journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Methodology for automatic detection of lung nodules in computerized tomography images

João Rodrigo Ferreira da Silva Sousa[a], *Aristófanes Corrêa Silva*[a,*], *Anselmo Cardoso de Paiva*[a], *Rodolfo Acatauassú Nunes*[b]

[a] *Federal University of Maranhão - UFMA, 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*

ARTICLE INFO

Article history:
Received 8 January 2009
Received in revised form
13 July 2009
Accepted 17 July 2009

Keywords:
Medical image
Computer-aided detection (CAD)
Lung nodules
Image processing
Computer tomography (CT)

ABSTRACT

Lung cancer is a disease with significant prevalence in several countries around the world. Its difficult treatment and rapid progression make the mortality rates among people affected by this illness to be very high.

Aiming to offer a computational alternative for helping in detection of nodules, serving as a second opinion to the specialists, this work proposes a totally automatic methodology based on successive detection refining stages.

The automated lung nodules detection scheme consists of six stages: thorax extraction, lung extraction, lung reconstruction, structures extraction, tubular structures elimination, and false positive reduction. In the thorax extraction stage all the artifacts external to the patient's body are discarded. Lung extraction stage is responsible for the identification of the lung parenchyma. The objective of the lung reconstruction stage is to prevent incorrect elimination of portions belonging to the parenchyma. Structures extraction stage comprises the selection of dense structures from inside the lung parenchyma. The next stage, tubular structures elimination eliminates a great part of the pulmonary trees. Finally, the false positive stage selects only structures with great probability to be nodule. Each of the several stages has very specific objectives in detection of particular cases of lung nodules, ensuring good matching rates even in difficult detection situations.

We use 33 exams with diversified diagnosis and slices numbers for validating the methodology. We obtained a false positive per exam rate of 0.42 and false negative rate of 0.15. The total classification sensitivity obtained, measured out of the nodule candidates, was 84.84%. The specificity achieved was 96.15% and the total accuracy of the method was 95.21%.

1. Introduction

Lung cancer has been attracting the attention of medical and scientific communities in the latest years because of its high prevalence allied with the difficult treatment. Statistics from 2008 indicate that lung cancer, throughout world, is the one that attacks the greatest number of people. In Brazil, a total of 27,270 new cases of the disease is estimated for this year, being smoking the main risk factor. The last global estimative, on the other hand, indicated the occurrence of 1,200,000 new cases of the disease only in the year 2000 [1].

* *Corresponding author at*: Federal University of Maranhão - UFMA, Electrical Engineering, Av. dos Portugueses, SN, Campus do Bacanga, Bacanga 65085-580, São Luís, MA, Brazil. Tel.: +55 98 21098832; fax: +55 98 21098841.
E-mail addresses: j.rodrigo.sousa@gmail.com (J.R.F. da Silva Sousa), ari@dee.ufma.br (A.C. Silva), paiva@deinf.ufma.br (A.C. de Paiva), rodolfoacatauassu@yahoo.com.br (R.A. Nunes).
0169-2607/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2009.07.006

More globally, the mortality rate related to lung cancer has been decreasing, but it is still high. Researches indicate that the 5-year survival of diagnosed patients varies between 13% and 21% in developed countries, but in emergent countries, like Brazil, it lies between 7% and 10% [1].

Identifying the disease still on its initial stages, can elevate the patient's 5-year survival chance up to 70% [2]. This way, it is known that precocious diagnosis is the best way to ensure the success of the treatment. Thus, there is a growing demand for methods that improve the existing mechanisms to aid it.

One of the main imaging modality for this kind of disease diagnostic is the computerized tomography (CT) of the patient's chest. In moderns CT scans, CT analysis demands great effort from the specialist, due to the great number of slices and fatigue becomes a recognized limitation to the early diagnosis.

One important aspect of the problem consists in the fact that the specialist diagnosis is mainly based on morphologic characteristics of the structure under analysis which can be verified only in 3D space. But the analysis of the result of a CT is performed, traditionally, through bi-dimensional images, representing the slices of the patient's body.

This tradeoff between what the specialist needs to observe and what is shown to him ends up requiring a mental reconstruction of the tri-dimensional aspects of the tissues under analysis, task which, besides complex and slow, can fall into significant mistakes.

This scenario demonstrates that there is a great demand for computational systems which can help with the tasks of detection and diagnosis of lung nodules, reason for which the number of published papers related to this matter increases every year [3].

Nevertheless, several challenges are involved in automatic segmentation of lung nodules. The first is related to the quality of the images. The use of images from old tomographs, with noises and low resolution, is very common. Besides, there is the problem of lack of standardization, as much in the file formats as in setup of parameters of the tomograph, generating exams with differing characteristics.

The most difficult problem, although, is the transfer of knowledge involved in the identification of a lung nodule to an automatic process. This involves, among other aspects, determining the numerically measurable characteristics to describe portions of an image, distinguishing those that are nodules from those that are not.

The task of distinguishing nodules from normal structures can also be especially difficult because, under the computational viewpoint, the nodules can be very similar to other elements of the lung parenchyma. Examples of this are the blood vessels, in which rates of absorption of X-rays (which origin the intensity of pixels in the exam) are comparable to those of the nodules.

Regarding morphology, the nodules can have the most diverse aspects: round, flat or spiculate. About the location, they occur all inside the lungs, inclusively on the peripheral area. This way, being aggregated to other normal structures makes difficult not only its detection, but the correct delimitation of its bounds, since it ends up being confused with elements around it.

Targeting the challenges above described, this work proposes an automatic methodology for the detection of lung nodules. The proposed methodology aims to contribute to the researches of mechanisms for diagnosis aiding, through the search for matching rates which are greater and more consistent than those presently obtained. Among the characteristics responsible for the high performance of this methodology, we have the progressive refining of nodule detection, which means that specific problems are solved in every stage, in order to achieve the objective of detection. Another important aspect of the proposed methodology is that it is complete. This is related to the fact that even in the most difficult cases, which are when the nodules are linked to the chest wall, aggregated to the bronchi or blood trees, or are the single ones, we achieved correct results. This result came from the implementation of specific strategies to detect and solve these specifics problems.

In addition, this work aims to give, besides correct results from the medical viewpoint, coherence in the obtained detections. As it concerns an automatic process with deterministic behavior, interference of the operator experience and precision on detection is avoided. This way, as the process does not depend on external parameters or user participation, its results must be consistent for the same image or image category. That is, detection results are expected to be reproducible.

Reproducibility is an important characteristic because it gives assurance to the results of segmentation, as it decreases subjectivism. If the specialist obtains different results each time he looks for automatic detection, he will lose confidence in the method because he would not be able to base the treatment on inconsistent information, which is under constant modification. On the other hand, reproducible results collaborate to acquisition of historical information and tendency analysis. Through the comparison of data from the lesion, obtained in different moments, the specialist obtains information which can have diagnostic as well as therapeutical implications.

This paper is organized in the following way: Section 2 presents a series of works concerning the matter of automatic detection of lung nodules and aims to demonstrate the present development stage of researches in this field. Section 3 describes the central points of the methodology being composed by subsections which correspond to its stages. Section 4, the contributions of the work are listed as well as the results we obtained. Finally, Section 5 presents the conclusions.

2. Related works

The development of medical images acquisition techniques, in particular computerized tomography (CT), which may furnish more detailed information about the human body, has increased the capability and fidelity in the diagnosing of many diseases. On the other hand, the dimensions of these images are becoming bigger and bigger, increasing the need for computer vision techniques that can make interpretation easier. This Section aims to provide an overview of literature in automatic CT image analysis in the lung region.

The work of Beigelman-Aubry et al. [4] presented evaluation of nodule detection and its response time when

performed by radiologists with and without use of a computerized system. The work showed that the system improves the sensibility of the detection, what raised the trust interval in 2%. Among the experiments with 109 patients, there was a nodule which was not detected by one of the radiologists, but was detected by the system. Besides, the use of the system decreases considerably the time required by the specialists to analyze the exams.

This way, nodule detection systems, have great importance in this process, despite they do not give the final diagnosis.

Nodule detection systems usually involve four steps: pre-processing, extraction of nodule candidates, reduction of false positives and classification. Pre-processing normally consists in restricting the search space, delimiting the lung, and reducing noises in the image. The region of the lung is segmented and nodule candidate objects are identified. Among these objects most of the non-nodule are discarded in the false positive reduction stage. The remaining objects are then classified into nodule and non-nodule. In some methods, the false positive reduction is performed after classification. Some works found in the literature involving these steps are presented next.

Armato and Sensakovic [5] showed the importance of adequate segmentation of lungs in computer aided detection and/or diagnosing systems. His studies indicated that up to 17% of lung nodules can be lost during lung segmentation if the algorithm is not adjusted to the task of nodule detection.

A great challenge is the segmentation of lungs affected by high density pathologies connected to their bounds. Due to the lack of contrast between these pathologies and the tissues adjacent to the lung, density-based methods fail in this region. In this case, it is necessary some edition technique, but, even so, part of the lung is normally lost [3].

Due to the large amount of air in the lung, its interior has dark tonality in CT images, differing from the region around it. This way, contrast between lung and neighbor tissues is the basis for most lung segmentation methods. Most methods is based on rules [6–8]. The lung region can be found by two ways [3]. The first one is by means of region growing starting at trachea. The second one, more usual, used thresholdings and restrictions in size and location.

To find nodule candidates, the main techniques used are multiple thresholding [9–12], mathematical morphology [13–16], clustering [17–20], analysis of connected elements in thresholded images [21,22], detection of circles in thresholded images [23] and use of emphasis filter with spherical structure elements [24–26].

In Osman et al. [27], for each slice, regions of interest (ROI) were found by using density values of the pixels and analyzing their eight directions. The joining of all slices formed 3D ROIs, which compared to a nodule model (template) allows identifying the nodules. Sensibility reached 100%, but the test data were restricted to six cases.

Retico et al. [28] proposed a system based on emphasis filters for spherical objects and a neural classification based on voxels of selected regions to reduce false positives. The system performance was evaluated in a set of data from 39 CT and reached 80–85% of sensibility and 10–13 FP/exam.

Bae et al. [29] developed a computer aided diagnosis (CADx) for high-resolution CT images (HRCT - high-resolution computed tomography) using bi-dimensional and tri-dimensional analysis algorithms. This technique was tested in eight lung cancer cases and obtained 95% of sensibility and 0.91 FP/slice.

To improve the sensibility of the detection, Li et al. [30] used an emphasis filter in the identification stage and, to reduce false positives, used a rule-bases classifier.

Having the nodule candidate objects been generated, characteristic features of these objects are calculated. Classifiers are then applied. These classifiers use the features to identify candidate objects either into the nodules set or into the non-nodule set.

Several techniques can be used as classifiers in the final stage of nodule detection: based on either rules or linear classifiers [31,32,24], by combining models (template matching) [33], analysis of the nearest cluster [13,15], support vector machine [34–36], neural networks [37–39] and Bayesian classifier [40,41]. The features mostly used for classification are features based on the density of voxels, description of shapes, spatial relation and size information.

da Silva Sousa et al. [42] proposed a set of three morphological features specially developed for characterization of lung nodules with which matching rates of 100% were achieved using support vector machine, despite this work used a small database.

In some works, the classifier presents good sensibility, but also a high number of false positives. This way, techniques have been looked for, in order to reduce this number after the identification that, in some cases, work as filters before classification.

Armato et al. [9] presented a methodology for the detection of lung nodules with just the pre-processing stages, candidates detection and classification. Nodule candidates were found by multiple thresholding and, next, using shape and density attributes and discriminant linear analysis, the classification detected 70% of the nodules indicated by specialists and three false positives per slice in average (approximately 80–90 false positives per exam). In later papers, Armato and co-authors focus in rules to reduce the number of false positives: rule-based [43,44], discriminant analysis [44,45] and neural networks [44,46]. The best result obtained by these techniques was of 80.3% in detection rate win 4.8 false positives per exam against 27.4 without false positives reduction [46].

Saita et al. [22] added to the nodules detection methodology proposed by Oda et al. [21] a false positives reduction stage.

Lee et al. [47] added the false positive reduction stage to the nodules detection method initially proposed by Lee et al. [31]. To do this, they added five density attributes and adjusted the thresholding parameters to the original model. The sensibility continued the same in 72.4% but the FP rate decreases from 30.8 to 5.5 per exam.

False positives reducing is important, because, even if sensibility keeps unaltered, the radiologist's final amount of work is reduced.

3. Methodology

This section presents, under the form of a sequence of stages, the procedures proposed to perform the detection of lung nodules in a CT in an incremental manner. This incremen-

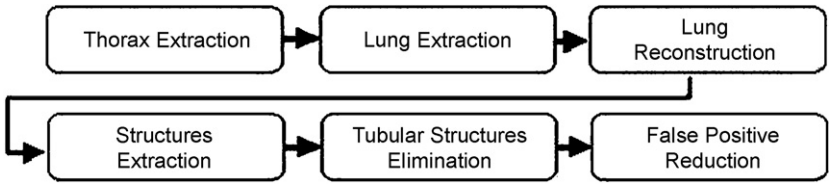

Fig. 1 – Methodology stages.

tal approach is presented in counterpoint to others which try to identify the possible nodules in one single step. Another important aspect of the methodology is the adoption of specific strategies for nodule detection in particular conditions, such as nodules linked to the chest wall, aggregated to the bronchi or blood trees, and the single ones.

The proposed methodology corresponds to the application of several successive stages of processing to CT images, eliminating portions of them which do not correspond to interest areas, in this case, lung nodules. As we can see in Fig. 1, the thorax extraction initially occurs by the use of regions growing [48], described in Section 3.1. In this stage all the artifacts external to the patient's body are discarded.

In the sequence occurs the lung extraction (Section 3.2), responsible for the identification of the lung parenchyma. In this stage the region growing algorithm is used again, being responsible for removing thoracic wall and mediastinum.

The next stage, lung reconstruction (Section 3.3), is important to prevent incorrect elimination of portions belonging to the parenchyma. This stage is strategic for peripheral nodules detection and is fundamentally based on the rolling-ball algorithm.

The next step is the parenchyma structures extraction (Section 3.4). This step comprises the selection of dense structures from inside the lung parenchyma with the use of region growing. After that, the tubular structures elimination stage (Section 3.5) eliminates a great part of the pulmonary trees based on their morphological characteristics.

The final stage is the false positive reduction (Section 3.6). In this stage, SVM classify the structures coming from the previous stage, selecting only those with indicatives of being actual lung nodules. Fig. 2 presents a CT slice consecutively submitted to this process.

3.1. *Thorax extraction*

The process is started with thorax extraction. This stage comprises the removal of all artifacts external to the patient's body, among which are bed sheets, the air that involves him and the surface on which he lies as example of the items numbered in Fig. 2(a).

These structures are identified by a 2D region growing algorithm, that use four seeds for each slice, initiated on its four corners. The similarity criterion for the algorithm is based on

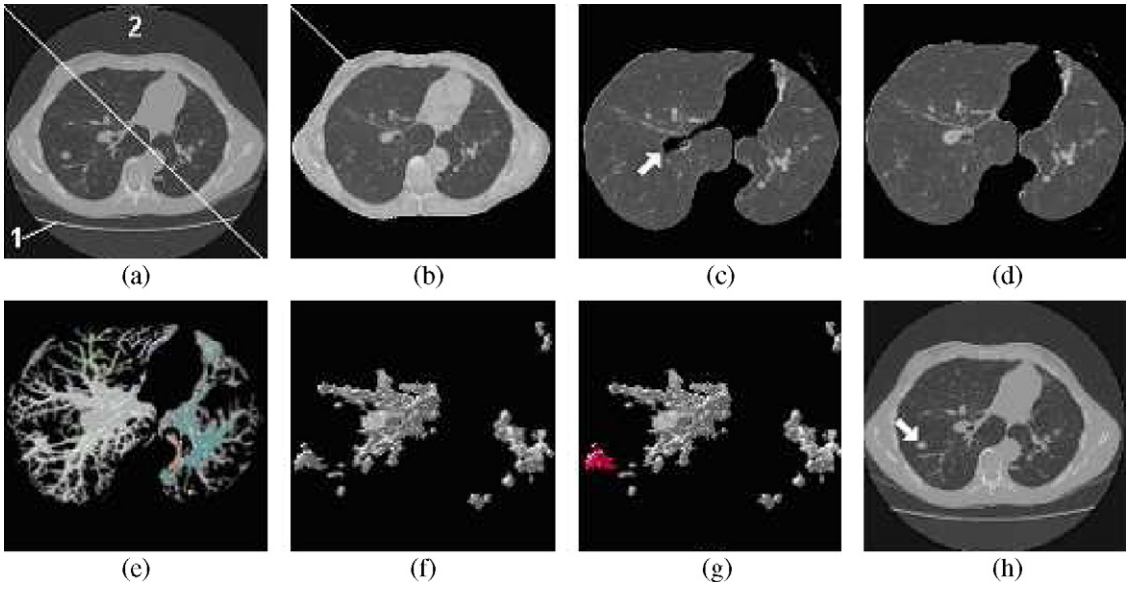

Fig. 2 – Automatic lung nodule detection sequence. (a) Eliminates of all artifacts external to the patient's body, identified with 1 and 2 in the figure. (b) Removal of thorax remaining just the parenchyma. (c) Shows an example of the internal lung region and the thoracic wall erroneously eliminated. (d) Reconstructed parenchyma with rolling-ball algorithm. (e) 3D visualization of the remaining structures after threshold application and identified with different colors. (f) 3D visualization of the structures after tubular elimination. (g) Shows the correct identification of a lung nodule among other normal lung structures which came from the previous stage. (h) Presents the same nodule identified in the original tomography image by an arrow.

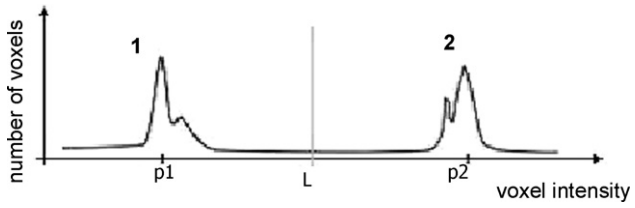

Fig. 3 – Diagonal histogram of a CT exam slice.

gray tones of the voxels, for great part of the external region of the thorax (which we want to identify) is formed by low intensity voxels.

The gray tone range used, however, is dynamically determined through the analysis of the histogram of the voxels that constitute the main diagonals of all slices, cumulatively, as in Fig. 2(a). In a typical histogram of a thorax CT image, seen in Fig. 3, there are two well -defined peaks: peak 1 is formed by low intensity voxels that constitute the lung parenchyma and the thorax external region; peak 2 is formed mainly by voxels from the thoracic wall and mediastinum.

This way, the central point that separates the peaks of the histogram is taken as a threshold and the region growing algorithm will identify all the voxels that have values that are smaller than the threshold, which are associated to the artifacts that we wish to identify. The L threshold is obtained by $L = (p1 + p2)/2$.

The need for association between the region growing and the thresholding algorithms comes from the fact that the first one, by itself, would not be able to distinguish the low intensity voxels from the thorax external region from those belonging to the lung parenchyma, resulting in the elimination of both. Another reason is that even in the thorax external region there are regions with high density, as we can see in the structure identified with number 1 in Fig. 2(a).

This way, the region growing algorithm, having the threshold as the stopping criterion, is able to identify only the thorax external dark voxels, not affecting those belonging to the parenchyma, because there is a range of high intensity voxels which constitute the thoracic wall, avoiding the union of these two regions. That is, even with similar intensity values, the regions are disconnected and the seed put in one region does not propagate to the other.

However, as already said, in the thorax external region there are more dense structures too. In order to treat these cases, the region growing algorithm was altered to allow the identification of these structures. The modification consisted in the inclusion of a spatial tolerance that allows the region growing to penetrate in areas outside the similarity criterion. This tolerance, however, is decreased as the growing goes on, making only relatively affiliated structures to be removed. The extra growing is limited by tolerance, which is the maximum distance that it can expand beyond the normal region. A five-pixel tolerance value was experimentally found.

The result of the application of this detection stage to a CT can be seen in Fig. 2(b), on which the voxels associated to the thorax external region have value zero, which means that these structures will not be considered in the processing of the following steps.

3.2. *Lung extraction*

The objective of lung extraction is to identify the thoracic wall and mediastinum voxels, making possible the work on the next stages just with the region which forms the pulmonary parenchyma. That is achieved again with use of the region growing algorithm, this time, however, identifying the high-intensity voxels with values greater than the threshold and with no need for tolerance.

The threshold value used in the step is calculated with basis on the diagonal histogram, as already commented. For this case only the voxels which were not discarded in the previous stage are considered. The resultant histogram is similar to the one presented in Fig. 3, being the threshold obtained in the same way.

The initial growing seed, as can be seen in Fig. 2(b), is released in the first voxel of the thorax situated over the main diagonal of each slice. The final result, after the growing and elimination of the high intensity voxels can be seen in Fig. 2(c).

The voxels which do not belong to the pulmonary parenchyma area are eliminated by giving them intensity value zero.

3.3. *Lung reconstruction*

Occasionally the lung extraction stage erroneously eliminates some voxels which belong to the pulmonary parenchyma. These mistakes can lead to elimination, inclusively, of possible nodules, inducing an error in detection. This way, the reconstruction stage has great importance for preservation of peripheral nodules.

Fig. 2(c) shows an example of an internal lung region and the thoracic wall erroneously eliminated together, due to its high intensity voxels.

In order to perform the reconstruction of the incorrectly eliminated lung outline, a previous knowledge about the object which is being segmented is used. It is known that the lung is an organ with a soft outline, without re-entrances. This way, any hole or abrupt discontinuity found on its outline is a strong indication of failure in the perimeter and it must be reconstructed.

In order to recover the correct lungs outlines, this stage uses the rolling-ball Algorithm [18], a mathematical morphology technique based on closing operations executed with a circular structuring element, whose radius, in this specific case, was of 30 pixels.

The size of the radius of the structure element is directly related to the maximum size of the irregularities that the rolling-ball algorithm will discard in reconstruction. For too small radiuses, bigger nodules (in the bounds) will not be recognized and will be discarded with the mediastinum. If the radius is too big, the reconstruction will start altering the very bounds of the lung, being even able to join both lungs into one. The ideal size was found through several tests for the exam database used.

3.4. *Parenchyma structures extraction*

The previous stages had the main objective of detecting the pulmonary region, but only in this stage, in fact, the search

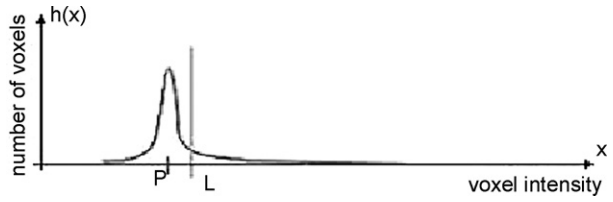

Fig. 4 – Typical histogram of a CT slice just with the pulmonary parenchyma.

for lung-internal regions occurs. This stage is performed in two steps: the first one identifies and removes the less dense parenchyma tissue out from the image, keeping only its internal structures; the second one isolates each of the tri-dimensional structures found so that they can be individually processed.

The elimination of less dense tissues is performed by means of a thresholding process. The proper threshold is once more obtained from the volume voxels histogram, being considered only those parenchyma-internal. This time, however, the histogram graphic shows that there are not well-defined peaks anymore, as shown in Fig. 4. This is the result of the decreasing on the number of high-intensity voxels related to the low-intensity ones.

As consequence, it becomes more difficult to choose the proper threshold for elimination, in such a way that its determination must be stricter than the previous ones.

This way, it was empirically verified that the threshold L is located on the inflection point of the histogram curve, right after the end of the p peak formed by the low-intensity voxels. L and p are represented in Fig. 4.

More formally, having the histogram been normalized between 0 and 100 in such a way that $h(p) = 100$, the L threshold is the smaller value that meets the conditions $L > p$ and $h(L-1) - h(L) \leq 1$. That is, the threshold is the exact point where the peak ends and the planner region that follows it starts. We consider this point as that where the derivative of the histogram reaches value 1. The value of L is sensitive. If it is shifted left just a little, the pulmonary structures get too broad. If it is shifted right they get too affiliated.

Finding this threshold, it is possible to eliminate all the voxels which are below this value, obtaining a result that is similar to the one seen in Fig. 2(e). We can notice on it that structures such as blood vessels, bronchi and nodules are preserved, while the major part of the parenchyma is suppressed.

Nevertheless, these structures need to be separated, individually, before the nodules can be identified. Each tri-dimensionally connected structure is so identified through a region growing algorithm which starts in each voxel of the structures that are not isolated yet. The result of this stage is that every tri-dimensionally connected region can be individually processed from this point. Fig. 2(e) shows each tri-dimensional structure identified with distinguished colors. Each color was randomly chosen and has no special meaning (For interpretation of the references to color in this sentence, the reader is referred to the web version of the article.).

3.5. *Tubular structures elimination*

We observed that among the objects identified by the 3D connectivity property exist structures that correspond to the bronchial and vascular trees. Besides, there are cases where each nodule is connected to one or more of these structures. This creates a problem for the detection of these nodules, generating the need for identifying the bronchial and vascular trees of the pulmonary parenchyma so that distinguishing these trees from possible nodules can be possible.

Blood vessels are, as a rule, tubular. The depth of the medial axis varies very gradually, inclusively in ramifications. In other words, blood vessels have thickness almost constant in a certain location. Nodules have totally different characteristics. As they are compact structures, they present an abrupt increase in the depth of the medial axis. This is perceived more clearly in spiculated nodules. The process consists in verifying to which of both patterns the structures match better. With this objective, observing the structures to be identified, we use an analysis based on their skeleton. This is possible since they resemble very much their medial axis, obtained by means of the 3D skeletonization algorithm proposed in [36].

The idea behind skeleton-based segmentation is to detect regions, or voxel sequences belonging to the medial axis of the structures, which show low variation of the average depth, that is, the average distance from the medial axis in relation to the border of the object. This pattern is characteristic of tubular structures, such as those ones which are necessary to be eliminated. For this we defined thresholds capable of making this distinction. It is important to stress that the first threshold measures the intensity of the thickness variation of the structure along the medial axis, that is, it is the derivative of the thickness in the axle. The second threshold performs the percent relation between the thickness variation and the very thickness in order to estimate its representativeness.

As an example of what has been said, one can notice in Fig. 5(a) that the highlighted branch probably represents a blood vessel. The observation of the average depth of its medial axis on Fig. 5(b) allows us to notice its low variation.

On the other hand, the structure indicated in Fig. 5(c) has great potential to be a nodule aggregated to a vessel. The same analysis on the average depth of the voxels of its skeleton indicates a very different pattern. As in Fig. 5(d), we notice an abrupt increase of values from the nodule border, behavior that can be used for detection.

The bifurcations among the vessels possibly present an increase in the depth of the medial axle, but this increase, besides being small when compared with the diameter of the vessel, is gradual. On the other hand, in the case of aggregated nodules, the increase in the depth of the media axle is much more abrupt and intense. With the correct balance of cutoff thresholds it is possible to come to a stage that results in few false positives or false negatives, with a good sensibility. Anyway, errors generally occur in this stage, making necessary the posterior stage of reducing false negatives and false positives, which, in our case, was based on SVM.

For each individual structure, the skeleton is calculated. After that, all of its segments are scanned sequentially. During the scan of each segment the maximum value of depth is selected and its neighborhood with the same pattern is also

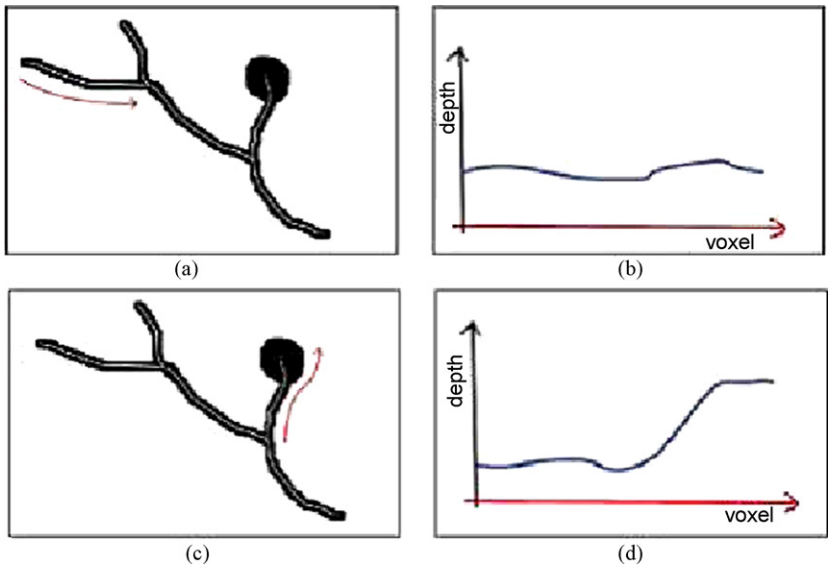

Fig. 5 – Analysis of the medial axis of some structures.

selected. The selection of the neighborhood must consider the average depth of the adjacent medial voxels and the variation from one to another, in sequence.

After the region is selected, it is previously evaluated. A very large rate between the length of the selected part of the branch and its thickness clearly indicates a tubular region. However, a great thickness in relation to the length indicates a compact structure, possibly a nodule.

Fig. 6 shows the pseudocode responsible for analyzing the skeleton branches of each structure and determining the regions with nodule characteristics, which will be considered candidates for the next stage.

This process allows solving a difficult case of lung nodule detection, which is when they adhere to the pulmonary trees, making the correct delimitation of its bounds more difficult.

The structures with excessively affiliated shape to be a lung nodule are readily eliminated. The remaining structures are considered candidates to lung nodules and pass to the next stage. An example can be seen in Fig. 7(a) where we can notice a nodule connected to several blood vessels. Fig. 7(b), on the other hand, presents the same region after the elimination of these vessels.

3.6. *False positives reduction*

False positives reduction is the stage in which the detection is refined by eliminating the false lung nodules. For that, we used the support vector machine (SVM) [49] previously trained to recognize the true nodules with basis on a series of descriptive characteristics. This work used characteristics commonly used in other works [50,34] and [51] with the same objective, but with new characteristics as well, especially developed for describing lung nodules and distinguishing them from other pulmonary structures.

The complete list of the studied characteristics is geometry (spherical disproportion, spherical density, pondered radial distance, sphericity, elongation, Boyce-Clark radial shape index), texture (contrast, energy, entropy, homogeneity, moment), histogram (average, standard deviation, skewness, kurtosis, energy, entropy), gradient (average, standard devia-

```
for each isolated structure
        generates structure skeleton
        scans the branches of the generated skeleton
        eliminatesVessels (branch);

function: eliminatesVessels (branch)
{
        gets the maximum intensity voxel of the branch
        gets the neighborhood of the maximum that maximizes the declivity
        if neighborhood characterizes a nudule
           considers neighborhood as a nodule
        else
        {
           eliminates neighborhood for it is a vessel
           eliminatesVessels (rest of the branch);
        }
}
```

Fig. 6 – Pseudocode for the elimination of tubular structures.

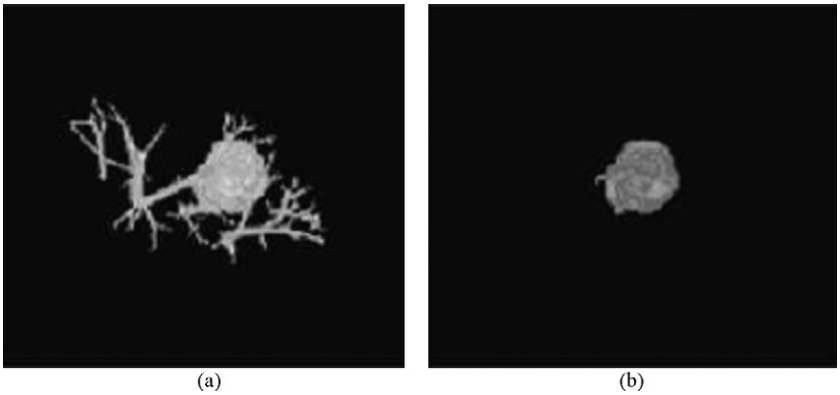

Fig. 7 – Tubular structures elimination.

tion, skewness, kurtosis, energy, entropy) and spatial (location of the candidate). The equations and details for all those measurements can be found in [36] or in AppendixA.

The set of characteristics extracted from every candidate origins a vector which characterizes them. As each characteristic, however, bears on one isolate aspect of the candidate, it occurs that many of them are in different units and frequently in disproportional scales.

To minimize the complexity of the model and speed up the process, we attempted to select a subset of characteristics which are more significant for classification. We empirically tested several subsets of characteristics and verified which one had the best performance. The starting model had 24 variables and after selecting the best subset, there were 8 variables left: geometry (spherical disproportion, spherical density), histogram (standard deviation, skewness, entropy), gradient (standard deviation, kurtosis), and spatial (location of candidate).

The adoption of the vector, such as obtained after calculating these characteristics would cause some of them to be overestimated by the SVM classifier due to the numerically greater value, while others, because they vary in smaller intervals, would be underestimated. This way, the characteristics vector must be normalized so that all the characteristics have the same representativeness.

After all candidates have been completely measured and described, each one by a normalized characteristics vector, these vectors are passed to the SVM, which uses the previous knowledge, obtained by the analysis of other seemingly cases, to identify the real nature of each candidate, recognizing them as lung nodules or as normal lung structures. As SVM kernel, we used the radial basis function. The library LIBSVM [24] was used for training and validation of the SVM classifiers.

Fig. 2(g) shows the correct identification of a lung nodule among other normal lung structures which came from the previous stage. Fig. 2(h) presents the same nodule identified in the original tomography image by an arrow.

4. Results and discussion

The images used herein were provided by Fernandes Figueira Institute and Pedro Ernesto University Hospital – both from Rio de Janeiro city – for a CAD tool development project. They were obtained from different real patients, comprising a total of 33 nodules (23 benign and 10 malignant ones).

The images were acquired with a Helical GE Pro Speed tomograph under the following conditions: tube voltage 120 kVp, tube current 100 mA, image size 512×512 pixels, voxel size $0.67\,\text{mm} \times 0.67\,\text{mm} \times 1.0\,\text{mm}$. The images were quantized in 12 bits and stored in DICOM format [52]. All the images were obtained in apnea after deep inspiration.

It is important to stress that the CT exams were performed without contrast injection, which may be clinically used in order to improve diagnosis but also causes some morbidity and occasional mortality due to allergic complications.

It is also necessary to highlight that the nodules were previously diagnosed by physicians and that the final diagnosis of benignity or malignancy was further confirmed by histopathological exam of the surgically removed specimen or by radiological 3-year stability. This also explains the reduced size of the present sample.

In our work the size of the nodules is considered to be the maximum diameter of the sphere that involves the most distant points in axes *xy* or *z*. According to this definition, the mean diameter of the benign nodules was 23.72 mm (standard deviation 13.34) and the mean diameter of the malignant nodules was 40.93 mm (standard deviation 17.86). There are some nodules that have a diameter equal to 3 mm or less in the *xy* axis (the most common definition of a nodule), but in the *z* axis their diameter is larger than 3 mm. The data set used contains malignant nodules with mean diameter larger than those of benign ones, what very common.

The general characteristic of malignant nodules being larger than benign ones is known and normally found in specialized literature [53]. But this does not mean that there is a cutoff diameter to separate malignant from benign nodules. In the studies, usually only percentages are given. In a revision study of patients with either screening-detected or incidentally detected lung nodules, the prevalence of malignancy was about 6–28% in nodules that measured from 5 to 10 mm in diameter, and 64–82% in nodules that measured over 20 mm in diameter [54].

In our work the smallest malignant nodule had 12 mm of diameter and the smallest benign one had 7 mm, but there

were malignant nodules that presented, for example, diameters of 28 and 29 mm while there were benign nodules with diameters of 35 and 36 mm. In spite of diameter being a general reference and being included in the Bayesian method to aid in the distinction between benign and malignant nodules, it does not have a decisive value by itself [55].

The evaluation of the matching statistics was done using the cross-validation process [56] on the sample. It takes, at each iteration, a group of two arbitrary elements for test and the others for training, that is, each nodule candidate that was part of the test base was, occasionally, and in different iterations, used for training or validation, according to the group for which it was selected in each moment. This kind of validation allowed the use of all of the segmented structures to obtain a more consistent training and a more precise validation.

In order to evaluate the methodology performance, two measurement criteria were adopted. The first one was the adoption of statistics related to nodule candidates classification, the second one was the false positive and negative medium rates per exam.

In order to evaluate the classification capability of the nodule candidates, their sensitivity, specificity and accuracy were considered. Sensitivity is defined by $TP/(TP + FN)$, the specificity by $TN/(TN + FP)$ and the accuracy by $(TP + TN)/(TP + TN + FP + FN)$, where TP is the number of correct positive classification cases, TN is the total number of negative classification cases which are also correct, FP and FN are the erroneously classified cases, positive and negative, respectively.

The false positive and false negative per exam rates were measured as well, which are more significant performance measurements than the previous ones in CAD evaluation because they depend equally on detection and classification. The false positives per exam rate is given by $FE = FP/n$, where n is the number of exams used in tests. The false negative per exam rate is given by $FS = FN/n$.

As the result of the tests and according to the above described criteria, we found a false positive per exam rate of 0.42 and 0.15 of false negatives. These results indicate a good detection power of the methodology as a whole, but it can be improved.

The total sensitivity of the classification obtained, measured on the nodule candidates, was 84.84%. The specificity achieved was 96.15% and the total accuracy of the method was of 95.21%.

Referring to the error distribution over the set of exams, we can notice that more than a half of the exams (57%) were correctly classified, that is, without occurrence of false positives or negatives. On the other hand, we can notice that the occurrence of more than one error in the same exam was 12% of our database.

Comparatively, in [51] the author obtained a rate of 1.25 false cases per exam using a CAD mainly based on geometric characteristics. Chang [24] reached 0.88 false positives per exam and 100% of sensitivity through sphericity tests, but used a reduced database with just eight exams. Zhang reached 83.9% of accuracy and 3.5 false positive per exam with his method [39].

Among the metrics used for describing lung nodule candidates, those that showed to be more efficient were geometry (spherical disproportion, spherical density), histogram (standard deviation, skewness, entropy), gradient (standard deviation, kurtosis), spatial (location of candidate). These, then, were chosen to form the characteristics vector since they showed better results after a performance analysis among many others tested.

It was noticed that the measurements related to the gray tone, as well as to texture and gradient were not as descriptive as expected. It was observed that, generally, due to low resolution of the images obtained by present CT exams, the particularities of the tissue, which could identify nodules and non-nodules, are not totally acquired. This way, and following the tendency of the related literature, the geometrical characteristics play the main role in detection.

Vertical gradients in CT represent variations of lung densities in function of gravitational influence in vascular tree. This finding is more pronounced when CT images are obtained in expiration phase closed to residual volume (RV) and less when they are obtained in inspiration phase at total lung capacity [57]. In the present work all images were obtained in apnea after deep inspiration and in this way is expected that the vertical gradients do not arise a special problem to the present automatic program. However, changing in interstitial lung component or increasing in vascular hydrostatic pressure caused by some associated lung disease may cause detection problems. In this case, automatic program probably will be affected at threshold phase by an increase of density around the nodule. To evaluate this possibility, it will be necessary in future studies include a selected group of patients with specific patterns like fibrosing alveolitis or cardiac disease. Recently, it has been proposed that morphological aspects of respiratory and vascular trees can also contribute to vertical gradients beside gravitational force [58].

It was verified that the time necessary for a complete detection varies, depending on the case under analysis and the number of slices which compound it. The average time stood around 8 min in a computer with an Athlon XP 2000+ processor and with 1.5 GB of memory.

4.1. *Case study – successful detection*

This section presents the illustrated step-by-step details of a successful detection case performed on an exam on which the nodule was detected by a specialist according to Fig. 8. This process is described with basis on the volume, originally obtained through reconstruction of the slices, as in Fig. 8(b).

Fig. 8(c) is the result of the thorax segmentation through elimination of external artifacts. Fig. 8(d) is the result of segmentation of the parenchyma and its reconstruction to prevent elimination of peripheral nodules.

The elimination of the soft tissues of the parenchyma allows the obtainment of just the denser structures inside it. The result of this stage can be seen in Fig. 8(e). Next, in Fig. 8(f), these structures are isolated and only those resembling a nodule are selected as candidates.

The classification result was correct. SVM did not consider the other structures and considered as nodule only the structure emphasized in Fig. 8(g) and (h).

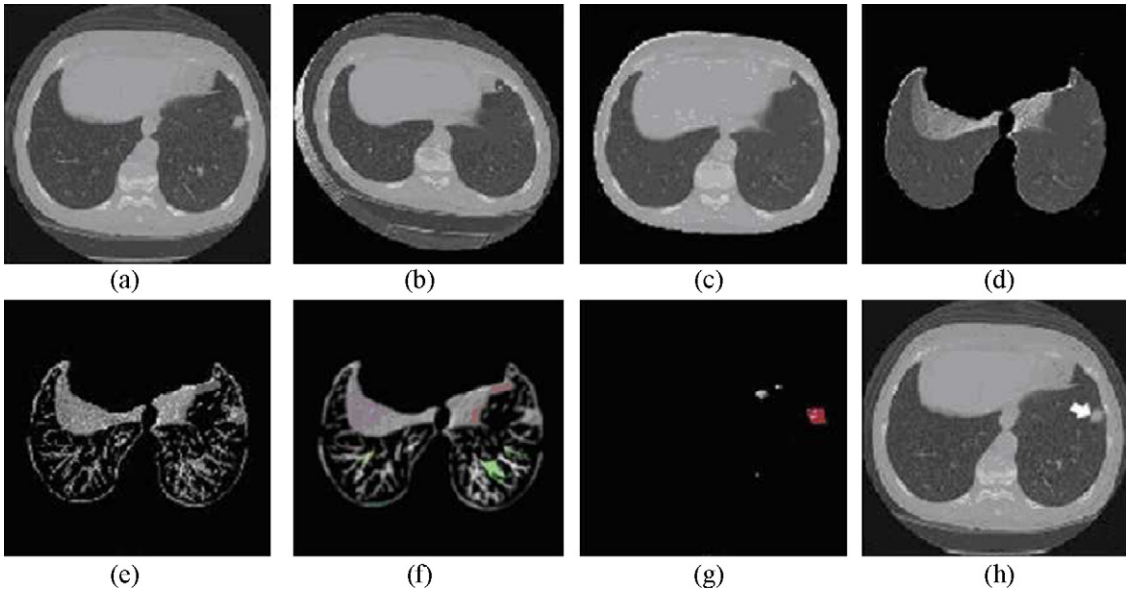

Fig. 8 – Case study – successful detection. (a) and (b) are 2D and 3D visualization of the original CT exam. (c) Eliminates of all artifacts external to the patient's body. (d) Removes thorax just remaining the parenchyma and shows the reconstructed parenchyma. (e) 3D Visualization of the remaining structures after threshold application. (f) 3D Visualization of the structures after tubular elimination. (g) Shows lung candidates before SVM classification. (h) Shows correctly detected nodule in the original CT.

4.2. *Case study – unsuccessful detection*

As not all nodules were correctly detected, this section shows one case where there was a detection failure. As far as possible an analysis on the causes which could have led to omission will be done.

The process starts the same way as the successful case, demonstrated in the previous section. Fig. 9(a) shows the initial volume and Fig. 9(b) shows the segmented thorax.

The parenchyma extraction stage, however, as can be seen in Fig. 9(c), incorrectly eliminated some internal structures,

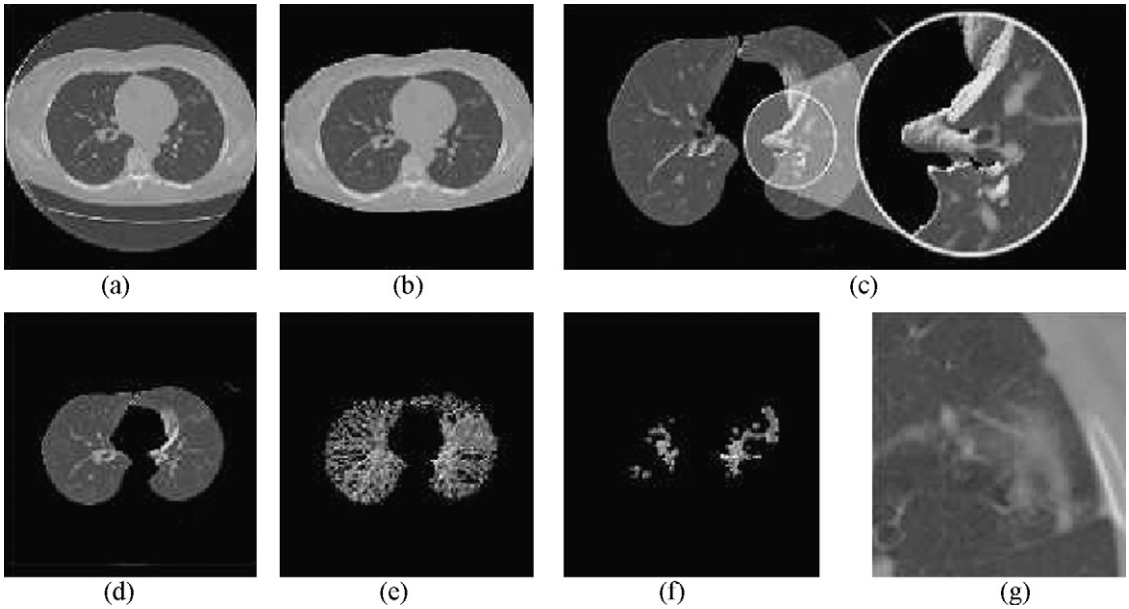

Fig. 9 – Case study – unsuccessful detection. (a) 2D visualization of the original CT exam. (b) Eliminates of all artifacts external to the patient's body. (c) Removal of thorax just remaining the parenchyma and shows the incorrectly eliminated internal structures, making necessary the reconstruction of the lungs perimeter. (d) Shows the reconstructed parenchyma. (e) 3D Visualization of the remaining structures after threshold application. (f) Shows that the elimination of the blood vessels resulted in the elimination of part of the nodule, damaging the detection. (g) Shows the original CT and as the contour nodule is very fuzzy.

making necessary the reconstruction of the lungs perimeter. As Fig. 9(d) illustrates, the reconstruction step was successful and the structures were recovered.

Fig. 9(e) shows the result of segmentation of lung internal structures. In this case, differently from the previous, occurs a large concentration of vessels all over the parenchyma, making nodule identification more difficult.

In fact, Fig. 9(f) shows that the elimination of the blood vessels resulted in elimination of part of the nodule, damaging the detection.

Through Fig. 9(f) we can guess the reason which led to failure in the detection of this spiky points nodule. It can be noticed that the nodule outlines are not well defined, making its clear delimitation more difficult.

Another factor of possible confusion is the large vascularization of the area, which, in addition to the formless aspect of the nodule, caused it to be eliminated with the tubular structures.

Nodule volume precision is an important feature to determine the malignity or benignity of the nodule. In this work we focus on the identification of the presence or absence of nodules, and in the case of presence, we focus on determining its location. We did not worry about the volume determination precision. But we also made a simple and manual test, comparing the automatic segmentation using the methodology proposed in this paper and a semi-automatic process made by a specialist. Comparing the results we find an average error rate of 15% that we believe is too high. Unfortunately we cannot give further details about the comparison because we did not perform a deep analysis. Therefore, the proposed segmentation must be also improved to be more adequate to some cases, as for example the case where the nodule border is not well defined.

5. Conclusion

This work proposes a methodology for automatic detection of lung nodules; a stage that precedes the diagnosis properly said, but that demands effort and experience from the specialist.

The methodology proposed in this work, besides automatic, is based on simple algorithms, most of them of easy implementation and, in general, fast. This gives to the methodology the quickness and efficiency necessary for processing a large volume of data, always taking into consideration the time restrictions related to any process that involves human health.

As previously described, the stage of separating the nodule from pulmonary trees is strongly based on a skeletonization process. Algorithms in this category have been used in diagnose-aiding tools, especially in the characterization of morphology with the objective of verifying the malignancy of the nodule [59], but the approach used here, relating them to detection, represents an innovation with satisfactory results.

Finally, new geometric characteristics were especially developed to evaluate the morphology of pulmonary structures in the classification of nodules and non-nodules. These geometric characteristics, numerically measurable, which were projected for this specific problem, showed results that are better than others of general use.

The inexistence of user-defined parameters and the absence of human participation in the detection of the nodule or random components reinforce the segmentation reproducibility. Independently on who is operating the software or on his experience, the result is always the same. This is also a characteristic that allows the batch processing of a large number of different exams. It is possible to select all of the exams that we want to process and obtain in a single step the results for all of them, with no intervention, needing only the posterior validation of a specialist.

In addition, the methodology is composed of consecutive stages that gradually produce the final result of segmentation. The low coupling of these stages allows their easy expansion with the objective of improving results or even to treat different cases and situations.

As a result of the executed tests, we verified a false positive per exam rate of 0.42 and a false negative rate of 0.15. The total sensitivity of the classification obtained, measured out of the nodule candidates, was 84.84%. The specificity achieved was 96.15% and the total accuracy of the method was 95.21%.

The matching rates discussed demonstrate that there is the technical viability for implantation of the methodology. Concerning the needs for it, statistics related to lung cancer clearly indicate that methods for helping in precocious diagnosis of lung nodule may increase the patient's survival chances.

Due to the high sensitivity per exam, this tool has triage exam characteristics, that is, belongs to the first set of exams to be required, which identify the suspicious cases, but need to be confirmed later, by more strict exams, in this case, the medical analysis.

Lung cancer, in the same way it occurs in the greatest part of the world, has a considerable prevalence in Brazil. It is the second commoner kind of cancer among men and the forth commoner among women in most part of Brazilian regions.

Since precocious diagnosis represents a considerable increase in the patient's survival chances, the proposed methodology promotes this increase, as it is shown as a very useful tool for the specialist in the attempt to anticipate more and more the nodule identification.

Another point is that the public network of hospitals in some places suffers from the lack of specialists. The resources to increase the staff, however, are also limited. Redirecting qualified craft of the available specialists to less repetitive tasks may mean making better use of their skills. One step in that direction is to use the proposed methodology in the preliminary analysis of CT exams, being the specialist just in charge of validating the result.

Finally, the proposed methodology also is a financially attractive solution because it works on simple microcomputers, many of which are already available in the hospitals. Large investments in infrastructure would not be necessary for its implantation.

Conflicts of interest

There are no conflicts of interest.

Appendix A. Features used in this study

A.1. Geometric features

A.1.1. Spherical disproportion

Spherical disproportion measures the regularity of the borders of an object. It is given by Eq. (A.1), 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

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

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

A.1.2. Spherical density

The spherical density measures how much compact an object is. It is given by Eq. (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 M_x, M_y, M_z are the coordinates of the object's mass center

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

A.1.3. Pondered radial distance

The pondered radial distance measures the flatness degree of an object, being obtained by Eq. (A.4), 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

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

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

A.1.4. Sphericity

Sphericity measures how much the shape of the object approximates of a spherical shape. It is obtained through Eq. (A.6), where V is the volume of the object and A is its area

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

A.1.5. Elongation

Elongation measures the elongation or asymmetry degree of an object. It is calculated through Eq. (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

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

A.1.6. Boyce–Clark radial shape index

This feature measures the regularity of the shape of an object. It is obtained through Eq. (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

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

A.2. Texture features

The texture features were calculated on the co-occurrence matrix of the volume, which is given by Eq. (A.9), where υ is the function that gives one of the 26 tri-dimensional 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

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

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

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

A.2.4. Homogeneity

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

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] INCA Estimativas da Incidencia e Mortalidade por Cancer no Brasil, available in: http://www.inca.gov.br/estimativas/2003/versaofinal.pdf, 2003.

[2] F.M. Navarro, Estudo Comparativo dos pacientes com Cancer de Pulmao, tabagistas e nao Tabagistas, Monografia, Universidade Federal do Oeste do Parana, 2003.

[3] I. Sluimer, A. Schilham, M. Prokop, B. van Ginneken, Computer analysis of computed tomography scans of the lung: a survey, IEEE Transactions on Medical Imaging 25 (4) (2006) 385–405, doi:10.1109/TMI.2005.862753.

[4] C. Beigelman-Aubry, P. Raffy, W. Yang, R.A. Castellino, P.A. Grenier, Computer-aided detection of solid lung nodules on

follow-up MDCT screening: evaluation of detection, tracking, and reading time, American Journal of Roentgenology 189 (4) (2007) 948–955, ISSN:1546–3141, doi:10.2214/AJR.07.2302.
[5] S.G. Armato, W.F. Sensakovic, Automated lung segmentation for thoracic CT: impact on computer-aided diagnosis, Academic Radiology 11 (2004) 1011–1021.
[6] S. Hu, E.A. Hoffman, J.M. Reinhardt, Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images, IEEE Transactions on Medical Imaging 20 (2001) 490–498.
[7] B. Zheng, J.K. L., G.S. Maitz, B.E. Chapman, C.R. Fuhrman, R.M. Rogers, F.C. Sciurba, A. Perez, P. Thompson, W.F. Good, D. Gur III, A simple method for automated lung segmentation in X-ray CT images, SPIE 5032 (2003) 1455–1463, doi:10.1117/12.480290.
[8] J.K. Leader, B. Zheng, R.M. Rogers, F.C. Sciurba, A. Perez, B.E. Chapman, S. Patel, C.R. Fuhrman, D. Gur, Automated lung segmentation in X-ray computed tomography: development and evaluation of a heuristic threshold-based scheme, Academic Radiology 10 (11) (2003) 1224–1236.
[9] S.G. Armato, M.L. Giger, C.J. Moran, J.T. Blackburn, K. Doi, H. MacMahon, Computerized detection of pulmonary nodules on CT scans, Radiographics 19 (5) (1999) 1303–1311.
[10] J.P. Ko, M. Betke, C.T. Chest, Automated nodule detection and assessment of change over time—preliminary experience, Radiologic Clinics of North America 218 (2001) 267–273.
[11] B. Zhao, D. Yankelevitz, Two-dimensional multi-criterion segmentation of pulmonary nodules on helical CT images, Medical Physics 26 (6) (1999) 889–895.
[12] B. Zhao, M.S. Ginsberg, R.A. Lefkowitz, L. Jiang, C. Cooper, L.H. Schwartz, Application of the LDM algorithm to identify small lung nodules on low-dose MSCT scans, SPIE 5370 (2004) 818–823, doi:10.1117/12.535558.
[13] T. Ezoe, H. Takizawa, S. Yamamoto, A. Shimizu, T. Matsumoto, Y. Tateno, T. Iimura, M. Matsumoto, Automatic detection method of lung cancers including ground-glass opacities from chest X-ray CT images, SPIE 4684 (2002) 1672–1680.
[14] C.I. Fetita, F. Preteux, C. Beigelman-Aubry, P. Grenier, 3D automated lung nodule segmentation in HRCT, Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003 2878 (2003) 626–634.
[15] M. Tanino, H. Takizawa, S. Yamamoto, T. Matsumoto, Y. Tateno, T. Iinuma, A detection method of ground glass opacities in chest X-ray CT images using automatic clustering techniques, 1728–1737, 561–577, in: M. Sonka, J.M. Fitzpatrick (Eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 5032 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 2003.
[16] 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 (2004) 347–352.
[17] K. Kanazawa, Y. Kawata, N. Niki, H. Satoh, H. Ohmatsu, R. Kakinuma, M. Kaneko, N. Moriyma, K. Eguchi, Computer-aided diagnosis for pulmonary nodules based on helical CT images, Computerized Medical Imaging and Graphics 22 (1998) 157–167.
[18] M.N. Gurcan, B. Sahiner, N. Petrick, H.P. Chan, E.A. Kazerooni, P.N. Cascade, L.M. Hadjiiski, Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system, Medical Physics (2002) 2552–2558.
[19] M. Kubo, K. Kubota, N. Yamada, Y. Kawata, N. Niki, K. Eguchi, H. Ohmatsu, R. Kakinuma, M. Kaneko, M. Kusumoto, K. Mori, H. Nishiyama, N. Moriyama, CAD system for lung cancer based on low-dose single-slice CT image, in: M. Sonka, J.M. Fitzpatrick (Eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 4684 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 2002, pp. 1262–1269.
[20] N. Yamada, M. Kubo, Y. Kawata, N. Niki, K. Eguchi, H. Omatsu, R. Kakinuma, M. Kaneko, M. Kusumoto, H. Nishiyama, N. Moriyama, ROI extraction of chest CT images using adaptive opening filter, in: M. Sonka, J.M. Fitzpatrick (Eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 5032 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 2003, pp. 869–876, doi:10.1117/12.483540.
[21] T. Oda, M. Kubo, Y. Kawata, N. Niki, K. Eguchi, H. Ohmatsu, R. Kakinuma, M. Kaneko, M. Kusumoto, N. Moriyama, K. Mori, H. Nishiyama, Detection algorithm of lung cancer candidate nodules on multislice CT images, in: M. Sonka, J.M. Fitzpatrick (Eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 4684 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 2002, pp. 1354–1361.
[22] S. Saita, T. Oda, M. Kubo, Y. Kawata, N. Niki, M. Sasagawa, H. Ohmatsu, R. Kakinuma, M. Kaneko, M. Kusumoto, K. Eguchi, H. Nishiyama, K. Mori, N. Moriyama, Nodule detection algorithm based on multislice CT images for lung cancer screening, in: J.M. Fitzpatrick, M. Sonka (Eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 5370 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 2004, pp. 1083–1090, doi:10.1117/12.534826.
[23] R. Wiemker, P. Rogalla, A. Zwartkruis, T. Blaffert, Computer-aided lung nodule detection on high-resolution CT data, in: M. Sonka, J.M. Fitzpatrick (Eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 4684 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 2002, pp. 677–688.
[24] S. Chang, H.D. Emoto, L.A. Metaxas, Pulmonary micronodule detection from 3-D chest CT, Lecture Notes in Computer Science-Medical Image Computing and Computer-Assisted Intervention 3217 (2004) 821–828.
[25] Q. Li, K. Doi, New selective nodule enhancement filter and its application for significant improvement of nodule detection on computed tomography, in: J.M. Fitzpatrick, M. Sonka (Eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 5370 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 2004, pp. 1–9, doi:10.1117/12.535802.
[26] D.S. Paik, Computer aided interpretation of medical images, Ph.D. Thesis, Stanford University, 2002.
[27] O. Osman, S. Ozekes, O.N. Ucan, Lung nodule diagnosis using 3D template matching, Computers in Biology and Medicine, 37 (8) (2007) 1167–1172, ISSN: 0010–4825, doi:10.1016/j.compbiomed.2006.10.007.
[28] A. Retico, P. Delogu, M.E. Fantacci, I. Gori, A. Preite Martinez, Lung nodule detection in low-dose and thin-slice computed tomography, Computers in Biology and Medicine, 38 (4) (2008) 525–534, ISSN: 0010–4825, doi:10.1016/j.compbiomed.2008.02.001.
[29] K.T. Bae, J.-S. Kim, Y.-H. Na, K.G. Kim, J.-H. Kim, Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists detection performance, Radiology 236 (2005) 286–293.
[30] 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, Academic Radiology 15 (2) (2008) 165–175.

[31] Y. Lee, T. Hara, H. Fujita, S. Itoh, T.I., Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique, 2001.
[32] Y. Mekada, T. Kusanagi, Y. Hayase, K. Mori, J. ichi Hasegawa, J. ichiro Toriwaki, M. Mori, H. Natori, Detection of small nodules from 3D chest X-ray CT images based on shape features, CARS (2003) 971–976.
[33] M.S. Brown, J.G. Goldin, R.D. Suh, M.F. McNitt-Gray, J.W. Sayre, D.R. Aberle, Lung micronodules: automated method for detection at thin-section CT-initial experience 1, Radiology 226 (2003) 256–262.
[34] X. Lu, G.-Q. Wei, J.Z. Qian, A.K. Jain, Learning-based pulmonary nodule detection from multislice CT data, CARS (2004) 1356.
[35] W. Mousa, M. Khan, Lung nodule classification utilizing support vector machines, in: ICIP02, III, 2002, pp. 153–156.
[36] J.R.F. da Silva Sousa, Metodologia Automatica para Deteccao de Nodulos Pulmonares, Master's Thesis, Universidade Federal do Maranhao, 2007.
[37] K. Suzuki, Z. Shi, J. Zhang, Supervised enhancement of lung nodules by use of a massive-training artificial neural network (MTANN) in computer-aided diagnosis (CAD), in: ICPR08, 2008, pp. 1–4.
[38] S.-C.B. Lo, L.-Y. Hsu, M.T. Freedman, Y.M.F. Lure, H. Zhao, Classification of lung nodules in diagnostic CT: an approach based on 3-D vascular features, in: Nodule Density Distributions, and Shape Features, The International Society for Optical Engineering, Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, San Jose, California -EUA, 2003, pp. 183–189.
[39] X. Zhang, G. McLennan, E.A. Hoffman, M. Sonka, Computerized detection of pulmonary nodules using cellular neural networks in CT images, in: Medical Imaging 2004: Image Processing, San Diego, CA, USA, 2004.
[40] A. Farag, A. El Baz, G. Gimel'farb, R. Falk, Detection and recognition of lung nodules in spiral CT images using deformable templates and Bayesian post-classification, in: ICIP04 V, 2004, pp. 2921–2924.
[41] C.C. McCulloch, R.A. Kaucic, P.R. Mendonca, D.J. Walter, R.S. Avila, Model-based detection of lung nodules in computed tomography exams 1: thoracic computer-aided diagnosis, Academic Radiology 11 (3) (2004) 258–266.
[42] J.R.F. da Silva Sousa, A.C. Silva, A.C. de Paiva, Lung structure classification using 3D geometric measurements and SVM, in: CIARP, Iberoamerican Congress on Pattern Recognition, 2007, pp. 783–792.
[43] S.G. Armato, M.L. Giger, H. MacMahon III, Automated detection of lung nodules in CT scans: preliminary results, Medical Physics 28 (2001) 1552–1561.
[44] H. Arimura, S. Katsuragawa, K. Suzuki, F. Li, J. Shiraishi, S. Sone, K. Doi, Computerized scheme for automated detection of lung nodules in low-dose CT images for lung cancer screening, Academic Radiology 11 (2004) 617–629.
[45] S.G.A. Li, F. Giger, M.L. MacMahon, H. Sone, S. Doi III, Lung cancer: performance of automated lung nodule detection applied to cancers missed in a CT screening program, Radiology, 225 (2002) 685–692.
[46] K. Suzuki, Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography, Medical Physics 30 (2003), 1602-+, doi:10.1118/1.1580485.
[47] Y. Lee, D.-Y. Tsai, T. Hara, H. Fujita, S. Itoh, Ishigaki, Improvement in automated detection of pulmonary nodules on helical X-ray CT images, in: J.M. Fitzpatrick, M. Sonka (Eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 5370 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 2004, pp. 824–832, doi:10.1117/12.536162.
[48] R.C. Gonzalez, R.E. Woods, Digital Image Processing, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, ISBN: 0201180758, 2001.
[49] V.N. Vapnik, Statistical Learning Theory, Wiley-Interscience, ISBN: 0471030031, URL http://www.amazon.ca/exec/obidos/redirect?tag=citeulike04-20&path=ASIN/0471030031, 1998.
[50] G. Agam, S.G.A. Iii, C. Wu, Vessel tree reconstruction in thoracic CT scans with application to nodule detection, IEEE Transactions on Medical Imaging 24 (4) (2005) 486–499.
[51] K. Peldschus, P. Herzog, S.A. Wood, J.I. Cheema, P. Costello, U.J. Schoepf, Computer-aided diagnosis as a second reader: spectrum of findings in CT studies of the chest interpreted as normal, Chest 128 (2005) 1517–1523.
[52] D.A. Clunie, DICOM Structured Reporting, PixelMed Publishing, Pennsylvania, 2000.
[53] M.K. Gould, J. Fletcher, M.D. Iannettoni, W.R. Lynch, D.E. Midthun, D.P. Naidich, D.E. Ost, Evaluation of patients with pulmonary nodules: when is it lung cancer?, ACCP Evidence-Based Clinical 132 (3) (2007) 108S–130S, Practice Guidelines (2nd ed.). Chest. Supplement.
[54] M.M. Wahidi, J.A.M.D. Govert, R.K. Goudar, M.K. Gould, D.C. McCrory, Diagnosis and management of lung cancer, ACCP Evidence-Based Clinical 132 (3) (2007) 94S–107S, Practice Guidelines (2nd Edition). Chest. Supplement.
[55] K. Nakamura, H. Yoshida, R. Engelmann, H. MacMahon, S. Katsuragawa, T. Ishida, K. Ashizawa, K. Doi, Computerized analysis of the likelihood of malignancy in solitary pulmonary nodules with use of artificial neural networks, Radiology 214 (2000) 823–830.
[56] M. Mullin, R. Sukthankar, Complete cross-validation for nearest neighbor classifiers, in: Proceedings of the 17th International Conference on Machine Learning, ICML, 2000.
[57] A.B. Millar, D.M. Denison, Vertical gradients of lung density in healthy supine men, Thorax 44 (6) (1989) 485–490, doi:10.1136/thx.44.6.485.
[58] I. Galvin, G.B. Drummond, M. Nirmalan, Distribution of blood flow and ventilation in the lung: gravity is not the only factor, British Journal of Anaesthesia (2007) aem036 doi:10.1093/bja/aem036, http://bja.oxfordjournals.org/cgi/content/abstract/aem036v1.
[59] A.C. Silva, Algoritmos para Diagnostico Assistido de Nodulos Pulmonares Solitarios em Imagens de Tomografia Computadorizada, Ph.D. Thesis, PUC-Rio, 2004.