Detection of Lung Nodules in Digital Chest Radiographs Using Artificial Neural Networks: A Pilot Study

Yuzheng C. Wu, Kunio Doi, and Maryellen L. Giger

Radiologists can fail to detect up to 30% of pulmonary nodules in chest radiographs. A back-propagation neural network was used to detect lung nodules in digital chest radiographs to assist radiologists in the diagnosis of lung cancer. Regions of interest (ROIs) that contained nodules and normal tissues in the lung were selected from digitized chest radiographs by a previously developed computer-aided diagnosis (CAD) scheme. Different preprocessing techniques were used to produce input data to the neural network. The performance of the neural network was evaluated by receiver operating characteristic (ROC) analysis. We found that subsampling of original 64- × 64-pixel ROIs to smaller 8- × 8-pixel ROIs provides the optimal preprocessing for the neural network to distinguish ROIs containing nodules from false-positive ROIs containing normal regions. The neural network was able to detect obvious nodules very well with an A_z value (area under ROC curve) of 0.93, but was unable to detect subtle nodules. However, with a training method that uses different orientations of the original ROIs, we were able to improve the performance of the neural network to detect subtle nodules. Artificial neural networks have the potential to serve as a useful classifier to help to eliminate the false-positive detections of the CAD scheme.
Copyright © 1995 by W.B. Saunders Company

KEY WORDS: artificial neural networks, digital chest radiography, lung nodules, receiver operating characteristic (ROC) analysis, computer-aided diagnosis (CAD).

IN CHEST RADIOGRAPHY, radiologists can fail to detect pulmonary nodules in chest radiographs in up to 30% of cases with actually positive findings.[1] A number of approaches[2-4] have been reported for the detection of lung nodules from digital chest radiographs using artificial neural networks (ANNs).

We had previously reported an approach[5] to detect lung nodules by ANNs using computer-extracted image features as input data. The neural network performed well in eliminating some of the false-positive detections reported by a computer-aided diagnosis (CAD) scheme.[6-9] As an alternative approach, we applied an ANN to detect lung nodules from digital chest radiographs based on image data directly in this pilot study. In this approach, the detection of lung nodules and elimination of false-positive detections were performed from a different prospective in which feature extraction was not explicitly involved. A potential advantage of this approach would be that the neural network might be able to use most of the information in the original images to distinguish nodules from normal regions instead of relying only on the limited number of extracted image features.

MATERIALS AND METHODS

Region of Interest (ROI) Selection

Fifty-six ROIs were selected from conventional, posteroanterior chest radiographs for our database. The chest radiographs were digitized by a Fuji optical drum scanner[10] using 0.1-mm pixel size and 10-bit quantization. The images were subsequently averaged into a 512- × 512-pixel matrix with an effective pixel size of 0.6 mm. The matrix size of the selected ROIs were 64 × 64 pixels (38.4 × 38.4 mm²). A demonstration of ROI selection is shown in Fig 1 in which an ROI containing a nodule and an ROI containing a normal region of lung are shown. Of the 56 ROIs selected, sixteen of them contained obvious nodules, another 24 ROIs contained subtle nodules, and another 16 ROIs, defined as false-positive ROIs, contained normal chest areas that had been incorrectly reported as positive by the CAD scheme. The obvious nodules were generally of higher contrast and more uniform in terms of their sizes as compared with the subtle nodules. Figure 2 shows all of the 56 ROIs selected. The size of nodules range from 10 to 20 mm. The presence of pulmonary nodules was verified by computed tomography scans or radiographic follow-up, whereas the absence of nodules was confirmed by the consensus of two chest radiologists.

Neural Networks

A three-layer, feed-forward neural network was used in this study. The training algorithm and implementation of the neural network were discussed in our other report[5] and will not be repeated here.
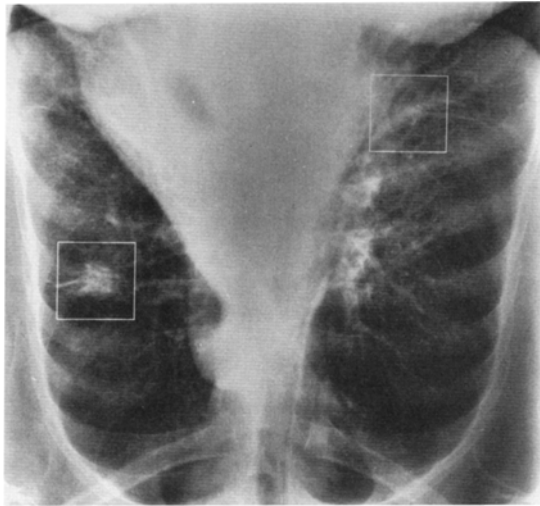
Fig 1. Demonstration of ROI selection in a digital chest radiograph. Two ROIs, one containing a nodule (above the heart in the left lung) and another containing normal region of lung (next to the heart in the right lung), are shown in the figure.

Preprocessing of the Input Data

Background-trend correction. Because the ROIs were selected from different regions of chest radiographs, the background structure in the ROIs were, in general, different from one to another. To isolate signals (ie, nodules) from the nonuniform background, we applied a background-trend correction technique[5,11] on the selected ROIs. The 56 ROIs, after background-trend correction, are shown in Fig 3.

Fourier transform. The ROIs were selected in such a way that nodules were approximately centered in each of the ROIs. However, the shape of a nodule was not always rounded and the size in different directions varied considerably. The input data entered into the neural network thus depend on the relative location of a nodule in an ROI and the shape of the nodule. Fourier transformation provides a way to centralize the spatial information in the frequency domain. Thus, Fourier transformation was examined as one of the preprocessing techniques in this study.

Spatial subsampling of image data. An ROI with a size of 64×64 pixels contains 4,096 pixels. When all of the data are presented to a neural network, the amount of information can be overwhelming to a neural network of simple structure such as the one used in this study. Therefore, the neural network may not be able to use the essential information effectively to identify the major characteristics of each input image. Thus, we subsampled the original ROIs into smaller matrix sizes, ranging from 32×32 to 2×2 pixels. The subsampling was performed uniformly throughout the ROI such that the relative geometric shape of a nodule candidate was unchanged. The new matrix size was smaller than the original ROI size to eliminate too much detailed information. However, the new matrix size was kept large enough so that enough information was preserved for the neural network to recognize a nodule. Pixel values of the subsampled ROIs were used as input to the neural network. The number of input units in a neural
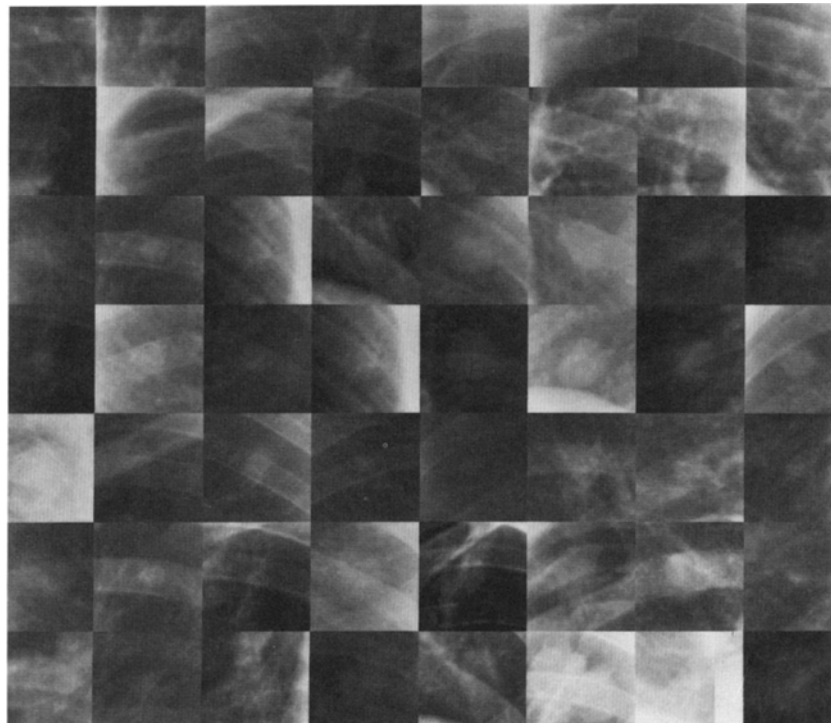

Fig 2. Fifty-six original ROIs of size 64×64 are shown. First two rows are normal ROIs; the next two rows are ROIs containing obvious nodules; the last three rows are ROIs containing *subtle nodules.*
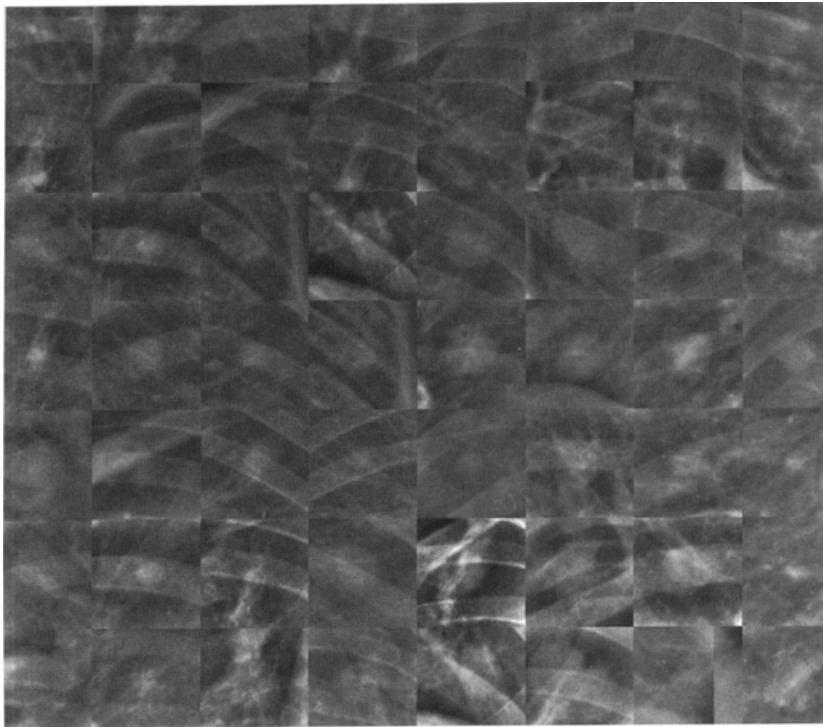
Fig 3. Background-trend corrected ROIs shown in Fig 2.

network was equal to the number of pixels in an input ROI, eg, 1,024 for ROIs of size 32 × 32 pixels.

Square-spiral method.   The image data or the power spectrum of Fourier transform can be entered into the neural network in different ways. The conventional approach is the line method, in which the pixel values in an image are entered into the network line by line. Another approach, in which the pixels of an image are entered into the network along a squared-spiral line, as shown in Fig 4,
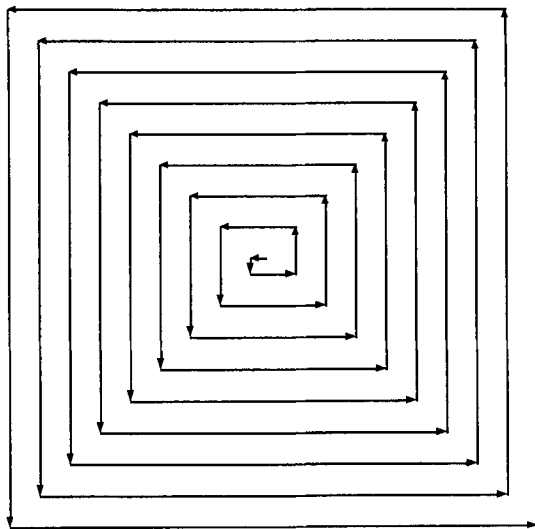

Fig 4. The spiral method of entering image data to the neural network is shown.

starting from the center of an ROI, was studied. The value of each pixel along the squared-spiral line is presented to an input unit of ANN. The total number of input units equals the number of pixels in an ROI. The spiral method emphasizes the central part of an ROI, allowing the pixels near the center of an ROI to be entered as a group into the neural network without the pixels near the edges. The performance of ANN using the spiral method was compared with that using the conventional line method.

Evaluation of Performance of Classifiers

In general, decision-making performance in terms of correct detections or classifications depends on the threshold value used. In this study, receiver operating characteristic (ROC) analysis[12-15] was used to objectively assess the overall performance of the neural network as a classifier by varying a threshold value. To estimate the ability of the neural-network classifier to generalize from the cases on which they were trained and make decisions concerning cases that had not been included in the training, we used a round-robin method—also known as the leave-one-out method—to evaluate the performance of the neural network. The applications of ROC analysis and the round-robin method have been discussed in our previous paper published in this journal.[5]

RESULTS

Effect of Matrix Size on the Performance of the Neural Network

We subsampled the original ROIs with 64 × 64 pixels into smaller matrix sizes. The ROIs of

smaller matrix size were then used to train and test the neural network. The performance of the neural network was evaluated using the round-robin method with 32 ROIs: sixteen false-positive ROIs and 16 containing obvious nodules. Figure 5 shows the A_z values obtained with different matrix sizes. It appears that the neural network performed the best when the matrix size of 8 × 8 pixels was used ($A_z = 0.93$). For the matrix size of 13 × 13 pixels or larger, the neural network was unable to distinguish between obvious nodules and false-positive ROIs.

In general, the neural network is expected to perform better with larger input matrix sizes, provided that sufficient training cases are available, and large numbers of hidden layers and hidden units are used. However, in this study, the number of cases available for training of the neural network is relatively small. The use of a very large structure of neural network is also limited by the computation time. Therefore, the performance of the neural network is better with an input matrix size of 8 × 8 pixels than those with input matrix sizes of 10 × 10 pixels and 11 × 11 pixels, as shown in Fig 5.

The failure of the neural network with large input matrix sizes may be caused by the fact that a relatively small number of hidden units and small number of iterations were used in the training process of the neural network. It is possible that with a significant increase in the size of neural-network structure (hidden units and connecting weights) and number of training iterations, the performance of the neural network with input matrix size larger than 13 × 13 pixels may be able to reach or exceed the level achieve by using matrix size 8 × 8 pixels shown in Fig 5. However, the increase in size of the neural network structure would require a substantial increase in the number of ROIs in the training database. A long computation time may be required to train the neural network. However, once the training process is completed, the time needed to classify and make decisions for new cases should be much less.

Performance of the Neural Network in Detecting Obvious Nodules and Subtle Nodules

The neural network was applied to distinguish between the ROIs containing obvious nodules and the false-positive ROIs, and to distinguish between ROIs containing subtle nodules and the same false-positive ROIs. All of the ROIs were subsampled from the matrix size of 64 × 64 to 8 × 8. The round-robin and jackknife methods were used to evaluate the performance of the neural network. The neural network performed very well in detecting obvious nodules with an A_z value of 0.93 evaluated by the round-robin method. The A_z value obtained using the jackknife method was only 0.76, probably because of the insufficient number in the training sets. Note that the round-robin method generally provides a greater A_z value than the jackknife method because more training data are used with the round-robin method, which requires more CPU time. However, the neural network was unable to detect subtle nodules, yielding an A_z value of 0.50 (random guessing) when evaluated by the round-robin method. The ROC curves representing the performances of the neural network are shown in Fig 6.

Rotation and Reflection of Input ROIs

To improve the effect of training with a limited number of input data on the performance of the neural network, we increased the number of training data by using images with different orientations.[16] Each of the ROIs, after being subsampled into the matrix size of 8 × 8 pixels, was rotated in increments of 90°, and
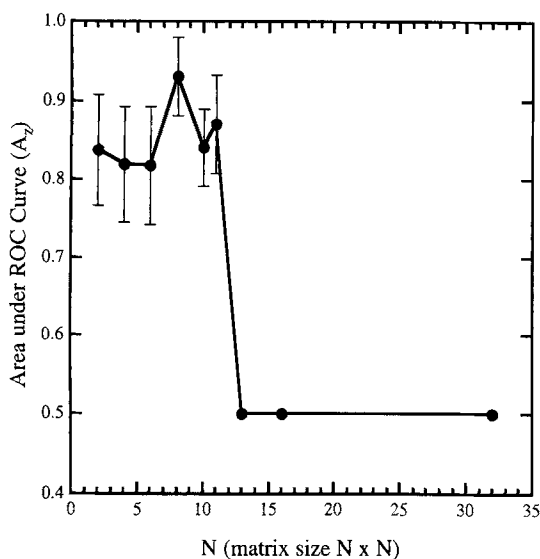

Fig 5.   Effect of matrix size on the performance of the neural network tested by using the round-robin method.
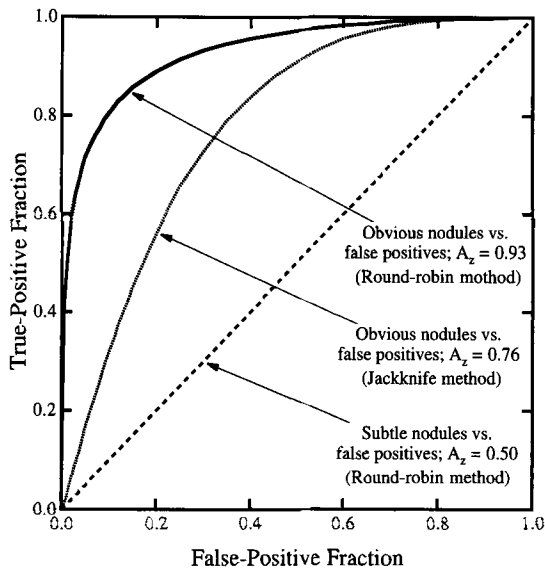
Fig 6.   Performance of the neural network for detecting lung nodules of different subtleties based on image data tested by using the round-robin and jackknife methods; 24 subtle nodules and 16 obvious nodules were used.

each rotation was also followed by a reflection around the rotation line. As a result, the number of ROIs in the training sets was increased eightfold. The round-robin method was used to evaluate the performance of the neural network in distinguishing obvious and subtle nodules from false positives using rotation and reflection of ROIs for the training data. The performance of the neural network in detecting subtle nodules was improved with an A_z value of 0.66 as opposed to that of 0.50 before the use of rotation and reflection. The A_z value of the neural network in detecting obvious nodules was 0.93, the same as the previous value obtained without the rotation and reflection of the ROIs using the round-robin method. Note that in the training process for the obvious nodules, only the ROIs containing nodules were rotated and reflected, whereas the ROIs containing false positives were simply duplicated eight times to keep the ratio of the total number of training data for nodules to that for false positives unchanged. The reason for not including the rotation and reflection of the false-positive ROIs was to avoid creating some orientations that would never exist in reality, which will be discussed later. These ROC curves of the neural network for detecting obvious and subtle nodules are shown in Fig 7.

Detection of Lung Nodules Using Other Preprocessing Techniques

Fourier transformation. Fourier transform was performed on ROIs after background-trend correction without subsampling. The power spectra of the ROIs, defined as the square of the absolute value of the Fourier transformation, were used as input data to train and test the neural network. However, the neural network was not able to distinguish obvious nodules from normal areas ($A_z = 0.5$ evaluated by round-robin method) using the power spectrum as input data.

Input to the neural network by square-spiral method. We compared the performances of the neural network for distinguishing ROIs containing obvious nodules and false positives (normal areas) using two different methods for entering input data to the neural network. In one method, the image data of ROIs were entered to the neural network by the square-spiral method as shown in Fig 4. In the other method, the conventional line method was used. The performance of the neural network was evaluated by using the round-robin method. In theory, the performance of the neural network should not be affected by the way image data is entered into the network, as long as the method is consistent for all ROIs, because there are no lateral connections in the feed-forward neural
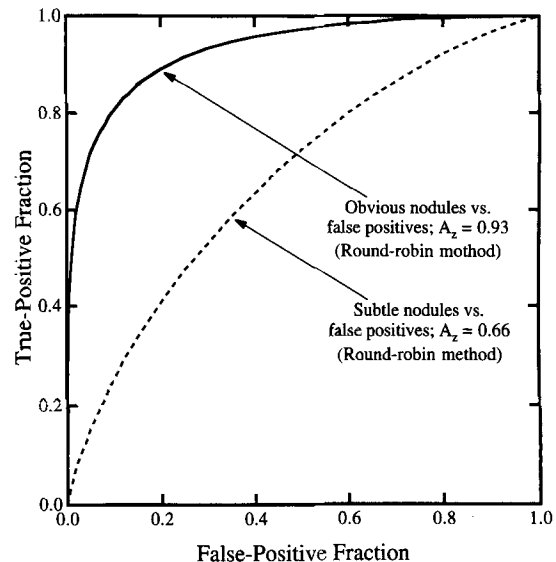

Fig 7.   Performance of the neural network for detecting lung nodules of different subtleties after rotation and reflection of the training data tested by using the round-robin method.
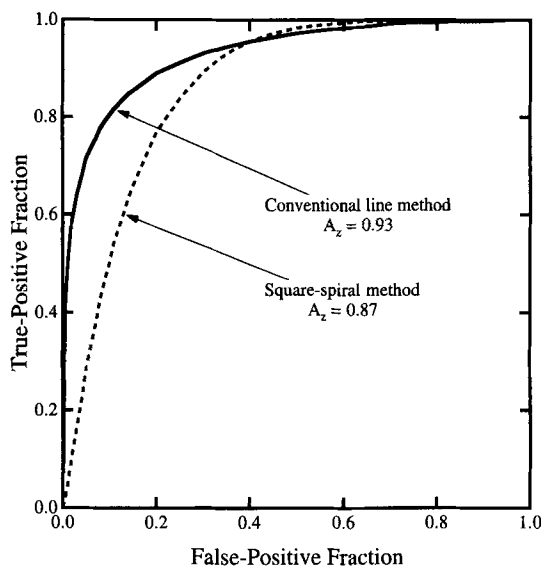
Fig 8. Performances of the neural network for detecting obvious nodules using the conventional line and square-spiral methods for entering data to the neural network, as evaluated by the round-robin method.

network. However, as shown in Fig 8, the performance of the neural network using the conventional line input method was slightly better than that using the square-spiral input method. We do not yet fully understand the causes of this phenomenon. It is important to notice that we are dealing with limited number of ROIs in the database and a relatively small size of neural-network structure. The neural networks with different ways of entering input data may not reach the same learning status after the same number of iterations. Another reason for the variation in the performance of the neural network may be that even though the square-spiral method emphasizes the central portion of a ROI where a nodule is usually located, the pixels along the boundary of a nodule are separated on the square-spiral profile. Therefore, the neural network was not able to use the boundary information efficiently to detect nodules.

DISCUSSION

The neural network performed very well in detecting obvious nodules using the spatial subsampling method, but failed to detect subtle nodules. The obvious nodules were not only higher in contrast but also more uniform in their

sizes and shapes. Therefore, the input pattern of obvious nodules after subsampling can be described as signals with high density at the center of an ROI. The subtle nodules, on the other hand, showed a somewhat random pattern to the neural network because of large variations in size, shape, and intensity (pixel value) of nodules, and thus totally confused the neural network. It seems that neural networks with simple structure such as the one used in this study may not be able to deal with complicated spatial patterns with satisfactory proficiency. The neural networks with algorithms and structure of shift-invariant properties may be better suited for this type of detection task. An application of a shift-invariant neural network for detection of microcalcifications is being performed in our labs. Our preliminary results have been very promising. The shift-invariant neural network achieved an A_z value of 0.91 in detecting clustered microcalcifications,[17] even with the jackknife method. The detection of nodules will not be the same as the detection of microcalcifications because nodules are generally of lower frequency objects and have lower contrast. However, there are many similarities in these two detection tasks. The success of the shift-invariant neural network for detecting microcalcifications is good indication that it can also be an effective tool in the detection of lung nodules.

By rotating and reflecting the original ROIs, we generated an extended database that not only was increased by 8 times in number, but also included seven additional orientations that were not present in the original database. These additional orientations helped to improve the generalizing ability of the neural network. As a result, the performance of the network to detect subtle nodules was improved.

One of the potential drawbacks resulting from the rotation and reflection is that some structures such as ribs in the false-positive ROIs are generally in horizontal or diagonal directions, and would never exist in certain orientations. To avoid this situation, we rotated and reflected only the nodules in the classification of nodule candidates. This method is practical and useful to create a sufficient database because the false positives can be easily obtained, whereas the number of nodules available to

train the neural network is usually very limited in clinical applications.

The normal ROIs used in this study actually contained false-positive detections as reported by the CAD scheme for detection of lung nodules. Even though they are normal regions in chest radiographs, many false positives have radiographic appearances that are similar to lung nodules. In general, it is more difficult to distinguish nodules from false positives than to distinguish nodules from other normal regions. The neural network was able to distinguish obvious nodules from the false-positive detections and to distinguish subtle nodules from the false positives to a certain extent with rotation and reflection of original ROIs. Therefore, the neural network, with appropriate training, can be used potentially as a classifier to help eliminating some false-positive detections of the CAD scheme.

ACKNOWLEDGMENT

REFERENCES

1. Forrest J, Friedman P: Radiologic errors in patients with lung cancer. West J Med 134:485-490, 1981

2. Garg S, Floyd C, Ravin C: Neural network localization of pulmonary nodules on digital chest radiographs. SPIE Medical Imaging, Image Processing 2167:623-629, 1994

3. Lo S, Freedman M, Lin J, et al: Automatic lung nodule detection using profile matching and back-propagation neural network techniques. J Digit Imaging 6:48-54, 1993

4. Chiou Y, Fleming Lure Y, Ligomenides P, et al: Shape feature analysis using artificial neural networks for improvements of hybrid lung nodule detection (HLND) system. SPIE Medical Imaging, Image Processing 1898:609-617, 1993

5. Wu Y, Doi K, Giger M, et al: Reduction of false positives in computerized detection of lung nodules in chest radiographs using artificial neural networks. J Digit Imaging 7:196-207, 1994

6. Giger M, Doi K, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography, III: Automated detection of nodules in peripheral lung fields. Med Phys 15:158-166, 1988

7. Giger M, Doi K, MacMahon H, et al: Pulmonary nodules: Computer-aided detection in digital chest images. Radiographics 10:41-51, 1990

8. Giger M, Ahn K, Doi K, et al: Computerized detection of pulmonary nodules in digital chest images: Use of morphological filters in reducing false-positive detections. Med Phys 17:861-865, 1990

9. Yoshimura H, Giger M, Doi K, et al: Computerized scheme for the detection of pulmonary nodules A nonlinear filtering technique. Invest Radiol 27:124-129, 1992

10. Ishida M, Kato H, Doi K, et al: Development of a new digital radiographic image processing. Proc SPIE 347:42-48, 1982

11. Katsuragawa S, Doi K, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography: Detection and characterization of interstitial disease in digital chest radiography. Med Phys 15:311-319, 1988

12. Metz CE: ROC methodology in radiologic imaging. Invest Radiol 21:720-733, 1986

13. Metz C: Some practical issues of experimental design and data analysis in radiological ROC studies. Invest Radiol 24:234-245, 1989

14. Metz C, Wang P, Kronman H: A new approach for testing the significance of differences between ROC curves measured from correlated data, in Deconinck F (ed): Information Processing in Medical Imaging. The Hague, The Netherlands, Nijhoffx, 1989, pp 432-445

15. Metz C: Quantification of failure to demonstrate statistical significance: The usefulness of confidence intervals. Invest Radiol 28:59-63, 1993

16. Stafford R, Beutel J, Mickewich D, et al: Application of neural networks to computer aided pathology detection in mammography. SPIE Medical Imaging, Physics of Medical Imaging 1896:341-353, 1993

17. Zhang W, Doi K, Giger ML, et al: Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network. Med Phys 21:517-524, 1994