Artificial Intelligence in Medicine 60 (2014) 165–177

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

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

Automatic detection of solitary lung nodules using quality threshold clustering, genetic algorithm and diversity index

Antonio Oseas de Carvalho Filho [a], Wener Borges de Sampaio [a], Aristófanes Corrêa Silva [a,*], Anselmo Cardoso de Paiva [a], Rodolfo Acatauassú Nunes [b], Marcelo Gattass [c]

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

ARTICLE INFO

Article history:
Received 4 June 2013
Received in revised form 10 October 2013
Accepted 8 November 2013

Keywords:
Quality threshold
Genetic algorithm
Support vector machine
Lung cancer
Nodule detection
Computer-aided detection
Medical image

ABSTRACT

Objective: The present work has the objective of developing an automatic methodology for the detection of lung nodules.
Methodology: The proposed methodology is based on image processing and pattern recognition techniques and can be summarized in three stages. In the first stage, the extraction and reconstruction of the pulmonary parenchyma is carried out and then enhanced to highlight its structures. In the second stage, nodule candidates are segmented. Finally, in the third stage, shape and texture features are extracted, selected and then classified using a support vector machine.
Results: In the testing stage, with 140 new exams from the Lung Image Database Consortium image collection, 80% of which are for training and 20% are for testing, good results were achieved, as indicated by a sensitivity of 85.91%, a specificity of 97.70% and an accuracy of 97.55%, with a false positive rate of 1.82 per exam and 0.008 per slice and an area under the free response operating characteristic of 0.8062.
Conclusion: Lung cancer presents the highest mortality rate in addition to one of the smallest survival rates after diagnosis. An early diagnosis considerably increases the survival chance of patients. The methodology proposed herein contributes to this diagnosis by being a useful tool for specialists who are attempting to detect nodules.

1. Introduction

Lung nodules are potential manifestations of lung cancer, and their early detection is essential for diagnosis. The most crucial factor related to the occurrence of this type of cancer is the past use of tobacco. In most populations, lung cancer cases related to tobacco use represent 80% or more of the total number of cases [1]. In comparison to non-smokers, smokers have an approximately 20–30 times greater risk of developing cancer. In general, the incidence rates of lung cancer in a particular country directly reflect the consumption of tobacco [2].

The detection of lung nodules is still a challenging task [3]. Lung nodules are difficult to detect using computed tomography (CT), as their densities may be equal to those of other structures or other issues; for example, they may have low density or a small size in a complex anomaly area (connected to vessels and lung edges). Another factor that hinders detection is that specialists have a large number of CT images to analyze. This process is repetitive and exhausting, and the specialists' attention might falter, resulting in an analysis mistake, especially when there are other anomalies in the image. Therefore, this type of analysis is often subject to errors [4].

To minimize such errors, a large amount of research has been conducted to improve computer-aided detection (CAD) and computer-aided diagnosis (CADx). These systems provide a second opinion, helping the radiologist in the interpretation of exams and indicating the diagnosis of lung nodules [5]. Table 6 list some works that propose methodologies for the detection of lung nodules.

The present work introduces a method for automatic detection of solitary lung nodules. The proposed methodology comprises the following stages: (1) acquisition of images from the Lung Image Database Consortium image collection (LIDC-IDRI) and division into training and test sets for validating the methodology; (2) and (3) preprocessing, where the pulmonary parenchyma is segmented and filters are applied to make the nodules more visible; (4) clustering of structures inside the pulmonary parenchyma, which are very similar to lung nodules, using the quality threshold (QT) clustering

* Corresponding author: Tel.: +55 9832728243; fax: +55 9832728841.
E-mail addresses: antoniooseas@gmail.com (A.O. de Carvalho Filho), wenersampaio@gmail.com (W.B. de Sampaio), ari@dee.ufma.br (A.C. Silva), paiva@deinf.ufma.br (A.C. de Paiva), rodolfoacatauassu@yahoo.com.br (R.A. Nunes), mgattass@tecgraf.puc-rio.br (M. Gattass).

0933-3657/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.artmed.2013.11.002

algorithm; (5) extraction and selection of the most suitable features based on shape and texture; and (6) finally, the segmented structures are classified as nodules and non-nodules. In this stage, we use a micro-genetic algorithm, a variant of the genetic algorithm, to find the best training model and support the vector machine for the final classification.

We believe we bring contributions to this area in the following aspects: (a) the use of the QT algorithm for segmentation of structures that resemble a lung nodule, (b) the use of the genetic algorithm to automatically determine the best training model to be used in the classification stage of the methodology (this process is usually determined manually, and much time is spent determining the best model), and (c) the use of texture measurements based on the diversity index to characterize the nodule.

This paper is divided as follows: In Section 3, we present all the steps of the automatic segmentation of lung nodules, explaining in detail the techniques used for segmentation of the pulmonary pleura and contrast enhancement, the quality threshold algorithm, and the extraction of shape and texture features from the mass candidates, which compose the proposed methodology and evolutionary methodology. In Section 4, we analyze all the tests obtained with the application of the methodology. Finally, in Section 5, we present our final remarks about this work.

2. Related work

The literature available offers acknowledged studies dealing with the problem of automatic segmentation of pulmonary nodules. The following is a summary of some works with the same objective as ours.

Dehmeshki et al. [6] present a method called genetic algorithm template matching for automatic detection of lung nodules. The computation of the fitness function is based on the geometric shape of the voxel, and then combined with the global distribution of the nodule's intensity. The authors report a rate of 14 false positives per exam.

Opfer and Wiemker [7] show that the result of the proposed method's sensitivity is directly related to the size of the nodule, presenting sensitivity of 89% for nodules with a diameter equal to or larger than 4 mm and a sensitivity of 60% for cases smaller than 4 mm.

The methodology presented by Tong et al. [8] follows these steps: (1) segmentation of the pulmonary parenchyma, where detection of suspect nodule candidates is performed using a smoothed Gaussian function for noise reduction, and then a selective filter technique with a Hessian matrix is applied; (2) feature extraction; and (3) classification. The authors obtained a sensitivity of 95%.

The clustering performed by the algorithm proposed by Nie et al. [9] makes use of mean-shift clustering and convergence index resources, forming a CAD system that yields promising results in the detection of lung nodules, with a sensitivity of 89%.

The work by yan Jiang and yu Cheng [10] presents a CAD system with a new approach to overcome lung nodule segmentation problems in two main stages. (1) The transition region is determined by wavelet coefficient statistical resources, and (2) the precise boundaries of the nodule are segmented based on an improved version of the two joint level methods. To validate the methodology, 47 slices were used, and there was a sensitivity of 80%.

A robust and automatic algorithm for lung nodule segmentation is proposed by Sun et al. [11]. A mean-shift estimation method is applied to lung nodule segmentation. This work presents a method for choosing a new bandwidth. The new width determination was used to analyze the K–L divergence rule. The method was evaluated using 95 slices containing 36 nodules, presenting a sensitivity of 100% and 3 false positives per slice.

The work by Dehmeshki et al. [12] describes an algorithm for segmenting different types of lung nodules. It is based on region growing and fuzzy connectivity. Region growing takes place inside a volume mask, which is created by a first application of a local adaptive segmentation algorithm. The initial results achieved a sensitivity of 84%.

The work presented by Miyake et al. [13] illustrates the use of a technique based on temporal subtraction. For the automatic detection of lung nodules, a new method based on artificial neural networks and a temporal subtraction image was developed. To validate the proposed methodology, the authors applied the method to 6 cases, achieving sensitivity of 80.5% for lung nodules smaller than 20 mm in size, with 7.5 false positives per exam.

The detection scheme proposed by Ye et al. [5] was developed to detect two types of nodules: solid and ground-glass opacity ones. The proposed method was trained and validated on a set of 108 exams, presenting a sensitivity of 90.2%.

The work developed by Netto et al. [14] uses a grouping technique based on neural networks, known as growing neural gas, in the extraction of pulmonary structures stage. To reduce false positives, a support vector machine (SVM) was used to classify suspect structures as nodules or non-nodules. Twenty-nine exams from the Lung Image Database Consortium (LIDC) were used. They were divided into training and testing images, a sensitivity of 85.93% and a false positive rate of 0.138 per exam were achieved.

The architecture of the CAD system described by Messay et al. [15] for lung nodule detection was evaluated using LIDC images. In total, 84 exams containing 143 nodules with sizes between 3 and 30 mm were used. This work presents a combination of intensity thresholding and morphological processing to detect nodule candidates. The stepwise technique is used to determine the subset with optimum features. Finally, the authors report a rate of 3 false positives per exam and an average sensitivity of 82.66% for the nodules, using 40 selected features.

The methodology developed by Xiaomin et al. [16] is divided into three stages. In the first stage, a 2D multi-scale filter is used. In stage 2, blob-shaped nodules and non-nodules are differentiated. In stage 3, the shape features of each region are extracted, and a classifier based on automated rules to reduce false positives is applied. The method was applied to 30 exams and presented a sensitivity of 100% and false positive rate of 8.4 per exam.

The CAD proposed by Tan et al. [17] includes innovations such as the use of a new classifier of selective features described as feature-deselective neuro-evolving augmenting of topologies. This classifier is based on neural networks and genetic algorithms. Two other classifiers are also used, namely fixed-topology artificial neural networks and SVM. The process executed by the CAD consists of four steps: preprocessing, detection of nodule candidates, feature selection, and classification. The model was validated by the sensitivity computation (87.5%).

The objective of the study conducted by Suárez-Cuenca et al. [18] was to investigate the usefulness of different methods to combine the classifier to improve the performance of a CAD system in the detection of lung nodules. With a sensitivity of 80%, the number of false positives per case for the six individual classifiers was 6.1 for linear discriminant analysis (LDA), 19.9 for quadratic discriminant analysis, 8.6 for artificial neural networks, 23.7 for SVM-dot, 17.0 for SVM-poly, and 23.35 for SVM-ANOVA. The number of false positives by case for the five combination methods was 3.4 for the majority rule, 6.2 for average, 5.7 for product, 9.7 for neural network, and 28.1 for the likelihood ratio method.

The combination of three CAD systems developed by Italian MAGIC-5 is presented in Camarlinghi et al. [19]. The first is based on a 3D highlight algorithm. The second is based on the normal

intersection of the pleura surface and an approach based on neural voxel to reduce false positives. In total, 138 exams were used to validate the methodology. A sensitivity of 85% was achieved using a combination of the three CADs analyzed, with a rate of 25 false positives per exam.

The combination of techniques presented by Farag et al. [20] is what sets this work apart. The techniques are scale-invariant feature transform (SIFT), local binary pattern (LBP), principal component analysis (PCA), and LDA. The first combinations were PCA–SIFT and PCA–LBP, which presented a sensitivity of 85%. The combination of LDA–SIFT and LDA–LBP achieved the same sensitivity.

Chama et al. [21] present a methodology that uses mean-shift followed by techniques based on geometric properties, such as region of interest (ROI), created from the symmetric centroid map of two normal subjects. To validate the performance, 429 images (133 normal and 296 abnormal) from the LIDC-IDRI and Interstitial Lung Disease (ILD) databases were used. The proposed method achieved sensitivities and specificities of 97% and 99% (normal images) and 83% and 99% (abnormal images), respectively.

The methodology developed by Sivakumar and Chandrasekar [22] aims to develop a lung nodule detection system by means of segmentation, using fuzzy clustering models and classification with SVM. This methodology uses three types of kernels (linear, polynomial and radial basis function (RBF)) for SVM. The RBF kernel presents better results, with 80.36% accuracy, 76.47% specificity, and 82.05% sensitivity.

3. Materials and methods

In this section, we present the methodology used for automatic detection of lung nodules. Fig. 1 displays all the stages of the methodology, from image acquisition to the final stage of classification of nodule candidates. The methodology comprises six major stages, which are subdivided into minor stages.

3.1. Stage 1: image acquisition

The database used in this work is the LIDC-IDRI [23], which is available on the internet as a result of the association between the Lung Image Database Consortium and the Image Database Resource Initiative. All images are in DICOM format [24] and have 16 bits per voxel. Their dimension is 512×512 with a variable number of slices per exam. The X-ray CT images were acquired by different tomographers, which makes the detection of lung nodules more difficult.

This image database has a file containing information about the marking of the nodules, as performed by four specialists. The markings of the nodule contours and their characteristics are made only for nodules between 3 mm and 30 mm. For nodules smaller than 3 mm, the only markings refer to their center of mass. In addition the markings, the files also provide information about the properties of the lung nodules: subtlety, internal structure, calcification, sphericity, spiculation, texture and malignancy.

The image database contains 1012 CT exams, but two factors made some of them, more precisely 212 exams, inappropriate for our methodology. The first factor is related to exams that do not present nodules equal to or larger than 3 mm, as they do not include the aforementioned properties. The second factor is the divergence of information found in the marking file of an exam versus the information present in the DICOM header of the same exam, which invalidates the marking. Therefore, our methodology was applied to 800 exams.

3.2. Stage 2: extraction and reconstruction of the pulmonary parenchyma

To start the process, it is important to be aware that there are secondary structures (trachea, bronchi, etc.) that must be removed before the main segmentation. The steps presented in this stage were based on the methodology proposed by Netto et al. [14], which describes a model for segmentation of structures that are not relevant for the process of automatic detection of lung nodules (as is the case of these secondary structures), so that only the pulmonary parenchyma passes to the following stage. This stage is composed of three sub-stages: (2.1) Extraction of the pulmonary parenchyma; (2.2) Reconstruction of the parenchyma; (2.3) Isolation of the parenchyma. Further details of these steps can be found in other works by our research group [14,25].

3.3. Stage 3: enhancement of the structures inside the pulmonary parenchyma

As we previously mentioned, sometimes the complexity and the low contrast of the CT may hinder the action of CAD systems. There studies that employ enhancement techniques to minimize the effects of this problem, as seen in Camarlinghi et al. [19] and Xin-Jiang and Shengdong-Nie [26].

First, we applied a quadratic enhancement for selective contrast increase. The application of this enhancement causes an increase in the amount of noise in the image, affecting the performance and quality of stage 4. To overcome this problem, we used two techniques intended to smoothen the image by eliminating or reducing the amount of noise. The first technique is the Gaussian Filter and the second is the median filter [27]. Fig. 2 illustrates the application of these techniques.

3.4. Stage 4: segmentation of nodule candidates

After stages 1, 2 and 3, we had the exams ready for the application of the algorithm for segmentation of parenchymal structures.

3.4.1. Stage 4.1: segmentation by clustering

The quality threshold (QT) algorithm was applied to the set of alternative data. This algorithm was originally proposed for segmentation of genes of the human chromosome [28,29] and has two main advantages: efficiency (or speed) and its easy implementation.

QT clusters the points whose Euclidean distances to their starting point are smaller than or equal to the quality threshold, the limiting parameter in the formation of clusters. This process is called formation of candidate clusters. We used a threshold of 90 for the whole methodology. This threshold was chosen because it presented better results in the cases used. To illustrate the analysis, Fig. 3 shows the results of segmentations with threshold values that are lower (50) and higher (150) than 90, as well as the threshold used in our method (90). Fig. 3A presents a parenchyma where QT partially segments the nodule. In this example, a threshold of 50 was used. Fig. 3B presents the result of the segmentation for a threshold of 90, which yielded the correct segmentation of that particular nodule. Finally, Fig. 3C shows the segmentation result for a threshold of 150, where, beside the nodule, a different structure with similarity close to the nodule's was also segmented.

Next, from among all the other candidates, we chose the one with the highest score to be the actual cluster. Finally, we removed the points that form this candidate from the processing. This process of forming and choosing the best candidates was repeated until all points were processed.

The cluster may present the problem of having disconnected parts in its formation. Fig. 4 shows an example of such problem,

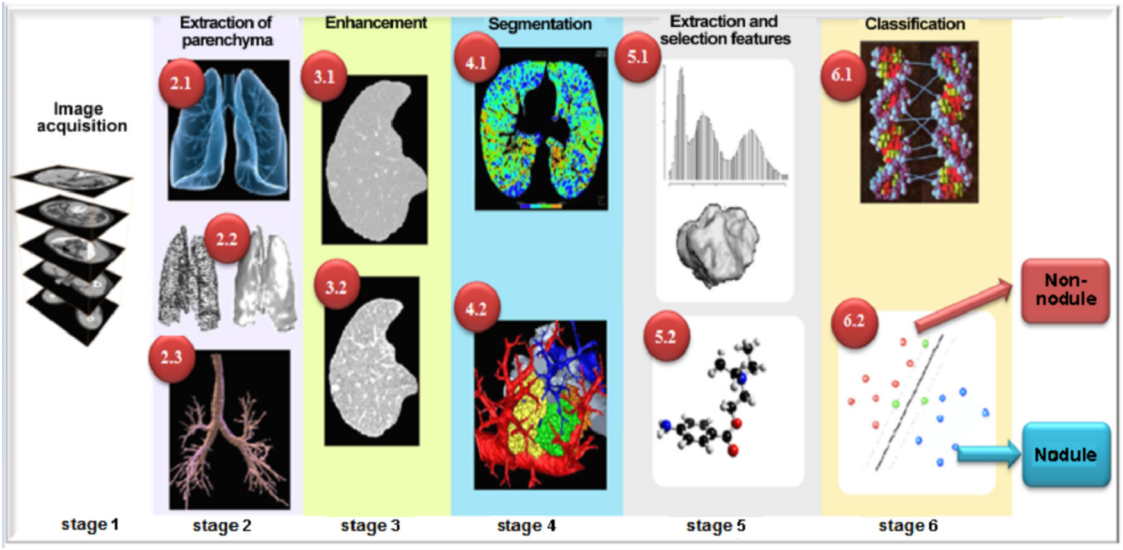

Fig. 1. Scheme of the proposed methodology for automatic detection of lung nodules.

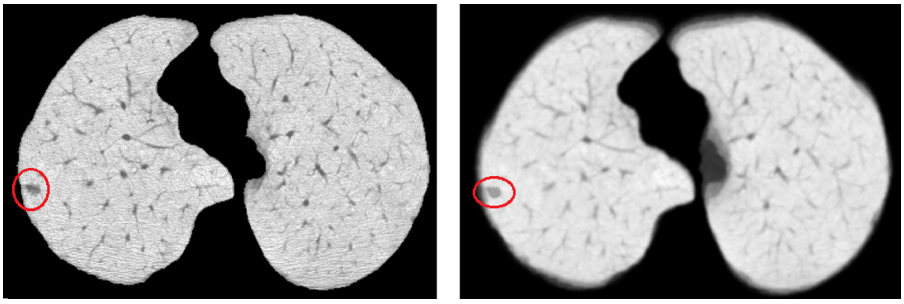

Fig. 2. Enhanced image and smoothed image.

with clusters painted in yellow inside a red circle. This occurs because we use the similarity between the intensities of voxels to form the cluster. To solve this problem, we adopted the use of a region growing algorithm.

3.4.2. Stage 4.2: region growing

The region growing algorithm was used to solve the problem of clusters with disconnected parts. This algorithm isolates all the structures individually for future analysis. This analysis involves computing the diameter of each structure. Structures with diameters between 3 and 30 mm pass to the next stage. Fig. 5 shows an example of this stage.

3.5. Stage 5: feature extraction and selection

Segmentation will produce nodules and non-nodules. The correct extraction of features that distinguish nodules from non-nodules is a key step for the following stages of the methodology.

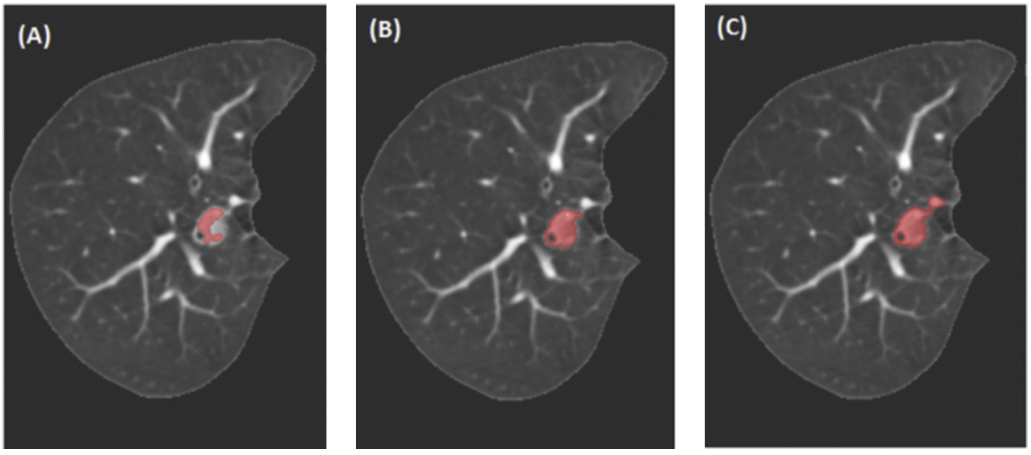

Fig. 3. Results of QT application with thresholds (A) 50, (B) 90, and (C) 150.

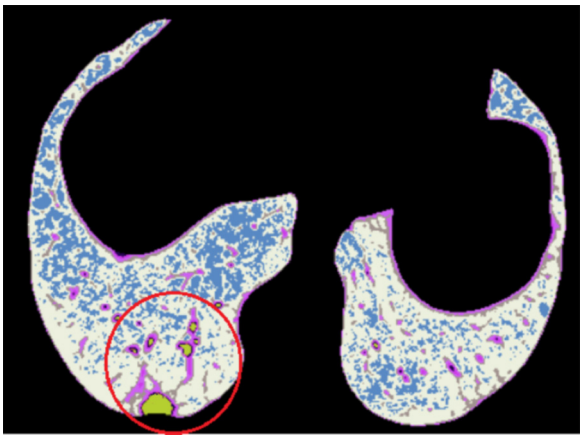

Fig. 4. Cluster with disconnected part(s). (For interpretation of the references to color in this text, the reader is referred to the web version of this article.)

3.5.1. Stage 5.1: features extraction

The extracted features are based on shape and texture. The shape measures (spherical disproportion, spherical density, sphericity, weighted radial distance, elongation and Boyce-Clark radial shape index) assess the geometry of each structure, looking for the most spherical ones, as nodules are more spherical than other structures. We chose these measures because they had been investigated in previous studies by our group and achieved satisfactory results. A more detailed description of the measurements can be found in da Silva Sousa et al. [25].

A better discrimination of textures can be provided by statistics. In this work, we used measures computed through the histogram and the co-occurrence matrix, based on Netto et al. [14]. With these measures, we can evaluate the case of structures that have spherical properties similar to nodules but which do not have matching texture features. The extraction of the five measures (contrast, energy, entropy, homogeneity and momentum) from the co-occurrence matrix (M) of the nodule candidate is done through Eq. (1), where v is the function that supplies one of the 26 neighbors of a voxel according to the index α, and P_{xyz} is the value of the voxel with coordinates x, y, z. In addition to the mentioned measures, we use six other measures based on the histogram (mean, standard deviation, obliquity, kurtosis, energy and entropy) and ten measures based on diversity indexes (Simpson's and Shannon's indexes).

$$M(i,j) = count(P_{x,y,z}, v(P_{x,y,z}, \alpha))|P_{x,y,z} = i, v(P_{x,y,z}, \alpha) = j, \alpha \in \{1, \ldots, 26\} \quad (1)$$

Diversity analysis, used in ecology to measure the biodiversity of an ecosystem, seeks to identify the distribution of a group of species and its interpellations. Diversity refers to the variety of species in a given community or habitat. Biodiversity is the relationship between the number of species (richness), the distribution pattern of individuals in their species (uniformity), and the domain of one or more species over others (dominance). All of these features can be measured and investigated by means of indexes, which are usually classified according to the coverage of the location (alpha) or among several habitats (beta) [30].

In general, diversity indexes can be used to measure the diversity of a population where each member belongs to a single group or species. We defined the voxels as being individuals, with their tonalities representing the sets of species present in the ecosystem. Thus, in this work we want to investigate Simpson's and Shannon's indexes as texture descriptors, that is, the distribution and relationship of the voxels in the ROI volume.

- Shannon's index: Originating in information theory, Shannon's index assumes that the individuals are randomly sampled from an infinitely large community and that all species are represented in the sample [30]. This index is given by Eq. (2).

$$H' = -\sum_{i=1}^{S} p_i \ln p_i \quad (2)$$

where H' is Shannon's index, S is the number of species present in the sample, p_i is the probability that an individual belongs to species i, which is given by Eq. (3):

$$p_i = \frac{n_i}{N} \quad (3)$$

where p_i is the probability of selecting an individual from species i, n_i is the number of individuals of species i, and N is the total number of individuals in the sample.

The values obtained through Shannon's index (H') vary from 0 to $\log S$, respectively when there is only one species, and when all species are represented by the same number of individuals.

- Simpson's index: This is an index that estimates the probability that two individuals of the same species are picked from the same sample. This was the first index to be used in ecological studies and expresses the dominance of a certain species in the community. Dominance is when one or a few species hold most of the individuals in the community. Simpson's index is given by:

$$D = \sum_{i=1}^{S} p_i^2 \quad (4)$$

where p_i is obtained through Eq. (3).

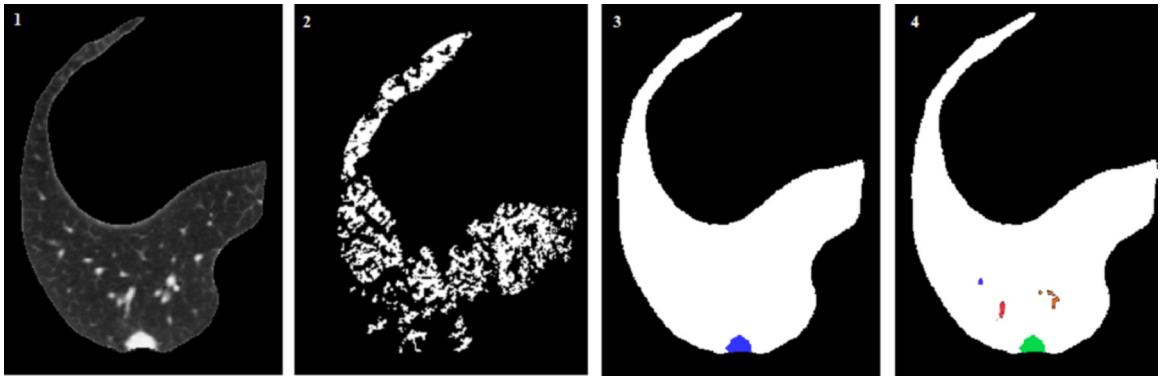

Fig. 5. (1) Original image. (2) Larger cluster isolated. (3) Smaller cluster isolated. (4) Final result.

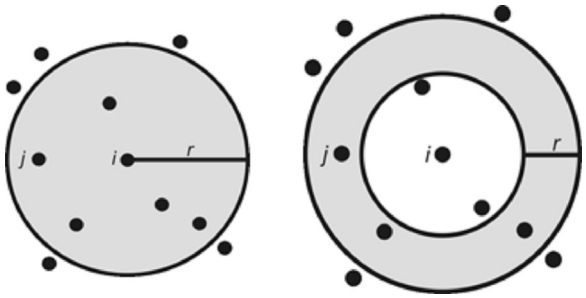

Fig. 6. Ring-based and sphere-based approaches used in the application of Shannon's and Simpson's indexes.

In this form, Simpson's index is considered a biased estimator because the computation of the probability of picking two individuals from the same species, performed considering sample replacement, is indicated for infinitely large communities [30].

The most suitable relation for finite communities consists of computing the probability of picking two individuals of the same species without sample replacement and is given by the following:

$$D = \sum_{i=1}^{S} \frac{n_i(n_i - 1)}{N(N - 1)} \quad (5)$$

where n_i is the number of individuals of species i, and N is the size of the sample. As opposed to Shannon's index, Simpson's index gives more weight to common species [30]. This means that a slight increase in the number of individuals of a rare species causes slight alterations in the value of Simpson's index. In the case of Shannon's index, any change in the relation between individuals and species causes a considerable variation in its output.

These two indexes were applied to each nodule candidate, using ring-based and sphere-based approaches, as shown in Fig. 6. Six circles and four rings were extracted and equally divided between the two indexes.

3.5.2. Stage 5.2: feature selection

After extraction, we performed the selection of the features that best discriminate the nodule and non-nodule classes. For this, we used the stepwise discriminant analysis technique [31].

3.6. Stage 6: selection of the best model and classification

Stage 5 selected the best features, so that the classifier could discriminate the structures as nodules and non-nodules. This stage supplies the definition of the best training model and finally the classification.

3.6.1. Stage 6.1: selection of the best model

The creation of a suitable model for result validation is necessary. For this, we used a genetic algorithm (GA). The literature presents a wide variety of methods for controlling the size of the population, the genetic diversity, and for improving the search methods. Dubey and Mastorakis [32] present a variant of the GA, called micro-genetic algorithm (MGA). The MGA uses a reduced starting population, 50 individuals at most [33]. With such a small population, the model tends to converge quickly to a local optimum. When this happens, the best individual is kept in the population, and all others are randomly generated again. This process is repeated until the stop criterion is reached. According to Ribas et al. [34], the MGA has demonstrated good performance in comparison to traditional GAs.

In this work, we used the MGA to generate a training set that best classifies new test cases. Its choice was intended to minimize the long run-time needed to compute the fitness of one single individual. The use of traditional GAs would be unfeasible, as the number of individuals per generation is too large.

The best individuals are selected based on the steps below, and as illustrated in Fig. 7:

1. From each ROI (nodule and non-nodule) obtained in the segmentation (Section 3.4.1), a feature vector is extracted (texture and geometry – Section 3.5.1). The set of all feature vectors makes up the analysis base (AB).
2. Each element in array A1 contains the position of a feature vector extracted from a nodule-ROI in the AB.
3. Each element in array A2 contains the position of a feature vector extracted from a non-nodule-ROI in the AB.
4. The mutation and crossing genetic operators modify the values contained in A1 and A2. However, the values cannot be repeated. Elements of A1 cannot be present in A2 or vice versa. That is, the mutation and crossing genetic operators do not mix data from both arrays.
5. For each individual of a generation, the feature vectors selected in A1 and A2 will be trained by the support vector machine (SVM) [31] machine-learning algorithm. The other feature vectors from the AB, which were not selected, will be tested also using the SVM. The fitness of each individual is computed through the sum of specificity, sensitivity and accuracy. This process is repeated until the fitness of the best individual is the same for 300 consecutive generations.
6. At the end of evolution, the best training model is formed by all feature vectors whose positions are contained in A1 and A2 of the individual with best fitness in the last generation.
7. Finally, the selected model is validated. This is performed by means of a classification where all remaining nodules in the analysis base are used, by observing a proportion of adding three non-nodules for every nodule. Sensitivity, specificity and accuracy are computed to evaluate the effectiveness of the model.

3.6.2. Stage 6.2: classification

With the best individual supplied by the GA, we define the model to perform the validation and tests using SVM [31]. SVM is a powerful, state-of-the-art algorithm with strong theoretical foundations based on the Vapnik-Chervonenkis theory. SVM has strong regularization properties. Regularization refers to the generalization of the model to new data. This was the main reason for choosing this classifier in our work. The accuracy of an SVM model is largely dependent on the selection of kernel parameters such as C and γ for a radial basis function (RBF). We used the LibSVM [35] software to estimate these two parameters. The values of C and γ are estimated based on the individuals selected in stage 6.1.

To evaluate the performance of our method, we computed sensitivity, specificity and accuracy. Sensitivity is given by TP/(TP + FN). Specificity is obtained by TN/(TN + FP), and accuracy is given by (TP + TN)/(TP + TN + FP + FN), where TP is true positive, TN is true negative, FP is false positive and FN is false negative. In this manner, the nodules that were correctly computed are reported as true positives.

The performance of computer-based detection techniques can be summarized using receiver operating characteristic (ROC) curves [36]. A ROC curve indicates the true positive rate (sensitivity) as a function of the false positive rate (1 – specificity). Region-based analysis can be summarized using free response operating characteristic (FROC) curves [37]. This is similar to ROC analysis, except that the false positive rate on the x-axis is replaced by the number of false positives per image. In this case, a definition of a detected region is needed, and a typical approach expects a 50% overlap

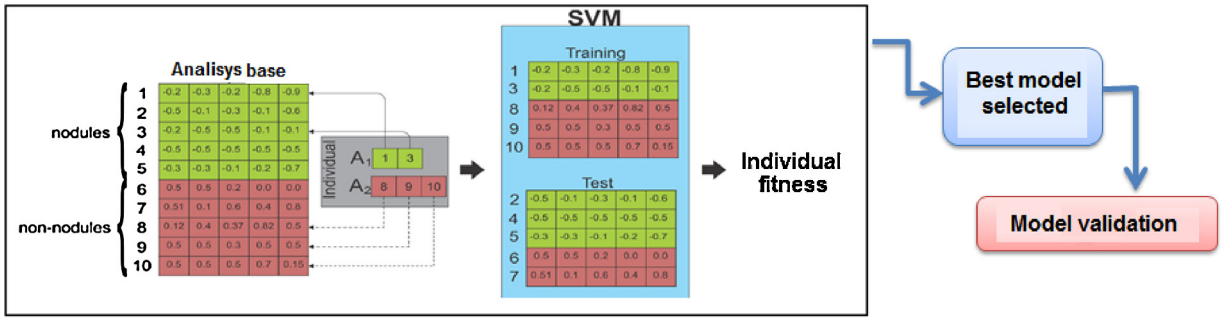

Fig. 7. Illustration of how to choose the best training model.

between the annotated and the detected regions to indicate a true positive.

4. Results and discussion

In this section, we present the results achieved using the proposed methodology for automatic detection of lung nodules described in Section 3. The strategy for result analysis is the following: (1) acquire and define the image database used for training, validating and testing the methodology; (2) demonstrate how the micro-genetic algorithm defines the most suitable model and then its training and validation with the SVM; and (3) display the test results (in these tests, new CT scans are used and the metrics explained in Section 3.6.2 are calculated); (4) display and analyze the application of the methodology to four cases of success and failure; and (5) finally, compare the results with the results of related works. Fig. 8 allows a better visualization of the flow of all tests and analyses performed in this work.

4.1. Image acquisition

The images used to test and validate the proposed method come from the LIDC-IDRI database [23]. The method was applied to 800 exams. For the purpose of testing and validation, the set of images was divided into two parts: 80% (640 exams) were used for defining and validating the model, and the remaining 20% (160 exams) were used for testing.

The division of the images was based on the degree of subtlety rating as defined by LIDC-IDRI radiologists. This degree is associated with how difficult it is to detect the nodule, taking into account its size, contrast and location. The subtlety rating measures the difficulty of nodule detection on a scale from 1 to 5; i.e., the lower the subtlety rating is, the higher the difficulty of detection. In our study, we ensured the presence of nodules with a subtlety ranging from 1 to 5. The choice of the 20% of test exams was made at random, ensuring the presence of all degrees of subtlety.

Based on the markings made by specialists, in our tests, we consider as a nodule the region inside these markings, as they may be larger than the actual nodule.

4.2. Training and validation

At the end of stages 2 and 3 (Sections 3.2 and 3.3), the segmentation of the structures inside the parenchyma was performed by QT (Section 3.4.1). The only parameter needed for its execution is a quality threshold, which delimits the process of forming candidate clusters. The threshold used in the whole method was 90.

The process of forming clusters presents a problem when structures that are part of the same cluster are totally disconnected. To solve this problem, we used a region growing algorithm (Section 3.4.2). After applying this algorithm, we calculated the diameter of each structure. Structures with a diameter equal to or higher than 3 mm and equal to or lower than 30 mm pass to the next stage.

At the end of stage 4, the extraction and selection of the most suitable features are performed. The shape and texture measures, described in Section 3.5.1, were extracted from each structure segmented in the previous stage.

When stage 5.1 was concluded, a total of 27 features were obtained from each segmented structure. Then, we selected the features that best described the nodule and non-nodule classes. The selection was made through the application of the stepwise discriminant analysis (Section 3.5.2). We selected 17 features in total.

For a good classification result, it is important to rely on a good model that best represents the whole image database. This model results from the classifier learning. Based on the results achieved with this model, it is possible to determine whether the classifier is capable of generalization by giving positive responses to new cases to be tested.

To create this model, we used MGA (Section 3.6.1) to select the individuals who best represent the whole data sample, thus guaranteeing the selection of the best model.

From this crossing, an SVM (Section 3.6.2) classifies the individuals who form the starting population, serving as training base for the classifier and being tested with the other individuals. The classification uses the RBF, and the values estimated for C and λ were 512 and 0.00712343, respectively.

The MGA result is directly related to the image database, i.e., the number of individuals must be proportional to the diversity of the analyzed data. Based on this fact, the proposed methodology used a proportion of 1–3. This means that for each nodule, 3 non-nodules were inserted in the training sample. This was because of the high number of non-nodules (high diversity of samples) in the training set compared with the number of nodules. With these proportions, the classifier was able to distinguish both classes considered in the classification more accurately.

The results of stage 5.2 include 1480 individuals (nodules and non-nodules) extracted from 355 exams, with an average of 1734 generations converging to the final result. Table 1 shows the number of suitable individuals selected by the MGA.

The validation of the best defined model yielded a sensitivity of 81.82%, specificity of 97.73% and accuracy of 92.42%.

Table 1
Result of selection by MGA.

	Total individuals	Selected individuals
Exams	600	355
Nodules	458	370
Non-nodules	47,067	1110

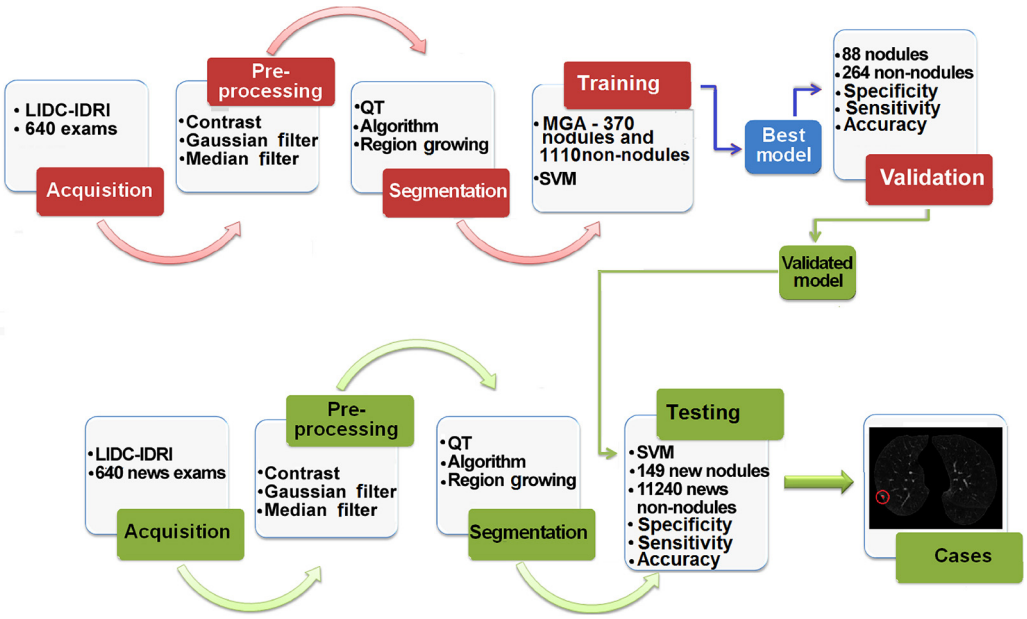

Fig. 8. Results analysis flow.

4.3. Tests

To conduct the tests, we used a set of 160 exams containing 182 nodules. From this total, 7 nodules were lost during the parenchyma extraction (stage 2), as they were connected to the thoracic wall and had a relatively small diameter (between 3 and 4 mm), low contrast and very similar texture to the parenchyma tissue. The enhancement (stage 3) eliminated 3 small nodules that were very close to the edge of the parenchyma. In this case, the error resulted from the formation of an edge with the application of the enhancement.

Of the remaining 172 nodules, stage 4 successfully segmented 167 nodules and was unable to segment 5 nodules. The observed errors are associated with small nodules and to their contrast, as some nodules were segmented together with other nearby structures with textures similar to that of a vessel. This led the QT to group these structures in the same cluster.

The stage of cluster isolation using the region growing algorithm made mistakes in 18 nodules, passing a total of 149 nodules to the classification stage. The errors in this stage were also associated to the size and contrast of the nodules. We also observed that, in this stage, the discarded nodules had markings made by specialists regarding the number of slices between 1 and 2; i.e., the nodule was detected, but, in some cases, only one slice contained the segmented structure, so the region growing algorithm discarded them.

All cases of eliminated nodules in the above-mentioned stages had subtlety values between 1 and 2, meaning that they were the most difficult to detect, according to the LIDC-IDRI. In Section 4.6, we show cases with the mentioned errors. Table 2 shows the number of segmented nodules for the five degrees of subtlety.

The results achieved during the tests demonstrate that the classifier is capable of effectively differentiating between lung nodules and other structures. This is proven by the sensitivity of 85.91%, specificity of 97.70%, accuracy of 97.55%, FROC of 0.8062 and rate of false positives of 1.82 per exam (total of 140) and 0.008 per slice (total of 29,680). These results were achieved using the measures described in Section 3.5.1. If we do not use the diversity indexes in this stage of the methodology, the values decrease for sensitivity (65.36%), specificity (85.16%) and accuracy (84.97%), and the rate of false positives increases (6.8 per exam). This demonstrates the importance of these indexes for our method. We also performed tests without reducing the features (Section 3.5), and the best achieved results were a sensitivity of 80.67%, specificity of 93.45%, accuracy of 92.18% and false positive rate of 2.48 per exam. By reducing the features, we achieved higher evaluation metrics.

The time spent for the whole method to be applied to 248 slices (average number of slices in the 800 exams from LIDC-IDRI used) appears promising, in that it required 13.56 min. However, the region growing stage (stage 4.2) requires improvement as this is the most time-consuming stage. The performance of the QT algorithm (stage 4.1) was very good, as it performed segmentation in a relatively short time, considering the amount of information that is processed. Table 3 shows the total and average run-time of each stage of the methodology.

4.4. Comparison of results obtained with and without MGA selection

To evaluate the use of the MGA approach for selecting the best classification model, we performed two experiments. The first

Table 2
Number of nodules based on subtlety.

Subtlety	Success	Failure	Total
1	2	15	17
2	5	18	23
3	34	0	34
4	42	0	42
5	66	0	66
Total	149	33	182

Table 3
Average run-time of each stage and total run-time of the methodology.

Stage	Average run-time per slice (min)	Average run-time per exam (min)
2	0.0049	1.23
3	0.0018	0.47
4.1	0.0015	0.39
4.2	0.041	10.30
5 and 6	0.005	1.34
Total	0.055	13.56

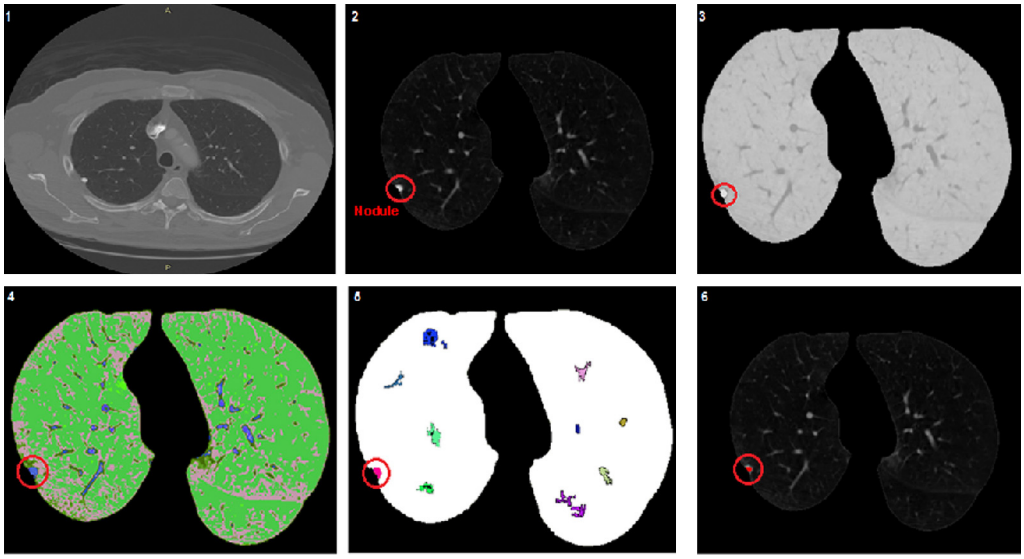

Fig. 9. Successful results of the method applied to exam 1.

Table 4
Experiment 1: results using all the available training database data.

Models	Sensitivity (%)	Specificity (%)	Accuracy (%)
1	69.3	93.57	83.33
2	68.7	93.86	83.56
3	68.3	93.95	83.35
4	67.7	94.5	83.7
5	69	93.43	82.2

experiment was performed using all the training database data available (47,525 nodules candidates) and generated five models changing the SVM parameters C and λ. Then, was used the same test database used previously in Section 4.3. As seen in Table 4, the amount of non-nodules present in the training test has a direct influence in the classifier learning performance, making it more sensible to non-nodules candidates.

The second experiment was performed generating five models using the same amount of nodules candidates that were used previously in Section 4.2. For the experiment, we used a number of randomly defined candidates. Table 5 shows the results for this experiment with the same database test used in the paper previously. The results matched the results obtained using the MGA. The sensitivity is improved, but the specificity is reduced.

In Section 4.3, we may see that using MGA, we obtained a specificity of 85.91%, a specificity of 97.7% and an accuracy of 97.55%. That is, using the MGA to select the best training set, we achieved even better results and selected the best candidates for the training base for each class.

Table 5
Experiment 2: results using the same quantity of nodules candidates that were used with MGA training.

Models	Sensitivity (%)	Specificity (%)	Accuracy (%)
1	74.4	91.32	81.98
2	72.3	92.61	82.21
3	73.3	91.92	82.06
4	73.7	92.1	82.12
5	72.9	92.82	82.29

4.5. Successful cases

In this section, we present the successful detection of lung nodules in two exams. The first contains a small nodule and the second has a juxtapleural nodule.

4.5.1. Exam 1

The selected exam had markings indicating the presence of a juxtapleural nodule. It has higher visibility in slice 105 and has a diameter of approximately 5.4 mm and estimated volume of 82.84 mm^3. Fig. 9-1 presents the original image. Fig. 9-2 shows the results of stage 2. Fig. 9-3 displays the results of stage 3. The results of stage 4.1 are shown in Fig. 9-4. Fig. 9-5 shows the results of stage 4.2. Fig. 9-6 presents the nodule.

4.5.2. Exam 2

In the second example, the exam had markings indicating the presence of a nodule. It has higher visibility in slice 75 and has a diameter of approximately 3.2 mm and estimated volume of 14.09 mm^3. Fig. 10-1 presents the original image. Fig. 10-2 shows the results of stage 2. Fig. 10-3 shows the results of stage 3. The results of stage 4.1 can be seen in Fig. 10-4. Fig. 10-5 shows the results of stage 4.2. Fig. 10-6 presents the detected nodule.

4.6. Failed cases

Next, we present two cases in which the method failed and identify in which stage the error occurred.

4.6.1. Exam 1

Fig. 11-1 shows the original image. Fig. 11-2 displays the results of stage 3 with an error where a nodule with a diameter of 3.8 mm appears connected to the edge. The nodule is inside the red circle.

4.6.2. Exam 2

In the second example, we have Fig. 12-1 showing the original image. Fig. 12-2 shows the results of stage 2. Fig. 12-3 shows the results of stage 3. Fig. 12-4 shows the results of stage 4.1. Fig. 12-5 presents the results of stage 4.2. Finally, Fig. 12-6 presents the actual nodule and the false positive. In this case, the method failed in stage 4.1, as only one of the two nodules was detected. The other failure is in stage 4.2, which discarded the small segmented structure.

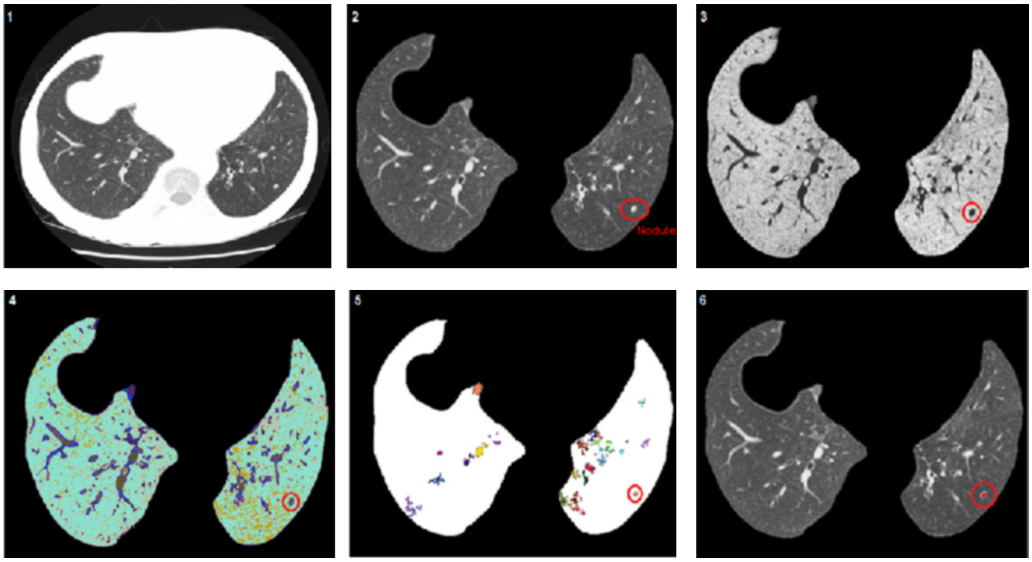

Fig. 10. Successful results of the method applied to exam 2.

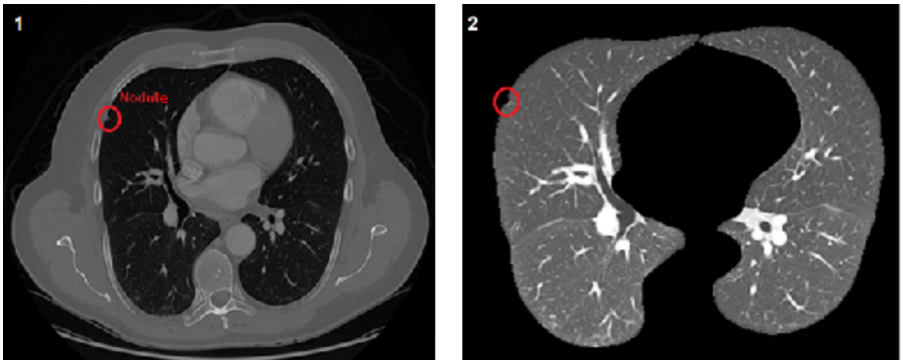

Fig. 11. Failed method for exam 1. (For interpretation of the references to color in this text, the reader is referred to the web version of this article.)

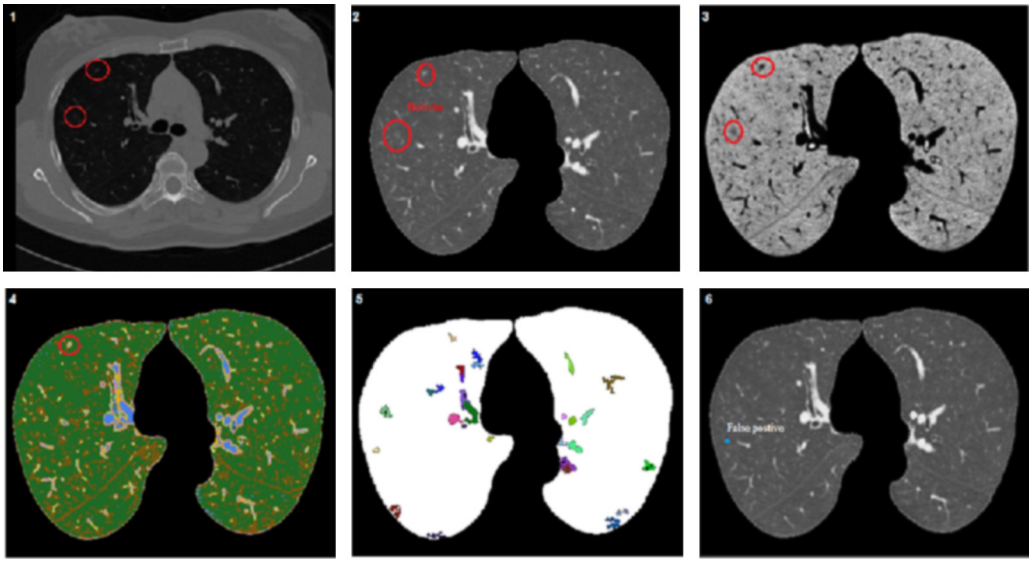

Fig. 12. Failure of method for exam 2.

Table 6
Comparison of results with related works. *Legend*: P, proprietary; L, LDCT; E, ELCAP and LI, LDCT-IDRI.

Work	Database	Number of exams	SE	SP	AC	FROC	FP/exam
[5]	P	108	90.2%	–	–	–	8.2
[6]	P	70	85%	–	–	–	14
[7]	L	93	89% = 4 mm 60% < 4 mm	–	–	–	2
[8]	P	8 943 slices	95% –	– –	– –	– –	0.91 –
[9]	P	39 slices	89%				
[10]	P	47 slices	80% > 40pixel	–	–	–	–
[11]	P	95 slices	100%	–	–	–	3/slice
[12]	P	343	84%	–	–	–	–
[13]	P	6	80.5		–	–	7.5
[14]	L	29	85.93%	91%	91%	–	0.138
[15]	L	30	100%	–	–	–	8.4
[16]	L	143	82.66%	–	–	–	3
[17]	L	125	87.50%	–	–	–	4
[18]	L	85	80%	–	–	–	6.1
[19]	L	138	85%	–	–	–	25
[20]	E	50	86%	97%	–	–	–
[21]	LI	429 slices	83%	99%	–	–	–
[22]	LI	59	82.05%	76.47%	80.36%	–	–
Proposed method	LI	140 29,680 slices	85.91%	97.7%	97.55%	0.8062	1.82 0.008/slice

Both failures dealt with exams with very low contrast, which makes detection more complex.

4.7. Result comparison

Table 6 presents the results in terms of sensitivity (SE), specificity (SP), accuracy (AC), FROC, and number of false positives per exam (FP/exam) achieved by the proposed methodology as compared with the results from other related works (Section 2). We must emphasize that a reliable comparison would only be possible if all works used the same exam database, the same exams for training and tests, and other such parameters.

4.8. Discussion

We evaluated the proposed methodology applied to a set of 800 exams from the LIDC-IDRI database, divided into training/validation and testing in proportions of 80% and 20%, respectively. The experimental results lead to the following conclusions:

1. The time spent for the whole method appears promising, as it takes, on average, 13.56 min for processing of the whole set. However, stage 4.2 requires improvement, as it demands a greater proportion of run-time. We highlight the performance of the QT algorithm (stage 4.1), which performs segmentation in a relatively short time, considering the amount of information to be processed. This run-time may vary (increase or decrease), depending on the threshold.
2. In some cases, the methodology failed. This is particularly the case in the set of nodules with subtleties 1 and 2, which are involved in a larger number of detection errors. This may be due to their low contrast, minimum diameter (3 mm), original markings by specialists in just one slice, and connection with other structures, such as vessels.
3. Some of the methodologies cited in Table 6 use sensitivity as a metric for detection performance, and the proposed methodology achieved good sensitivity, 85.91%, which is comparable to the best methodologies in the literature. This metric is important to assess how capable the methodology is at detecting nodules.

   Only a few methodologies use specificity as a metric for validating performance, which prevents us from knowing the ability of other methodologies to correctly detect non-nodules. The proposed methodology achieved a specificity of 97.7%, proving its high performance in reducing false positives.

   Regarding accuracy, which computes the total performance of the methodology, a value of 97.55% was achieved, thus proving the good ability to discriminate between nodules and non-nodules.

   Finally, the rate of false positives per exam was 1.82. Only the methodology proposed by Opfer and Wiemker [7] used FROC to measure results. In our work, we achieved a value of 0.8062 for FROC. Therefore, in quantitative terms, the proposed method presents, in addition to the very promising results listed above, cases of success in detecting small nodules (between 3 and 10 mm in diameter), juxtapleural nodules, as well as nodules between 10 and 30 mm.
4. Two techniques deserve attention in this methodology, and they are also considered the main contributions of this work. The first is the use of QT algorithm for segmentation, and the second is the use of a genetic algorithm to select the individuals who provide the best model for training. These two techniques displayed good performance and good run-time in performing their tasks.
5. Another contribution, the use of Simpson's and Shannon's diversity indexes combined with other texture and shape measurements, led to a significant reduction in the number of false positives. We believe that these two indexes had good performance and deserve more investigation.
6. In stage 2, as developed by Netto et al. [14], requires improvements to fix the errors that occurred in some cases where the nodules were connected to the thoracic wall.
7. The failures found in the region growing algorithm (stage 4.2) must also be fixed, either by changing the parameters to overcome the problems found or by using another technique that proves to be more efficient for the isolation task, which is the critical point of the methodology.
8. Finally, tests with other shape and texture measurements are needed to decrease the rate of false positives.

5. Conclusion

In the present work, we developed a methodology for the automatic detection of lung nodules. The methodology is divided into the following stages: acquisition, extraction and reconstruction of

the pulmonary parenchyma, QT-based segmentation of nodule candidates, extraction of measures based on shape and texture, and selection of the model and classification.

It is important to emphasize that the LIDC-IDRI is an extremely complex and diversified database, i.e., it contains countless different cases of lung nodules. This database is formed through a partnership between the Lung Image Database Consortium and the Image Database Resource Initiative and therefore has exams that were acquired through several different tomographers, which presents difficulties for detection using CAD systems. The following properties may be obstacles for detection: different contrasts between tomographers, smaller or larger spacing between slices, and, in some cases, divergences between the header of the exams and the annotations made by specialists.

The results achieved in this work prove that the developed methodology has an acceptable performance. With the model resulting from the MGA, new cases could be successfully classified for testing the methodology. In the testing stage, with 140 new exams, with a ratio of 80% for training and 20% for testing, good results were achieved, as indicated by the sensitivity of 85.91%, specificity of 97.70% and accuracy of 97.55%, with a false positive rate of 1.82 per exam and 0.008 per slice, and FROC of 0.8062.

To increase the rates that measure the performance of the methodology, corrections and improvements of the aspects mentioned in Section 4.8 are necessary. These corrections and improvements will make the methodology neater and more efficient, thus increasing its contribution.

Lung cancer stands out among cancers because it presents the highest mortality rate, in addition to having one of the smallest survival rates after diagnosis (five years for approximately 14–20% of patients). An early diagnosis can result in a considerable increase in the survival chance of patients. The methodology proposed herein contributes to this and can be a useful tool for specialists who are attempting to detect nodules.

It should also be noted that public health networks sometimes lack specialists. The funding to increase the number of professionals is also limited. Redirecting qualified personnel to work in less repetitive tasks may be a better use of their abilities. One of the steps toward this would be to use the proposed methodology for a preliminary analysis of thorax computed tomography exams, only leaving the task of validating the results to a physician.

Finally, the proposed methodology is also an affordable solution, as it can be implemented using ordinary computers available in many hospitals.

Acknowledgements

The authors acknowledge CAPES, CNPq and FAPEMA for financial support, and the National Cancer Institute and the Foundation for the National Institutes of Health and for their critical role in the creation of the free publicly available LIDC-IDRI Database used in this study.

References

[1] Chen W, Li Z, Bai L, Lin Y. NF-kappaB in lung cancer, a carcinogenesis mediator and a prevention and therapy target. Frontiers in Bioscience (Landmark Edition) 2011;16:1172–85, http://dx.doi.org/10.2741/3782. ISSN: 1093-4715.

[2] do Câncer IN. Estimativas da incidência e mortalidade por câncer no Brasil; 2013. Available at: http://www.inca.gov.br/estimativa/2012/ [accessed 01.07.13].

[3] Chan TF, Vese LA. Active contours without edges. IEEE Transactions on Image Processing 2001;10(2):266–77, http://dx.doi.org/10.1109/83.902291. ISSN: 1057-7149.

[4] Leef 3rd J, Klein J. The solitary pulmonary nodule. Radiologic Clinics of North America 2002;40(1):123–43, http://dx.doi.org/10.1056/NEJMcp012290, ix.

[5] Ye X, Lin X, Dehmeshki J, Slabaugh G, Beddoe G. Shape-based computer-aided detection of lung nodules in thoracic CT images. IEEE Transactions on Biomedical Engineering 2009;56(7):1810–20, http://dx.doi.org/10.1109/TBME.2009.2017027. ISSN: 0018-9294.

[6] Dehmeshki J, Ye X, Casique MV, Lin X. A hybrid approach for automated detection of lung nodules in CT images. In: Kovacevic E, Meijering J, editors. IEEE international symposium on biomedical imaging: nano to macro, 2006. 3rd. Institute of Electrical and Electronics Engineers; 2006. p. 506–9, http://dx.doi.org/10.1109/ISBI.2006.1624964.

[7] Opfer R, Wiemker R. Medical imaging 2007: image perception, observer performance, and technology assessment. San Diego, CA. In: Society of Photo-Optical Instrumentation Engineers (SPIE) conference series, vol. 6515, SPIE proceedings. 2007., http://dx.doi.org/10.1117/12.708210.

[8] Tong J, Da-zhe Z, Ying W, Xin-hua Z, Xu W. Computer-aided lung nodule detection based on CT images. In: Zhixiang Li MH, Takeuchi H, editors. 2007 IEEE/ICME international conference on complex medical engineering. Institute of Electrical and Electronics Engineers; 2007. p. 816–9, http://dx.doi.org/10.1109/ICCME.2007.4381854.

[9] Nie S-D, Chen Z-X, Li L-H. A CI feature-based pulmonary nodule segmentation using three-domain mean shift clustering. In: Tang YY, editor. IEEE international conference on wavelet analysis and pattern recognition, 2007. ICWAPR'07, vol. 1. Institute of Electrical and Electronics Engineers; 2007. p. 223–7, http://dx.doi.org/10.1109/ICWAPR.2007.4420699.

[10] yan Jiang H, yu Cheng Z. Segmentation of pulmonary nodules based on statistic features of wavelet coefficients and dual level sets. In: Zhixiang Li MH, Takeuchi H, editors. IEEE/ICME international conference on complex medical engineering, 2007. CME 2007. Institute of Electrical and Electronics Engineers; 2007. p. 648–51, http://dx.doi.org/10.1109/ICCME.2007.4381817.

[11] Sun S-S, Li H, Hou X-R, Kang Y, Zhao H. Automatic segmentation of pulmonary nodules in CT images. In: Lei Y, editor. The 1st international conference on bioinformatics and biomedical engineering, 2007. ICBBE 2007. Institute of Electrical and Electronics Engineers; 2007. p. 790–3, http://dx.doi.org/10.1109/ICBBE.2007.206.

[12] Dehmeshki J, Amin D, Valdivieso M, Ye X. Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach. IEEE Transactions on Medical Imaging 2008;27(4):467–80, http://dx.doi.org/10.1109/TMI.2007.907555. ISSN: 0278-0062.

[13] Miyake N, Kim H, Itai Y, Tan JK, Ishikawa S, Katsuragawa S. Automatic detection of lung nodules in temporal subtraction image by use of shape and density features. In: Liao P, Bin-Yih, Shi, editors. Fourth International conference on innovative computing, information and control (ICICIC), 2009. Institute of Electrical and Electronics Engineers; 2009. p. 1288–92, http://dx.doi.org/10.1109/ICICIC.2009.118.

[14] Netto SMB, Silva AC, Nunes RA, Gattass M. Automatic segmentation of lung nodules with growing neural gas and support vector machine. Computers in Biology and Medicine 2012;42(11):1110–21, http://dx.doi.org/10.1016/j.compbiomed.2012.09.003.

[15] Messay T, Hardie RC, Rogers SK. A new computationally efficient CAD system for pulmonary nodule detection in CT imagery. Medical Image Analysis 2010;14(3):390–406, http://dx.doi.org/10.1016/j.media.2010.02.004.

[16] Xiaomin P, Hongyu G, Jianping D. Computerized detection of lung nodules in CT images by use of multiscale filters and geometrical constraint region growing. In: Lu W, Sun L, Zhang X, Sun J, editors. 2010 4th international conference on bioinformatics and biomedical engineering (iCBBE). Institute of Electrical and Electronics Engineers; 2010. p. 1–4, http://dx.doi.org/10.1109/ICBBE.2010.5517771. ISSN 2151-7614.

[17] Tan M, Deklerck R, Jansen B, Bister M, Cornelis J. A novel computer-aided lung nodule detection system for CT images. Medical Physics 2011;38(10):5630–45, http://dx.doi.org/10.1118/1.3633941.

[18] Suárez-Cuenca JJ, Guo W, Li Q. Automated detection of pulmonary nodules in CT: false positive reduction by combining multiple classifiers. In: Medical Imaging 2011: Computer-Aided Diagnosis, Lake Buena Vista, FL, USA. 2011., http://dx.doi.org/10.1117/12.878793.

[19] Camarlinghi N, Gori I, Retico A, Bellotti R, Bosco P, Cerello P, et al. Combination of computer-aided detection algorithms for automatic lung nodule identification. International Journal of Computer Assisted Radiology and Surgery 2012;7(3):455–64, http://dx.doi.org/10.1007/s11548-011-0637-6.

[20] Farag A, Ali A, Graham J, Farag A, Elshazly S, Falk R. Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose CT scans of the chest. In: Wright XLM, Steve P, editors. Medical imaging 2011: Computer-aided diagnosis. 2011. p. 169–72, http://dx.doi.org/10.1109/ISBI.2011.5872380. ISSN 1945-7928.

[21] Chama CK, Mukhopadhyay S, Biswas PK, Dhara AK, Madaiah MK, Khandelwal N. Automated lung field segmentation in CT images using mean shift clustering and geometrical features; 2013, http://dx.doi.org/10.1117/12.2007910.

[22] Sivakumar S, Chandrasekar C. Lung nodule detection using fuzzy clustering and support vector machines. International Journal of Engineering and Technology (IJET) 2013;5(11):179–85, doi:10.3-513-361.

[23] Samuel GA, McLennan G, Bidaut L, Michael FM-G, Meyer CR, Reeves APea. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Medical

Physics 2011;38(2):915–31, http://dx.doi.org/10.1118/1.3528204. ISSN: 0094-2405.
[24] Pianykh OS. Digital imaging and communications in medicine (DICOM): a practical introduction and survival guide. Berlin Heidelberg, Berlin, Germany: Springer-Verlag; 1998, http://dx.doi.org/10.1007/978-3-540-74571-6.
[25] da Silva Sousa JRF, Silva AC, de Paiva AC, Nunes RA. Methodology for automatic detection of lung nodules in computerized tomography images. Computer Methods and Programs in Biomedicine 2010;98(1):1–14, http://dx.doi.org/10.1016/j.cmpb.2009.07.006.
[26] Xin-Jiang S-N. Segmentation of pulmonary nodule in CT image based on level set method. In: Zhi-Pei Liang K-CC, Kainz W, editors. The 2nd international conference on bioinformatics and biomedical engineering, 2008. ICBBE 2008. Institute of Electrical and Electronics Engineers; 2008. p. 2698–701, http://dx.doi.org/10.1109/ICBBE.2008.1006.
[27] Gonzalez RC, Woods RE. Digital image processing. 2nd ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.; 1992. ISBN: 0201508036.
[28] Jiang D, Tang C, Zhang A. Cluster analysis for gene expression data: a survey. IEEE Transactions on Knowledge and Data Engineering 2004;16:1370–86, http://dx.doi.org/10.1109/TKDE.2004.68.
[29] Xu R, Wunsch D. Survey of clustering algorithms. IEEE Transactions on Neural Networks 2005;16(3):645–78, http://dx.doi.org/10.1109/tnn.2005.845141. ISSN: 1045-9227.
[30] Magurran AE. Measuring biological diversity. MA, USA: Wiley-Blackwell; 2004. ISBN 978-0-632-05633-0.
[31] Duda RO, Hart PE. Pattern classification and scene analysis. New York, USA: Wiley-Interscience Publication; 1973.
[32] Dubey M, Mastorakis NE. Design of genetic algorithm based fuzzy logic power system stabilizers in multimachine Power system. In: Mukhopadhyay S, editor. Joint international conference on power system technology and IEEE power India conference, 2008. POWERCON 2008, ICS'09. Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS); 2009. p. 523–32, http://dx.doi.org/10.1109/ICPST.2008.4745227. ISBN: 978-960-474-097-0.
[33] Toscano Pulido G, Coello Coello C. The micro genetic algorithm 2: towards online adaptation in evolutionary multiobjective optimization. In: Fonseca C, Fleming P, Zitzler E, Thiele L, Deb K, editors. Evolutionary multi-criterion optimization, Faro, Portugal. Lecture notes in computer science, vol. 2632. Berlin, Heidelberg: Springer; 2003. p. 252–66, http://dx.doi.org/10.1007/3-540-36970-8_18. ISBN: 978-3-540-01869-8.
[34] Ribas PC, Yamamoto L, Polli HL, Arruda L, Neves-Jr F. A micro-genetic algorithm for multi-objective scheduling of a real world pipeline network. Engineering Applications of Artificial Intelligence 2013;26(1):302–13, http://dx.doi.org/10.1016/j.engappai.2012.09.020. ISSN: 0952-1976.
[35] Chang C-C, Lin C-J. LIBSVM – a library for support vector machines; 2013. Available at: http://www.csie.ntu.edu.tw/cjlin/libsvm/ [accessed 01.07.13].
[36] van Erkel A, Pattynama P. Receiver operating characteristic (ROC) analysis: basic principles and applications in radiology. European Journal of Radiology 1998;27(2):88–94, http://dx.doi.org/10.1016/S0720-048X(97)00157-5.
[37] Khurd P, Liu B, Gindi G. Ideal AFROC and FROC observers. IEEE Transactions on Medical Imaging 2010;29(2):375–86, http://dx.doi.org/10.1109/TMI.2009.2031840. ISSN: 0278-0062.