Published in final edited form as:
Acad Radiol. 2008 February ; 15(2): 165–175.

Computerized Detection of Lung Nodules in Thin-Section CT Images by use of Selective Enhancement Filters and an Automated Rule-Based Classifier

Qiang Li, **Feng Li**, and **Kunio Doi**
Department of Radiology, University of Chicago 5841 S. Maryland Avenue, Chicago, Illinois 60637

Abstract

We have been developing a computer-aided diagnostic (CAD) scheme for lung nodule detection in order to assist radiologists in the detection of lung cancer in thin-section computed tomography (CT) images. Our database consisted of 117 thin-section CT scans with 153 nodules, obtained from a lung cancer screening program at a Japanese university (85 scans, 91 nodules) and from clinical work at an American university (32 scans, 62 nodules). The database included nodules of different sizes (4-28 mm, mean 10.2 mm), shapes, and patterns (solid and ground-glass opacity (GGO)). Our CAD scheme consisted of modules for lung segmentation, selective nodule enhancement, initial nodule detection, feature extraction, and classification. The selective nodule enhancement filter was a key technique for significant enhancement of nodules and suppression of normal anatomic structures such as blood vessels, which are the main sources of false positives. Use of an automated rule-based classifier for reduction of false positives was another key technique; it resulted in a minimized overtraining effect and an improved classification performance. We employed a case-based four-fold cross-validation testing method for evaluation of the performance levels of our computerized detection scheme. Our CAD scheme achieved an overall sensitivity of 86% (small: 76%, medium-sized: 94%, large: 95%; solid: 86%, mixed GGO: 89%, pure GGO: 81%) with 6.6 false positives per scan; an overall sensitivity of 81% (small: 69%, medium-sized: 91%, large: 91%; solid: 79%, mixed GGO: 88%, pure GGO: 81%) with 3.3 false positives per scan; and an overall sensitivity of 75% (small: 60%, medium-sized: 88%, large: 87%; solid: 70%, mixed GGO: 87%, pure GGO: 81%) with 1.6 false positives per scan. The experimental results indicate that our CAD scheme with its two key techniques can achieve a relatively high performance for nodules presenting large variations in size, shape, and pattern.

Keywords

nodule detection; computer-aided diagnosis; CAD; CT scan; rule-based classifier

1. Introduction

Lung cancer is the leading cause of deaths among all types of cancer in the U.S [1]. The number of deaths it causes is greater than the total number of deaths resulting from colon cancer, breast cancer, and prostate cancer combined. Some evidence suggests that early detection of lung cancer may allow for timely therapeutic intervention, which in turn results in a more favorable

Phone: (773) 834-5096, Fax: (773) 702-0371 Email: qiangli@uchicago.edu.

prognosis for the patients. Therefore, screening programs for early detection of lung cancer have been attempted in the U.S. and Japan by use of computed tomography (CT) [2,3]. In a screening program with CT, radiologists must read a large number of images, and they are likely to overlook some lung cancers. Therefore, a computer-aided diagnostic (CAD) scheme for nodule detection [4,5], which provides radiologists with the locations of nodule candidates, would be particularly useful for reduction of detection errors in the early detection of cancer in thoracic CT scans.

CAD schemes for lung nodule detection were developed first for chest radiographs [6] and then for thick-section CT images [7-13]. The typical performance of current CAD schemes in thick-section CT is an 80-90% sensitivity with 1-2 false positives per section, which translates into tens of false positives per CT scan. The majority of false positives are caused by blood vessels and other normal anatomic structures [10,12]. Because of the relatively large section thickness (5-10 mm), CAD schemes for nodule detection in thick-section CT generally detect nodules on a section-by-section basis. Since most of the processing steps, such as nodule segmentation and feature extraction, are performed on two-dimensional (2D) section images, they are considered to be 2D.

Since 2001, some investigators have reported their efforts in the development of CAD schemes for lung nodule detection in thin-section CT images [14-22]. Because the section thickness in thin-section CT is small, three-dimensional (3D) image processing and analysis techniques become applicable. CAD schemes with good performance levels achieve a sensitivity of about 90% with 5-10 false positives per scan. As in thick-section CT, the majority of false positives are caused by blood vessels and other normal anatomic structures [17]. A major disadvantage of some current CAD schemes is the use of a relatively small database, with a small number of nodules of ground-glass opacity (GGO).

To achieve a high performance level for a computerized detection scheme for lung nodule, it is important to employ new effective techniques in major steps of the detection scheme, including initial nodule identification and false positive reduction; it is also important to use a relatively large database including both solid nodules and nodules with GGO. In this study, we utilized a selective nodule enhancement filter [23] for improving the detection sensitivity in the initial nodule identification step; we used an automated rule-based classifier [24] for substantial reduction of false positives in the false positive reduction step; and we employed a relatively large database for reliable estimation of the performance level of our CAD scheme. We believe that the two new techniques considerably improved the performance level of our CAD scheme, and that the use of a relatively large database enabled us to estimate the performance level of our CAD scheme reliably.

2. Materials

The IRB approval for this research project was obtained before it started. A nodule is defined as a focal lesion inside lungs that is 30 mm and smaller. We employed two thin-section CT databases obtained from an American university and a Japanese university. The first database consisted of 32 cases with 62 lung nodules collected at the American university by use of a multi-detector-row CT (MDCT) (LightSpeed QX/I, GE Medical Systems, Milwaukee, WI) from June 2002 to February 2003. The nodule cases were identified based on the radiology report and/or pathology report whenever available. We excluded cases with severe diffuse lung disease because these cases were not targets of this study. We also excluded cases with more than 6 nodules, because these nodules were very likely to be metastatic, and they generally had characteristics that were similar. In addition, nodules that were too small (<4 mm) or too large (>30 mm) were not included because these lung nodules are clinically non-significant [25]. Calcified nodules were also excluded from our database. For the MDCT examinations, various

protocols were used for the CT scans. The exposure ranged from 40 to 360 mAs, and the tube potential was either 120 or 140 kVp. Various collimation and various reconstruction intervals were employed; the ranges of collimation and reconstruction intervals were 1.25 - 5.0 mm and 1.0 - 2.0 mm, respectively. Each section had a matrix size of 512×512, a pixel size between 0.50 and 0.75 mm, and 4096 gray levels in Hounsfield units (HU). An expert chest radiologist (22 years of experience) identified the location of each nodule by use of software developed in our laboratory. The locations of nodules were employed for evaluation of our nodule detection scheme in MDCT scans.

The second database was obtained from an annual screening program for early detection of lung cancer with CT at a Japanese university [1]. The nodules found in the screening program had been confirmed either as malignant by biopsy or surgery, or as benign by biopsy, surgery, or follow-up examinations showing no growth for 2 or more years. The database consisted of 76 CT studies with confirmed cancers and 160 studies with confirmed benign nodules. Among them, we excluded calcified nodules; we also excluded CT scans that had either too few slices (<10) or discontinuous slices. Finally, we retained 85 high-resolution CT (HRCT) scans with 41 confirmed cancers (mean size: 12.6 mm, range: 6-20 mm) and 50 confirmed benign nodules (mean size: 9.5 mm, range: 4-21 mm). For these examinations, a HRCT scanner (HiSpeed Advantage, GE Medical Systems, Milwaukee) was used with a standard tube current (200 mA), a 1-mm collimation, and a 0.5-mm reconstruction interval. Each section had a matrix size of 512×512, a pixel size of 0.29 or 0.39 mm, and 4096 gray levels in HU. Unfortunately, each HRCT examination in this database included only a portion of a lung containing nodules; therefore, it was not appropriate to use cases in this database for calculating the false-positive rate per scan. For this reason, the nodules in this database were employed for determining the sensitivity of nodule detection for our CAD scheme, but the false positives reported from this database were discarded and were not used for determining the rate of false positives. Thus, only the false positives reported from the first database were employed to calculate the false-positive rate.

The combined database with cases from both databases included 117 cases with 153 nodules, and was utilized for training and testing of our CAD scheme for nodule detection in thin-section CT. In order to assess the characteristics of nodules included in the combined database, an expert chest radiologist measured the size of each nodule. The radiologist first measured the lengths of a long axis and a short axis of each nodule in a “central” section, in which it appeared with its largest size. The average of the lengths of the two axes was defined as being the size of the nodule. Figure 1 shows the distribution of the nodule sizes in our database. There were 68 (44.4%) small (4 - 8 mm), 52 (34.0%) medium-sized (9 - 13 mm), and 33 (21.6%) large nodules (14 mm and above). The mean and standard deviation of nodule size were 10.2 mm and 4.7 mm, respectively; the sizes ranged from 4 mm to 28 mm. It is obvious that our database contained nodules with a relatively wide range of sizes.

In addition to nodule size, the pattern of a lung nodule, such as pure GGO, mixed GGO, or solid nodules, is another major characteristic; therefore, we also determined the fractions of nodules with GGO in our database. Based on the consensus of three radiologists, our database included 101 (66%) solid nodules and 52 (34%) nodules with GGO, including 36 (24%) with mixed GGO and 16 (10%) with pure GGO.

3. Methods

3.1. Overall scheme of our computerized detection technique

Figure 2 is a diagram of our CAD scheme for nodule detection in thin-section CT. The first step was to segment lung regions of interest from other regions, such as muscle, fat, bone, mediastinum, and background (air) outside the body. All subsequent processing steps were

applied to the interior of the segmented lung regions. We then employed a selective nodule enhancement filter to simultaneously enhance nodules (spherical structures) and suppress normal anatomic structures included in the lung regions, such as blood vessels and airway walls. In this step, we also employed two other filters for selective enhancement of linear structures such as blood vessels and planar structures such as airway walls. Because of the unique characteristics of our selective nodule enhancement filter, we could then identify most nodules together with many non-nodules (false positives) from the nodule-enhanced images by use of a thresholding technique. Also included in this step was a constrained 3D region-growing technique for accurate segmentation of each nodule in the original CT images. Next, based on the initial region and the grown region, we determined 18 features from the original, nodule-, vessel-, and airway-wall-enhanced images, as well as an image of the shape index and an image of the curvedness. Finally, we developed and applied a fully automated rule-based classifier for analyzing various features in order to remove most non-nodules while retaining most true nodules. The nodule candidates that survived the rule-based classifier were used for determining the sensitivity and false-positive rate of our CAD scheme.

3.2. Lung segmentation

First, for use of the 3D image-processing and analysis technique, we employed a tri-linear interpolation technique to make the size of the voxels equal to 1 mm in each of three dimensions. We then separated lung regions from other regions and background outside the body for each of the CT sections by use of a thresholding technique. A pixel with a CT value between -400 HU and -1000 HU was considered to be located inside the lung region, and was thus assigned a value of 1; otherwise, the pixel was assigned a value of 0. After the thresholding, the lung regions, together with background (air) regions outside the body, were separated from other regions inside the body. In each section, we removed air regions outside the body by discarding regions of value 1 that were attached to image boundaries. The regions that survived this processing were considered to be lung regions.

However, nodules connected to the pleura might have been excluded from the lung regions since the CT values for the pixels inside the nodule were outside the range -400 HU and -1000 HU. To correct this type of segmentation error, we first tracked the contour of the lung region, shown as a solid curve in Fig. 3. From a starting contour point, we then scanned counter-clockwise all points on the entire contour one by one. For each current contour point ‘A’, we then scanned clockwise all contour points between the starting point and point ‘A’ until we found a point ‘B’ that satisfied the following conditions: (1) the distance between points ‘A’ and ‘B’ is less than 30 mm, (2) all the pixels on a straight line ‘AB’ connecting points ‘A’ and ‘B’, shown as a dotted line in Fig. 3, have a value of 0 in the lung-segmented image, and (3) the maximum distance between the straight line ‘AB’ and all the contour points between points ‘A’ and ‘B’ on the contour is larger than 2/3 of the distance between points ‘A’ and ‘B’. When such a point ‘B’ was found, we assumed that we had found a juxtapleural object and we replaced all the contour points between points ‘A’ and ‘B’ by the dotted straight line ‘AB’. By doing so, we included a juxtapleural object inside the lung region. If such a point ‘B’ could not be found, we simply proceeded to the next contour point of ‘A’ and considered it as the new current point. From the new current point, we repeated the above procedure until all contour points were checked.

3.3. Image enhancement by use of three selective filters

To achieve a high sensitivity and specificity for initial nodule detection, we applied a multi-scale selective filter to the original image for simultaneous enhancement of nodules and suppression of normal anatomic structures such as blood vessels in 3D image space [23]. Figure 4 shows the maximum intensity projection of (a) the original and (b) the nodule-enhanced images for two cancers by use of our multi-scale enhancement filters. It is apparent in Fig. 4

(b) that the low-contrast cancer and the blood vessel-attached cancer were enhanced, that the second cancer was separated from the blood vessels, and that most of the blood vessels were eliminated. It is thus much easier to detect nodules in the nodule-enhanced image than in the original image. In fact, we also applied to the original image two selective filters for enhancement of blood vessels and airway walls [23], which were employed for extracting useful features for removal of non-nodules caused by blood vessels and airway walls.

3.4. Initial identification and region growing of nodule candidates

We then thresholded the nodule-enhanced images with a fixed, empirically selected value of 40 for separating nodule candidates from anatomic structures inside the lung regions. This simple thresholding technique worked very well because our selective nodule enhancement filter removed the majority of non-nodule objects with CT values similar to those of nodules. A 3D connected-component labeling technique was then employed for identifying all of the isolated objects. The labeling algorithm identified many small objects that were mainly due to noise and other small non-nodule structures. Therefore, objects with an effective diameter smaller than 2.5 mm were considered to be non-nodules and were eliminated hereafter. Other objects were considered to be initial nodule candidates.

By comparing Figs. 4(a) and 4(b), it is apparent that the nodules appear slightly smaller in the nodule-enhanced image than in the original image. Therefore, for each nodule candidate, we developed a 3D constrained region-growing technique to segment it accurately in the original CT images. First, the 3D region of each initial nodule candidate segmented from the nodule-enhanced image was employed as a seed region for the constrained region growing in the original CT images. We then calculated the mean and standard deviation of CT values for voxels inside the seed region. Next, we added some voxels to the seed region if (a) these voxels were adjacent to (in terms of 26-neighborhood in 3D image space) the seed region and (b) their CT values were within a range defined by the mean CT value of the seed region plus or minus two standard deviations. To prevent nodule regions from leaking into adjacent structures such as blood vessels, the above region growing process was constrained to 5 repetitions, i.e., a maximum of 5 mm growth. In each repetition, the mean and standard deviation of the current grown region were re-calculated.

3.5. Feature determination

Based on the initial nodule region and the grown nodule region, we then extracted 18 features for each of the nodule candidates from the original image, from the nodule-, blood vessel-, and airway wall-enhanced images obtained with our selective enhancement filters [23], and from the two images for the shape index and curvedness based on the differential characteristics of the isointensity surface [26]. The 18 features extracted were:

1. Six features based on nodule shape: effective diameter, degree of compactness and irregularity, determined in the initial and the grown nodule regions.
2. Twelve features based on voxel value: mean and standard deviation of voxel values inside the grown region, each extracted from the original CT image, from the nodule-, blood vessel-, and airway wall-enhanced images [23], and from the two images for the shape index and curvedness [26].

The compactness was defined as $36\pi V^2/S^3$, where V and S represent, respectively, the volume and the surface area of the segmented nodule region. The compactness achieves a maximum value of 1 for a sphere, and a value between 0 and 1 for other shapes. The irregularity was defined as $1-(\pi d^2/S)$, where the effective diameter d was defined as the diameter of a sphere whose volume is equal to that of a segmented nodule region. The irregularity achieves a minimum value of 0 for a sphere, and a value between 0 and 1 for other shapes.

3.6. False positive reduction by use of an automated rule-based classifier

The rule-based classifier probably has been the most frequently used one since the early days of development of various CAD schemes, including the detection and diagnosis of breast masses and microcalcifications [27], lung nodules [6], colonic polyps [28], melanoma [29], heart disease [30], mesothelial lesions [31], and bone disease [32]. The existing rule-based classifiers are generally designed manually, and therefore, they often lead to a large overtraining effect, i.e., a large difference between the estimated performance levels of training and testing [24].

A rule in CAD schemes typically consists of two steps, i.e., selection of a feature and selection of a cutoff threshold. As most existing rule-based classifiers do, our previous prototypical rule-based classifier employs for a rule only a simple feature, which is either a clinical parameter or a feature directly determined from various medical images [24]. Composite features, which are generally a linear combination of many simple features, may provide a higher performance than does the sequential utilization of individual simple features. To create a linear composite feature, a subset of simple features must be determined first for each rule. Starting from a single feature that achieved maximum separation between the two classes, a new feature was added to the currently selected feature subset if its addition was "optimal" among all available features to be added, in terms of Wilks' lambda and the corresponding F-value [33,34]. Similarly, a feature was deleted from the currently selected feature subset if its deletion was "optimal" among all features available for deletion. With this sequential iterative procedure, we selected a feature subset for determining an optimal composite feature.

For a selected specific subset of simple features, we employed a separation metric as a criterion for determining the "optimal" composite feature for separating nodules from non-nodules (a separation metric is defined as the ratio of the between-class and the within-class scatter matrices [35]). The optimal composite feature is the linear combination of simple features that maximizes the separation metric.

For this optimal composite feature x, it has been proven that the cutoff threshold must pass through one of the nodules if one wishes to eliminate overtraining in the process of threshold selection [24]. By use of the gain defined below, we developed a technique for automatically determining which nodule the threshold should pass through. When the threshold was set to be the kth smallest feature value among all nodules (k=1,2,...,K, where K is the maximum number of nodules allowed to be sacrificed), we calculated the ratio of the number of non-nodules removed to the number of nodules sacrificed. Thus, for the K possible thresholds, we determined K ratios, among which one achieved the maximum value. This maximum value was defined as the gain of the feature, and the corresponding threshold was considered to be the optimal threshold for the feature. The Appendix gives the pseudo-code for the automated rule-based classifier with composite features. In each loop (steps 3-6) of the Appendix, a rule is constructed. After a rule has been applied, the training dataset and the number K of nodules allowed to be sacrificed are updated dynamically. The algorithm terminates when K=0, i.e., when it reaches a predetermined sensitivity. Nodule candidates surviving this rule-based classifier were considered to be the final nodule candidates.

3.7. Evaluation of the CAD scheme for nodule detection

The overall performance of our CAD scheme for nodule detection was evaluated by comparing the computer-identified locations with the pre-determined locations of the centers of nodules. To verify whether a true nodule in a dataset was detected, we first determined a reported nodule candidate that was nearest to the center of the true nodule identified by a radiologist. If the distance between the center of the candidate and that of the true nodule was less than 12 mm, the nodule was considered to be a detected one; otherwise, it was considered as one missed by

the computerized detection scheme. The distance threshold of 12 mm was selected empirically based on our experiences in nodule detection. In addition to the above criterion, we visually confirmed for each true nodule that the center of the candidate found by our scheme was indeed inside the region of the true nodule. We thus determined the detection sensitivity by use of the detected nodules. All other nodule candidates were considered to be non-nodules, and they were employed for calculation of the false-positive rate.

We employed a four-fold case-based cross-validation method for evaluating our CAD scheme. The entire database was first randomly partitioned into four folds, each of which contained about 29 cases, including 8 (1/4 of 32) cases from the first database and about 21 (1/4 of 85) cases from the second database. Each fold was then employed for testing, and the other three folds were employed for training our CAD scheme. When all four folds were used for testing our CAD scheme, a four-fold cross-validation was completed, and a sensitivity and a false-positive rate per case were calculated based on the testing results of the four folds. For reducing the variability in the estimated performance levels, the case-based cross-validation method was repeated 10 times. Because our classifier is fully automated, we specified in each trial four design sensitivities {80%, 85%, 90%, 95%} for training the automated rule-based classifier. Consequently, in each trial, we determined four operating points (the pairs of sensitivity and false-positive rate) for training and testing. At each operating point, we then calculated the mean values of sensitivity and false-positive rate in 10 trials, and we employed the mean sensitivity and false-positive rate for generating an empirical free-response receiver operating characteristic (FROC) curve [36] by connecting consecutive mean operating points with straight lines.

4. Results

4.1. Results of initial nodule detection

The results of initial nodule detection usually demonstrate how effective the steps of preprocessing, enhancement, and initial nodule identification methods are. Of the existing publications on computerized detection schemes in thin-section CT, only one, by Zhao et al. [18], provided such an initial detection result, although their study was based on simulated instead of real nodules. Zhao et al. detected 94.4% of nodules (251 out of 266 nodules) with 906 false positives per scan, whereas in our scheme, we detected 98.7% of nodules (151 of 153) with 140.2 false positives per scan (4486 false positives for 32 CT scans in the Chicago database). Figure 5 shows (a) three low-contrast nodules with GGO that were successfully identified by our initial nodule detection technique and (b) the only two nodules that our initial detection technique failed to detect. The two missed nodules were approximately 11 mm and 12 mm in diameter and both were low-contrast nodules with pure GGO.

4.2. Results of final nodule detection

Figure 6 shows the mean FROC curves for training and testing the CAD scheme. For clarity, the maximum number of false positives per scan in Fig. 6 was set to 25 rather than 140 (the number of false positives per scan after initial nodule detection). Please note that the FROC curve for training is located above that for testing, which indicates that overtraining bias occurred in our CAD scheme for nodule detection. However, the extent of overtraining bias in the automated rule-based classifier appears limited because the overtraining bias in the step of cutoff threshold selection was completely eliminated for the automated rule-based classifier [24].

Figure 7 shows the mean FROC curves obtained by testing our CAD scheme for the nodules in the American dataset and the Japanese dataset, and for all nodules. All three curves were obtained from the same testing experiments by use of the four-fold cross-validation evaluation

method; however, in FROC curves for the nodules in the American dataset and the Japanese dataset, the sensitivities were determined by use of the nodules in the American dataset and the Japanese dataset only. Please also note that only the false positives in the American dataset were employed to calculate the false positive rates in the three FROC curves. The performance level for the nodules in the American dataset was lower than that in the Japanese dataset, since the American dataset contained many subtle small nodules, which were more difficult to detect than the medium-sized and large nodules.

Figure 8 shows the mean FROC curves obtained by testing our CAD scheme for the small nodules (4-8 mm), the medium-sized ones (9-13 mm), the large ones (>14 mm), and all nodules. It is obvious that the performance level for small nodules was considerably lower than those for medium-sized and large nodules.

Figure 9 shows the respective mean FROC curves obtained for the solid nodules, the mixed GGO nodules, the pure GGO nodules, and for all nodules. The sensitivity for solid nodules was quite low, since 82.4% (56 out of 68) of the small nodules were solid; the other ones (17.6%) were those with GGO.

Table 1 lists the means and standard deviations of sensitivities and false-positive rates per case for different groups of nodules at four operating points, obtained with ten trials of the four-fold case-based cross-validation method. At the first operating point, our CAD scheme achieved an overall sensitivity of 75% (small: 60%, medium-sized: 88%, large: 87%; solid: 70%, mixed GGO: 87%, pure GGO: 81%) with 1.6 false positives per scan. At the second operating point, the CAD scheme achieved an overall sensitivity of 81% (small: 69%, medium-sized: 91%, large: 91%; solid: 79%, mixed GGO: 88%, pure GGO: 81%) with 3.3 false positives per scan. At the third operating point, the CAD scheme achieved an overall sensitivity of 86% (small: 76%, medium-sized: 94%, large: 95%; solid: 86%, mixed GGO: 89%, pure GGO: 81%) with 6.6 false positives per scan.

5. Discussion

CAD schemes for lung nodule detection in thin-section CT images have been developed by some investigators [14-22]. A major disadvantage of some current CAD schemes is the use of a relatively small database. The maximum number of cases in the current CAD schemes was 56, employed in a CAD scheme developed by Ge et al. [22]. Another problem with the current CAD schemes is the use of a small number of GGO nodules. Most current CAD schemes employed solid nodules only. Only McCulloch et al. [18] clearly stated that their database included 8 GGO nodules.

In this study, we employed a significantly larger database (117 cases) than the databases used in other CAD schemes. Our database contained 52 GGO nodules, which is also significantly larger than those used in other CAD schemes. We believe that the use of a relatively large database with more GGO nodules enables us to evaluate the performance levels of our CAD scheme more reliably. In addition, we employed a combined database collected from two universities. Such combination of datasets from different sites is advantageous because it makes our CAD scheme more robust and stable.

In terms of techniques, we attempted to develop a CAD scheme for lung nodule detection in thin-section CT by use of two key techniques, i.e., a selective nodule enhancement filter [23] and an automated rule-based classifier with composite features [24]. The selective nodule enhancement filter is very useful for simultaneously enhancing low-contrast spherical nodules and suppressing linear and planar anatomic structures such as blood vessels and airway walls [23], which are the main sources of false positives in our CAD scheme. This technique allowed us to detect low-contrast nodules and blood vessel-attaching nodules with a high sensitivity

and a low false-positive rate, which is evidenced by our high performance level for initial nodule identification. The other key technique is an automated rule-based classifier with composite features, which provides a high performance in distinguishing between nodules and non-nodules by minimizing the overtraining biases [24]. With the application of these two techniques, our CAD scheme achieved a good performance for our relatively large database including not only solid nodules, but also nodules with GGO.

Generally it is very difficult to compare the performance levels of different CAD schemes because they may use different databases with different sizes and different levels of difficulty. In spite of this, we tried to carefully compare the performance levels of our CAD scheme with those of other CAD schemes, and we ask the readers to cautiously interpret the comparison results. We found that some current CAD schemes employed CT scans with only simulated nodules [18]. Some employed CT scans with only partial lungs and used the number of false positive per slice (instead of per scan) as their false positive rate [17,22]. Some did not describe how their CAD schemes were evaluated with certain testing methods such as resubstitution, leave-one-out, or cross-validation [16,21]. We believe that comparison of our CAD scheme with these CAD schemes is inappropriate. In the remaining two studies, Paik et al. [20] used a leave-one-out method to evaluate the performance levels of their CAD scheme based on a database of 8 CT scans with an unknown number of solid nodules, and they achieved a sensitivity of 80% with 1.3 false positives per scan or a sensitivity of 90% with 5.6 false positives per scan; McCulloch et al. [19] used a cross-validation method to evaluate the performance levels of their CAD scheme based on a database of 50 CT scans with 43 nodules (35 solid and 8 GGO nodules), and they achieved a sensitivity of 70% with 8.3 false positives per scan. In this study, we employed a cross-validation method to evaluate the performance levels of our CAD scheme based on a database of 117 CT scans with 153 nodules (101 solid and 52 GGO nodules), and we achieved a sensitivity of 81% with 3.3 false positive per scan, or a sensitivity of 86% with 6.6 false positive per scan. Considering the size and the complexity of our database, we believe that our CAD scheme achieved relatively high performance levels compared with other CAD schemes.

Our CAD scheme has its disadvantages. For instance, the Japanese dataset included only confirmed malignant and confirmed benign nodules. Such confirmed nodules are generally larger than nodules to be encountered in screening program for lung cancer. The use of these larger nodules would artificially increase the sensitivity for nodule detection, and thus introduce an optimistic bias in the estimated sensitivity. The CT scans in the Japanese dataset included only part of lungs, thus the false positives in the Japanese dataset were not employed to calculate the false positive rate per CT scan. Therefore, we do not know the false positive rate for the Japanese dataset. In addition, we excluded CT scans with severe diffuse lung diseases in the American dataset, which would also introduce an optimistic bias in the estimated false positive rate.

6. Conclusion

We have developed a CAD scheme for nodule detection in thin-section CT images that employs two key techniques, i.e., a multi-scale, selective nodule enhancement filter and an automated rule-based classifier. The selective nodule enhancement filter enhanced nodules and considerably suppressed normal anatomic structures such as blood vessels and airway walls; as a result, it significantly improved the performance of the initial nodule detection. The automated rule-based classifier was “optimal” in terms of overtraining effects, and it achieved good classification results, due to the use of composite features. Our CAD scheme for nodule detection achieved good results for a relatively large database of thin-section CT scans.

Acknowledgments

This work was supported by USPHS grants.

Appendix: Optimal rule-based classifier based on linear composite features

Input: A training dataset including N nodules and a design sensitivity S to be achieved.
Output: A list of rules based on composite features.

1. Determine the maximum number of nodules allowed to be sacrificed by using K=int[(1-S)(N+1)], where int[z] represents a function to round off z to the nearest integer.
2. Do 3-6 below until K=0.
3. Determine a feature subset based on Wilks' lambda and the corresponding F-value.
4. Transform the feature subset into a single linear composite feature by maximizing the separation metric.
5. For the composite feature, use gain to determine an optimal threshold for the construction of a rule.
6. Use the rule to update the training dataset by removing non-nodules and sacrificing k nodules, and update K by letting K = K-k.

References

1. Greenlee RT, Murray T, Bolden S, Wingo PA. Cancer statistics, 2000. CA Cancer Journal of Clinicians 2000;50:7–33.
2. Sone S, Takashima S, Li F, Yang Z, Honda T, Maruyama Y, Hasegawa M, Yamanda T, Kubo K, Hanamura K, Asakura K. Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet 1998;351:242–245. [PubMed: 9457093]
3. Henschke CI, McCauley DI, Yankelevitz DF, Naidich DP, McGuinness G, Miettinen OS, Libby DM, Pasmantier MW, Koizumi J, Altorki NK, Smith JP. Early lung cancer action project: overall design and findings from baseline screening. Lancet 1999;354:99–105. [PubMed: 10408484]
4. Doi K. Overview on research and development of computer-aided diagnostic schemes. Seminars in Ultrasound, CT, and MRI 2004;25:404–410.
5. Li Q, Li F, Suzuki K, Shiraishi J, Abe H, Engelmann R, Nie YK, MacMahon H, Doi K. Computer-aided diagnosis in thoracic CT. Seminars in US, CT, and MRI 2005;26:357–363.
6. Giger ML, Doi K, MacMahon H. Image feature analysis and computer-aided diagnosis in digital radiography. 3 Automated detection of nodules in peripheral lung fields. Med Phys 1988;15:158–166. [PubMed: 3386584]
7. Giger ML, Bae KT, MacMahon H. Computerized detection of pulmonary nodules in CT images. Invest Radiol 1994;29:459–465. [PubMed: 8034453]
8. Reeves AP, Koitis WJ. Computer-aided diagnosis of small pulmonary nodules. Semin Ultrasound CT MR 2000;21:116–128. [PubMed: 10776884]
9. Ko JP, Betke M. Chest CT: automated nodule detection and assessment of change over time: preliminary experience. Radiology 2001;218:267–273. [PubMed: 11152813]
10. Lee Y, Hara T, Fujita H, Itoh S, Ishigaki T. Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. IEEE Trans Med Imag 2001;20:595–604.
11. Brown MS, McNitt-Gary MF, Goldin JG, Suh RD, Sayre JW, Aberle DR. Patient specific models for lung nodule detection and surveillance in CT images. IEEE Trans Med Imag 2001;20:1242–1250.
12. Gurcan MN, Sahiner B, Petrick N, Chan HP, Kazerooni EA, Cascade PN, Hadjiiski L. Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system. Med Phys 2002;29:2552–2558. [PubMed: 12462722]
13. Wormanns D, Fiebich M, Saidi M, Diederich S, Heindel W. Automatic detection of pulmonary nodules at spiral CT: clinical application of a computer-aided diagnosis system. Eur Radiol 2002;12:1052–1057. [PubMed: 11976846]
14. Li F, Novak C, Qian J, Kohl G, Naidich DP. Automatic detection of lung nodules from multi-slice low-dose CT images. Proc SPIE Med Imag 2001;4322:1828–1835.
15. Wiemker R, Rogalla P, Zwartkruis A, Blaffert T. Computer aided lung nodule detection on high resolution CT data. Proc SPIE Med Imag 2002;4684:677–688.

16. Lawler LP, Wood SA, Pannu HS, Fishman EK. Computer-assisted detection of pulmonary nodules: preliminary observations using a prototype system with multidetector-row CT data sets. J. Digital Imaging 2003;16:251–261.

17. Brown MS, Goldin JG, Suh RD, McNitt-Gary MF, Sayre JW, Aberle DR. Lung micronodules: automated method for detection at thin-section CT - initial experience. Radiology 2003;226:256–262. [PubMed: 12511699]

18. Zhao B, Gamsu G, Ginsberg MS, Jiang L, Schwarts LH. Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm. J of Applied Clinical Medical Physics 2003;4:248–260.

19. McCulloch CC, Kaucic RA, Mendonca PRS, Walter DJ, Avila RS. Model-based detection of lung nodules in computed tomography exams. Acad Radiol 2004;11:258–266. [PubMed: 15035515]

20. Paik DS, Beaulieu CF, Rubin GD, Acar B, Jeffrey RB Jr. Yee J, Dey J, Napel S. Surface normal overlap: A computer-aided detection algorithm with application to colonic polyps and lung nodules in helical CT. IEEE Trans on Med Imag 2004;23:661–675.

21. Bae KT, Kim JS, Na YH, Kim KG, Kim JH. Pulmonary nodules: Automated detection on CT images with morphologic matching algorithm - Preliminary results. Radiology 2005;236:286–294. [PubMed: 15955862]

22. Ge ZY, Sahiner B, Chan H-P, Hadjiiski LM, Cascade PN, Bogot N, Kazerooni EA, Wei J, Zhou C. Computer-aided detection of lung nodules: False positive reduction using a 3D gradient field method and 3D ellipsoid fitting. Med Phys 2005;32:2443–2454. [PubMed: 16193773]

23. Li Q, Sone S, Doi K. Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans. Med Phys 2003;30:2040–2051. [PubMed: 12945970]

24. Li Q, Doi K. Analysis and minimization of overtraining effect in rule-based classifiers for computer-aided diagnosis. Med Phys 2006;33:320–328. [PubMed: 16532936]

25. Henschke CI, Yankelevitz DF, Naidich DP, McCauley DI, McGuinness G, Libby DM, Smith JP, Pasmantier MW, Miettinen OS. CT screening for lung cancer: Suspiciousness of nodules according to size on baseline scans. Radiology 2004;231:164–168. [PubMed: 14990809]

26. Thirion JP, Gourdon A. Computing the differential characteristics of isointensity surfaces. Computer Vision and Image Understanding 1995;61:190–202.

27. Markey MK, Lo JY, Tourassi GD, Floyd CE. Self-organizing map for cluster analysis of a breast cancer database. Artif Intell Med 2003;27:113–127. [PubMed: 12636975]

28. Yoshida H, Masutani Y, MacEneaney P, Rubin DT, Dachman AH. Computerized detection of colonic polyps at CT colonography on the basis of volumetric features: pilot study. Radiology 2002;222:327–336. [PubMed: 11818596]

29. Shao S, Grams RR. A proposed computer diagnostic system for malignant melanoma (CDSMM). J Med Syst 1994;18:85–96. [PubMed: 7964215]

30. Hedan B, Ohlin H, Rittner R, Edenbrandt L. Acute myocardial infarction detected in the 12-lead ECG by artificial neural networks. Circulation 1997;96:1798–1802. [PubMed: 9323064]

31. Robutti F, Betta PG, Donna A, Pavesi M. A morphometrically-based classification rule for the diagnosis of primary mesothelial lesions. J Pathol 1990;162:57–60. [PubMed: 2231194]

32. Pietka E. Computer-assisted bone age assessment based on features automatically extracted from a hand radiograph. Comput Med Imaging Graph 1995;19:251–259. [PubMed: 7641169]

33. Aoyama M, Li Q, Katsuragawa S, Doi K. Automated computerized scheme for distinction between benign and malignant solitary pulmonary nodules on chest images. Med Phys 2002;29:701–708. [PubMed: 12033565]

34. Aoyama M, Li Q, Katsuragawa S, Li F, Sone S, Doi K. Computerized scheme for determination of the likelihood measure of malignancy for pulmonary nodules on low-dose CT images. Med Phys 2003;30:387–394. [PubMed: 12674239]

35. Loew, M. Feature Extraction. In: Sonka, M.; Fitzpatrick, JM., editors. Handbook of Medical Imaging. 2. SPIE Press; Bellingham, WA: 2000. p. 273-338.

36. Metz CE. Some practical issues of experimental design and data analysis in radiological ROC studies. Invest Radiol 1989;24:234–245. [PubMed: 2753640]

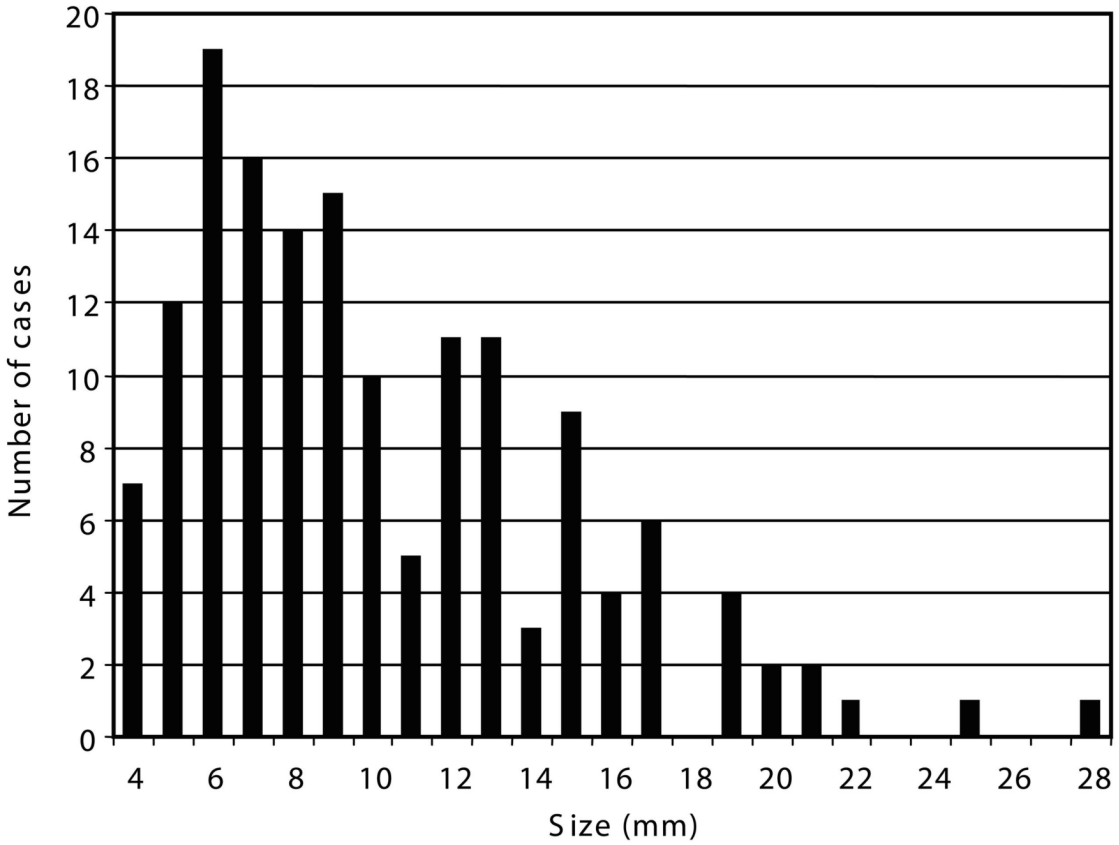

Fig. 1.
Distribution of nodule sizes in our database. The database contained nodules with a relatively wide range of sizes. There were 68 (44.4%) small (4 - 8 mm), 52 (34.0%) medium-sized (9 - 13 mm), and 33 (21.6%) large nodules (14 mm and above) in the database.

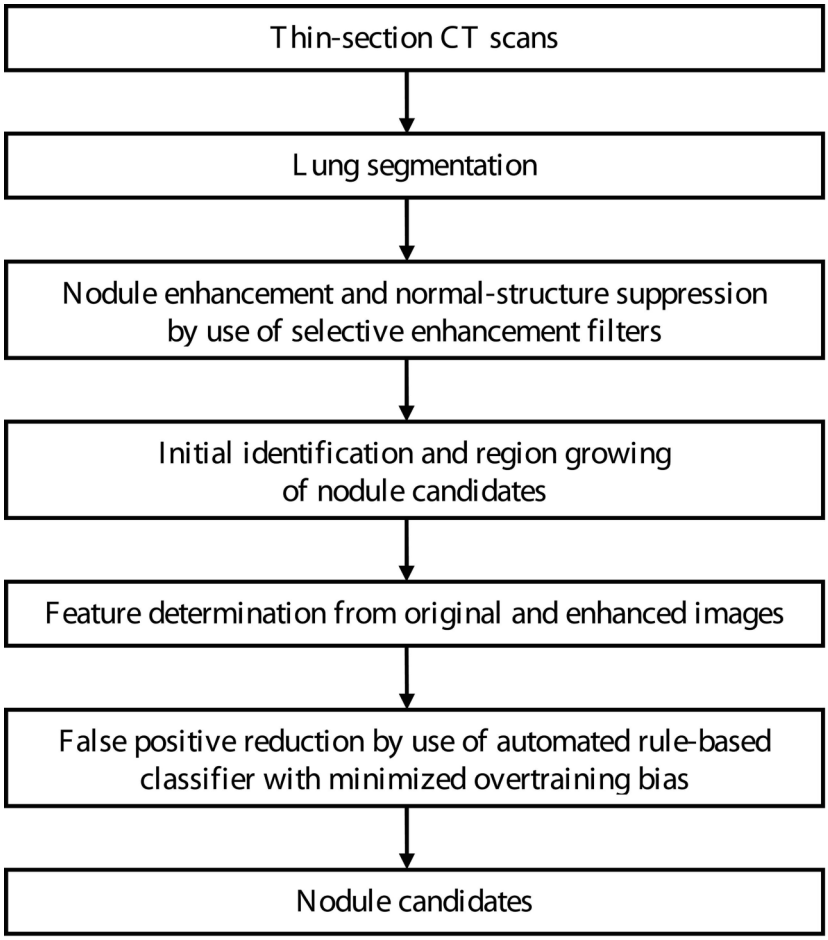

Fig. 2.
Overall scheme of the computerized detection technique.

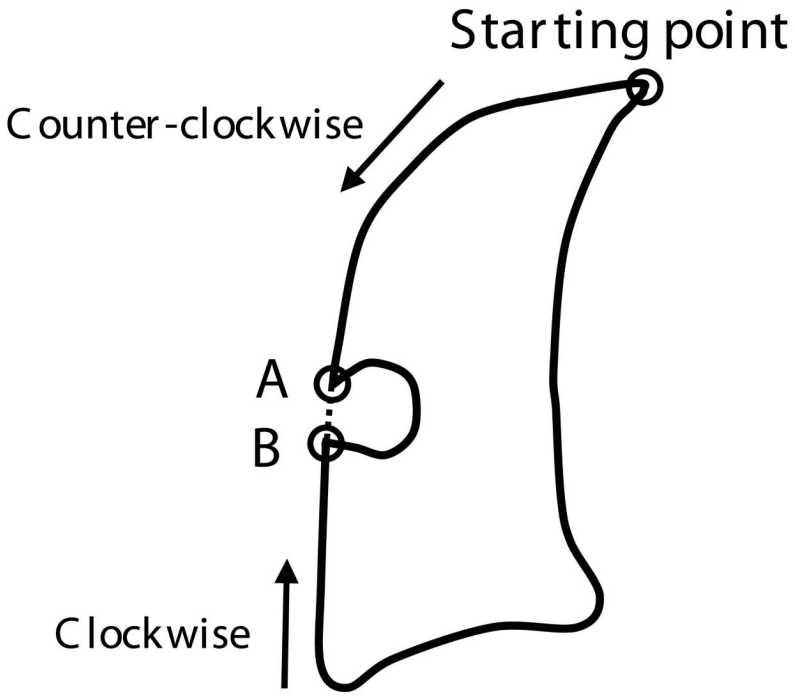

Fig. 3.
Schemetic illustration for inclusion of a juxtapleural object.

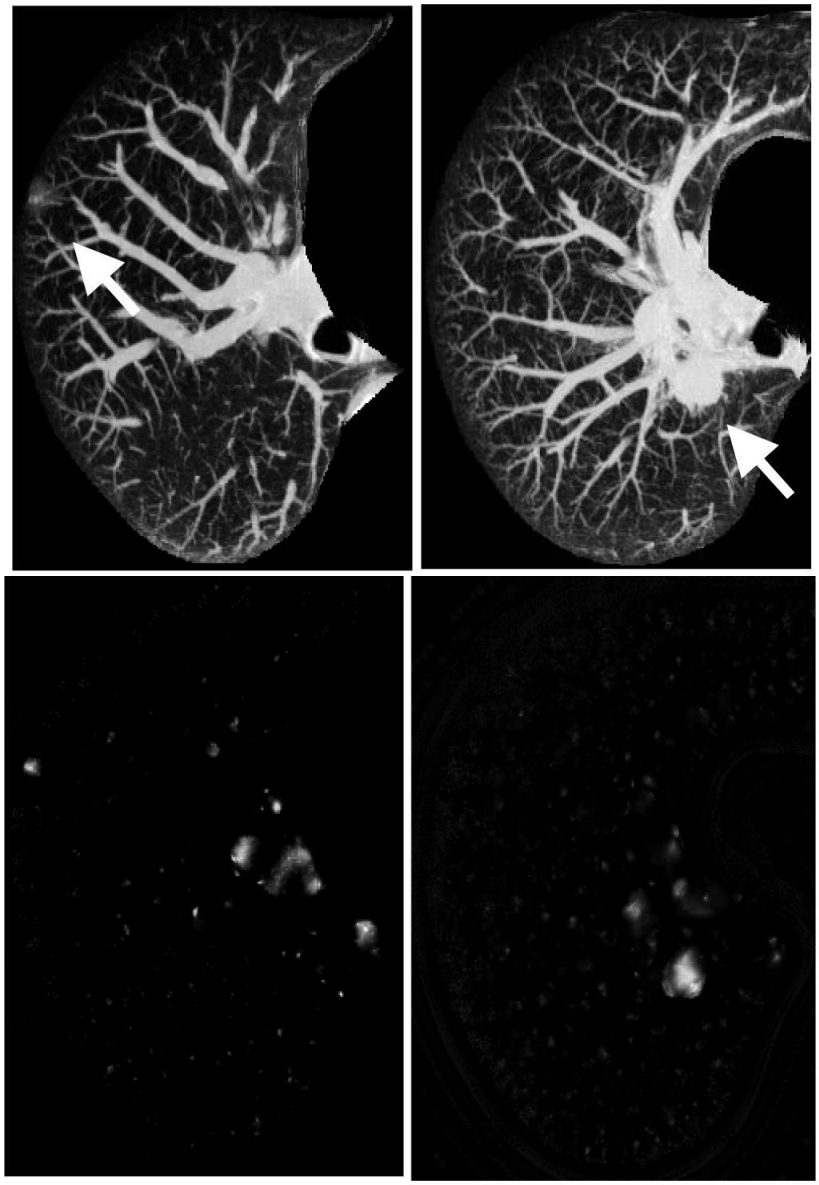

Fig. 4.
Maximum intensity projection of (a) two 3D original images with nodules identified by arrows and (b) nodule-enhanced images.

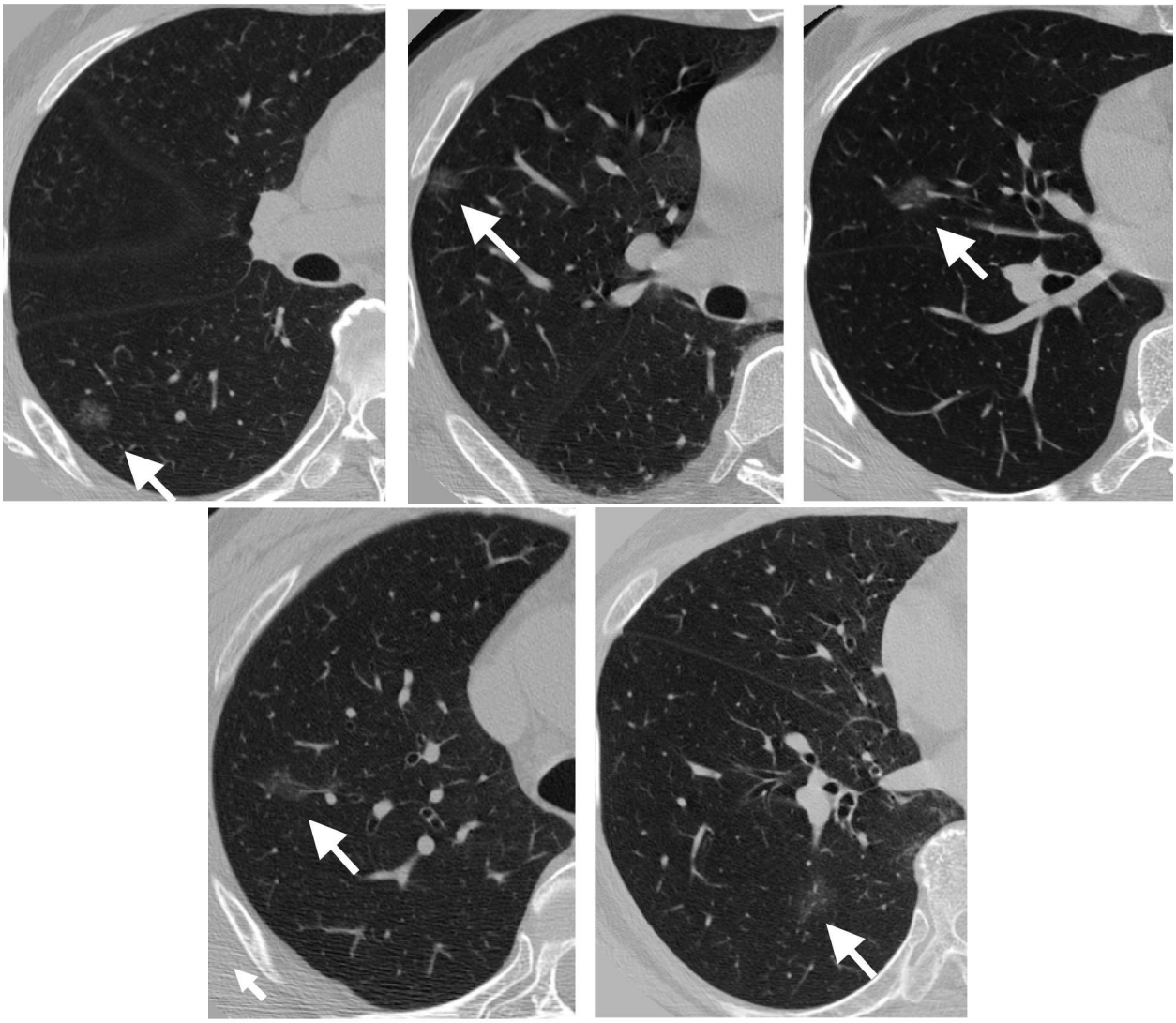

Fig. 5.
(a) Three low-contrast nodules with GGO that were successfully identified, and (b) the only two nodules that were missed by our initial nodule detection technique.

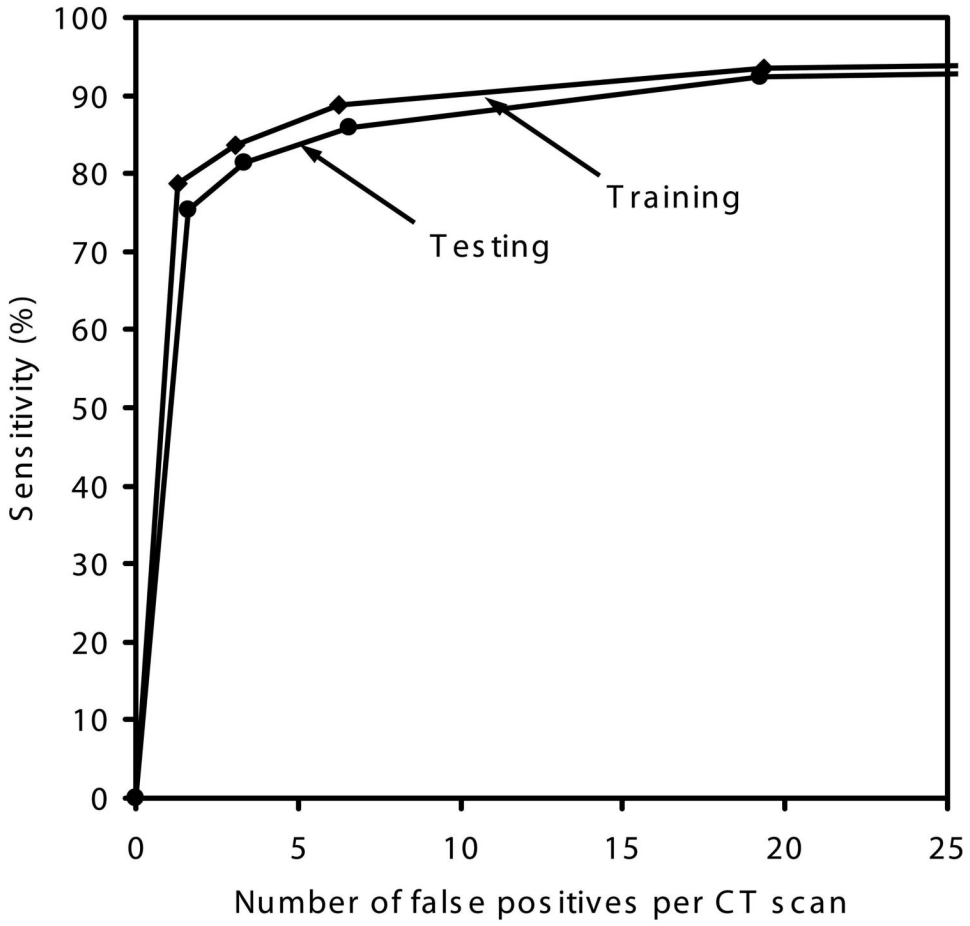

Fig. 6.
Mean FROC curves for training and testing of our CAD scheme.

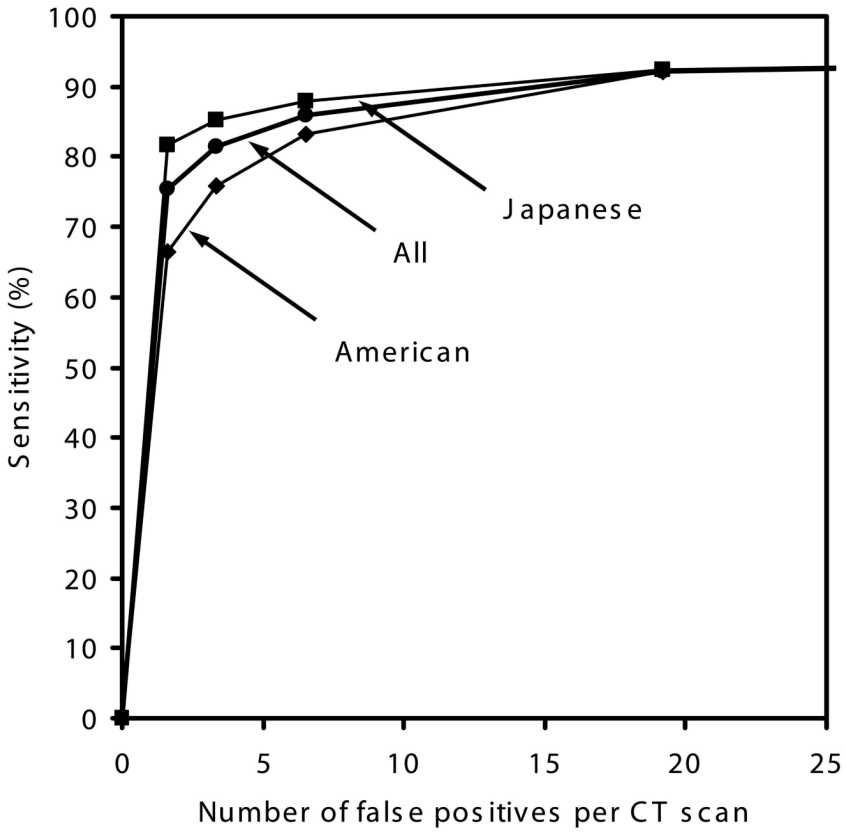

Fig. 7.
Mean FROC curves obtained from the testing of our CAD scheme for the nodules in the American dataset and the Japanese dataset.

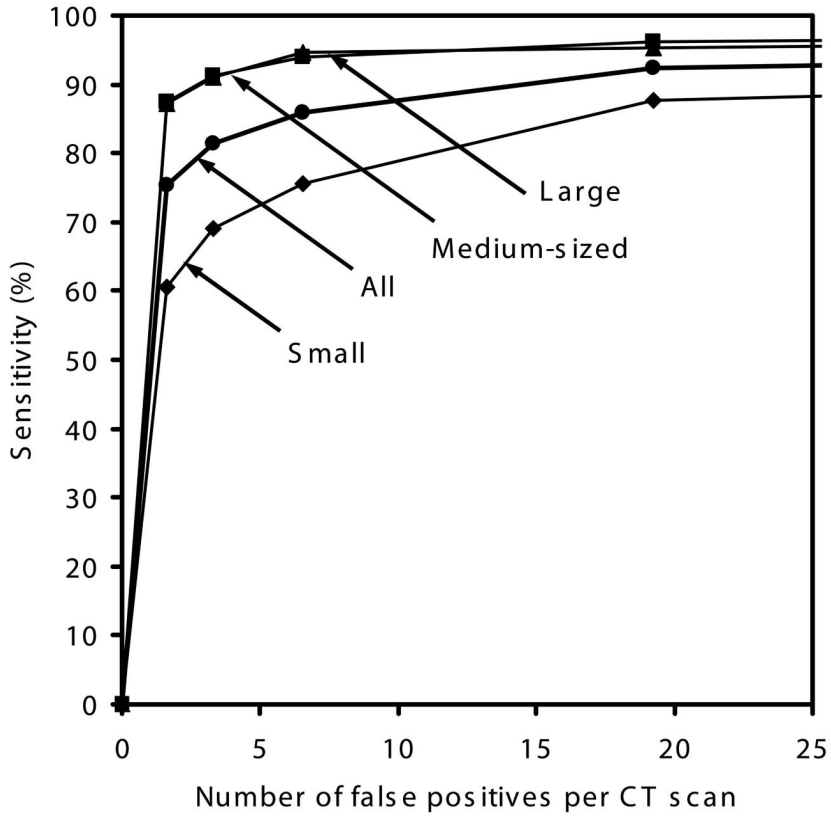

Fig. 8.
Mean FROC curves obtained from the testing of our CAD scheme for the small nodules (<9 mm), the medium-sized nodules (9-13 mm), the large nodules (>13 mm), and all nodules.

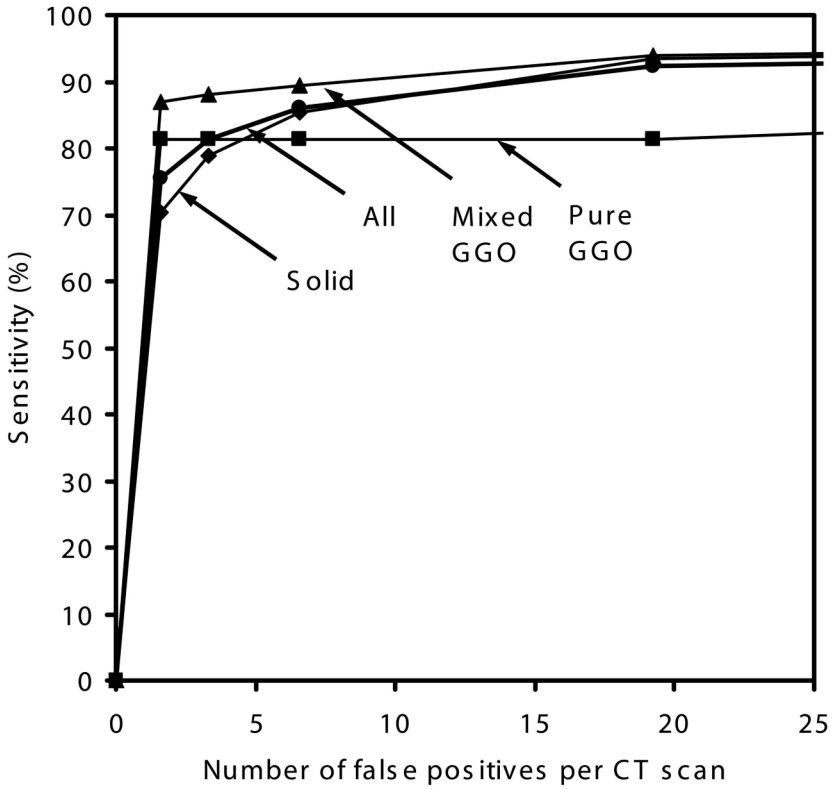

Fig. 9.
Mean FROC curves obtained from the testing of our CAD scheme for the solid nodules, the mixed GGO nodules, the pure GGO nodules, and all nodules.

Table 1

Means ± standard deviations of sensitivities and false-positive rates per case for different groups of nodules at four operating points, obtained with ten trials of the four-fold case-based cross-validation method. The number below the title of each column indicates the number of nodules in a specific nodule group

Operating points	Sensitivities (%) for different groups of nodules							FP rates per case
	All (n=153)	Small (n=68)	Medium-sized (n=52)	Large (n=33)	Solid (n=101)	Mixed GGO (n=36)	Pure GGO (n=16)	
1	75.4 ± 1.6	60.4 ± 2.3	87.5 ± 2.6	87.2 ± 3.1	70.3 ± 2.5	86.9 ± 1.3	81.3 ± 0	1.6 ± 0.2
2	81.3 ± 1.6	69.0 ± 2.4	91.3 ± 1.6	90.9 ± 2.3	78.9 ± 2.2	88.1 ± 1.3	81.3 ± 0	3.3 ± 0.3
3	86.0 ± 1.5	75.5 ± 3.0	94.0 ± 1.9	94.7 ± 2.1	85.5 ± 2.0	89.4 ± 1.2	81.3 ± 0	6.6 ± 0.4
4	92.3 ± 1.6	87.8 ± 3.7	96.2 ± 0.0	95.3 ± 1.6	93.4 ± 2.1	93.9 ± 1.2	81.3 ± 0	19.2 ± 1.7