

---

# False-positive Reduction in Computer-aided Diagnostic Scheme for Detecting Nodules in Chest Radiographs by Means of Massive Training Artificial Neural Network<sup>1</sup>

Kenji Suzuki, Junji Shiraishi, Hiroyuki Abe, Heber MacMahon, and Kunio Doi

---

**Rationale and Objective.** We developed a technique that uses a multiple massive-training artificial neural network (multi-MTANN) to reduce the number of false-positive results in a computer-aided diagnostic (CAD) scheme for detecting nodules in chest radiographs.

**Materials and Methods.** Our database consisted of 91 solitary pulmonary nodules, including 64 malignant nodules and 27 benign nodules, in 91 chest radiographs. With our current CAD scheme based on a difference-image technique and linear discriminant analysis, we achieved a sensitivity of 82.4%, with 4.5 false positives per image. We developed the multi-MTANN for further reduction of the false positive rate. An MTANN is a highly nonlinear filter that can be trained with input images and corresponding teaching images. To reduce the effects of background levels in chest radiographs, we applied a background-trend-correction technique, followed by contrast normalization, to the input images for the MTANN. For enhancement of nodules, the teaching image was designed to contain the distribution for a “likelihood of being a nodule.” Six MTANNs in the multi-MTANN were trained by using typical nodules and six different types of non-nodules (false positives).

**Results.** Use of the trained multi-MTANN eliminated 68.3% of false-positive findings with a reduction of one true-positive result. The false-positive rate of our original CAD scheme was improved from 4.5 to 1.4 false positives per image, at an overall sensitivity of 81.3%.

**Conclusion.** Use of a multi-MTANN substantially reduced the false-positive rate of our CAD scheme for lung nodule detection on chest radiographs, while maintaining a level of sensitivity.

**Key Words.** Computer-aided diagnosis; lung nodule; cancer; chest radiography; artificial neural network.

© AUR, 2005

---

Chest radiography is the most frequently used imaging examination technique for chest diseases because of its low cost, simplicity, and low radiation dose. Chest radiography has been used for the detection of lung cancer because some

evidence suggests that early detection of lung cancer may allow a more favorable prognosis (1,2). Radiologists, however, may fail to detect lung nodules in chest radiographs in up to 30% of cases in which nodules are visible in retrospect (3,4). Computer-aided diagnostic (CAD) schemes for nodule detection on chest radiographs have been investigated (5) because the computer can improve radiologists' detection accuracy (6,7).

A number of researchers have developed CAD schemes for detecting lung nodules in chest radiographs (8–23). Giger et al. (9,10) developed a CAD scheme based on a thresholding technique together with a rule-based classifier, and Wu et al. (11), Xu et al. (12), and Li et al. (13) im-

---

**Acad Radiol** 2005; 12:191–201

<sup>1</sup> From the Department of Radiology, The University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637 (K.S., J.S., H.A., H.M., K.D.). Received June 8, 2004; revision requested November 11; revision received November 17; revision accepted November 17, 2004. This study was supported by USPHS Grants No. CA62625 and CA98119. **Address correspondence to** K.S. e-mail: suzuki@uchicago.edu

© AUR, 2005

doi:10.1016/j.acra.2004.11.017

proved the performance of the CAD scheme by incorporating an artificial neural network (ANN) and linear discriminant analysis, an adaptive thresholding technique, and a multiple-template matching technique, respectively. Sankar et al. (14) reported a CAD scheme based on the segmentation of nodule candidates in which they used a dynamic programming technique and image feature analysis. Lo et al. (15) developed a CAD scheme based on a convolution neural network (16), and Lin et al. (17) improved the performance of the scheme by incorporating two-level convolution neural networks. Carreira et al. (18) devised a CAD scheme based on the detection of nodule candidates with normalized cross-correlation images and classification of candidates in curvature space, and Penedo et al. (19) improved the performance of the scheme by incorporating two-level ANNs that employed cross correlation teaching images and input images in curvature peak space. Coppini et al. (20) developed a CAD scheme based on biologically inspired ANNs with fuzzy coding. Other researchers, including Lampeter et al. (21), Floyd et al. (22), and Mao et al. (23) reported on CAD schemes to which they applied various techniques for improving performance.

An observer performance study by Kobayashi et al. (7) has demonstrated the utility of a CAD scheme for detecting lung nodules in chest radiographs. Recently, Kakeda et al. (24) showed that radiologists' performance was improved by the use of a commercial CAD system applied to chest radiographs. A major challenge in current CAD schemes for chest radiography is to achieve a low number of false-positive results, because a wide variety of normal structures resemble nodules in chest radiographs (25). It is difficult to detect nodules overlapping with normal anatomic structures such as ribs and vessels, and the majority of false-positive findings are caused by ribs and/or vessels (12,25), which can reduce the sensitivity as well as the specificity of a CAD scheme, and could also reduce radiologists' confidence in the CAD scheme. In addition, a large number of false positives would reduce radiologists' efficiency. Our purpose in this study was to develop a technique for reducing false-positive results in a CAD scheme for detecting lung nodules in chest radiographs by using a multiple massive-training artificial neural network (multi-MTANN).

## MATERIALS AND METHODS

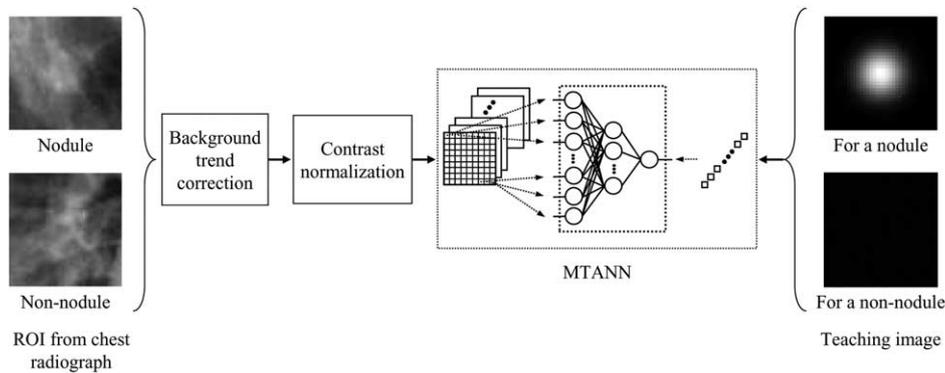
### Database

The database used in this study consisted of 91 chest radiographs containing 91 solitary pulmonary nodules with subtlety ratings of subtle, relatively obvious, and obvious

from the Digital Image Database developed by the Japanese Society of Radiological Technology (JSRT) (26), which is a publicly available database. The chest radiographs were collected from 14 medical institutions. The absence or presence of nodules in the chest radiographs was confirmed by computed tomography (CT). The locations of all nodules were confirmed by three chest radiologists. The criteria for inclusion of radiographs in the database were: (1) absence of nodules larger than 35 mm, (2) absence of suspicious nodules that were not confirmed by CT examination, (3) no more than one nodule per patient, and (4) absence of nodules with margins that could not be confirmed by radiologists. The chest radiographs were digitized with a 0.175-mm pixel size, a matrix size of 2,048 x 2,048, and a 12-bit gray-scale level. The sizes of nodules ranged from 8.9–29.1 mm, with an average size of 17.4 mm. The database contained 64 malignant nodules and 27 benign nodules, which were confirmed by histologic or cytologic examination or by follow-up imaging. For reducing noise and increasing computational efficiency, the size of the chest radiographs was reduced to 512 x 512 pixels with a 10-bit gray-scale level through the use of averaging.

### Existing CAD Scheme

Our current CAD scheme (27) for detecting lung nodules in chest radiographs consists of four steps, as follows: (1) pre-processing based on a difference-image technique (9,10,12), (2) identification of initial nodule candidates by use of a multiple gray-level thresholding technique, (3) grouping of initial nodule candidates, and (4) the use of rule-based and linear-discriminant classifiers for reducing false-positive results. The difference-image technique is a technique for enhancing lung nodules and suppressing normal background structures. The difference image was obtained by subtraction of the nodule-suppressed image from the nodule-enhanced image. Initial nodule candidates were identified in the difference images through the use of a multiple gray-level thresholding technique. The initial nodule candidates were classified into 13 groups according to their detection threshold levels. Eight image features for each group were calculated, consisting of the effective diameter (9), degree of circularity (9), degree of irregularity (9), growth rate of the effective diameter (10), growth rate of the degree of circularity (10), growth rate of the degree of irregularity (10), run length (10), and contrast (9) in the original image and the difference image. These eight image features were used as input to the rule-based and linear discriminant classifiers. With our current CAD scheme, a sensitivity of 82.4% (75/91), with 4.5 (410/91) false positives per image, was achieved for the database of 91 chest radiographs. We used



**Figure 1.** Architecture and training of an MTANN for distinguishing nodules from non-nodules in chest radiographs. In this procedure, a background-trend-correction technique followed by contrast normalization is applied to an ROI. The pixel values in sub-regions extracted from an ROI in a chest radiograph are entered as input to the MTANN. The MTANN outputs single pixels corresponding to the center pixels in the input sub-regions. The teaching image for a nodule contains a 2D gaussian distribution of gray levels, and that for a non-nodule contains zero gray levels.

the 75 true positives (nodules) and the 410 false positives (non-nodules) for testing our scheme for reducing false-positive results in this study.

### MTANN

Suzuki et al. have investigated supervised nonlinear image-processing techniques based on artificial neural networks (ANNs), known as “neural filters,” for reducing the quantum mottle in X-ray images (28), a “neural edge enhancer” for enhancing desired edges in noisy images (29,30), and a “neural edge detector” for the supervised detection of subjective edges traced by cardiologists (31). They have developed training methods (32,33), design methods (34,35), and an analysis method (36) for these techniques. They (37,38) have also extended these techniques to accommodate various pattern-classification tasks, and have developed an MTANN and applied it for reducing false-positive results in computerized detection of lung nodules in low-dose CT. It is not clear, however, whether this approach is applicable to images made with a different modality, such as chest radiography, because a chest radiograph is a projection image whose characteristics differ from those of an axial image such as that acquired by CT (eg, a nodule can overlap with ribs in a chest radiograph, whereas a nodule and ribs are separated in a CT image).

The MTANN is a highly nonlinear filter that can be trained by the use of input images and corresponding teaching images. The MTANN consists of a modified multilayer ANN (39,40) that is capable of operating on image data directly, as shown in Figure 1. First, a region of interest (ROI) including a nodule or non-nodule is extracted automatically from a chest radiograph on the basis of the location of the nodule or non-nodule identified by our CAD scheme. The

inputs to the MTANN are pixel values in a sub-region,  $R_S$ , of the ROI. The output of the MTANN is a continuous value that corresponds to the center pixel in the sub-region, and is »

$$f(x, y) = NN\{g(x - i, y - j) \mid i, j \in R_S\}, \quad (1)$$

where  $f(x, y)$  is an estimate for the teaching value,  $x$  and  $y$  are the indices of coordinates,  $NN\{\bullet\}$  is the output of the modified multilayer ANN, and  $g(x, y)$  is an input pixel value. The output image is obtained by scanning of an input image with the MTANN.

### Preprocessing of ROIs

The background trends in the ROIs in a chest radiograph are, in general, different from those at other locations in the same image, those in another patient’s image, and those in an image acquired under different acquisition conditions. In order to reduce these effects, we applied a background-trend-correction technique (41) to the ROIs in our database images. The background-trend-correction technique is a technique that involves subtracting, from the original ROI, a two-dimensional (2D) surface that is fitted to gray levels in the ROI. We used a 2D  $n$ th order polynomial as the 2D surface, as follows:

$$F^n(x, y) = \sum_{k=1}^{n+1} \sum_{m=1}^k a_{(k-1)k2+m} x^{k-m} y^{m-1}, \quad (2)$$

where  $a_k$  is the  $k$ th coefficient, and  $k$  and  $m$  are variables. The coefficients of the 2D polynomial are determined by

use of the least-squares method. The background trend in the ROI is corrected by the following equation:

$$g_B(x, y) = g(x, y) - F^n(x, y). \quad (3)$$

Contrast normalization is then performed on the background-trend-corrected ROI. All pixel values in the ROI are divided by the average pixel value in a circle region,  $R_C$ , represented by

$$g_C(x, y) = \frac{g_B(x, y)}{\sum_{x, y \in R_C} g_B(x, y) / N}, \quad (4)$$

where  $N$  is the number of pixels in  $R_C$ . The diameter of the circle region was determined to be 40 pixels, which corresponds to the maximum size of nodules to be detected. The pixel values  $g_C(x, y)$  of the ROI are normalized such that a pixel value of  $-200$  is zero and a pixel value of  $200$  is one, with these numbers corresponding, respectively, to the mean for the minimum pixel values in the ROIs and the mean for the maximum pixel values in the ROIs.

### Training of the MTANN

The training of the MTANN for distinguishing between nodules and non-nodules in chest radiographs is illustrated in Figure 1. For distinguishing between nodules and non-nodules, the teaching image is designed to contain the distribution for a "likelihood of being a nodule" (ie, the teaching image for nodules contains a 2D gaussian distribution of gray levels with standard deviation  $\sigma_T$ , and that for non-nodules contains zero gray levels. The training region  $R_T$  in the input image is divided on a pixel-by-pixel basis into a large number of overlapping sub-regions. The MTANN is trained by presenting each of the input sub-regions together with each of the corresponding teaching single pixels. The error to be minimized by training is defined by

$$E = \frac{1}{2P} \sum_p \{T^{(p)} - f^{(p)}\}^2, \quad (5)$$

where  $p$  is a training pixel number,  $T^{(p)}$  is the  $p$ th training pixel in the training regions  $R_T$  in the teaching images,  $f^{(p)}$  is the  $p$ th training pixel in the training region  $R_T$  in the output images, and  $P$  is the number of training pixels. The training region may be determined so as to cover nodules sufficiently. The MTANN is trained by a modified back-propagation algorithm (39,40), which was de-

rived for the modified multilayer ANN in the same way as the original back-propagation algorithm (42,43). After training, the MTANN is expected to generate a higher likelihood value for a nodule, and a lower value for a non-nodule.

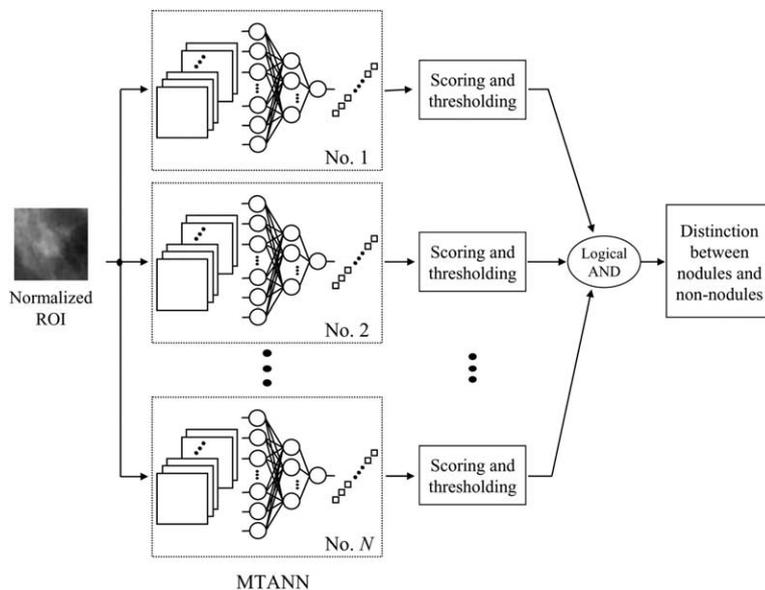
### Multi-MTANN

In order to distinguish between nodules and various types of non-nodules, we have extended the capability of a single MTANN and developed a multi-MTANN (37). The architecture of the multi-MTANN is shown in Figure 2. The multi-MTANN consists of several MTANNs arranged in parallel. Each MTANN is trained independently by use of the same nodules and a different set of non-nodules. Each MTANN acts as an expert system for distinguishing nodules from non-nodules of a specific non-nodule type.

When an original image for the  $s$ th nodule candidate is entered into the  $n$ th trained MTANN, the output image for the  $s$ th nodule candidate is obtained by scanning of the original image with the trained MTANN. The distinction between a nodule and a non-nodule is determined by use of a score defined from the output image of the  $n$ th trained MTANN, described as

$$s_{ns} = \sum_{x, y \in R_E} f_G(\sigma_n; x, y) \times f_{ns}(x, y), \quad (6)$$

where  $S_{ns}$  is the score of the  $n$ th trained MTANN for the  $s$ th nodule candidate,  $R_E$  is the region for evaluation,  $f_{ns}(x, y)$  is the output image of the  $n$ th trained MTANN for the  $s$ th nodule candidate, whose center corresponds to the center of  $R_E$ ,  $f_G(\sigma_n; x, y)$  is a 2D gaussian function with the standard deviation  $\sigma_n$  where its center corresponds to the center of  $R_E$ , and  $n$  is the MTANN number in the multi-MTANN. A higher score would indicate a nodule and a lower score would indicate a non-nodule. The distinction between a nodule and the specific type of non-nodule is determined by thresholding of the score with a different threshold for each trained MTANN, because the appropriate threshold for each trained MTANN may be different according to the type of non-nodule used for training it. The thresholds may be determined so as not to remove any nodules, but to eliminate the imaging of non-nodules to the maximum extent possible. The outputs of the expert MTANNs are combined by use of the logical multiplication (AND) operation in such a way that each of the trained MTANNs eliminates no nodules but removes some of the specific types of non-nodule for which the MTANN was trained.



**Figure 2.** Architecture of a multi-MTANN consisting of several MTANNs arranged in parallel. Each MTANN is trained by use of a different type of non-nodule, but with the same nodules. Each MTANN acts as an expert system for distinguishing nodules from a specific type of non-nodule. The output of each MTANN is integrated by the logical multiplication (AND) operation.

## RESULTS

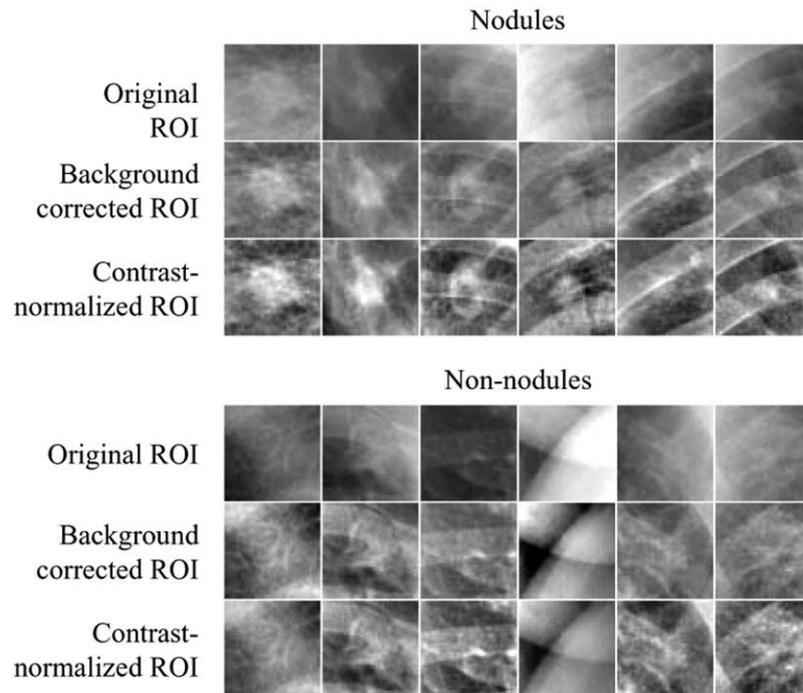
### Training

Figure 3 illustrates the effect of background-trend correction and contrast normalization on the distinction of nodules from non-nodules. Because the variations in the backgrounds and degrees of contrast of the original ROIs are large, as shown in Figure 3, it is difficult to distinguish nodules from non-nodules in these original ROIs. Background-trend correction and contrast normalization substantially reduce the variations in the backgrounds and degrees of contrast, as also shown in Figure 3. The contrast of nodules with their backgrounds is relatively constant in the processed ROIs. It is apparent that the distinction between nodules and non-nodules in the processed ROIs is better than that in the original ROIs.

We classified false positives (non-nodules) reported by our CAD scheme into six groups by using a method for selecting training samples for a multi-MTANN (38). With this method, training samples for each MTANN were determined on the basis of the rankings of scores in the free-response receiver-operating characteristic (FROC) (44) space. As training samples for each MTANN, we used 12 typical nodules selected by one of the authors (K.S.) and 12 non-nodules determined by use of the selection method (38) from each

of the 6 CAD-reported false-positive groups. Figure 4 shows 4 of the 12 training samples for nodules and non-nodules for six MTANNs in the multi-MTANN. These ROIs (45 x 45 pixels of each ROI are shown in the figure) were extracted from chest radiographs. The six sets of non-nodules included: (1) low-contrast ribs with lung vessels, (2) soft-tissue opacities with higher gray levels, (3) right ribs, (4) small round opacities, (5) left ribs or ribs with horizontal orientations, and (6) soft-tissue opacities. Each set was used for training the MTANN, with the number in the multi-MTANN (eg, set 1) corresponding to the MTANN number (number 1).

A three-layer structure was used as the structure of each MTANN because any continuous mapping can be realized approximately by three-layer ANNs (45,46). The size of the sub-region  $R_S$  of the MTANN, the standard deviation  $\sigma_T$  of the 2D gaussian distribution, and the size of the training region  $R_T$  in the teaching image were determined empirically to be 9 x 9 pixels, 5.0 pixels, and 19 x 19 pixels, respectively. We determined the number of hidden units of the MTANN by using a method for designing the structure of an ANN (34). The method is a sensitivity-based pruning method (ie, the sensitivity to the training error was calculated when a certain unit was removed experimentally, and the unit with the smallest training error was removed). Removing the redundant



**Figure 3.** Effect of background-trend correction and contrast normalization. Illustrations of original ROIs including nodules extracted from chest radiographs, the background-trend-corrected ROIs for these nodules, the contrast-normalized ROIs for the nodules, original ROIs including non-nodules, the background-trend-corrected ROIs for the non-nodules, and the contrast-normalized ROIs for the non-nodules.

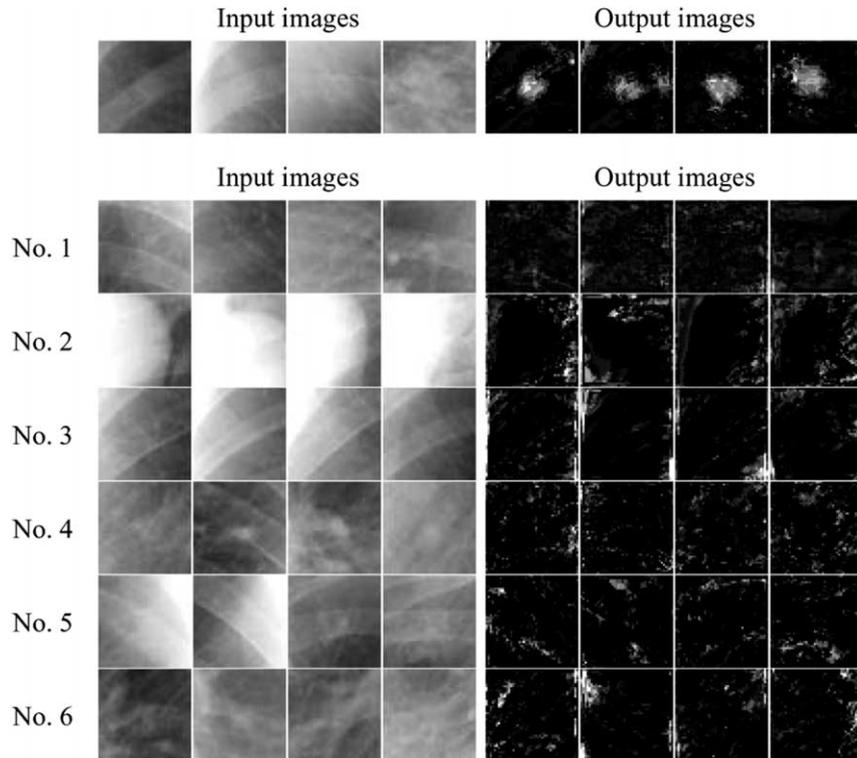
hidden units and retraining the MTANN for recovering the potential loss from this removal were performed alternately, resulting in a reduced structure from which redundant units were removed. As a result, the number of hidden units was determined to be 20. Thus, the numbers of units in the input, hidden, and output layers were 81, 20, and 1, respectively. With the parameters given above, the training of each MTANN in the multi-MTANN was performed 500,000 times. The training took a CPU time of 29.8 hours on a PC-based workstation (CPU: Pentium IV, 1.7 GHz; Dell Inc., Round Rock, TX), while the time for applying the trained MTANN to nodule candidates was negligibly small.

### Evaluation

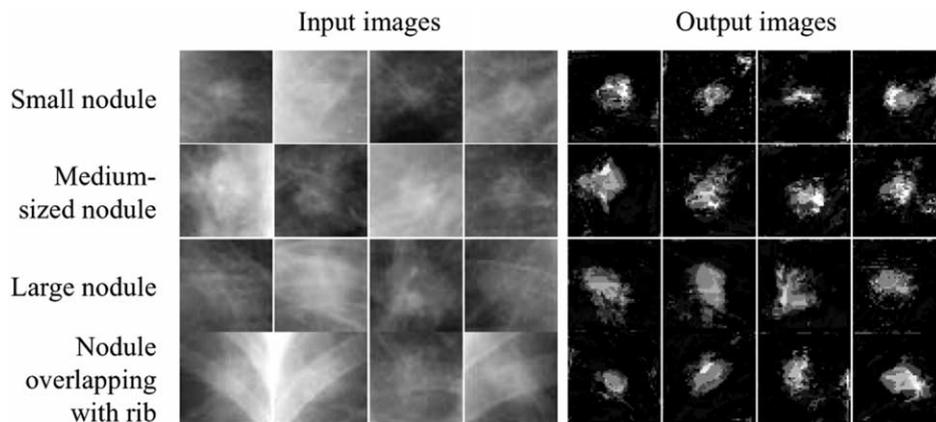
Figure 4 shows the output images of the trained MTANNs for the cases used in training. The nodules in the output images are represented by the light distribution near the centers of ROIs, whereas the non-nodules in the output images are largely dark around the center. This result indicated that the MTANNs had properly learned the training images. Figure 5 shows non-training nodules, and the corresponding output images of the trained single MTANN 1. Various nodules are represented by light distributions. Figure

6 shows input images and the corresponding output images of each of the six MTANNs for non-training cases. The nodules in the output images of each MTANN were represented by light distributions near the centers of the nodules, whereas non-nodules in the corresponding group for which the MTANN was trained were mainly dark around the center in the output images, as expected. It is apparent that the distinction between nodules and non-nodules in the output images of the MTANN is superior to that in the original images.

We applied the trained multi-MTANN to the 75 true positives (nodules) and 410 false positives (non-nodules) produced by our CAD scheme. The scoring method was applied to the output images of the MTANNs, where the standard deviation  $\sigma_n$  was determined empirically to be within the range from 4.5–7.7. Figure 7 shows the relationships between the scores for the images from two MTANNs in the multi-MTANN. The results show that each MTANN could remove different non-nodules without removing any true-positive nodule; accordingly, use of the multi-MTANN could eliminate various non-nodules. For example, if a threshold of 0.64 is used for thresholding the scores from MTANN 1 (Fig. 7a, vertical axis), all nodules can be distinguished from about 30% of non-nodules. If a threshold of



**Figure 4.** (a) Training samples of nodules (top left four images) and the corresponding output images of the trained MTANN 1, and (b) training samples of non-nodules (four images for each MTANN) and the corresponding output images of the trained MTANNs.



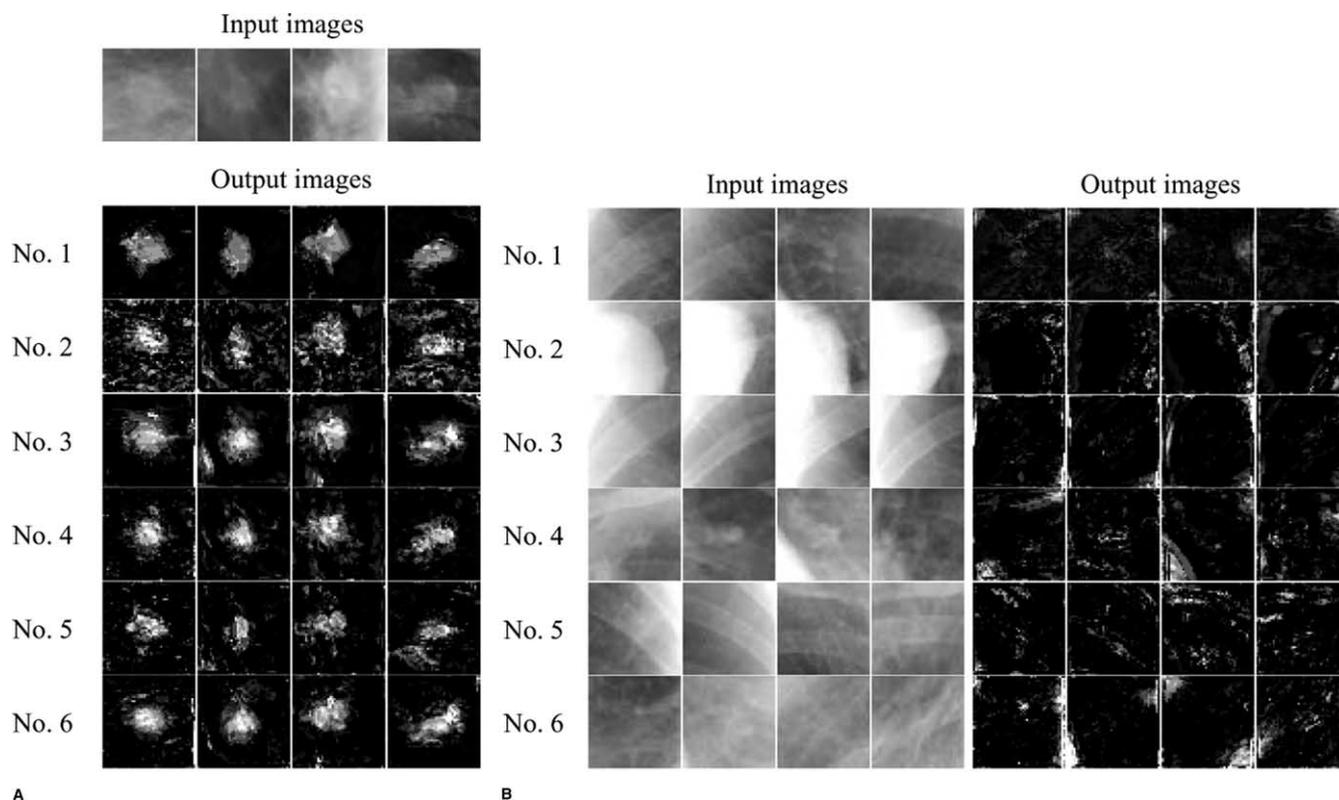
**Figure 5.** Illustrations of various non-training nodules and the corresponding output images of the trained MTANN 1.

0.49 is used for thresholding the scores from MTANN 2 (Fig. 7a, horizontal axis), another 10% of non-nodules can be removed without the loss of any nodules. By combining these thresholds, 40% of non-nodules can be removed. Thus, the use of 6 different thresholds for thresholding the scores from 6 MTANNs could eliminate many non-nodules.

The performance of the multi-MTANN was evaluated by FROC curves, as shown in Figure 8. The FROC curve expresses an overall sensitivity as a function of the number of false positives per image at a specific operating point on the

curve. With the multi-MTANN, the number of false positives was reduced, and at a certain operating point on the FROC curve, it was reduced to 31.7% (130/410), with a reduction of 1 true positive. The multi-MTANN reduced the false-positive rate of our original CAD scheme from 4.5 to 1.4 (130/91) false positives per image, at an overall sensitivity of 81.3% (74/91).

In order to investigate the effect of the background-trend correction on the performance of the MTANN, we conducted experiments under four conditions: (1) no



**Figure 6.** (a) Non-training nodules and the corresponding output images of the trained MTANNs, and (b) non-training non-nodules and the corresponding output images of the trained MTANNs.

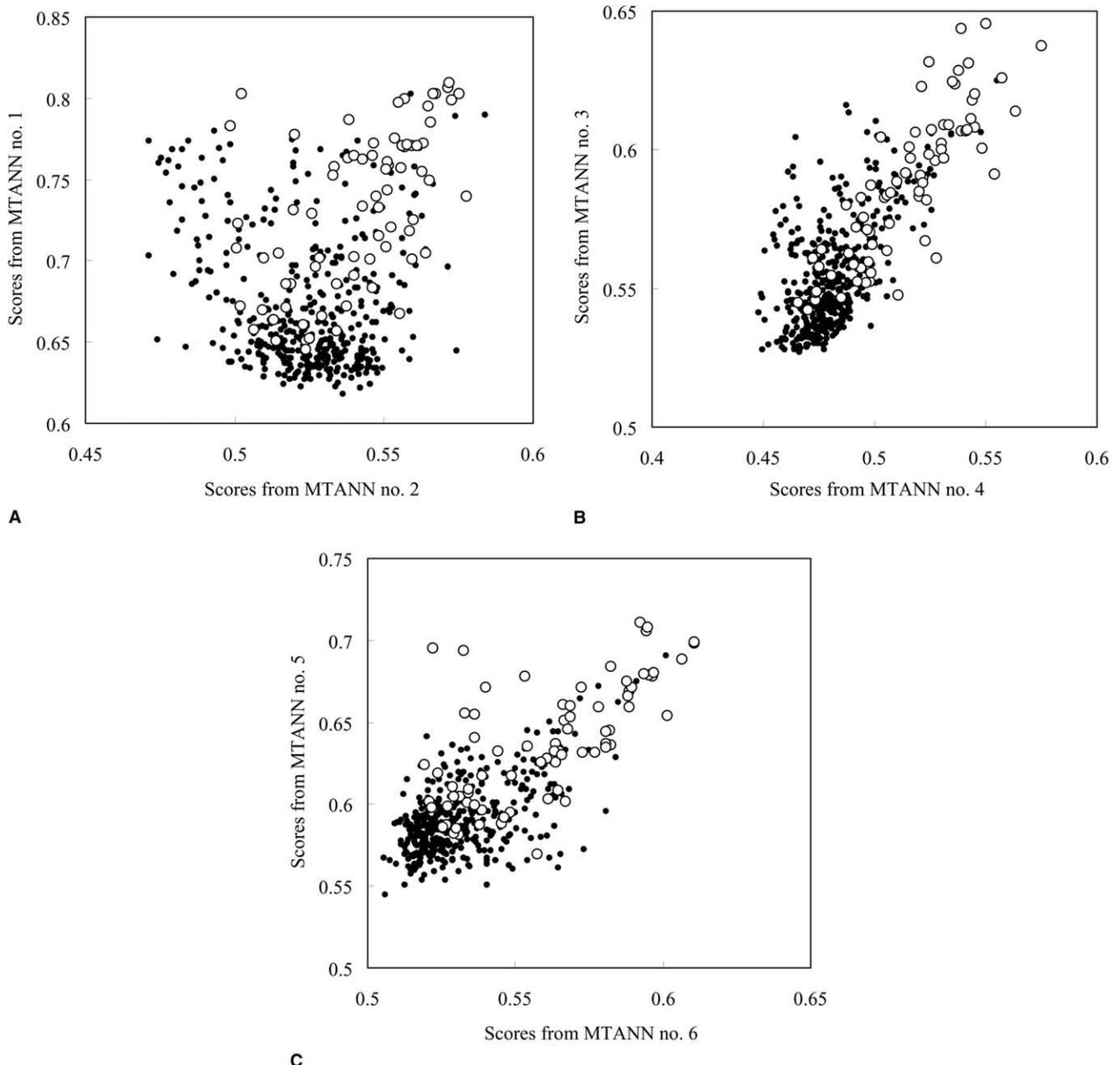
background-trend correction, (2) use of the first-order polynomial in the background-trend correction, (3) use of the second-order polynomial, and (4) use of the third-order polynomial. The performance of the MTANN was evaluated by ROC analysis (47,48). Figure 9 shows the ROC curves for the MTANNs under the four conditions. The result showed that the background-trend correction was required for the MTANN for chest radiographs. Statistical analyses of ROC curves showed that differences between MTANNs without correction and each of the other three conditions, and also between the first-order polynomial and the third-order polynomial, were statistically significant, but that differences between other combinations were not. The values of  $A_z$ , the area under the ROC curve (49), for the MTANN with use of the background-trend correction with the first-order polynomial and for the MTANN with the second-order polynomial were very similar (both  $A_z = 0.75$ ). However, the performance of the MTANN with the first-order polynomial at higher sensitivities was better than that with the second-order polynomial. Therefore, we used the first-order polynomial in the background-trend correction.

To investigate the generalization ability (performance for non-training cases) of the multi-MTANN, we evaluated its

performance with non-training cases alone (ie, the training samples for 12 nodules and 72 non-nodules were excluded from the evaluation). Because the distinction between nodules and non-nodules in the output images of the multi-MTANN for training cases shown in Figure 4 was similar to that for the non-training cases shown in Figure 6, the performance of the multi-MTANN for non-training cases would be similar to that for training cases. We defined the number of false positives per image in this evaluation as

$$FPS = \frac{RFP}{NIMG \times \frac{TFP - FPT}{TFP}}, \quad (7)$$

where  $RFP$  is the number of remaining false positives after application of a false-positive reduction method,  $NIMG$  is the number of images in the complete database,  $TFP$  is the total number of false positives reported by the CAD scheme, and  $FPT$  is the number of false positives used as training samples, with the result that the use of some false positives for training the multi-MTANN does not artificially reduce the overall false-positive rate. The performance of the multi-MTANN for non-training cases



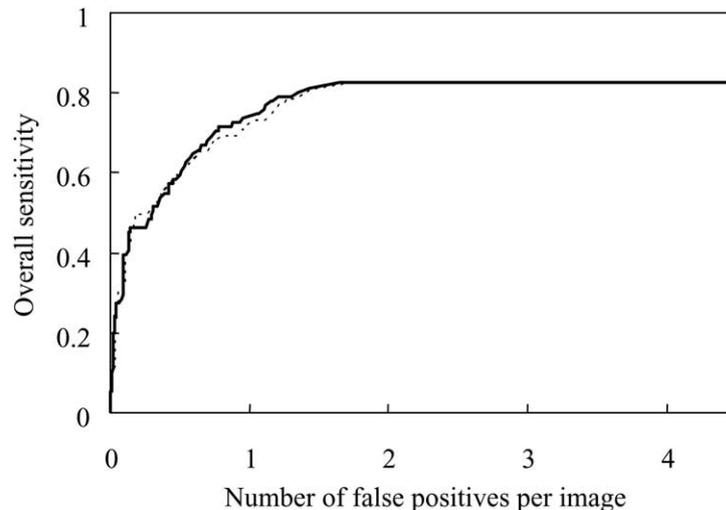
**Figure 7.** Relationships between scores from two MTANNs in the multi-MTANN for nodules (white circles) and non-nodules (black circles). (a) Relationship between MTANNs 1 and 2, (b) relationship between MTANNs 3 and 4, and (c) relationship between MTANNs 5 and 6.

was similar to its performance for the complete database, as shown in Figure 8.

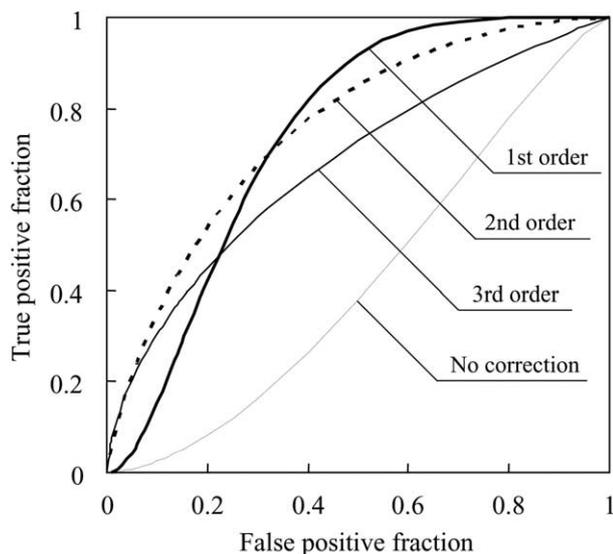
## DISCUSSION

We used 75 true-positive nodules from our original CAD scheme for testing MTANNs in this study. The performance of MTANNs with the false negatives (16 nodules) in our original scheme is of interest. Because

the nodules that produced false-negative results were relatively small and of low-contrast, it might be difficult for MTANNs to detect these false-negative nodules. However, the results of our experiments with CT images (37) indicated a relatively high level of performance of MTANNs on the database, including false-negative nodules that were small and of low-contrast. Therefore, we believe that the performance of MTANNs for such false-negative nodules would be



**Figure 8.** FROC curve (thick solid curve) of the multi-MTANN consisting of 6 MTANNs for 75 true positives (nodules) and 410 false positives (non-nodules), and FROC curve (dotted curve) of the multi-MTANN for non-training cases (ie, the training samples were excluded from the evaluation). The FROC curve of the multi-MTANN indicates an 81.3% overall sensitivity and a reduction in the false-positive rate from 4.5 to 1.4 per image.



**Figure 9.** Effect of background-trend correction on the performance of an MTANN. The ROC curves are shown of the MTANNs for four conditions: (1) no background-trend correction, (2) use of the first-order polynomial in the background-trend correction, (3) use of the second-order polynomial in the background-trend correction, and (4) use of the third-order polynomial in the background-trend correction.

comparable to their performance on the true-positive nodules in the present study.

Use of the multi-MTANN technique substantially reduced the false-positive rate with our CAD scheme for lung-nodule detection in chest radiographs, while preserving a high level of sensitivity.

#### ACKNOWLEDGMENT

The authors are grateful to Feng Li, MD, PhD, for her clinical advice; to Qiang Li, PhD, Hidetaka Arimura, PhD, Hotaka Takizawa, PhD, and Roger Engelmann, MS, for their valuable suggestions; to Charles E. Metz, PhD, for the use of the LABROC5 program; and to Mrs. Elisabeth F. Lanzl for improving the manuscript. K. Doi and H. MacMahon are shareholders in R2 Technology, Inc., Sunnyvale, CA, and K. Doi is a shareholder in a Deus Technologies, Inc., Rockville, MD.

#### REFERENCES

1. Heelan RT, Flehinger BJ, Melamed MR, et al. Non-small-cell lung cancer: results of the New York screening program. *Radiology* 1984; 151:289-293.
2. Sobue T, Suzuki R, Matsuda M, et al. Survival for clinical stage I lung cancer not surgically treated. *Cancer* 1992; 69:685-692.
3. Austin JHM, Romney BM, Goldsmith LS. Missed bronchogenic carcinoma: radiographic findings in 27 patients with a potentially resectable lesion evident in retrospect. *Radiology* 1992; 182:115-122.
4. Shah PK, Austin JHM, White CS, et al. Missed non-small cell lung cancer: radiographic findings of potentially resectable lesions evident only in retrospect. *Radiology* 2003; 226:235-241.
5. MacMahon H, Doi K, Chan HP, et al. Computer-aided diagnosis in the chest. *J Thoracic Imaging* 1990; 5:67-76.
6. Abe K, Doi K, MacMahon H, et al. Computer-aided diagnosis in chest radiography: analysis of results in a large clinical series. *Invest Radiol* 1993; 28:987-993.
7. Kobayashi T, Xu XW, MacMahon H, et al. Effect of a computer-aided diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs. *Radiology* 1996; 199:843-848.
8. Ginneken BV, Romeny BTH, Viergever MA. Computer-aided diagnosis in chest radiography: a survey. *IEEE Trans Med Imag* 2001; 20:1228-1241.

9. 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.
10. Giger ML, Doi K, MacMahon H, et al. Pulmonary nodules: computer-aided detection in digital chest images. *Radiographics* 1990; 10:41-51.
11. Wu Y, Doi K, Giger ML, et al. Reduction of false positives in computerized detection of lung nodules in chest radiographs using artificial neural networks, discriminant analysis, and a rule-based scheme. *J Digital Imag* 1994; 7:196-207.
12. Xu XW, Doi K, Kobayashi T, et al. Development of an improved CAD scheme for automated detection of lung nodules in digital chest images. *Med Phys* 1997; 24:1395-1403.
13. Li Q, Katsuragawa S, Doi K. Computer-aided diagnostic scheme for pulmonary nodule detection in digital chest radiographs: elimination of false-positives by using a multiple-templates matching technique. *Med Phys* 2001; 28:2070-2076.
14. Sankar P, Sklansky JA. Gestalt guided heuristic boundary follower for X-ray images of lung nodules. *IEEE Trans Pattern Anal Machine Intell* 1982; 4:326-331.
15. Lo SC, Lou SL, Lin JS, et al. Artificial convolution neural network techniques and applications for lung nodule detection. *IEEE Trans Med Imag* 1995; 14:711-718.
16. Lin JS, Hasegawa A, Freedman M, et al. Differentiation between nodules and end-on vessels using a convolution neural network architecture. *J Digital Imag* 1995; 8:132-141.
17. Lin JS, Lo SC, Hasegawa A, et al. Reduction of false positives in lung nodule detection using a two-level neural classification. *IEEE Trans Med Imag* 1996; 15:206-217.
18. Carreira MJ, Cabello D, Penedo MG, et al. Computer-aided diagnosis: automatic detection of lung nodules. *Med Phys* 1998; 25:1998-2006.
19. Penedo MG, Carreira MJ, Mosquera A, et al. Computer-aided diagnosis: a neural-network-based approach to lung nodule detection. *IEEE Trans Med Imag* 1998; 17:872-880.
20. Coppini G, Diciotti S, Falchini M, et al. Neural networks for computer-aided diagnosis: detection of lung nodules in chest radiograms. *IEEE Trans Inform Technol Biomed* 2003; 7:344-357.
21. Lampeter W, Wandtke J. Computerized search of chest radiographs for nodules. *Invest Radiol* 1986; 21:384-390.
22. Floyd C Jr, Patz E, Lo J, et al. Diffuse nodular lung disease on chest radiographs: a pilot study of characterization by fractal dimension. *AJR Am J Roentgenol* 1996; 167:1185-1187.
23. Mao F, Qian W, Gaviria J, et al. Fragmentary window filtering for multi-scale lung nodule detection: preliminary study. *Acad Radiol* 1998; 5:306-311.
24. Kakeda S, Moriya J, Sato H, et al. Improved detection of lung nodules on chest radiographs using a commercial computer-aided diagnosis system. *AJR Am J Roentgenol* 2004; 182:505-510.
25. Matsumoto T, Yoshimura H, Doi K, et al. Image feature analysis of false-positive diagnoses produced by automated detection of lung nodules. *Invest Radiol* 1992; 27:587-579.
26. Shiraishi J, Katsuragawa S, Ikezoe J, et al. Development of a digital image database for chest radiographs with and without a lung nodule: ROC analysis on radiologists' performance in detection of pulmonary nodules. *AJR Am J Roentgenol* 2000; 174:71-74.
27. Shiraishi J, Abe H, Engelmann R, et al. Characteristics of image database on the performance of computer-aided diagnosis for the detection of pulmonary nodules in chest radiographs. *Proc SPIE Med Imag* 2003; 5032:177-182.
28. Suzuki K, Horiba I, Sugie N, et al. Neural filter with selection of input features and its application to image quality improvement of medical image sequences. *IEICE Trans Inform Syst* 2002; E85-D:1710-1718.
29. Suzuki K, Horiba I, Sugie N. Neural edge detector—a good mimic of conventional one yet robust against noise. *Lect Notes Comput Sci* 2001; 2085:303-310.
30. Suzuki K, Horiba I, Sugie N. Neural edge enhancer for supervised edge enhancement from noisy images. *IEEE Trans Pattern Anal Machine Intell* 2003; 25:1582-1596.
31. Suzuki K, Horiba I, Sugie N, et al. Extraction of left ventricular contours from left ventriculograms by means of a neural edge detector. *IEEE Trans Med Imag* 2004; 23:330-339.
32. Suzuki K, Horiba I, Sugie N. Training under achievement quotient criterion. *Neural Networks Signal Proc* 2000; 10:537-546.
33. Suzuki K, Horiba I, Sugie N. Simple unit-pruning with gain-changing training. *Neural Networks Signal Proc* 2001; 11:153-162.
34. Suzuki K, Horiba I, Sugie N. Designing the optimal structure of a neural filter. *Neural Networks Signal Proc* 1998; 8:323-332.
35. Suzuki K, Horiba I, Sugie N. A simple neural network pruning algorithm with application to filter synthesis. *Neural Proc Lett* 2001; 13:43-53.
36. Suzuki K, Horiba I, Sugie N. Efficient approximation of a neural filter for quantum noise removal in X-ray images. *IEEE Trans Signal Proc* 2002; 50:1787-1799.
37. Suzuki K, Armato III SG, Li F, et al. Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose CT. *Med Phys* 2003; 30:1602-1617.
38. Suzuki K, Armato III SG, Li F, et al. Effect of a small number of training cases on the performance of massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose CT. *Proc SPIE Medical Imaging (SPIE MI)* 2003; 5032:1355-1366.
39. Suzuki K, Horiba I, Ikegaya K, et al. Recognition of coronary arterial stenosis using neural network on DSA system. *Systems and Computers in Japan* 1995; 26:66-74.
40. Suzuki K, Horiba I, Sugie N, et al. Computer-aided diagnosis system for coronary artery stenosis using a neural network. *Proc SPIE Med Imag* 2001; 4322:1771-1782.
41. Katsuragawa S, Doi K, Nakamori N, et al. Image feature analysis and computer-aided diagnosis in digital radiography: Effect of digital parameters on the accuracy of computerized analysis of interstitial disease in digital chest radiographs. *Med Phys* 1990; 17:72-78.
42. Rumelhart DE, Hinton GE, Williams RJ. Learning representations of back-propagation errors. *Nature* 1986; 323:533-536.
43. Rumelhart DE, Hinton GE, Williams RJ. Learning internal representations by error propagation. In: Rumelhart DE, McClelland JL, and the PDP Research Group (eds.), *Parallel Distributed Processing*. Cambridge, MA: MIT Press, 1986; 1:318-362.
44. Chakraborty DP, Winter LHL. Free-response methodology: alternate analysis and a new observer-performance experiment. *Radiology* 1990; 174:873-881.
45. Funahashi K. On the approximate realization of continuous mappings by neural networks. *Neural Networks* 1989; 2:183-192.
46. Barron AR. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Trans Inform Theor* 1993; 39:930-945.
47. Metz CE. ROC methodology in radiologic imaging. *Invest Radiol* 1986; 21:720-733.
48. Metz CE, Herman BA, Shen JH. Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. *Stat Med* 1998; 17:1033-1053.
49. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; 148:839-843.