethical approval (CAAE: 12385713.4.0000.5336, protocol: 339574). Dataset is from 37 subjects, 20 females (8 benign and 12 malignant), and 17 males (9 benign, and 8 malignant), with ages varying between 41, and 83 years. Subject's ages were 62 ± 9 y and 63 ± 12 y for malignant and benign nodules, respectively.

CT images were acquired on a Siemens Somatom Emotion 16-channel CT, using 120 kV, 241 mAs, 2.0 mm slice thickness, 0.9 pitch, and a 512 x 512 pixels matrix with a 1.185 pixel/mm resolution.

Nodules were divided into two groups: (1) malignant nodules and (2) benign nodules, segmented from CT images. The nodules' anatomopathological results were used as diagnosis gold standard for the classification model analysis.

Malignant nodules were pathologically characterized as levels II and III adenocarcinoma, and squamous cell carcinomas. Benign nodules were pathologically identified as hamartoma, lipoma, and adenoma.

2.2. Nodules Segmentation

CT lung nodules were segmented by trained radiologists, without any pre-processing steps, to avoid tissues of no interest in the quantitative analysis. Segmented images of lung nodules were organized in symmetric matrices depending on the larger diameter of the nodule.

2.3. Features Extraction

An in-house code, implemented using MATLAB R2012b, was developed to calculate GLCM and Haralick texture features from 2D segmented nodules (lesion with the largest diameter) and the whole 3D lung nodules. GLCM was calculated based on 8-neighbourhood pixels for 2D nodules, and 26-neighbourhood pixels for 3D nodules. Before applying Haralick descriptors, zero-pairs in GLCM were discarded to avoid high fluctuations in the features.

Haralick texture features were based on the mathematical descriptors of the original article[8]. The Haralick texture descriptors used in this paper were: angular second moment (ASM); contrast; correlation; dissimilarity; energy; entropy; homogeneity; variance; GLCM mean.

2.4. Classification Model

Binary logistic regression, after principal component analysis (PCA) with leave-one-out cross-validation (LOOCV), was used to classify benign from malignant nodules, based in the previously described Haralick features extracted from the 2D and 3D nodule images.

Data were normalized by scaling each feature between 0 and 1, independently. PCA was applied due to the small number of images compared to the number of features. Components were selected if they could explain more than 75% of data variance. These components were used in binary logistic regression with $p<0.005$. Statistical analyses were performed using SPSS software version 17.0 (IBM, Armonk, NY, USA).

2.5. Performance analysis of the classification model

We have computed the sensitivity, specificity, and accuracy of each method to produce a quantitative analysis of the predictive model.

Equations 1, 2, and 3, show sensitivity, specificity, and accuracy definitions, respectively:

$$Sensitivity = \frac{TP}{TP + FN} * 100\% \quad (1)$$

$$Specificity = \frac{TN}{TN + FN} * 100\% \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100\% \quad (3)$$

A true positive (TP) classification occurs when a malignant nodule is detected as malignant, and false positive (FP) when a malignant nodule is classified as a benign one. True negative (TN) classification occurs when a benign nodule is pointed out as benign, while false negative (FN) occurs when a malignant nodule is classified as benign[15].

3. Results

Table 1 shows the mean and standard deviation (sd) of all Haralick texture descriptors calculated on 2D and 3D images of the segmented nodules in CT lung scans.

Table 1 - Mean and standard deviation (mean ± sd) of Haralick texture descriptors.

Texture descriptors	2D		3D	
	Malignant	Benign	Malignant	Benign
ASM	0.29 ± 0.20	0.13 ± 0.07	0.008 ± 0.004	0.01 ± 0.01
Contrast	21000 ± 20000	22000 ± 20000	23000 ± 17000	40000 ± 27000
Correlation	0.96 ± 0.04	0.95 ± 0.04	0.69 ± 0.13	0.75 ± 0.12
Dissimilarity	43 ± 27	63 ± 31	73 ± 36	115 ± 52
Energy	0.5 ± 0.2	0.35 ± 0.10	0.09 ± 0.02	0.10 ± 0.05
Entropy	4.9 ± 1.5	6 ± 1	10.2 ± 0.9	10.4 ± 0.9
Homogeneity	0.5 ± 0.2	0.37 ± 0.08	0.05 ± 0.02	0.07 ± 0.07
Variance	234000 ± 63000	240000 ± 70000	44000 ± 31000	110000 ± 92000
GLCM mean	473 ± 170	545 ± 99	964 ± 92	821 ± 180

Source: The author (2019).

PCA showed that the use of two and three components can explain, for 2D lesion image, 85.8% and 96.25% of data variance. For 3D lesion image, two and three components can explain 72.4% and 91.7% of data variance.

Table 2 shows the performance of binary logistic regression in training and test datasets, using sensitivity, specificity, and accuracy, calculated using two and three principal components.

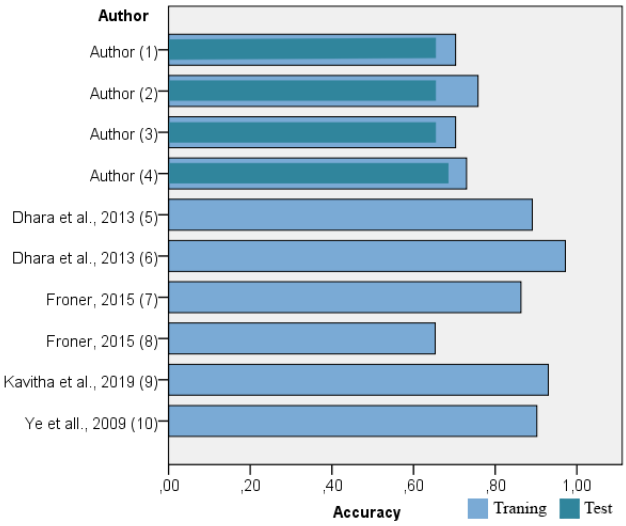
Table 2 - Classification performance with 2 and 3 principal components (PC) in training and test datasets.

		Training			Test		
Image	# of PC	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
2D	2	75%	69%	70%	75%	53%	63%
	3	75%	72%	76%	75%	53%	65%
3D	2	75%	69%	70%	65%	65%	65%
	3	80%	73%	73%	70%	64%	68%

Source: The author (2019).

Our results were compared to other studies[5,12–14], in order to evaluate the accuracy. Figure 1 shows the comparison between our results with other studies.

Figure 1 - Accuracy comparison between our work and the literature.



(1) Our data results for 2D images with 2 PC;
(2) Our data results for 2D images with 3 PC;
(3) Our data results for 3D images with 2 PC;
(4) Our data results for 3D images with 3 PC;
(5) Dhara et al., 2013[13], for 2D images of the nodule;
(6) Dhara et al., 2013[13], for 3D images of nthe odule;
(7) Froner, 2015[5], multivariate logistic regression method;
(8) Froner, 2015[5], accuracy results from expert radiologists based just in visual interpretation, within blind evaluation;
(9) Kavitha et al., 2019[12], with Haralick texture features calculates in 2D images
(10) Ye et al., 2009[14], with features extracted in 2D and 3D images.
Source: The author (2019).

4. Discussions

We investigated the performance of lung nodules classification using 2D and 3D CT images of segmented nodules using texture feature analysis. The classification method was the binary logistic regression and only two and three components were used after PCA.

Our results showed more than 63% of accuracy using 2D and 3D images to calculate the GLCM and texture features based on Haralick approach.

Figure 1 shows that our method is less accurate than others found in the literature. However, the higher values are found in methods based on AI algorithms, like Support Vector Machines[12,14] and Artificial Neural Networks[13], and bigger databases. It is worth mentioning that, within the models proposed in the literature, ours have a smaller database and a wide variety of benign and malignant nodule types. Moreover, our work used only texture features, while others used additional attributes in the classification model, such as intensity, morphology, and demographic data from the subjects.

It is important to emphasize that small radiology departments have no resources or support required to develop Artificial Intelligence algorithms in CAD systems. Binary logistic regression can be applied in any standard computer using statistical software.

Our results agree with Dhara *et. al*. (2013)[13], which uses neural networks to classify nodules based on Haralick texture features, where classification performance on 3D texture features performed better than 2D. Their higher accuracy (.80%) probably is related to the preprocessing and smoothing steps applied prior to the segmentation and classification.

Froner et al.[5] used the same CT lung database to evaluate the use of patient data (demography) and quantitative features from lung nodules to build a classification model. They analyzed several morphological, demographical, and texture features applied to a multivariate logistic regression model, compared to the visual interpretation of radiologists. Our results showed similar or slightly superior predictive values than the visual evaluation of lung nodules by expert radiologists (65%).

Our model showed an accuracy of more than 63% in CT lung nodules classification using two and three PC in binary logistic regression. The use of 2D nodule images with 2 and 3 PC showed better sensitivity (75%) than 3D texture, while 3D texture with 3 PC showed a better specificity (64%) and accuracy (68%) on nodule classification.

5. Conclusions

We investigated the performance of lung nodules classification using binary logistic regression classification, and 2D and 3D texture analysis from CT images. The method was able to produce similar or slightly superior predictive values than the visual interpretation by expert radiologists, considering the reduced number of images. The use of 3D images increased the accuracy and specificity of lung nodule classification. Further studies are required to improve the classification method.

The main limitation of this study was the small sample size, which restricted the number of attributes to be included in the classification and the lack of morphological features. For future work, a larger sample of CT lung images, and additional features will be included.

Acknowledgements

This study was financed in part by PUCRS and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

1. Doi K, MacMahon H, Giger ML, Hoffmann KR. Computer-aided diagnosis and its potential impact on diagnostic radiology. Comput Diagnosis Med imaging Amsterdam, Netherlands Elsevier Sci. 1999;11–20.
2. Giger ML, Suzuki K. Computer-Aided Diagnosis. Biomed Inf Technol [Internet]. 2008 Jan 1 [cited 2019 Mar 18];359–XXII. Available from: https://www.sciencedirect.com/science/article/pii/B9780123735836500207
3. Zia ur Rehman M, Javaid M, Shah SIA, Gilani SO, Jamil M, Butt SI. An appraisal of nodules detection techniques for lung cancer in CT images. Biomed Signal Process Control [Internet]. 2018 Mar 1 [cited 2019 Mar 18];41:140–51. Available from: https://www.sciencedirect.com/science/article/pii/S1746809417302811
4. Rubin GD. Lung nodule and cancer detection in computed tomography screening. In: Journal of Thoracic Imaging

[Internet]. 2015 [cited 2019 Mar 18]. p. 130–8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4654704/pdf/nihms714352.pdf

5. Froner APP. Caracterização de nódulos pulmonares em imagens de tomografia computadorizada para fins de auxílio ao diagnóstico [Internet]. Pontifícia Universidade Católica do Rio Grande do Sul; 2015 [cited 2019 Apr 10]. Available from: http://tede2.pucrs.br/tede2/handle/tede/6385
6. Madero Orozco H, Vergara Villegas OO, Cruz Sánchez VG, Ochoa Domínguez H de J, Nandayapa Alfaro M de J. Automated system for lung nodules classification based on wavelet feature descriptor and support vector machine. Biomed Eng Online [Internet]. 2015 Dec 12 [cited 2019 Mar 18];14(1):9. Available from: http://www.biomedical-engineering-online.com/content/14/1/9
7. Zayed N, Elnemr HA. Statistical Analysis of Haralick Texture Features to Discriminate Lung Abnormalities. Int J Biomed Imaging [Internet]. 2015 Oct 8 [cited 2019 Apr 3];2015:1–7. Available from: http://www.hindawi.com/journals/ijbi/2015/267807/
8. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. IEEE Trans Syst Man Cybern [Internet]. 1973 Nov [cited 2019 Apr 10];SMC-3(6):610–21. Available from: http://ieeexplore.ieee.org/document/4309314/
9. Han F, Wang H, Zhang G, Han H, Song B, Li L, et al. Texture Feature Analysis for Computer-Aided Diagnosis on Pulmonary Nodules. J Digit Imaging [Internet]. 2014 [cited 2019 Mar 22];28(1):99–115. Available from: https://link.springer.com/content/pdf/10.1007%2Fs10278-014-9718-8.pdf
10. Wei G, Cao H, Ma H, Qi S, Qian W, Ma Z. Content-based image retrieval for Lung Nodule Classification Using Texture Features and Learned Distance Metric. J Med Syst [Internet]. 2018 [cited 2019 Apr 15];42(1). Available from: https://doi.org/10.1007/s10916-017-0874-5
11. Franco MLN, Nunes LM, Froner APP, Silva AMM, Patrocínio AC. REDES NEURAIS ARTIFICIAIS APLICADAS NA CLASSIFICAÇÃO DE TUMORES PULMONARES [Internet]. [cited 2019 Apr 15]. Available from: http://www2.inca.gov.br/wps/wcm/connect/tiposdec
12. Kavitha MS, Shanthini J, Sabitha R. ECM-CSD: An Efficient Classification Model for Cancer Stage Diagnosis in CT Lung Images Using FCM and SVM Techniques. J Med Syst [Internet]. 2019 Mar 12 [cited 2019 Apr 15];43(3):73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30746555
13. Dhara AK, Mukhopadhyay S, Khandelwal N. 3D texture analysis of solitary pulmonary nodules using co-occurrence matrix from volumetric lung CT images. In: Novak CL, Aylward S, editors. International Society for Optics and Photonics; 2013 [cited 2019 Apr 15]. p. 867039. Available from: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2007016
14. Xujiong Ye, Xinyu Lin, Dehmeshki J, Slabaugh G, Beddoe G. Shape-Based Computer-Aided Detection of Lung Nodules in Thoracic CT Images. IEEE Trans Biomed Eng [Internet]. 2009 Jul [cited 2019 Apr 15];56(7):1810–20. Available from: http://ieeexplore.ieee.org/document/5073252/
15. Ramteke RJ, Y KM. Automatic Medical Image Classification and Abnormality Detection Using K- Nearest Neighbour. Int J Adv Comput Res [Internet]. 2012 [cited 2019 Apr 3];2(4):190–6. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.300.4729&rep=rep1&type=pdf

Contact:

Corresponding Author:
Luís Vinícius de Moura
PUCRS, Medical Image Computing Laboratory (MEDICOM)
Av. Ipiranga, 6681 Partenon - Porto Alegre/RS
CEP: 90619-900 - Phone: (51) 3320.7813
Email: *luis.moura.001@acad.pucrs*.br

Caroline Machado Dartora
PUCRS, Medical Image Computing Laboratory (MEDICOM)
Email: *caroline.dartora@acad.pucrs.br*

Prof. Dr. Ana Maria Marques da Silva
PUCRS, Medical Image Computing Laboratory (MEDICOM)
Email: *ana.marques@pucrs.br*