Computers in Biology and Medicine 42 (2012) 1110–1121

Contents lists available at SciVerse ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Automatic segmentation of lung nodules with growing neural gas and support vector machine

Stelmo Magalhães Barros Netto [a], Aristófanes Corrêa Silva [a,*], Rodolfo Acatauassú Nunes [b], Marcelo Gattass [c]

[a] Federal University of Maranhão - UFMA, Applied Computing Group - NCA, Av. dos Portugueses, SN, Campus do Bacanga, Bacanga 65085-580, São Luís, MA, Brazil
[b] State University of Rio de Janeiro - UERJ, São Francisco de Xavier, 524, Maracanã, 20550-900 Rio de Janeiro, RJ, Brazil
[c] Pontifical Catholic University of Rio de Janeiro - PUC-Rio, R. São Vicente, 225, Gávea, 22453-900 Rio de Janeiro, RJ, Brazil

ARTICLE INFO

Article history:
Received 12 November 2010
Accepted 5 September 2012

Keywords:
Medical image
Computer-aided detection (CAD)
Lung nodules
Growing neural gas
Image processing

ABSTRACT

Lung cancer is distinguished by presenting one of the highest incidences and one of the highest rates of mortality among all other types of cancer. Unfortunately, this disease is often diagnosed late, affecting the treatment outcome. In order to help specialists in the search and identification of lung nodules in tomographic images, many research centers have developed computer-aided detection systems (CAD systems) to automate procedures. This work seeks to develop a methodology for automatic detection of lung nodules. The proposed method consists of the acquisition of computerized tomography images of the lung, the reduction of the volume of interest through techniques for the extraction of the thorax, extraction of the lung, and reconstruction of the original shape of the parenchyma. After that, growing neural gas (GNG) is applied to constrain even more the structures that are denser than the pulmonary parenchyma (nodules, blood vessels, bronchi, etc.). The next stage is the separation of the structures resembling lung nodules from other structures, such as vessels and bronchi. Finally, the structures are classified as either nodule or non-nodule, through shape and texture measurements together with support vector machine. The methodology ensures that nodules of reasonable size be found with 86% sensitivity and 91% specificity. This results in a mean accuracy of 91% for 10 experiments of training and testing in a sample of 48 nodules occurring in 29 exams. The rate of false positives per exam was of 0.138, for the 29 exams analyzed.

1. Introduction

Patients with lung cancer have mean survival of five years, varying between 13% and 21% in developed countries and between 7% and 10% in emerging countries. Only in 2008, a total of 1.37 million deaths caused by lung cancer occurred throughout the world [1]. Official statistics by the Brazilian National Cancer Institute (INCA) for the same year registered that lung cancer caused the death of 26,990 Brazilians. The estimates of incidence in 2012, according to the INCA [2], pointed out that the number of new cases of cancer would be 27,320. Such an incidence is still the result of the high consumption of tobacco in the past and does not reflect the present scenario of reducing smoking habits in response to recent preventive actions [2].

Two causes of the shortened survival of lung cancer patients are the difficult early diagnosis, due to the lack of symptoms, and the poor prognostic when the disease is diagnosed at more advanced stages [3]. Given these characteristics, efforts have been made to improve the early diagnosis of lung cancer.

The detection of lung cancer at the earliest stages has been improved by non-invasive imaging techniques, such as radiography and computed tomography (CT) of the chest. Nevertheless, invasive techniques are still necessary to define the diagnosis, which is confirmed through the cytological or histopathological examination.

In this scenario, where the application of non-invasive techniques gains special relevance, many computational tools have been employed to support the specialist physician, such as computer-aided detection/diagnosis (CAD/CADx), developed from image processing and computational vision techniques.

From digital images generated by the process of acquiring the CT, it is possible to identify lung nodules and execute a series of measurements on them. The objective is to find some correlation between these measurements and the diagnosis of malignancy or benignity [4].

* Corresponding author. Tel.: +55 98 33018243; fax: +55 98 33018841.
E-mail addresses: nobnet2000@yahoo.com.br (S. Magalhães Barros Netto), ari@dee.ufma.br (A. Corrêa Silva), rodolfoacatauassu@yahoo.com.br (R. Acatauassú Nunes), mgattass@tecgraf.puc-rio.br (M. Gattass).

0010-4825
http://dx.doi.org/10.1016/j.compbiomed.2012.09.003

This work aims to detect lung nodules automatically through computed tomography images. For this, we have used the growing neural gas (GNG) algorithm to segment structures with features very similar to those of lung nodules. Then we have used the distance transform to separate the segmented structures, which are connected to blood vessels and bronchi. Finally, we have used a set of shape and texture features and support vector machine (SVM) to classify these structures as lung nodules.

This paper is divided as follows: in Section 2 we show some related works. Section 3 presents all steps of the automatic segmentation of the lung nodules, explaining in detail the use of techniques, such as growing neural gas, 3D distance transform, and the extraction of shape and texture features of the lung nodule candidates, which compose the proposed methodology. In Section 4 we analyze all the tests obtained with the application of the methodology. Finally, in Section 5, we present the final remarks about this work.

2. Related works

The use of CAD/CADx systems to assist specialist physicians has been increasing day by day [5]. Using these systems is interesting because of the easy way they show suspect radiologic artifacts and indicate possible diagnosis faster than a specialist alone. They are especially used in screening programs, in which the tasks of detection and diagnosis of lung nodules are performed continuously and repeatedly by a specialist physician.

Recent research has demonstrated the advantages of the use of CAD/CADx systems to help specialist physicians in the detection [6–8] and diagnosis of lung nodules [9]. Sahiner et al. [6] show through experiments that either the specialist or the CAD system alone presents lower performance for the detection of lung nodules than when a doctor is aided by a CAD system.

The need for obtaining a precise diagnosis of lung nodules to allow the identification of cancer at an earlier stage that could lead to cure or a greater survival has motivated many researchers to look for new forms of detection and diagnosis with the help of a computer [10–13].

The rationale behind these tools is the proposal of an aid to specialist physicians, whether to highlight suspect radiologic artifacts or to offer a second opinion to the diagnosis specialist. Works such as those of Jeong et al. [14] and Reeves et al. [15] have demonstrated well this task of detecting and diagnosing lung nodules.

The automatic segmentation of CT images is a hard task because there are many interfering elements in the images, such as noises generated in the acquisition process, the features of the images (thickness of the slices, dimensions of the pixels, etc.), the morphology of the nodules, their location, and the presence of neighbor structures with very close densities and directly connected.

The segmentation based on clustering was used in the automatic segmentation of lung nodules by Antonelli et al. [16], with fuzzy c-means, followed by the morphological analysis of the resulting structures.

The automatic segmentation of low-dose CT images (LDCT) was done by [17] in a framework developed for the task. Using information from probabilistic models created to control the evolution of a deformable model, this framework was able to segment the lung nodules with mean error of 0.96% and standard deviation of 1.1%, with respect to the form described by the specialist.

The work of Ozekes et al. [18], using segmentation based on rules and template matching trained with genetic algorithm, had very significant results. Their system achieved 93.4% sensitivity and 0.594 false positives per exam. Still based on template matching, but adding a genetic cellular neural network and threshold based on fuzzy rules, Ozekes et al. [19] achieved 100% sensitivity and a rate of 13.375 false positives per case.

Under the same perspective, through selective enhancement filters and segmentation based on rules, Li et al. [9] obtained 86%, 81%, and 75% sensitivity, with false positives of 6.6, 3.3, and 1.6 per exam, respectively, in a scheme of fourfold cross-validation.

Pu et al. [20] used a signed distance field in the whole set of images and detected the maximums found as potential nodules, which were then ranked according to their distance to the medial axis, obtained by means of the clustering technique and the application of marching cubes algorithm. A total of 52 tests were performed on a proprietary base that contained 184 lung nodules and whose results were 95.1% sensitivity and a mean of 1200 suspect voxels per exam.

Fiebich et al. [5] used low dose and active contour images as well, combined with a multi-thresholding algorithm, starting at points that had maximum resemblance to a sphere and obtained from the auto values in each voxel. The researchers managed to detect 58% of the nodules in this base, with a rate of 1.38 false positives.

Works such as that of Ye et al. [21] used five features, containing intensity information, shape index, and 3D spatial location. Their work shows that the mean overlapping of the nodules segmented by this method was near 0.81.

da Silva Sousa et al. [22] developed a complete system for automatic segmentation in multiple stages. Each stage of the system was responsible for segmenting significant portions of the volume of CT images. The whole process was comprised of six stages: extraction of the thorax, extraction of the lung, pulmonary reconstruction, extraction of structures, elimination of tubular structures, and reduction of false positives. The results obtained by this system presented a rate of 0.42 false positives and 0.15 false negatives, with 100% sensitivity. The three first stages of this system have been used in our work to reduce the volume of interest.

In Lee et al. [23], a set of classifiers called random forest was used in two stages. The first one was the stage of detection of lung nodules, and the second one was the stage of reduction of false positives. The results obtained by this methodology were of 100% for true positives and 1.4 false positives per image.

Messay et al. [24] combined simple image processing techniques, such as intensity thresholding and morphological processing, to segment and detect structures that were lung nodule candidates. The authors used 245 features to determine the lung nodule candidates. However, they made a selection of these features and used them in two classifiers: the Fisher Linear Discriminant classifier and a quadratic classifier. The method was able to detect 92.8% of the structures, which were nodule candidates.

The CAD system proposed by Tan et al. [25] employs three classifiers; one of them is based on genetic algorithms and artificial neural networks, which is then compared to results from SVM and fixed-topology neural networks. The lung nodule detection method consists in using filters to highlight nodules and vessels, and divergence features to locate possible lung nodule candidates. Once the candidates are detected, invariant features in a gauge system are applied to the three classifiers. The results obtained with the fixed-topology neural network had sensitivity of 87.5%, with average of four false positives per exam for nodules with diameter larger than or equal to 3 mm.

Opfer and Wiemker [26] have shown how to evaluate the performance of CAD systems using the LIDC base, which has the ground truth given by four specialists. In their analyses, the authors showed that their CAD system is able to reach a detection rate of 89% of the lung nodules in the LIDC base, with average of

two false positives per exam for nodules with diameter larger than or equal to 4 mm.

The method proposed by Xiaomin et al. [27] is divided into three steps to detect lung nodules in CT images. Firstly, a 2D multi-scale filter is used to detect the lung nodule candidates in the images. Secondly, the authors discriminate nodules from non-nodules with a blob-like shape using a growing geometric restriction region. Finally, after the shape features of each region are extracted, a classifier based on automated rules to reduce false positives is applied.

Suárez-Cuenca et al. [28] selected six classifiers: linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), artificial neural network (ANN), and three types of support vector machines (SVM). They were applied separately and combined to the LIDC base, using 85 images with 110 lung nodules. The sensitivity achieved was 80%, and the number of false positives per case for each of the six classifiers was 6.1 for LDA, 19.9 for QDA, 8.6 for ANN, 23.7 for SVM-dot, 17.0 for SVM-poly, and 23.35 for SVM-ANOVA. When these were used in combination, the number of false positives per case was 3.4 for the majority-vote rule, 6.2 for the mean, 5.7 for the product, 9.7 for the neural network, and 28.1 for the likelihood ratio method.

Camarlinghi et al. [29] propose the combination of different CAD systems aiming to provide enhanced support for lung nodule identification. Their experiment was then compared to results of the individual systems by means of the ROC curve. The results suggest that the higher the number of CAD systems used in the detection, the higher the number of true positives, 65, and the lower the number of false positives, 139, over a LIDC base of 69 images containing 114 lung nodules.

All these works present good results concerning sensitivity. However, a few of them result in low presence of false positives, whether per exam or per image. The present work aims to show a methodology comprised of simple techniques to find more efficient results concerning sensitivity and seeking to reduce the number of false positives per image or per exam.

Some authors already studied the application of the growing neural gas algorithm for medical purposes with good results [30,31]. In this work, we analyze the use of this algorithm to segment lung nodules.

3. Detection of lung nodules

The methodology proposed for lung nodule detection consists of the acquisition of computed tomography images of the lung, the reduction of the volume of interest through techniques for extraction of the chest, extraction of the lung, and reconstruction of the original shape of the parenchyma, which is lost in the previous stages. After that, there is the application of growing neural gas (GNG, a competitive learning algorithm) to the resulting volume of interest. The aim is to constrain even more the structures that are denser than the pulmonary parenchyma (nodules, blood vessels, bronchi, etc.). The next stage is the separation of the structures resembling lung nodules from other structures, such as vessels and bronchi. Finally, the structures are classified as either nodule or non-nodule, through shape and texture measurements together with a machine learning algorithm called support vector machine (SVM). Fig. 1 shows the stages of the proposed methodology.

3.1. Reduction of the volume of interest (VOI)

The objective of this stage is to eliminate the structures that are part of the image (see labels 1 and 2 of Fig. 2(a)) and do not contribute to the detection of lung nodules. This way, only the volume of interest remains, that is, only the lungs are used in the subsequent stage of the methodology.

The reduction of the VOI was based on a previous work done by our group; more precisely, da Silva Sousa et al. [22], who also aimed to segment lung nodules automatically. A brief explanation of their methodology is exposed below.

3.1.1. Extraction of the chest

The extraction of the chest is done through the region growing algorithm in 2D, with a seed located in each of the four corners of the image, as exemplified in Fig. 3(a). The thresholds for the region growing algorithm were found through the process of thresholding of the histogram calculated along the diagonal of a single image. An example of the histogram found is given in Fig. 1(b), which also shows the threshold found. This figure illustrates each of the peaks taken as basis for calculating the threshold. The peak labeled as 1 refers to soft tissues and air (lower density), while the peak labeled as 2 refers to muscles and bones (higher density). The thresholds used by the 2D region growing algorithm are T1 and T2 values identified in Fig. 2(b).

The sequence of results in the extraction of the chest can be seen in Fig. 3: (a) the beginning of the segmentation at the four seeds, (b) the fast growing progress, (c) the complete result of the extraction of the chest, and (d) the chest after complete extraction.

3.1.2. Extraction of the lungs

The extraction of the lungs comprises the removal of the structures of the mediastinum and the thoracic wall (label 3 indicated in Fig. 2(a)) to isolate the parenchyma and the several internal structures it contains.

The process of extraction of the lungs is also a result of the region growing algorithm, with a threshold taken from the same diagonal histogram described in the extraction of the chest. Thus, the thresholds used are T2 and the maximum value of the

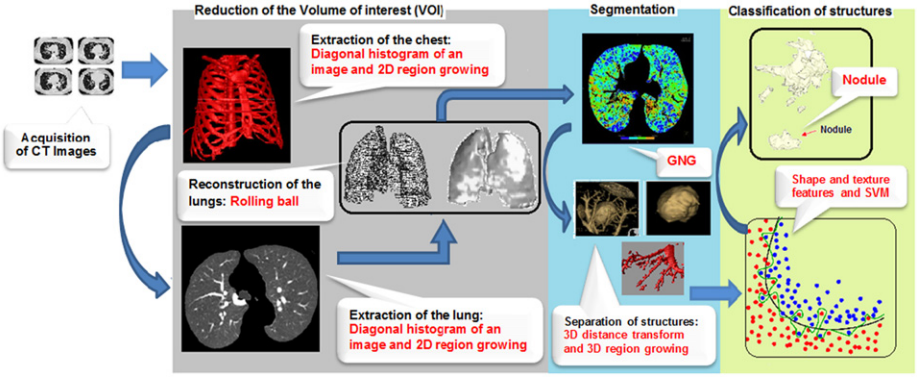

Fig. 1. Flow of the proposed methodology.

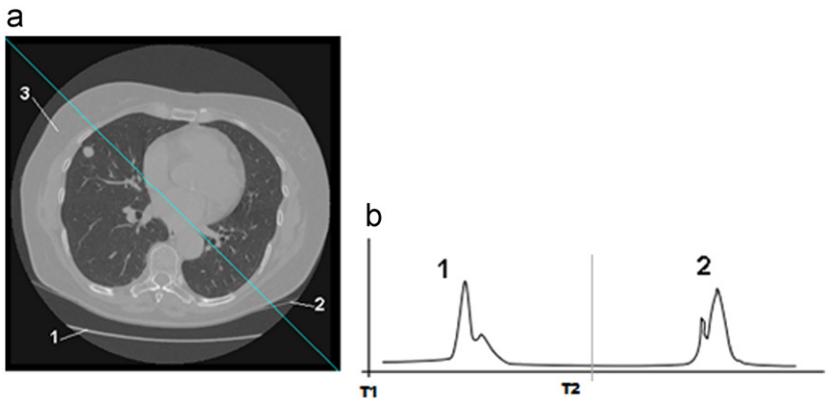

Fig. 2. (a) Example of image used in the calculation of the histogram along the diagonal line in blue. (b) Diagonal histogram of a slice of CT exam. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

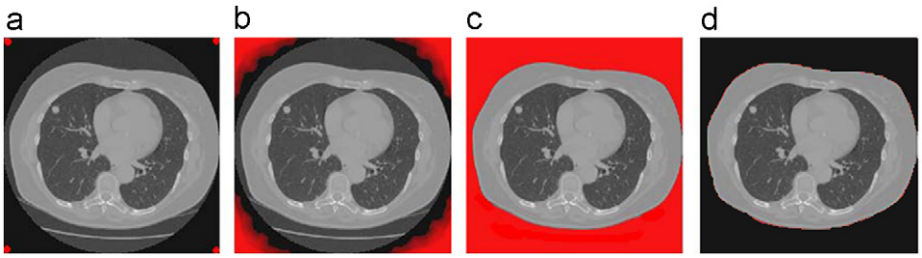

Fig. 3. Progress of the region growing for extraction of the chest. (a) The red dots indicate the beginning of the application of the algorithm. (b and c) The algorithm running. (d) The final result of this stage. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

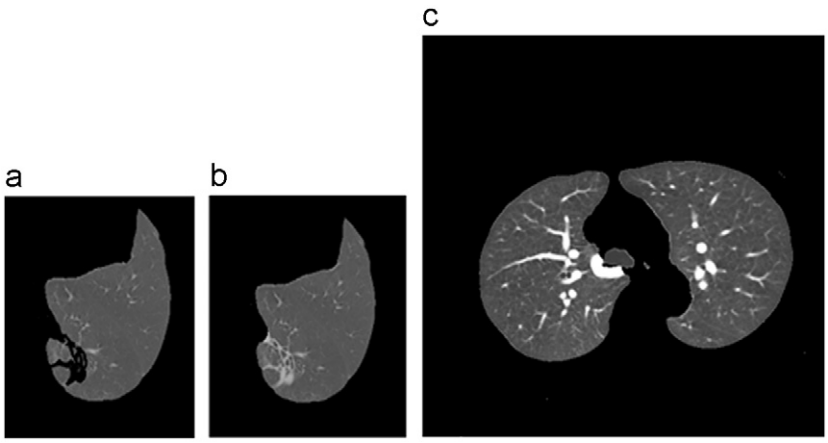

Fig. 4. (a) Right lung extracted. (b) Image of a reconstructed parenchyma. (c) Image resulting from the reduction of the volume of the left and right parenchyma.

histogram voxel. Nevertheless, the seed is started at a point of the diagonal that has higher density, that is, inside the thoracic wall. Fig. 4(a) shows one of the sides of the lung after the region growing algorithm.

A problem with this process is that nodules that are attached to the thoracic wall (juxtapleural nodules) are eliminated together with the thoracic wall. For this reason, the next stage is the reconstruction of the lungs to restore part of the structures of the parenchyma that are connected to the thoracic wall.

3.1.3. *Reconstruction of the lungs*

The reconstruction of the lungs is an important step that aims to recover lung nodules that are attached to the thoracic wall and were erroneously excluded at the lung extraction stage.

The reconstruction stage employs a technique called rolling ball, which uses morphological closing operations with a circular structuring element along the contour of the lung, causing the reconstruction of the concavities where this element cannot enter. The structuring element used is a (2D) disk with radius equal to 15. Fig. 4(b) exemplifies the result obtained when this technique is applied along the contour of the lung.

3.2. *Segmentation of lung nodules*

The segmentation of lung nodules is constituted by two stages. The first stage is the clustering of the dense structures through the application of growing neural gas (GNG). The second stage is the separation of the structures connected to the nodules by applying the 3D distance transform.

3.2.1. Application of growing neural gas

The growing neural gas (GNG) algorithm [32] is considered an optimization of the previous methods, such as Kohone's self-organizing maps [33] and the neural gas algorithm [34]. GNG is a clustering algorithm that works in an incremental manner, i.e., the number of clusters increases as the algorithm is executed, and in non-supervised mode, as the clusters are formed based only on the statistical features of the input data.

Firstly, the algorithm receives a set of input vectors of known dimensionality. Starting from two known nodes, the algorithm adds new nodes as it runs, following a set of rules. In doing so, it creates a graph in the same vectorial space as the input samples. Each of the nodes is a correspondent cluster that has a position in this dimensional space of the samples. The GNG algorithm can be used in quantization processes, in which each correspondent has to be interpreted as one of the code vectors. The topology of the graph found, after inserting all nodes, strongly reflects the probabilistic distribution of the input samples. In each iteration of the algorithm, a sample belonging to the mass of input data is chosen, according to the distribution of probabilities used in training. This sample is called input signal. The neuron closer to this sample is then chosen (competition stage). Then, the second closest neuron is determined, creating a connection between this neuron and the winner (cooperation stage). The winner and its neighbors are moved towards the input sample according to predefined parameters (adaptation stage). Further details on this technique can be found in [32].

In our methodology, GNG is used in the segmentation of the parenchyma structures using its adaptive capacity to find values of voxels that represent structures inside the parenchyma (high-density tissues) in an individual manner and not depending on the image base used.

The number of nodes in the network defines the number of clusters desired. Despite the fact that the GNG algorithm is not a supervised neural network, it is necessary to include the information of the cluster each node belongs to. The cluster classified as parenchyma is the one that has the most voxels, because the number of voxels of the parenchyma clearly exceeds the number of voxels of all structures it contains.

The parameters used in segmentation were empirically adjusted to achieve a better segmentation of the structures of the parenchyma and were taken as constant for the whole base. These parameters are: maximum number of nodes, number of iterations to insert a new unit, error decay rate for the unit and its neighbor with higher error, error decay rate for the other units, maximum age for an edge, adaptation rate of the winner, adaptation rate of the winner's neighbors.

The segmentation proper is performed by scanning each voxel of the volume and verifying which node of the GNG algorithm it belongs to.

3.2.2. Separation of structures

The separation of lung nodules from the structures (vessels, bronchi, etc.) that connect to them is an essential prerequisite for the task of detecting nodules. To perform the separation, we employ the distance transform together with the 3D region growing segmentation algorithm.

After applying GNG to the reduced volume, the structures are regrouped through the application of the 3D region growing algorithm, so to segment structures connected to each other. The resulting structures are not divided into elementary structures (such as nodules, bronchi, and vessels) at this point.

Shape is the most used feature for the separation of nodules from other structures that are connected to them. The simple fact that bronchi and vessels are more cylindrical (tubular) than lung nodules allows us to find shape-based metrics and distinguish

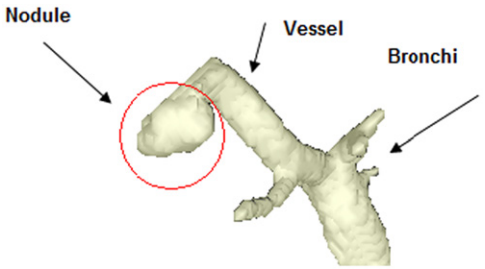

Fig. 5. Image of the surface of two structures: a nodule and the bronchioles.

nodules from those structures. Fig. 5 exemplifies the situation in which a lung nodule is attached to a blood vessel.

The distance transform gives us this key shape information. When applied along the structure resulting from the clustering with GNG and 3D region growing, the distance transform maps each voxel of the volume with the distance information with respect to the edge. With this depth information, it is possible to separate spherical structures (resembling the shape of nodules) from cylindrical ones (resembling the shape of vessels and bronchi).

The 3D distance transform [22] is calculated in each of the structures segmented in the previous stage, starting with value one, in all voxels of the edge, and is incremented as we move towards the inner voxels, until there is one or two voxels left at the same distance from the edge. In other words, voxels of a same layer have the same value for the distance transform. The calculation ends when we reach the most inner layer.

Tubular and elongated structures have constant values of the distance transform with small intervals along their greater length. An example of this statement is given in Fig. 6, where each voxel of a 2D projection of a volumetric structure has one value for the distance transform. We can also notice that structures with large concentration of nodules, such as nodules, have a high value of the distance transform. This large value highlights the spherical nature of the lung nodule.

Since the distance transform is calculated with respect to the edge of the structure, the values of maximum concentration of voxels are in the center of the several structures found in the bronchial and vascular trees. These points are possible lung nodules. Hence, the local maximums are found along all structures of the bronchial and vascular trees and used as seeds in the process of segmentation through 3D region growing. Fig. 6(a) exhibits two voxels, in red and blue, representing the local maximums found in the structure. However, depending on the geometry of the segmented structures, local maximums can appear in conglomerates and, in this case, they are treated as a single local maximum, but all of them are seeds of the region growing algorithm.

The separation of lung nodules from vessels, bronchi, etc., is done by exploring the feature of cylindrical shape of vessels and bronchi. They have a constant value along their length, and a simple 3D region growing is capable of isolating a nodule from structures connecting to it. The only requirement is to aggregate the voxels starting at the seed (local maximums of the distance transform) until the voxels on the edge, with values equal to one, with no repetition of the value of the distance transform, between the value of the seed and the voxel that will be aggregated. Fig. 6(a)–(c) demonstrates the progress of the application of the distance transform. Each seed also originates a particular structure, as illustrated in Fig. 6(d), and, in the regions where there is no growing, the breakage occurs.

At this point, this separation is done without knowing the classification of the structures (whether the structure is a nodule

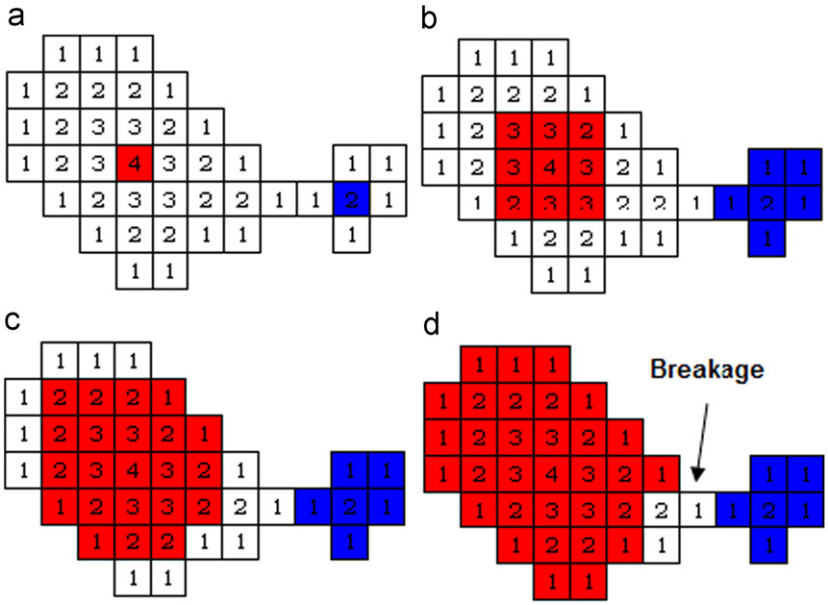

Fig. 6. Stages of the process of separation of lung nodules from other connected structures. (a) Local maximums of the structures are determined and become seeds. (b) Progress by the inner layer of the region growing. (c) Progress by the inner layer. (d) Both seeds originated structures, and the breakage occurs where there is no growing. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

or not), because this information is not required. The application of this method simply breaks down the whole bronchial tree into portions with high concentration of voxels, separating them from the most tubular regions. The result is a variety of separated structures with diverse shapes and sizes. This result may increase the complexity of the classification in the next stage of the methodology. To solve this problem, we apply a cut-off threshold. With this threshold, only the structures that have global maximum above value seven are identified and passed to the next stage. This threshold was determined empirically, by analyzing all of the nodules in the LIDC base selected to validate the methodology.

3.3. Extraction of features

The differentiation of lung nodules from other structures is performed through geometric features like spherical disproportion, spherical density, pondered radial distance, sphericity, elongation and Boyce-Clark radial shape index, and through texture measurements obtained from the Haralick features (contrast, energy, entropy, homogeneity, and momentum) and the histogram of the structures (mean, standard deviation, obliquity, kurtosis, energy, and entropy). The equations and details for all these measurements can be found in [22] and in Appendix A. All the measurements were scaled in the interval from -1 to 1, so they all have the same importance in the classification.

3.4. Classification of structures

To classify the structures coming from the previous stage of the methodology, we use the support vector machine (SVM) classifier [35,33]. The classification with SVM is performed through measurements extracted from geometry and texture. The core function used in the methodology is the radial basis function (RBF). In this work, we used the LIBSVM [36] library to classify candidates as nodule or non-nodule.

A training set was used for training the SVM classifier, and the remaining cases were used for testing and subsequent validation of the methodology. From the training set, it is possible to estimate the parameters C and γ of the RBF, which are applied to the training in order to generate the model used in the classification of the test cases.

The validation of the classification is given by the sensitivity (Se), specificity (Sp), accuracy (Ac), and also by the rate of false positives per exam (FP/exam) and per image (FP/image) [33].

4. Results and discussion

4.1. Materials

The image base used for validating the proposed methodology is from the Lung Imaging Database Consortium (LIDC) [37]. This base has 84 exams available. A nodule in this base is identified by four specialists who analyze the exam. Each specialist classifies the nodule as either malignant or benign (on a 5-level scale ranging from moderate to highly suspicious) and identifies it, marking its spatial location and contour on each exam. Contours are indicated for lung nodules larger than 3 mm, whereas only the geometric center is indicated for nodules smaller than 3 mm. Therefore, only 50 out of the 84 exams have contour information, and are used in the evaluation stage of the proposed methodology. In order to increase the number of nodules evaluated, we took into consideration all nodules whose contour was marked by at least one specialist (agreement level 1); i.e., if one specialist identified a region as a nodule, then this region will be part of our nodule list. In summary, we used 50 exams (gold standard) with 198 nodules identified by specialists to evaluate our methodology.

4.2. Definition of parameters

The parameters empirically determined and used in the GNG algorithm for separating the structures internal to the lung are maximum number of nodes equal to 2, number of iterations to insert a new node (λ) equal to 600, decay rate of the error for units with larger accumulated error (α) equal to 0.5, error decay rate for the remaining units (β) equal to 0.0005, maximum age of a connection ($a_{\max}$) equal to 88, winner's learning rate (ϵ_b) equal to 0.5, and learning rate of the neighbor connected to the winner (ϵ_n) equal to 0.005.

The number of iterations used in the segmentation was 100,000. However, the status of the network that had the smallest distortion usually occurred just before the maximum number of iterations was reached. This happens because of the value of the winner's learning rate, which makes it move at half the distance between the closest nodes.

It is important to stress that it was necessary to adjust the parameter C of the SVM algorithm because of sample misbalance. This means that the error variable will have a higher penalty when the nodules are incorrectly classified. The value of C was adjusted to 0.9, and the γ of the RBF function was adjusted to 0.0078125.

The total number of measurements used in the stage of classification of the segmented structures was 20, being five of co-occurrence matrix (contrast, energy, entropy, homogeneity, and momentum), seven of histogram (mean, standard deviation, obliquity, kurtosis, energy, and entropy), and eight of shape (spherical disproportion, spherical density, pondered radial distance, sphericity, elongation, volume, number of voxels of the edge, and Boyce-Clark radial shape index).

4.3. Application of the proposed method

The stages of the methodology concerning the reduction of the volume of interest and the application of the GNG algorithm were performed on the 50 exams. Fig. 8 illustrates the result of the application of these stages to a slice. After these stages were applied, 29 exams with 48 nodules remained. This is because in some exams the method was unable to precisely segment the lung parenchyma and nodules of small dimensions or very low densities.

In the stage of separation of the lung nodule from other structures, some errors occurred when the lung nodule was very large (Fig. 15) or very small. When a nodule is very large, it hardly has spherical shape and its texture is mistaken for other structures. When it is very small, a nodule resembles the continuation of other structures, like a blood vessel. In both cases, the separation of nodules from other structures was not successful.

Another fact that led the methodology to fail in some cases was the value of the filter based on the distance transform defined on tests. With a value of seven for the filter, some very small structures were not considered in the next stage (classification as nodule and non-nodule).

In order to validate the classification of the structures identified in the previous stage as nodules or non-nodules, we used 10-fold cross-validation. In the segmentation stage, using GNG, 48 nodules were found, as well as several structures considered non-nodules. To be precise, for each nodule identified, 35 non-nodules were found. This set was randomly divided into 30% for the model estimation (training) subset and 70% for the test subset. These sets were stratified in order to ensure that they always included nodule cases. This procedure was performed 10 times, and the average sensitivity, specificity and accuracy measures were obtained. To better analyze the behavior of this stage of the method, we used other training and validation subsets with percentages of 40/60, 50/50, 60/40 and 70/30, repeating the same form of evaluation. Table 1 shows the result of the tests for validating the classification of our method.

The average time needed to detect the nodules varied from 15 to 25 min. This time is almost double or triple the time taken by the detection of nodules in the work of da Silva Sousa et al. [22], 8 min. However, the image base used was different. The number of images used in this work is 2–4 times superior to the number of images used by [22], and the voxels have various dimensions, which explain the longer time taken by this methodology to perform the same task.

Table 1
Validation of the classification as nodule and non-nodule for 29 exams.

Train/test (%)	Average value of tests (%)		
	Sensitivity	Specificity	Accuracy
30/70	79.74	88.67	88.43
40/60	76.25	91.73	91.31
50/50	80.56	91.39	91.10
60/40	85.00	90.79	90.63
70/30	**85.93**	**90.79**	**90.65**
Standard deviation	**3.98**	**1.19**	**1.15**

4.4. Comparisons with other works

In order to put the quality of our work into perspective, Table 2 compares the method proposed herein to other methods for detecting lung nodules in CT images. Because the standard deviation among the results described in Table 1 was not significant, we opted for using the best result, i.e., the 70/30 train and test percentage, in the comparison with other works (Table 2). It is important to stress that a fair comparison of the cited methodologies (Section 2) would only be possible if all of the works had used the same group of images. Besides, there should be some standard parameters, such as resolution, bits per pixel, etc. Another factor

Table 2
Comparison of the performance of works about detection of nodules using the LIDC.

Works	LIDC	Se (%)	FP/image	FP/exam
Proposed method	**29 exams** **4949 images** **48 nodules**	**85.93**	**0.00082**	**0.138**
Ozekes et al. [18]	12 exams 276 images 153 nodules	93.4	0.594	–
Lee et al. [23]	32 exams 5721 images	100	1.4	–
Ozekes et al. [19]	16 exams 425 images	100	–	13.4
Pu et al. [20]	52 exams	81.5	–	6.5
Dolejsi et al. [38]	38 exams 223.6 images/exam	89.6	12.03	–
Messay et al. [24]	84 exams	82.66	–	3
Tan et al. [25]	399 exams	87.5	–	4
Opfer and Wiemker [26]	93 exams	89% $\geq$ 4 mm 60% $<$ 4 mm	–	2
Xiaomin et al. [27]	30 exams	100	–	8.4

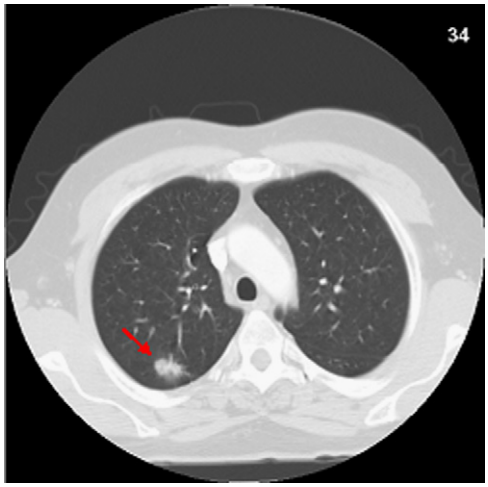

Fig. 7. Slice of a CT exam with the lung nodule indicated by the red arrow. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

that should be common to the works is the sample used, because all the methodologies should have used the same data for the training and testing stages.

4.5. Case studies

Below are some examples of cases, one in which the detection was correct, and two in which the detection was incorrect.

4.5.1. Correct detection

Correct detection occurs when all nodules of an exam are found in the stages of segmentation and classification. For this correct detection case, we demonstrate all the steps described in the methodology for the detection of lung nodules.

The exam in which the correct detection occurred contained just one lung nodule, as we can see in Fig. 7.

The first stage of nodule detection is the reduction of the VOI. Fig. 8(a) shows the result of the reduction of the VOI of the same image shown in Fig. 7.

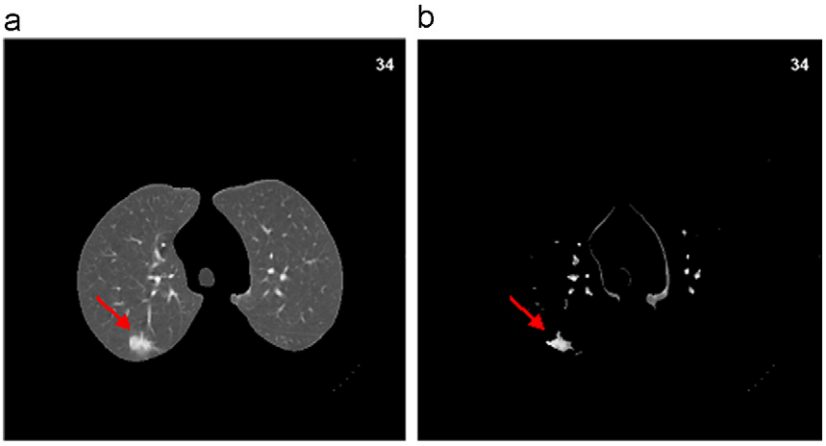

Fig. 8. (a) Picture of the lung of Fig. 7 after reduction of the VOI. (b) Picture of the same lung after application of GNG.

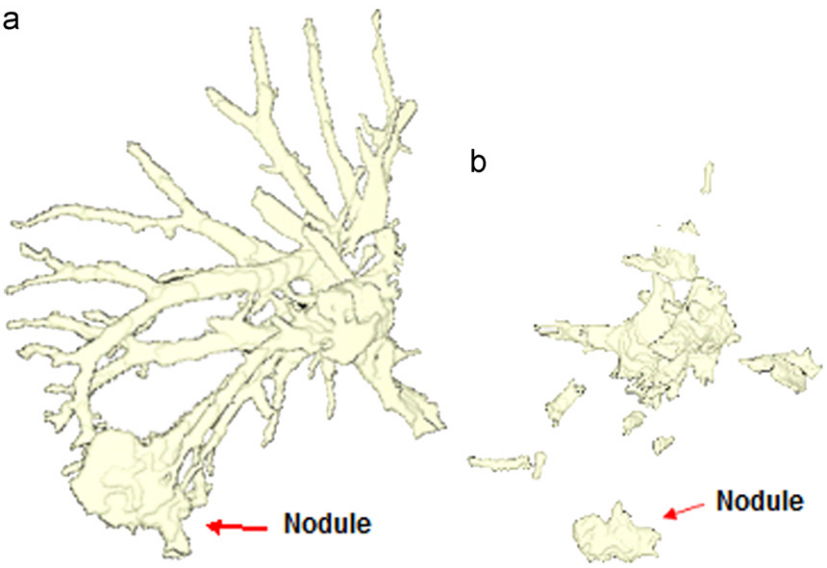

Fig. 9. (a) 3D picture of the surface of the nodule and the structures connecting to it. (b) Result of the separation of the structures by the distance transform.

The next stage is the application of GNG, so to segment only the structures inside the parenchyma. Fig. 8(b) illustrates the result of the application of GNG. As we can see in this same image, besides the nodule, other structures with similar density were segmented.

The next stage is the separation of the nodule from the structures connecting to it. Fig. 9(a) shows the surface of the same nodule shown in Fig. 7 connected to vessels. The result of the separation of the structures shown in Fig. 9(a) can be seen in Fig. 9(b).

The number of segmented structures was 52, being one of the structures the nodule, and the 51 remaining structures, non-nodules. All structures were correctly classified with SVM.

Fig. 10(a) shows the segmented nodule in red in one of the images of the volume. An example of the surface of the segmented nodule is given in Fig. 10(b). The same picture exemplifies the result of the separation of the nodule from the structures.

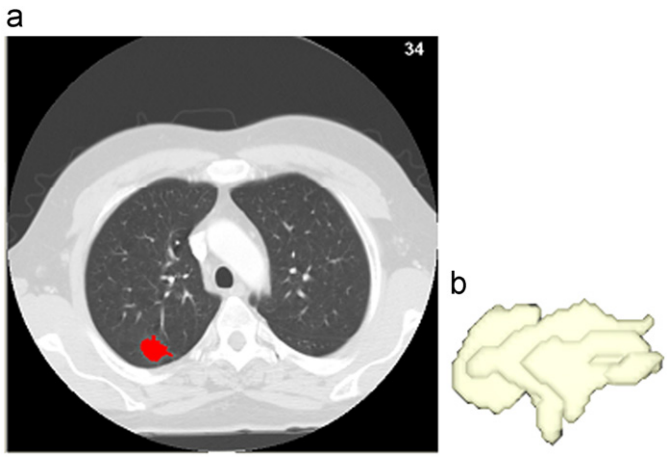

Fig. 10. (a) Slice of a CT exam with nodule in red. (b) Surface of the nodule segmented by the detection methodology. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

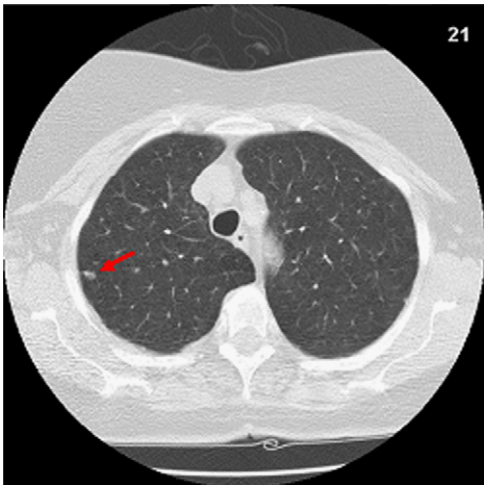

Fig. 11. Slice of a CT exam with its only nodule indicated by the red arrow. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

4.5.2. Incorrect detection

The first case of incorrect detection occurred when the segmented structure did not correspond to the nodule existing in the exam. The exam contained one single nodule, and the segmentation originated 88 structures, being 1 nodule and 87 non-nodules. Two of the 88 structures found were wrongly classified as nodules, and the actual nodule, indicated by the arrow in Fig. 11, was not found. Fig. 12(a) and (b) shows the structures wrongly classified as nodules, and Fig. 13 shows the surfaces of the false positives found.

In the second case of incorrect detection, the nodule was not detected in the classification. The exam contained two lung nodules, and they can be seen in Fig. 14(a) and (b), indicated by red arrows.

Fig. 15 shows the 3D surface of the structures found in the segmentation. In this picture, we can view a variety of structures whose shape is similar to that of the lung nodule.

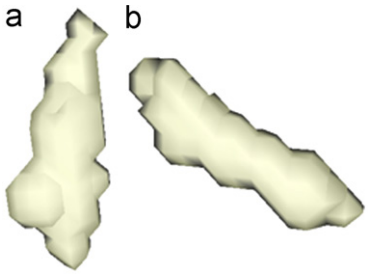

Fig. 13. 3D surface of the structures wrongly classified as nodules. (a) Structure in Fig. 12(a) segmented. (b) Structure in Fig. 12(b) segmented.

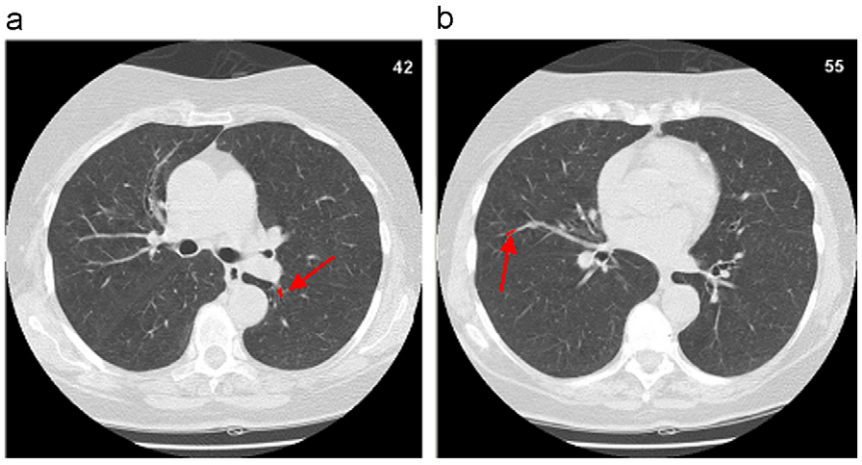

Fig. 12. (a) Slice of a CT exam with the first false positive in red. (b) Slice of a CT exam with the second false positive in red. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

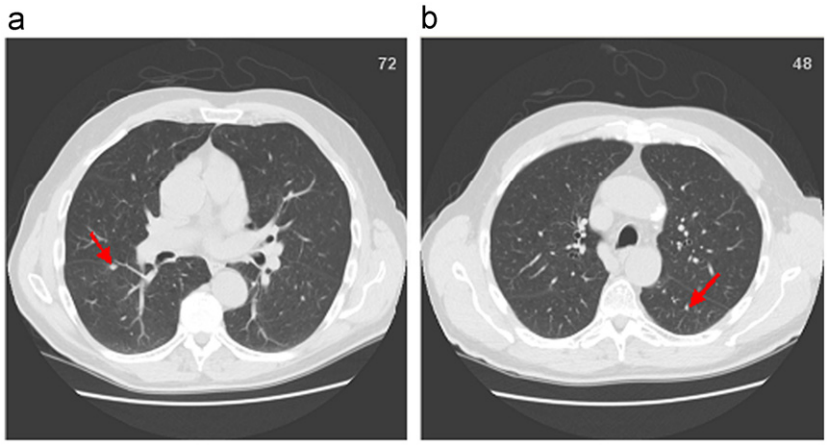

Fig. 14. (a) Slice of a CT exam with the first nodule indicated by a red arrow. (b) Slice of a CT exam with the second nodule indicated by a red arrow. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

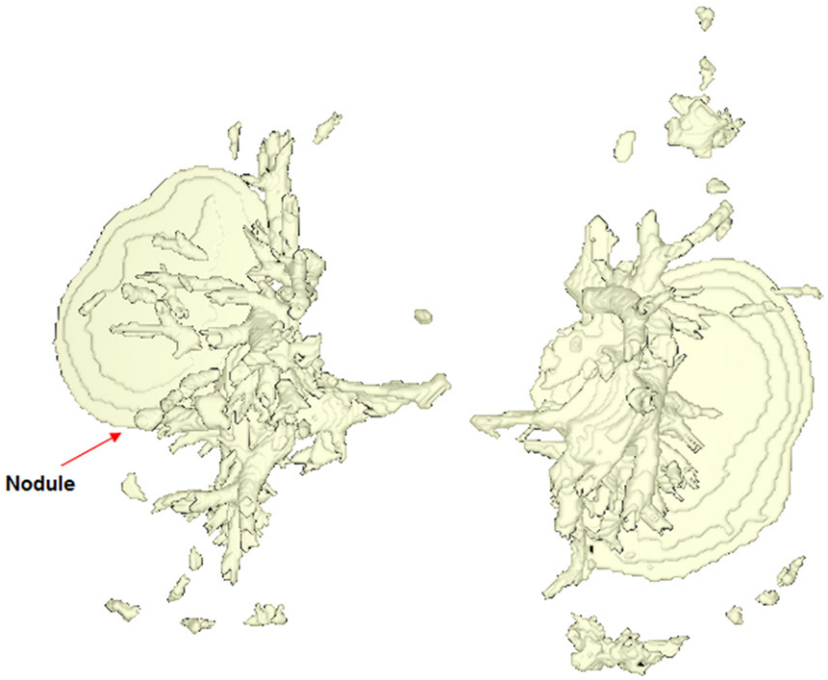

Fig. 15. Image of the 3D surface of the structures segmented in the exam of Fig. 14(a) and (b).

5. Conclusion

This work presented a methodology for automatic detection of lung nodules. The stages involved, from obtaining exams to indicating classified nodules, as well as the results obtained, demonstrate that this is a promising methodology.

Besides automatic, the methodology proposed herein is based on simple techniques, many of which are of easy implementation and usually fast. This makes the proposed methodology to be considered as fast and efficient for the processing of large volumes of data, always taking into consideration the time constraints inherent to any process that involves human health.

Besides, the methodology is composed by successive stages that gradually produce the final result of the detection. The low coupling of these stages allows their easy expansion to improve results or even treat different cases and situations.

The methodology showed to be sensible to the classification of structures due to the dimensions of the structures in the lung. We also noticed that the large number of structures generated in the segmentation stage and their similar shape and texture features make the classification more difficult. This problem must be solved, because the misbalance between the classes nodule and non-nodule makes this task difficult. Almost all the structures segmented and separated generate more elementary structures, which confuses the SVM algorithm in the classification task.

The testing stage resulted in 0.138 false positives per exam. The best sensitivity of the classification obtained with respect to the candidate nodules was 85.93%. The best specificity found was 90.79%, and the best accuracy of the method was 90.65%.

The rates discussed demonstrate that there is technical viability for the implementation of the methodology. With respect to the reasons for using it, the statistics related to lung cancer clearly indicate that this method can be helpful in the early diagnosis of the disease. Due to the high rates of sensitivity per exam, this tool can be useful in the screening process, that is, as part of the first set of exams to be requested, which indicate suspicious cases but need to be confirmed later by medical analyses.

Similar to what occurs in most parts of the world, lung cancer has a considerable prevalence in Brazil. It is the second most common type of cancer among men and the fourth among women, in many parts of the country. Early diagnosis represents a considerable increase in the patients' chance of survival. The proposed methodology contributes with it, showing to be a useful tool for specialists seeking to anticipate the detection of nodules.

Another point to be mentioned is that the public hospital network in some places suffers from lack of specialists, and the funding for increasing the number of involved people is limited. Redirecting qualified professionals to work on less repetitive tasks might reflect in better use of their abilities. One of the steps in this direction would be to use the proposed methodology for a preliminary analysis of CT exams, with a physician in charge of just validating the results.

Finally, the proposed methodology is also a financially attractive solution, because it runs on ordinary computers, which are available in many hospitals.

Conflict of interest statement

It can be observed that are no conflict of interest in the paper revision.

Acknowledgments

The authors acknowledge CAPES, CNPq, and FAPEMA for financial support.

Appendix A. Geometric and texture features

The measures below were obtained from da Silva Sousa et al. [22].

A.1. Geometric features

A.1.1. Spherical disproportion

$$D = \frac{A}{4\pi R^2} \tag{A.1}$$

$$R = \sqrt[3]{\frac{3V}{4\pi}} \tag{A.2}$$

where R is the radius of the sphere with the same volume V of the object obtained through Eq. (A.2). A is the area of the object.

A.1.2. Spherical density

$$E = \frac{100 \cdot n}{V} \tag{A.3}$$

where n is the number of voxels with tone value $p_{x,y,z}$ such that $p_{x,y,z} \neq 0$ and $(x-M_x)^2+(x-M_x)^2+(x-M_x)^2 \leq R^2$, V is the volume of the object, R is the estimated radius obtained through Eq. (A.2), and R is the coordinate of the M_x,M_y,M_z mass center of the object.

A.1.3. Pondered radial distance

$$Drp = R^{-1}\sum_{x,y,z} c_{x,y,z} \tag{A.4}$$

$$c_{x,y,z} = \frac{3}{4\pi}[(r_{x,y,z}+0.5)^3-(r_{x,y,z}-0.5)^3]^{-1} \tag{A.5}$$

where R is the estimated radius obtained through Eq. (A.2), $c_{x,y,z}$ is the pondering coefficient applied to each voxel through Eq. (A.5), and $r_{x,y,z}$ is the radial distance of the voxels with coordinates x,y,z.

A.1.4. Sphericity

$$Es = (6V^{\frac{2}{3}}\pi^3)A^{-1} \tag{A.6}$$

where V is the volume of the object, and A is its area.

A.1.5. Elongation

$$El = Ar_{\min}/Ar_{\max} \tag{A.7}$$

where $Ar_{\min}$ is the measurement of the smaller corner of the minimal box, while $Ar_{\max}$ is the measurement of the bigger corner.

A.1.6. Boyce-Clark radial shape index

$$Bc = \sum_{i=1}^{n}\left|\left(100 \cdot r_i \Big/ \sum_{k=1}^{n} r_k\right)-100/n\right| \tag{A.8}$$

where n is the number of voxels in the bounds of the volume, and r_i or r_k are the distances of specific border voxels to the mass center of the object.

A.2. Haralick features

$$Co(i,j) = \{quant(P_{x,y,z}, v(P_{x,y,z},\alpha))|P_{x,y,z} = i, v(P_{x,y,z},\alpha) = j, \alpha \in \{1,\ldots,26\}\} \tag{A.9}$$

where v is the function that gives one of the 26 tridimensional neighbors of a voxel according to the index α, and $P_{x,y,z}$ is the value of a voxel of coordinates x,y,z.

A.2.1. Contrast

$$Con = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} Co(i,j).(i-j)^2 \tag{A.10}$$

A.2.2. Energy

$$Ene = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} Co(i,j)^2 \tag{A.11}$$

A.2.3. Entropy

$$Ent = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} -Co(i,j) \cdot \log(Co(i,j)) \tag{A.12}$$

A.2.4. Homogeneity

$$Hom = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \frac{1}{1+(i-j)^2} \cdot Co(i,j) \tag{A.13}$$

A.2.5. Moment

$$Mom = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \frac{(Co(i,j))^2}{1+|i-j|} \tag{A.14}$$

References

[1] I.A. for Research on Cancer, Lung Cancer Incidence and Mortality Worldwide in 2008, 2012. URL ⟨http://globocan.iarc.fr/factsheets/cancers/lung.asp⟩.
[2] N.I. of Cancer (INCA), Estimativa 2012: incidência de câncer no Brasil, 2012. URL ⟨http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/pulmao/definicao⟨.
[3] S. Jamnik, I.L. Santoro, C. Uehara, Comparative study of prognostic factors among longer and shorter survival patients with bronchogenic carcinoma,

Pneumologia 28 (5) (2002) 245–249, http://dx.doi.org/10.1590/S0102-35862002000500002, ISSN: 0102-3586 (online).
[4] C.A. da Silva, A.C. Silva, S.M.B. Netto, A.C. de Paiva, G.B. Junior, R.A. Nunes, Lung nodules classification in CT images using Simpson's index, geometrical measures and one-class SVM, in: MLDM, 2009, pp. 810–822.
[5] C. Schneider, A. Amjadi, A. Richter, M. Fiebich, Automated lung nodule detection and segmentation, in: Proceeding of the SPIE, 72601T, vol. 7260, 2009, http://dx.doi.org.10.1117/12.811985. URL 〈http://anode09.isi.uu.nl/results/Schn09.pdf〉.
[6] B. Sahiner, H.-P. Chan, L.M. Hadjiiski, P.N. Cascade, E.A. Kazerooni, A.R. Chughtai, C. Poopat, T. Song, L. Frank, J. Stojanovska, A. Attili, Effect of CAD on Radiologists' detection of lung nodules on thoracic ct scans: analysis of an observer performance study by nodule size, Acad. Radiol.y 16 (12) (2009) 1518–1530, http://dx.doi.org/10.1016/j.acra.2009.08.006. (ISSN 1076-6332).
[7] R. Yuan, P.M. Vos, P.L. Cooperberg, Computer-aided detection in screening CT for pulmonary nodules, Am. J. Roentgenol. 186 (5) (2006) 1280–1287, http://dx.doi.org/10.2214/AJR.04.1969. URL 〈http://www.ajronline.org/cgi/content/abstract/186/5/1280〉.
[8] K. Marten, C. Engelke, Computer-aided detection and automated CT volumetry of pulmonary nodules, Eur. Radiol. 17 (2007) 888–901, http://dx.doi.org/10.1007/s00330-006-0410-3, ISSN: 0938-7994.
[9] Q. Li, F. Li, K. Doi, Computerized detection of lung nodules in thin-section CT images by use of selective enhancement filters and an automated rule-based classifier, Acad. Radiol. 15 (2) (2008) 165–175, http://dx.doi.org/10.1016/j.acra.2007.09.018, ISSN: 1076-6332.
[10] S. Matsuoka, Y. Kurihara, K. Yagihashi, H. Niimi, Y. Nakajima, Peripheral solitary pulmonary nodule: CT findings in patients with pulmonary emphysema1, Radiology 235 (1) (2005) 266–273, http://dx.doi.org/10.1148/radiol.2351040674. URL 〈http://radiology.rsna.org/content/235/1/266.abstract〉.
[11] A. Khan, P.G. Herman, P. Vorwerk, P. Stevens, K.A. Rojas, M. Graver, Solitary pulmonary nodules: comparison of classification with standard, thin-section, and reference phantom CT, Radiology 179 (2) (1991) 477–481. URL 〈http://radiology.rsna.org/content/179/2/477.abstract〉.
[12] N.F. Vittitoe, J.A. Baker, C.E. Floyd, Fractal texture analysis in computer-aided diagnosis of solitary pulmonary nodules, Acad. Radiol. 4 (2) (1997) 96–101, http://dx.doi.org/10.1016/S1076-6332(97)80005-0, ISSN: 1076-6332.
[13] K. Awai, K. Murao, A. Ozawa, M. Komi, H. Hayakawa, S. Hori, Y. Nishimura, Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists' detection performance, Radiology 230 (2) (2004) 347–352, http://dx.doi.org/10.1148/radiol.2302030049. URL 〈http://radiology.rsna.org/content/230/2/347.abstract〉.
[14] Y.J. Jeong, C.A. Yi, K.S. Lee, Solitary pulmonary nodules: detection, characterization, and guidance for further diagnostic workup and treatment, Am. J. Roentgenol. 188 (1) (2007) 57–68, http://dx.doi.org/10.2214/AJR.05.2131. URL 〈http://www.ajronline.org/cgi/content/abstract/188/1/57〉.
[15] Computer-aided diagnosis of small pulmonary nodules, Sem. Ultrasound CT MRI 21 (2) (2000) 116–128, ISSN 0887-2171, http://dx.doi.org.10.1016/S0887-2171(00)90018-0, the Solitary Pulmonary Nodule.
[16] M. Antonelli, G. Frosini, B. Lazzerini, F. Marcelloni, Automated detection of pulmonary nodules in CT scans, in: International Conference on Computational Intelligence for Modelling, Control and Automation, vol. 2, 2005, pp. 799–803 〈http://doi.ieeecomputersociety.org/101109CIMCA20051631566〉.
[17] A.E.-B.G. Gimelfarb, R.F.M.A. El-Ghar, Computer aided characterization of the solitary pulmonary nodule using volumetric and contrast enhancement features, Acad. Rodiol. 12 (2005) 1310–1319.
[18] S. Ozekes, Rule based lung region segmentation and nodule detection via genetic algorithm trained template matching, Istanbul Commer. Univ. J. Sci. 6 (11) (2007) 17–30.
[19] S. Ozekes, O. Osman, O.N. Ucan, Nodule detection in a lung region that's segmented with using genetic cellular neural networks and 3D template matching with fuzzy rule based thresholding, Korean J. Radiol. 9 (2008) 1–9.
[20] J. Pu, B. Zheng, J.K. Leader, X.-H. Wang, D. Gur, An automated CT based lung nodule detection scheme using geometric analysis of signed distance field, Med. Phys. 35 (8) (2008) 3453–3461. ISSN: 0094-2405. URL 〈http://www.biomedsearch.com/nih/automated-CT-based-lung-nodule/18777905.html〉.
[21] X. Ye, G. Beddoe, G. Slabaugh, Graph cut-based automatic segmentation of lung nodules using shape, intensity, and spatial features, in: The Second International Workshop on Pulmonary Image Analysis, 2009, pp. 103–113.
[22] J.a.R.F. da Silva Sousa, A.C. Silva, A.C. de Paiva, R.A. Nunes, Methodology for automatic detection of lung nodules in computerized tomography images, Comput. Methods Prog. Biomed. 98 (1) (2010) 1–14, http://dx.doi.org/10.1016/j.cmpb.2009.07.006, ISSN: 0169-2607.
[23] S. Lee, A. Kouzani, E. Hu, Automated identification of lung nodules, in: 2008 IEEE Tenth Workshop on Multimedia Signal Processing, 2008, pp. 497–502, http://dx.doi.org.10.1109/MMSP.2008.4665129.
[24] T. Messay, R.C. Hardie, S.K. Rogers, A new computationally efficient CAD system for pulmonary nodule detection in CT imagery, Med. Image Anal. 14 (3) (2010) 390–406.
[25] M. Tan, R. Deklerck, B. Jansen, M. Bister, J. Cornelis, A novel computer-aided lung nodule detection system for CT images, Med. Phys. 38 (10) (2011) 5630–5645, http://dx.doi.org/10.1118/1.3633941. ISSN: 0094-2405. URL 〈http://www.ncbi.nlm.nih.gov/pubmed/21992380〉.
[26] R. Opfer, R. Wiemker, Performance analysis for computer-aided lung nodule detection on LIDC data, in: Proceedings of the SPIE, vol. 6515, 2007, pp. 65151C–65151C-9. URL 〈http://link.aip.org/link/PSISDG/v6515/i1/p65151C/s1&Agg=doi〉.
[27] P. Xiaomin, G. Hongyu, D. Jianping, Computerized detection of lung nodules in CT images by use of multiscale filters and geometrical constraint region growing, in: 2010 Fourth International Conference on Bioinformatics and Biomedical Engineering (iCBBE), 2010, pp. 1–4, ISSN: 2151-7614, http://dx.doi.org.10.1109/ICBBE.2010.5517771.
[28] Suárez-Cuenca, W. Guo, Q. Li, Automated detection of pulmonary nodules in CT: false positive reduction by combining multiple classifiers, in: Proceedings of the SPIE, vol. 7963, 2011, pp. 796338–796338-6. URL 〈http://dx.doi.org/10.1117/12.878793〉.
[29] N. Camarlinghi, I. Gori, A. Retico, R. Bellotti, P. Bosco, P. Cerello, G. Gargano, E. Lopez Torres, R. Megna, M. Peccarisi, M. Fantacci, Combination of computer-aided detection algorithms for automatic lung nodule identification, Int. J. Comput. Assisted Radiol. Surg. (2011) 1–10, ISSN: 1861-6410, URL http://dx.doi.org.10.1007/s11548-011-0637-6.
[30] A. Angelopoulou, A. Psarrou, J.G. Rodríguez, K. Revett, Automatic landmarking of 2D medical shapes using the growing neural gas network, in: Proceedings of the First International Conference on Computer Vision for Biomedical Image Applications (CVBIA'05), Springer-Verlag, Berlin/Heidelberg, 2005, pp. 210–219, ISBN: 3-540-29411-2, 978-3-540-29411-5, URL http://dx.doi.org.10.1007/11569541_22.
[31] L. Oliveira Martins, A.C. Silva, A.C. De Paiva, M. Gattass, Detection of breast masses in mammogram images using growing neural gas algorithm and Ripley's K function, J. Signal Process. Syst. 55 (1–3) (2009) 77–90. ISSN: 1939-8018 〈http://dx.doi.org/10.1007/s11265-008-0209-3〉.
[32] B. Fritzke, Some Competitive Learning Methods, Technical Report, 1997. URL 〈http://dx.doi.org/10.1016/j.cmpb.2009.07.006〉.
[33] S. Haykin, Neural Networks: A Comprehensive Foundation, Bookman, Porto Alegre, 2001.
[34] S. Iwano, T. Nakamura, Y. Kamioka, T. Ishigaki, Computer-aided diagnosis: a shape classification of pulmonary nodules imaged by high-resolution CT, Comput. Med. Imaging Graph. 29 (7) (2005) 565–570, http://dx.doi.org/10.1016/j.compmedimag.2005.04.009, ISSN: 0895-6111.
[35] V.N. Vapnik, Statistical Learning Theory, Wiley-Interscience, 1998, ISBN: 0471030031.
[36] C. Chang, C. Lin, LIBSVM—A Library for Support Vector Machines, 2012. Available at 〈http://www.csie.ntu.edu.tw/cjlin/libsvm/〉.
[37] S.G. Armato III, G. McLennan, M.F. McNitt-Gray, C.R. Meyer, D. Yankelevitz, D.R. Aberle, C.I. Henschke, E.A. Hoffman, E.A. Kazerooni, H. MacMahon, A.P. Reeves, B.Y. Croft, L.P. Clarke, Lung image database consortium: developing a resource for the medical imaging research community1, Radiology 232 (3) (2004) 739–748, http://dx.doi.org/10.1148/radiol.2323032035. URL 〈http://radiology.rsna.org/content/232/3/739.abstract〉.
[38] M. Dolejsi, J. Kybic, M. Polovincak, S. Tuma, The lung TIME: annotated lung nodule dataset and nodule detection framework, in: Proceedings of the SPIE, 72601U, vol. 7260, 2009, http://dx.doi.org.10.1117/12.811645. URL 〈ftp://cmp.felk.cvut.cz/pub/cmp/articles/dolejsi/Dolejsi-SPIE2009.pdf〉.

Stelmo Magalhães Barros Netto received his bachelor's degree in electric engineering from the Federal University of Maranhão (UFMA), Brazil, in 2008. He is currently a professor at the same university. His research interests include signals and image processing, pattern recognition, and automation systems.

Aristófanes Corrêa Silva received his PhD in informatics from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Brazil, in 2004. He is currently a professor at the Federal University of Maranhão (UFMA), Brazil. He teaches image processing, pattern recognition, and programming language. His research interests include image processing, image understanding, medical image processing, machine vision, artificial intelligence, pattern recognition, and machine learning.

Rodolfo Acatauassú Nunes received his PhD in General Surgery—Thoracic Area from the Federal University of Rio de Janeiro (UFRJ), Brazil, in 1995. He is currently a professor of the General Surgery Department at the State University of Rio de Janeiro (UERJ), Brazil, where he has been in charge of the coordination of the thoracic surgery course since 2000. He belongs to a research group on thoracic oncology, and the quantification of thoracic images applied to general thoracic surgery practice is a specific topic of his interest.

Marcelo Gattass, took his PhD in 1982 from Cornell University and is a full professor at PUC-Rio's Computer Science Department. He currently teaches computer graphics and 3D computer vision for graduate students and data structures for undergraduates. He is also the director of Tecgraf/PUC-Rio, Computer Graphics Technology Laboratory, where he supervises several industry-cooperation projects in the areas of 3D modeling and visualization, geographic information systems, user interfaces, and web-based applications.