Automatic detection of pulmonary nodules:
Evaluation of performance using two different MDCT scanners

Miguel Souto Bayarri¹, Jorge Juan Suárez-Cuenca¹, Pablo G. Tahoces², Marie-Pierre Revel³, Damien Delhaye⁴, José M Carreira¹, Martine Remy-Jardin⁴, Jacques Remy⁴


Correspondence: Miguel Souto Bayarri. Address: Departamento de Radiología, Universidad de Santiago de Compostela Facultad de Medicina, C/ San Francisco 1, 15782 Santiago de Compostela. Telephone: 34-981-570-982. Email: miguel.souto@usc.es.

Received: May 6, 2012 Accepted: June 25, 2012 Published: December 1, 2012

DOI: 10.5430/jbgc.v2n2p55 URL: http://dx.doi.org/10.5430/jbgc.v2n2p55

Abstract
The purpose of this study was to evaluate the diagnostic performance of a computer-aided diagnosis (CAD) system, on the detection of pulmonary nodules in multidetector row computed tomography (MDCT) images, by using two different MDCT scanners. The computerized scheme was based on the ir is filter. We have collected CT cases of patients with pulmonary nodules. We have included in the study one hundred and thirty-two calcified and noncalcified nodules, measuring 4-30 mm in diameter. CT examinations were performed by using two different equipments: a CT scanner (SOMATOM Emotion 6), and a dual-source computed tomography system (SOMATOM Definition) (Siemens Medical System, Forchheim, Germany), with the following parameters: collimation, 6x1.0mm (Emotion 6); and 64×0.6mm (Definition); 100-130 kV; 70-110 mAs. Data were reconstructed with a slice thickness of 1.25mm (Emotion 6) and 1mm (Definition). True positive cases were determined by an independent interpretation of the study by three experienced chest radiologists, the panel decision being used as the reference standard. Free-response Receiver Operating Characteristic curves, sensitivity and number of false-positive per scan, were calculated. Our CAD scheme, for the test set of the study, yielded a sensitivity of 80%, with an average of 5.2 FPs per examination. At an average false positive rate of 9 per scan, our CAD scheme achieved sensitivities of 94% for all nodules, 94.5% for solid, 80% for non-solid, 84% for spiculated, and 97% for non-spiculated nodules. These encouraging results suggest that our CAD system, advocated as a second reader, may help radiologists in the detection of lung nodules in MDCT.

Key words
Computer aided diagnosis, Computer-aided diagnosis, Multidetector row computed tomography, Pulmonary nodule, Automatic detection of pulmonary nodules

1 Introduction
Pulmonary nodules are a frequent incidental finding at chest computed tomography [1]. Most nodules represent benign processes such as granulomas, but only fully calcified and fat-containing nodules can be assumed to be benign.
Unfortunately, a solitary nodule is also the most common presentation for asymptomatic bronchogenic carcinoma. Of interest is that early detection of lung cancer may allow for more timely therapeutic intervention and a more favorable patient prognosis [2]. Multidetector-row computerized tomography (MDCT) is at the moment the best method to detect pulmonary nodules (potential lung cancers). However, this technique generates large volumes of image data, that is, hundreds of images per patient, and can lead to fatigue for the radiologists [3, 4].

Computer-aided diagnosis (CAD) could provide a solution, its purpose being to direct the radiologist’s attention by indicating suspected focal opacities that may represent cancer on the image. It has been shown that CAD for lung nodule detection could increase radiologists’ performance [5-7]. Different approaches have been proposed for automatic detection of lung cancer on CT around the world: while some have evaluated different marketed CAD systems, other investigators proposed other methods and processing techniques [5-15].

We seek to develop an automatic system to help radiologists in early detection of lung nodules (potential lung cancer). In this manuscript, we describe and validate a computer-aided diagnosis (CAD) scheme for automated detection of pulmonary nodules in MDCT images. Therefore, the purpose of our study was twofold: first, to present our CAD scheme; and secondly, to evaluate the diagnostic performance of the CAD system for detection of pulmonary nodules.

2 Materials and methods

2.1 Study design and CT technique

The CAD scheme was applied on CT scans performed in sixty-three patients. The database consisted on 132 lung nodules from 45 males and 18 females. The criteria to determine whether a CT scan was eligible for inclusion in the database was as follows: 1) the nodules were independently identified by three experienced chest radiologists from the different institutions that collaborated in the project; and 2) only cases with complete agreement among the three radiologists were included in the study. That is, the three radiologists decided the true-positives, by reading the anonymized images of the 63 patients, and registering, a) the lobar location, b) the number, and c) the characteristics of the nodules. The coordinator of the project (MS), in an independent review, made an attempt to reach a consensus in each nodule, in order not to artificially increase sensitivity (by including only the most conspicuous nodules) and decrease specificity (by ignoring non-consensus true positives that could be detected by CAD). Nodules were classified as being located either in the peripheral zone (defined as the lung parenchyma within approximately 2cm of the pleural surfaces) or in the central zone (all remaining parenchyma excluding the peripheral regions) [4]. Twenty-eight (44%) of the 63 patients had one nodule, nineteen (30%) had two nodules, six (9%) had three, five (8%) had four, three (5%) had five, one (2%) had six, and one (2%) had seven. There were a total of 18695 CT images. Eighty-eight (66.7%) of the 132 nodules were smaller than 10mm in diameter, but larger than 4mm, 31 (23.5%) were 10mm or larger but smaller than 15mm, 7 (5.3%) were 15mm or larger but smaller than 20mm, and 6 (4.5%) were between 20 and 30mm. The upper limit on the number of nodules per scan has been set at seven; and the lower size of the nodules was 4mm [16]. Ground glass nodules were included. A summary of the database, characteristics and distribution of the nodules, is given in Tables 1-4.

Examinations were performed with a 6-slice CT scanner “SOMATOM-Emotion6”, and a dual-source computed tomography (DSCT) scanner “SOMATOM-Definition”, (Siemens Medical System, Forchheim, Germany). The study was approved by ethics committees in Santiago de Compostela, ES and Lille, FR. We also required informed consent from all patients. The CT protocol consisted on the acquisition over the entire thorax, obtained in the cranio-caudal direction, with the patients scanned in the supine position and after deep inspiration. The acquisition parameters were as follows: collimation, 6×1.0 (Emotion 6); and 64×0.6 (Definition); the exposure parameters were 70-110 mAs (depending on the patient’s weight) and 100-130 kV. Dual energy was not performed. Data were reconstructed: at 1.25mm (Emotion6) and 1mm (Definition) contiguous transverse CT scans of the entire thorax, viewed in both the mediastinal (window width, 350 HU, window center, 30 HU, soft reconstruction kernel) and lung parenchymal (window width, 1600 HU; window center,
-600 HU; high spatial frequency algorithm) window settings. The images were obtained in DICOM file formats directly from the CT modality. All patient data were removed from the images.

2.2 CAD scheme
The method is a four-step process:

1) Segmentation of the lung volume. A segmentation method, based on the application of an adaptive threshold, was used to extract the lung regions. First, a two dimensional segmentation process is performed slice by slice. Pixels are classified according to two categories: pixels corresponding to dense tissue and pixels belonging to the lower density regions. Adaptive thresholds that automatically converge to the optimum gray level value for each image were calculated [13]; it could not be selected a global threshold for all the images, because the gray level values varied on the different CT images, depending on the acquisition parameters. As a result, binary images containing the two previously described regions were obtained. Then, we applied three-dimensional region-labelling techniques for the elimination of remaining artefacts and the extraction of the lung volume. The resulting mask was used in the next steps for the nodule detection process. Figure 1b shows a representative result for the segmentation method.

2) Segmentation and feature extraction of the suspicious regions. To segment the suspicious regions, adaptive thresholds based on gray level values, were applied inside the lung volume. This was performed slice by slice. To reduce the noise, we applied a size test, and, as a result, we eliminated the very small regions, that corresponded to noise (Figures 1c-d). Next, we applied 3D-labelling techniques for the segmentation of the suspicious regions. 3D-feature analysis techniques were also applied over those regions. Suspicious regions were then characterized by extraction of features related to mean gray level value, position, volume, sphericity, and iris filter output.

Figure 1. Computer-aided detection images. (a) CT image with a nodule in the middle lobe (arrow); (b) lung mask; (c) binary image with suspicious regions after thresholding; (d) suspected regions after applying the size test; (e) iris filter output; (f) CAD marker for the final detection step.
3) Application of a filter, for enhancing the possible locations of the nodules. We have developed a 3D version of the iris filter\textsuperscript{[17]}. The iris filter enhances the rounded regions depending on the orientation of the gradient vectors of the region. If the gradient vectors point to the centre of the region, that region will be highly enhanced. In a CT slice, the nodules appear as rounded opacities of homogeneous nature, with a convex shape. Therefore, a nodule would appear in a map of gradient orientations as a circular region, with many gradients towards the centre. The filter transforms each pixel of the original image to another value, which is proportional to the mean angle of the radial direction with the gradient of pixels in a region surrounding the given pixel. The resulting image shows opacities enhanced over the background (Figure 1e).

4) Reduction of false positives. Initially, we applied a volume test to eliminate the small regions. Secondly, another test based on the 3D-iris filter output is also applied (regions with small iris filter output are suppressed). Finally, for the remaining regions, the previously calculated features were analyzed with a linear discriminant analysis (LDA), to discriminate between nodule and non-nodule regions, and therefore, to reduce false positives (software for windows SPSS Inc., Chicago, IL)\textsuperscript{[18]}. To reduce FP detections, we developed a multistage classifier to discriminate between nodules and FPs\textsuperscript{[13]}.

2.3 Statistical analysis and performance evaluation

We utilized FROC (Free-response Receiver Operating Characteristic) methodology. FROC method allows an arbitrary number of abnormalities per image, and permits the observer to indicate both the confidence levels and the locations of all the detected abnormalities. This method analyzes the performance of the system by calculating the area under the AFROC curve (A1), that shows the relationship between the fraction of nodules detected versus the probability of an image to produce one or more false positive findings, thus showing the diagnostic capability of the method. The analysis was the calculation of the FROC curve, which is defined as the representation of the fraction of nodules with true-positive ratings versus the mean number of the false positive responses per case\textsuperscript{[19, 20]}.

A total of 132 nodules were included by the independent review panel. Figure 1f shows the result of the application of the proposed CAD scheme to a CT slice. In order to evaluate the performance of the system, the data set was separated into two completely independent sets. One set was used for developing and training the different algorithms of our CAD scheme, while the second set was used to evaluate the performance of the system. The system was trained with the training set, that was employed to calculate the classification functions. The dataset was partitioned in such a way that approximately the 70% of the cases and nodules were included in the training set. The rest of the cases were employed to validate the system.

The performance of the system was calculated in terms of sensitivity and number of false positives per case, for both the training and the test set. Results, in terms of FROC curves, were are also depicted for the two independent data sets. The different points of the FROC curves were obtained by varying the score of the discriminant values in the classification functions.

3 Results

A CAD scheme for the detection of pulmonary nodules in MDCT was developed. Table 1 shows the results that were obtained at 80% sensitivity for both, the training and the test sets.

| Table 1. Summary of the database and performance, for the 132 pulmonary nodules, at a sensitivity of 80%. |
|------------------|------------------|------------------|------------------|------------------|
|                  | Training Set     |                  | Test Set         |                  |
|                  | Scanner          | patients         | images           | nodules          | FPs/case         | patients         | images           | nodules          | FPs/case         |
| Emotion6         | 14               | 4607             | 34               | 4.5              | 6                | 2015             | 15               | 3.8              |
| Definition       | 30               | 8328             | 58               | 2.6              | 13               | 3745             | 25               | 5.8              |
| Both             | 44               | 12935            | 92               | 3.2              | 19               | 5760             | 40               | 5.2              |
a) Training sets

Scanner Emotion 6: The training set consisted on 14 CT scans. It was composed of 4607 images with a total of 34 nodules. FROC curves (obtained after the reduction of FP detections by the LDA classifier) yielded a sensitivity of 80% at a 4.5 FP detection rate per case. The FROC curve is shown in Figure 2.

Scanner Definition: The training set consisted on 30 CT scans. It was composed of 8328 images with a total of 58 nodules. The FROC curve (see Figure 3) yielded a sensitivity of 80% at a 2.6 FP detection rate per case.

![Figure 2. Scanner Emotion6: FROC curves after applying FP reduction, for training and test sets.](image1)

![Figure 3. Scanner Definition: FROC curves after applying FP reduction, for training and test sets](image2)

b) Test sets

Scanner Emotion 6: The algorithm was applied to an independent test set. This set consisted on 6 CT scans (2015 images) with a total of 15 nodules. After applying the FP reduction, by using the triple-LDA-classifier, the FROC curve of our CAD system resulted in 3.8 FPs per case with a sensitivity of 80% (see Figure 2).
Scanner Definition: The independent test set, for this scanner, consisted on 13 CT scans (3745 images) with a total of 25 nodules. After applying the FP reduction stage, the FROC curve has shown 5.8 FPs per case with a sensitivity of 80% (see Figure 3).

Taking into account all the nodules of the study (including both training and test sets), our CAD scheme yielded a sensitivity of 94% (124/132) of the nodules, with an average of 9 FPs per examination.

Tables 2, 3 and 4 show the results for the detection, depending on the characteristics of the nodules, according to density, morphology, and position (when the sensitivity of the system was set to 94%).

Table 2. Density characteristics of nodules and detection

<table>
<thead>
<tr>
<th>Type of nodule (density)</th>
<th>Detected/total</th>
<th>Detected Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solid non calcified</td>
<td>99/106</td>
<td>93%</td>
</tr>
<tr>
<td>Solid and Calcified</td>
<td>21/21</td>
<td>100%</td>
</tr>
<tr>
<td>Ground Glass opacity</td>
<td>4/5</td>
<td>80%</td>
</tr>
</tbody>
</table>

Table 3. Morphology of nodules and detection

<table>
<thead>
<tr>
<th>Type of nodule (morphology)</th>
<th>Detected/total</th>
<th>Detected percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round or Oval</td>
<td>87/90</td>
<td>97%</td>
</tr>
<tr>
<td>Lobulated</td>
<td>11/11</td>
<td>100%</td>
</tr>
<tr>
<td>Spiculated</td>
<td>26/31</td>
<td>84%</td>
</tr>
</tbody>
</table>

Table 4. Distribution of nodules and detection

<table>
<thead>
<tr>
<th>Type of nodule (location)</th>
<th>Detected/total</th>
<th>Detected percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central</td>
<td>49/51</td>
<td>96%</td>
</tr>
<tr>
<td>Peripheral</td>
<td>75/81</td>
<td>93%</td>
</tr>
</tbody>
</table>

4 Discussion

Lung cancer is a leading cause of death in the world. Accurate staging is critical in the selection of treatment and in the evaluation of prognosis of patients with bronchogenic carcinoma. On MDCT scans (the most sensitive imaging modality for detection of early lung cancer), a pulmonary nodule (a potential cancer) appears as a rounded or irregular opacity, well or poorly defined, measuring up to 3 cm in diameter [21]. Over the last years, there has been a great deal of interest in developing computer-aided diagnosis (CAD) technology for early detection of lung cancer in both radiography and computed tomography (CT). One of the reasons for this is because an increasing number of pulmonary nodules are found incidentally, mainly on MDCT studies [1, 3]. Moreover, in several studies, investigators have shown that the detection rates for lung nodules using CAD software (as second reading), could be superior to those achieved by using conventional reading alone [5, 22].

In this work, we described and validated our experience on the development of a “homemade” computerized system to detect lung nodules on multidetector CT (MDCT) scans. Two different CT scanners from two different institutions were used in the study. One of the aims of our study was to present a robust algorithm, valid for different situations. A CAD scheme evaluation for lung nodule detection, using two different MDCT scanners, to our knowledge, has not been previously conducted.

Our scheme reached high sensitivity, showing that our version of the iris filter was accurate, that is, being acceptable as a processing method in the nodule detection process. Although our results were slightly better for the scanner Emotion 6,
comparing 6 and 64 detector row CT scans, there were not appear to be significant differences in performance. The eight nodules that the system missed out at a sensitivity of 94%, were nodules (between 5 and 14mm and) attached to a vessel; pulmonary nodules in close proximity to the vasculature can appear to be contiguous to the vessel and can be removed during the segmentation process. As a result, the improvement of the system must focus on this topic [4]. The location and morphology of the nodule did not affect the detection rate. Calcified nodules presented higher sensitivity values.

Conclusions about ground glass opacities would be hazardous because the number of this type of nodules in our database was small.

It is difficult to establish comparison among CAD schemes tested over different databases [5]; there should be an independent and standardized nodule database for all these studies; the algorithms for automated detection of the different studies are also different…; however, our results are in good agreement with those of Retico et al. [12], who reported 80% of sensitivity and 10 FPs/scan; or Belloti et al. [11] who presented 80% at 2.47 FPs/ using a neural network classifier to reduce the false positive rate. Pu et al. [10] have also presented a sensitivity of 81.5% with 6.5 FP identifications per examination. Other investigators have also presented similar results, such as, Riccardi et al. [14] have recently presented a sensitivity of 71% with 6.5 FP/patient, and Tan et al. [15] who reported 87.5% sensitivity at 4 FP detections per scan.

Usually, the number of false-positive findings is a major drawback of CAD systems. In our study, the false-positive rate was either 3.8 (Emotion6) or 5.8 (Definition) FP/case for a sensitivity of 80%; consequently, the false positive fraction needs to be improved; it is very important to keep low the number of false positives, because when the number of detections is higher than it should, this can lead to confusion of the radiologist. Further investigation is needed to reduce our false positive rate. In fact, reduction of false-positive findings will be the focus of our ongoing developments.

The nodules that were included in the study were between 4mm and 3cm in diameter, the upper limit corresponding with stage T1 tumors, according to the Union Internationale Contre le Cancer Classification, and to the nodule definition of the Fleischner society, as well [2, 16, 21, 23].

This study suffers from several limitations. Firstly, the study population was limited; with the three radiologists evaluating all the scans, the task was already very time consuming, however, the number of the examinations could have been increased. In addition, the examinations were randomly chosen and might not reflect a perfect average type of routine examination. Nodule cases were included despite the presence of other abnormalities on the scan, unless the other abnormality was spatially contiguous and interfered with visual interpretation [16]. And yet something more should have been taken into consideration: it would be interesting to study the impact of the system on a general population; and to compare our system with some of the different comercial CAD systems that are already in the market.

Besides, a relevant aspect, apart from the detection of the pulmonary nodules, is to characterize them, to measure both their size and volume, and to calculate their duplication time/nodule growth (for comparative studies). In fact, it is advisable to monitor the nodules for temporal changes; therefore, the CAD should analyze the internal density and marginal characteristics, as well as the 3D volumetric evaluation [3, 22]; this will be the objective of our future investigations.

Finally, although the system has been designed to avoid any dependence on the data set used in the investigation, it would be interesting to apply our algorithms over an independent database [16], because this would permit to develop and test our CAD method in a more robust and reliable fashion.

We will continue to explore different ways of combining visual and artificial intelligence techniques for our detection algorithms, but in conclusion, our computer-aided diagnosis scheme yielded encouraging results, suggesting that our CAD system appears to be potentially appropriate, and may help radiologists, as a second reader, in the detection of lung nodules in MDCT.
5 Conclusions
The CAD detection rate for pulmonary nodules is high; however, further investigation is needed to reduce our false positive rate. In its present version, our CAD system appears to be helpful as a second reader, in the detection of lung nodules in MDCT.

Acknowledgments
This work has been partially supported by the Xunta de Galicia (expte. nº PGIDIT06BF20802PR), and by the FIS (expte. nº PI060058) and (expte. nº PI080072).

References

