

How Can a Massive Training Artificial Neural Network (MTANN) Be Trained With a Small Number of Cases in the Distinction Between Nodules and Vessels in Thoracic CT?¹

Kenji Suzuki and Kunio Doi

Rationale and Objectives. To demonstrate that a massive training artificial neural network (MTANN) can be adequately trained with a small number of cases in the distinction between nodules and vessels (non-nodules) in thoracic computed tomography (CT) images.

Materials and Methods. An MTANN is a trainable, highly nonlinear filter consisting of a linear-output multilayer artificial neural network model. For enhancement of nodules and suppression of vessels, we used 10 nodules and 10 non-nodule images as training cases for MTANNs. The MTANN is trained with a large number of input subregions selected from the training cases and the corresponding pixels in teaching images that contain Gaussian distributions for nodules and zero for non-nodules. We trained three MTANNs with different numbers (1, 9, and 361) of training samples (pairs of the subregion and the teaching pixel) selected from the training cases. In order to investigate the basic characteristics of the trained MTANNs, we applied the MTANNs to simulated CT images containing various-sized model nodules (spheres) with different contrasts and various-sized model vessels (cylinders) with different orientations. In addition, we applied the trained MTANNs to nontraining actual clinical cases with 59 nodules and 1,726 non-nodules.

Results. In the output images for the simulated CT images by use of the MTANNs trained with small numbers (one and nine) of subregions, model vessels were clearly visible and were not removed; thus, the MTANNs were not trained properly. However, in the output image of the MTANN trained with a large number of subregions, various-sized model nodules with different contrasts were represented by light nodular distributions, whereas various-sized model vessels with different orientations were dark and thus were almost removed. This result indicates that the MTANN was able to learn, from a very small number of actual nodule and non-nodule cases, the distinction between nodules (spherelike objects) and vessels (cylinder-like objects). In nontraining clinical cases, the MTANN was able to distinguish actual nodules from actual vessels in CT images. For 59 actual nodules and 1,726 non-nodules, the performance of the MTANN decreased as the number of training samples (subregions) in each case decreased.

Conclusions. The MTANN can be trained with a very small number of training cases (10 nodules and 10 non-nodules) in the distinction between nodules and non-nodules (vessels) in CT images. Massive training by scanning of training cases to produce a large number of training samples (input subregions and teaching pixels) would contribute to a high generalization ability of the MTANN.

Key Words. Computer-Aided Diagnosis (CAD); Lung Nodule; Cancer; Thoracic CT; Artificial Neural Network.

© AUR, 2005

Acad Radiol 2005; 12:1333–1341

¹ From the Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, The University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637. Supported by USPHS Grants Nos. CA62625 and CA98119. Received January 10, 2005; revision received and accepted June 16. **Address correspondence to:** KS. e-mail: suzuki@uchicago.edu

© AUR, 2005

doi:10.1016/j.acra.2005.06.017

Lung cancer is the leading cause of cancer deaths among Americans (1). Low-dose helical computed tomography (LDCT) has been used for early detection of lung cancer (2–6). Radiologists, however, may fail to detect lung nodules in CT images that are visible in retrospect (7,8). Therefore, a computer-aided diagnostic (CAD) scheme for detecting lung nodules in CT images (9–15) has been investigated as a tool for improving radiologists' detection accuracy. A major problem with current CAD schemes is a relatively large number of false positives, which is likely to lower radiologists' efficiency in using a CAD scheme. Therefore, it is important to reduce the number of false positives as much as possible while a high sensitivity is maintained. It is difficult, however, to eliminate false positives without removal of any true-positive nodules, because variations in patterns of nodules and non-nodules are large (eg, there are various-sized nodules with different contrasts and various-sized lung vessels with different orientations in CT images; actually, the major source of false positives are lung vessels) (16).

Artificial neural networks (ANNs) have been applied for distinction between lesions and nonlesions (false positives) (17,18) and for distinction between malignant and benign lesions (19–22) in CAD schemes, and they have been shown to be useful for various CAD schemes (17–24). For achieving a high and reliable performance for nontraining cases, a large number of training cases (eg, 400–800 cases) are commonly required (25,26). If an ANN is trained with only a small number of cases, the generalization ability (performance for nontraining cases) will be lower (ie, the ANN may fit only the training cases); this is known as “overtraining” (or “overfitting”) (27). Because diagnostic radiology is progressing rapidly as technology advances, the timely development of CAD schemes is important. However, it is very difficult to collect a large number of abnormal cases for training, particularly for a CAD scheme with a new modality, such as lung cancer screening with multidetector-row CT (MDCT).

Massive training ANNs (MTANNs) have been developed for reducing the number of false positives in CAD schemes for LDCT images (16) and chest radiographs (28). With an MTANN, 54% of 1,726 false positives were removed without eliminating any of 58 true-positive nodules in a database of 63 LDCT scans containing 63 primary lung cancers (16). The MTANN was trained with only 10 nodules and 10 non-nodules (29), whereas other ANNs usually require training with a large number of cases because ANNs generally have a large number of

parameters to be determined. However, it was not clear how and why the MTANN can be trained with a small number of cases and can provide a high performance even for nontraining cases.

Our purpose in this study was to demonstrate and to verify that an MTANN can be trained with a small number of cases in the distinction between nodules and vessels (non-nodules) in a CAD scheme for detecting nodules in thoracic CT images.

MATERIALS AND METHODS

Database

Our database in this study consisted of 68 LDCT scans acquired from 68 patients who participated voluntarily in a lung cancer screening program between 1996 and 1999 in Nagano, Japan (2). The CT examinations were performed on a mobile CT scanner (CT-W950SR; Hitachi Medical, Tokyo, Japan). The CT scans were acquired with a low-dose protocol of 120 kVp, 25 mA or 50 mA, 10-mm collimation, and a 10-mm reconstruction interval at a helical pitch of 2. The pixel size was 0.586 mm or 0.684 mm. Each reconstructed CT slice had an image matrix size of 512×512 pixels, and the number of gray levels was 4,096. The number of CT slices per scan was 31 or 33. The 68 scans included 71 lung cancers that were determined by biopsy or surgery. The size (effective diameter) of the 71 cancers ranged from 6 mm to 24 mm, with a mean of 14 mm. These cancer cases included nodules in three different categories—pure ground glass opacity (GGO or non-solid) nodules (24% of nodules), mixed GGO (or part-solid) nodules (30%), and solid nodules (46%). A training set for the MTANNs used in this study included 10 LDCT scans containing 10 nodules obtained from our “missed” cancer database (8), in which 38 cancers were not reported or misreported during the initial clinical interpretation and were identified retrospectively.

Our CAD scheme

Our CAD scheme for detecting lung nodules in CT (30) consisted of a difference-image technique, a multiple gray-level-thresholding technique, extraction of image features, and a rule-based scheme. To summarize the methodology, lung segmentation was performed by use of thresholding. Nodules in the segmented lungs were enhanced by use of the difference-image technique. Nodule candidates were identified by application of the multiple

gray-level–thresholding technique to the nodule-enhanced images. Morphologic and gray-level–based features were extracted from nodule candidates. Each candidate was analyzed for distinction between a nodule and a non-nodule by use of a rule-based scheme with the extracted features. When our scheme was applied to the database, a sensitivity of 83.1% (59 of 71 nodules) with 25.4 false positives per scan (1,726/68) was achieved. We used the 59 true-positive nodules and 1,726 false positives (non-nodules) for evaluating MTANNs in this study.

MTANN

The architecture and the training of an MTANN are shown in Fig 1 (see Appendix for details). The MTANN is considered to be a trainable, highly nonlinear filter consisting of a linear-output multilayer ANN model (31–35) that is capable of operating on image data directly. The MTANN is trained with input images and the corresponding teaching images for enhancement of a specific type of opacities and suppression of other types of opacities. For enhancement of nodules and suppression of non-nodules in CT images, the teaching image contains the distribution for a “likelihood of being a nodule” (ie, the teaching image for a nodule contains a two-dimensional [2D] Gaussian function, and that for non-nodules contains zero). It is important to note that the teaching with “zero” for non-nodules is intended to remove non-nodules by use of the MTANN.

For training of the MTANN, a large number ($361 = 19 \times 19$) of overlapping subregions (9×9 pixels) is obtained by scanning pixel by pixel over the training region (27×27 pixels) in an input CT image. The pixel values in each subregion are entered as input to the MTANN. The output of the MTANN is a single pixel value, where the teaching single pixels are obtained by scanning pixel by pixel over the teaching image (19×19 pixels), and are entered as teacher to the MTANN. The MTANN is trained by presenting each of the input subregions together with each of the corresponding teaching single pixels. A modified back-propagation algorithm (31–35), which was derived for the linear-output multilayer ANN model in the same way as the back-propagation algorithm (36,37), is used for the training. The MTANN is trained by adjustment of the weights between layers iteratively so that the error between the output values and the teaching values becomes small.

After the training of the MTANN is completed, the output image is obtained by scanning of an input CT image with the trained MTANN. The trained MTANN is expected to provide higher values for nodules and lower values for non-nodules. The distinction between a nodule and a non-nodule is determined by use of a score defined as the product of the output image of the trained MTANN and a 2D Gaussian weighting function. This score represents the weighted sum of the estimate for the likelihood that the image contains a nodule near the center (ie, a higher score would indicate a nodule and a lower score a non-nodule).

Training for Nodule Enhancement and Vessel Suppression

For training cases of an MTANN, we used 10 different-sized nodules with various contrasts and 10 non-nodule images including medium and small vessels (Fig 2). The nodules in our database included three groups of nodule patterns (i.e., pure GGO, mixed GGO, and solid opacity) (8). The 10 nodules were selected from the three groups so that various-sized nodules with various contrasts were included. The size of the nodules ranged from small to large enough to cover the size range in our database. The false positives reported by our CAD scheme included a group of small or medium-sized vessels (about 70% of false positives) and another group of various opacities. The 10 vessels with relatively high contrast were selected from the group of small or medium-sized vessels, because the majority of vessels were of these sizes. The vessels were oriented in various directions, such as horizontal, vertical, and diagonal. Parameters such as the size of the subregion of the MTANN, the standard deviation of the 2D Gaussian function in the teaching image, and the size of the teaching image were determined by experimental analysis (16) to be 9×9 pixels, 5.0 pixels, and 19×19 pixels, respectively. We employed a three-layer structure for the MTANN, because it has been proved theoretically that a three-layer ANN can approximate any continuous mapping (38,39). The number of hidden units in the MTANN was determined to be 20 by use of a method for determining the structure of an ANN (40,41). Thus the numbers of input, hidden, and output units were 81, 20, and 1, respectively. With these parameters, the training of the MTANN was performed 500,000 times, and it converged with a mean absolute error of 0.112.

To gain insight as to how the MTANN can be trained only with a small number of cases, we trained MTANNs

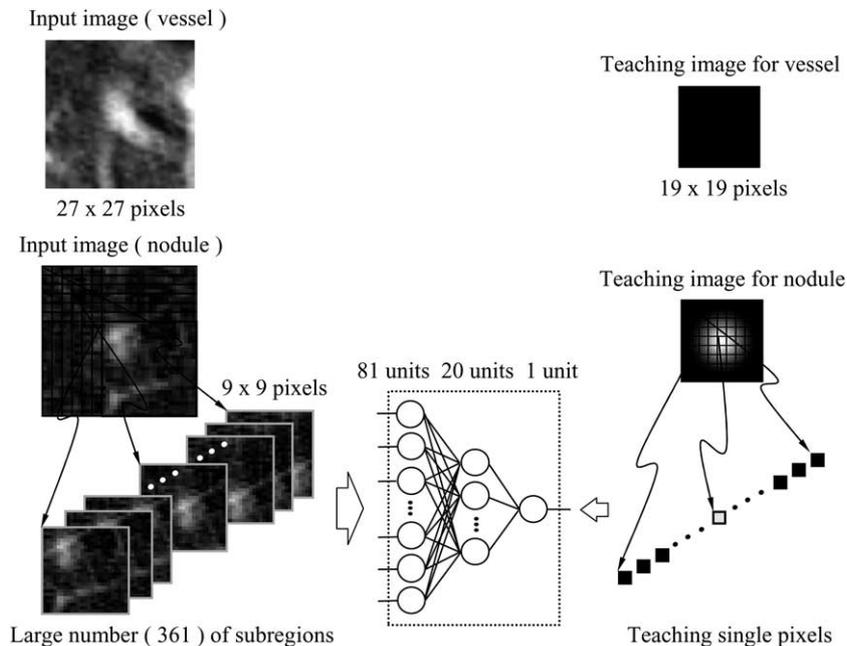


Figure 1. Architecture and training of a massive training artificial neural network (MTANN) for the distinction between nodules and non-nodules (vessels). The MTANN is trained with a large number of subregions selected from the input images and the corresponding teaching single pixels in the teaching images.

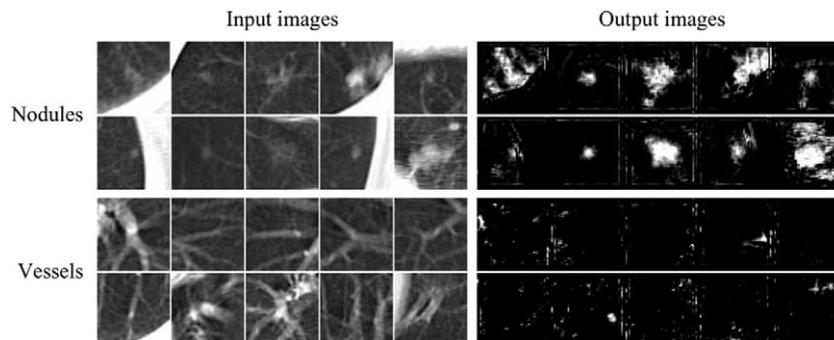


Figure 2. Ten nodules and 10 non-nodule images including vessels used for training a massive training artificial neural network (MTANN), and the corresponding output images of the trained MTANN. The nodules are various-sized with different contrasts. The non-nodule images include medium-sized and small vessels with various orientations, which are the majority of non-nodules in the lungs.

with two different numbers of subregions selected from the same training region (27×27 pixels): only one subregion (and the corresponding teaching pixel) for each case and nine (3×3) subregions, as shown in Fig 3A and 3B, respectively. The training of the MTANNs with one and nine subregions converged with mean absolute errors of 0.031 and 0.022, respectively. It should be noted that the use of one training sample for each case corresponds to a conventional method for training an ANN.

Simulated CT Images

To investigate the basic characteristics of the trained MTANNs, we created simulated CT images that contained model nodules and model vessels. A nodule was modeled as a sphere, and a vessel as a cylinder. The simulated images included various-sized model nodules (8.0 mm, 14.0 mm, and 20.0 mm in diameter) with low, medium, and high contrast (200 Hounsfield units [HU], 400 HU, and 600 HU), various-sized model vessels (2.0 mm,

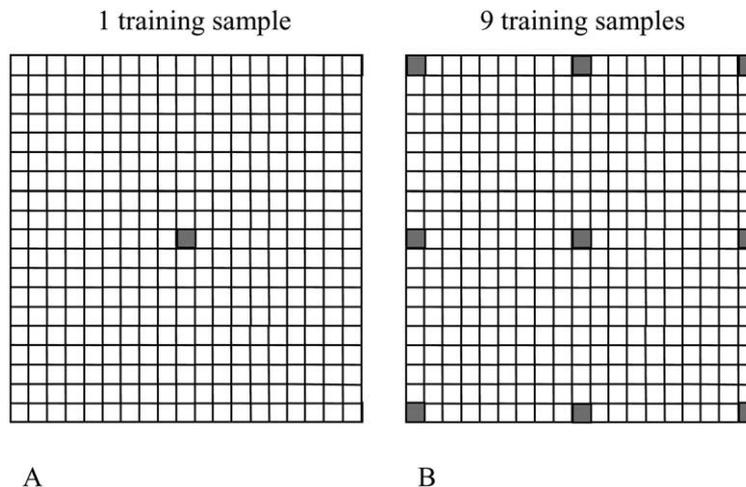


Figure 3. Illustration of two different training samples (pairs of an input subregion and the corresponding teaching single pixel) selected from the training region (27×27 pixels in an input image and 19×19 pixels in a teaching image). **(a)** One training sample in each case is used for training a massive training artificial neural network (MTANN). The training region in a teaching image is shown. The location of a teaching pixel is shown as a dark square that corresponds to the center pixel in the corresponding subregion in the input image. **(b)** Nine (3×3) training samples in each case are used for training another MTANN.

3.0 mm, and 4.0 mm in diameter) with different orientations such as horizontal, vertical, and diagonal, and model nodules overlapping with model vessels, as shown in Fig 3A. We created the same-sized model nodules with different contrasts, because solid opacity and GGO of the same size have different contrasts. The background level was -900 HU, which corresponds to the average background level in the lungs.

RESULTS

Figure 2 shows the input images used for training the MTANN and also the output images of the trained MTANN. It is apparent that the nodules are represented by light “fuzzy nodular” distributions in the output images, whereas the vessels are dark and thus “almost removed.” Figure 4 shows (A) the simulated CT image, (B) the output image of the MTANN trained with one training sample (a pair of an input subregion and a teaching pixel) in each case, (C) the output image of the MTANN trained with nine (3×3) training samples, and (D) the output image of the MTANN trained with 361 (19×19) training samples. In the output images of the MTANNs trained with one and nine training samples in Fig 4B and 4C, both model nodules and model vessels are clearly

recognizable, and it is important to note that vessels are not removed. Therefore, it is apparent in Fig 4B and 4C that the MTANNs were not trained properly with only one subregion and also with nine subregions. In the output image of the trained MTANN in Fig 4D, however, the various-sized model nodules with different contrasts are represented by light “nodular” distributions, whereas various-sized model vessels with different orientations are almost dark, and are thus removed. Therefore, it is apparent in Fig 4D that model nodules can be distinguished from model vessels. This result indicates that the MTANN was able to learn from a very small number of training actual cases (10 actual nodules and 10 actual vessel images) to enhance spherelike objects (model nodules) and suppress cylinder-like objects (model vessels), and that the trained MTANN would be robust against a change in scale and rotation. Thus the key to achieving a high generalization ability of the MTANN would be related to scanning over the training cases, producing a large number of training samples.

To investigate the performance for actual nodules and vessels, we applied the trained MTANN to nontraining cases. Figure 5 shows the output images of the trained MTANN, where various-sized actual nodules with different contrasts are represented by light “nodular” distribu-

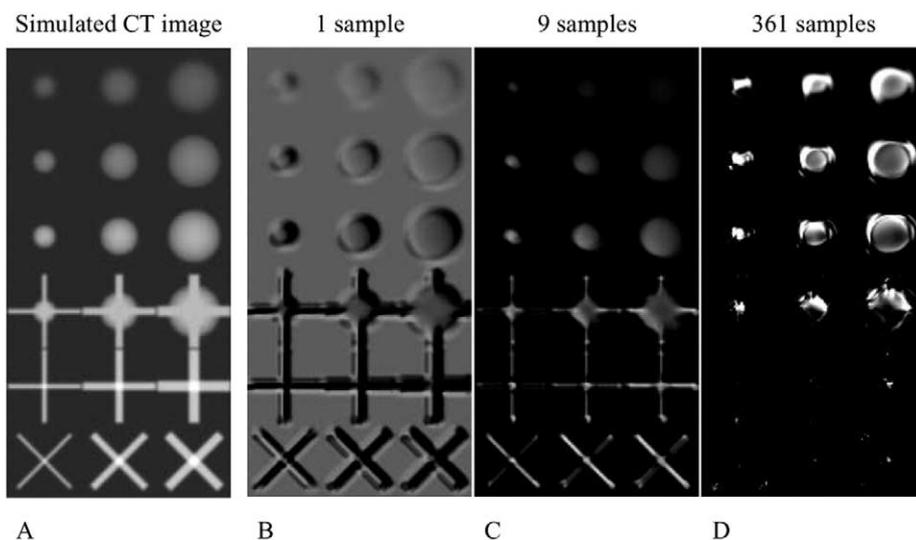


Figure 4. Simulated computed tomography image that contains various-sized model nodules with different contrasts and various-sized model vessels with different orientations, and the corresponding output images of the massive training artificial neural networks (MTANNs) trained with 10 nodules and 10 vessel images. **(a)** Input image for the MTANNs. **(b)** Output image of the MTANN trained with one training sample in each case. **(c)** Output image of the MTANN trained with nine (3×3) training samples. **(d)** Output image of the MTANN trained with 361 (19×19) training samples.

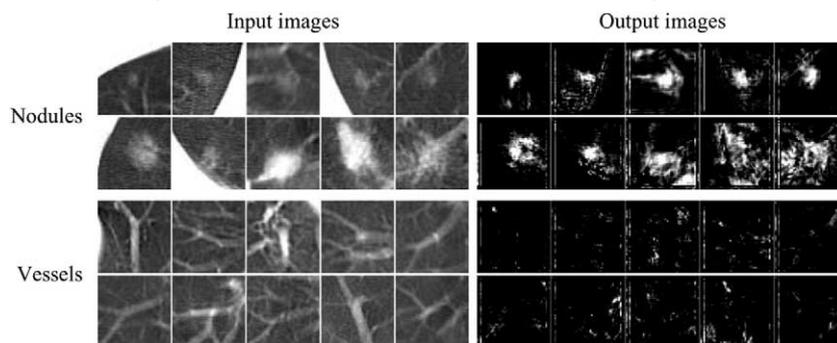


Figure 5. Illustrations of non-training actual nodules and vessels, and the corresponding output images of the massive training artificial neural network trained with 361 subregions.

tions, whereas medium-sized and small actual vessels with different orientations are almost eliminated. The performance of three MTANNs was evaluated by receiver operating characteristic (ROC) analysis (42,43). Figure 6 shows the ROC curve of each MTANN for distinction between nodules and non-nodules by use of nontraining cases of 59 true-positive nodules and 1,726 false positives (non-nodules). The Az values (areas under the ROC curve) (44) for the MTANNs that were trained with 1, 9, and 361 subregions were 0.60, 0.73, and 0.89, respectively. This result indicates that a large number of training samples would be essential in providing a high generalization ability of the MTANN, although the number of training cases was rather small.

DISCUSSION

The results of this study suggest that massive training with a large number of training samples allowed an MTANN to be trained with a very small number of cases (10 nodules and 10 non-nodule images). By dividing a case (image) into a large number of subregions (19×19), the MTANN can be trained not on a case basis, but on a subregion basis. A large number of training subregions (361 for each case) can include various parts of a nodule and various parts of vessels in different sizes and orientations. This would be the reason that the MTANN was robust against the change in scale and rotation, as shown in Fig 4, even when trained with a very small

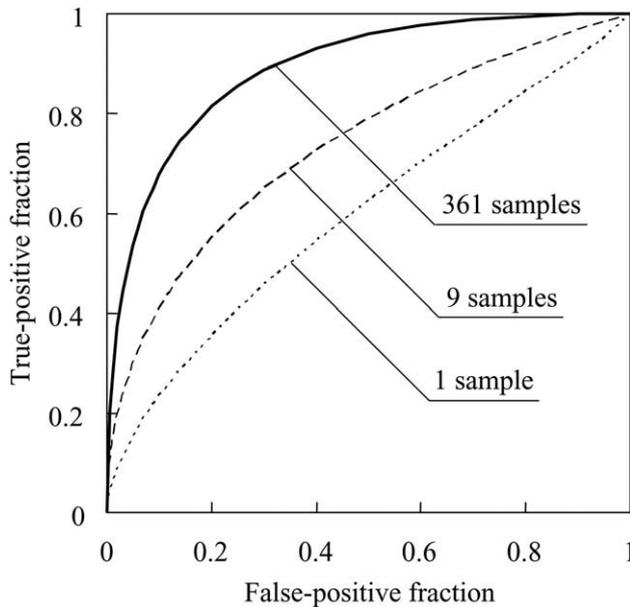


Figure 6. Receiver operating characteristic curves for massive training artificial neural networks (MTANNs) trained with different numbers of subregions for distinction between 59 true-positive nodules and 1,726 false positives (non-nodules). The Az values for the MTANNs trained with 1, 9, and 361 subregions were 0.60, 0.73, and 0.89, respectively.

number of cases. The massive training with a large number of training samples (361 samples \times 20 cases = 7,220 samples) allows an MTANN to avoid the “overfitting” problem (27) of ANNs. This problem often occurs when a number of training samples are too small for determining the parameters of an ANN. The required number of training samples is, in general, greater than the number of parameters of an ANN. For example, an ANN with a 9-9-1 structure (9 inputs \times 9 input-hidden-layer weights + 9 hidden-output-layer weights + 10 offsets = 100 parameters) required 400–800 training samples to achieve an adequate performance for non-training cases (25). The massive training with a large number of training samples (7,220), which are not completely independent, however, would contribute to the proper determination of the parameters (81 inputs \times 20 input-hidden-layer weights + 20 hidden-output-layer weights + 21 offsets = 1,661 parameters) of the MTANN, and avoid the overfitting problem. This is probably the reason that the MTANN had a high generalization ability, as shown in the results for the simulated CT images in Fig 4 and also those for non-training cases in Fig 5.

In this study, we investigated how an MTANN can be trained adequately with a small number of cases in the distinction between nodules and vessels. An

MTANN is, however, applicable to the distinction between nodules and other types of normal anatomic structures other than vessels in a CAD scheme. In fact, we have developed a multiple MTANN scheme (multi-MTANN) that consisted of nine MTANNs for removing nine types of false positives (non-nodules) (ie, five different types of vessels ranging from small to medium-sized, large vessels in the hilum, relatively large vessels with some opacities, soft-tissue opacities caused by the partial volume effect, and focal interstitial opacities) (16). Each of the MTANNs in the multi-MTANN was able to be adequately trained with a small number of cases (10 nodules and 90 non-nodules in total). The results in an independent test showed that the multi-MTANN removed 83% of the false positives generated by our initial CAD scheme with the loss of one true positive (ie, a classification sensitivity of 98.3%). With the multi-MTANN, the false-positive rate of our original CAD scheme was improved from 27.4 to 4.8 false positives per scan at an overall sensitivity of 80.3% for the database of 63 LDCT scans containing 71 nodules including 66 primary cancers. For testing the versatility of a multi-MTANN, we applied it to a different CAD scheme based on a difference-image technique (30). Our CAD scheme incorporating the multi-MTANN achieved a sensitivity of 83.5% with 5.8 false positives per scan for the database of 106 LDCT scans containing 109 cancers. Moreover, we have investigated the feasibility of a multi-MTANN for thin-slice CT, and developed a multi-MTANN for thin-slice CT images (slice thickness was 1.25 or 2.5 mm) acquired with a MDCT system (45). Our scheme incorporating the multi-MTANN trained with 10 nodules and 80 non-nodules achieved a sensitivity of 95.2% with 6.6 false positives per scan for the database of 32 MDCT scans containing 62 nodules. In addition, we have investigated the application of a multi-MTANN to the classification of abnormal lesions, and developed a multi-MTANN for the distinction between malignant and benign nodules in LDCT (46). The multi-MTANN trained with 10 malignant nodules and 60 benign nodules achieved an Az value of 0.882 for the database consisting of 76 primary cancers and 413 benign nodules. The multi-MTANN correctly identified 100% of malignant nodules as malignant, whereas 48% of benign nodules were identified correctly as benign. Thus MTANNs are versatile and can be adequately trained with a small number of cases in different tasks in CAD schemes.

ACKNOWLEDGMENTS

The authors are grateful to F. Li, MD, and H. MacMahon, MD, for their clinical advice; to H. Arimura, PhD, Q. Li, PhD, and J. Shiraishi, PhD, for their valuable suggestions; to C. E. Metz, PhD, for the use of the ROC-KIT program; and to Ms. E. F. Lanzl for improving the manuscript. K. Doi is a shareholder in R2 Technology, Inc., Sunnyvale, CA.

REFERENCES

- Jemal A, Murray T, Ward E, et al. Cancer statistics. 2004. *CA Cancer J Clin* 2004; 54:8–29.
- Sone S, Takashima S, Li F, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet* 1998; 351:1242–1245.
- Rusinek H, Naidich DP, McGuinness G, et al. Pulmonary nodule detection: low-dose versus conventional CT. *Radiology* 1998; 209:243–249.
- Henschke CI, McCauley DI, Yankelevitz DF, et al. Early lung cancer action project: overall design and findings from baseline screening. *Lancet* 1999; 354:99–105.
- Swensen SJ, Jett JR, Hartman TE, et al. Lung cancer screening with CT: Mayo Clinic experience. *Radiology* 2003; 226:756–761.
- Henschke CI, Yankelevitz DF, Naidich DP, et al. CT screening for lung cancer: suspiciousness of nodules according to size on baseline scans. *Radiology* 2004; 231:164–168.
- Gurney JW. Missed lung cancer at CT: imaging findings in nine patients. *Radiology* 1996; 199:117–122.
- Li F, Sone S, Abe H, et al. Lung cancer missed at low-dose helical CT screening in a general population: comparison of clinical, histopathologic, and imaging findings. *Radiology* 2002; 225:673–683.
- Giger ML, Bae KT, MacMahon H. Computerized detection of pulmonary nodules in CT images. *Invest Radiol* 1994; 29:459–465.
- Armato SG, Giger ML, MacMahon H. Automated detection of lung nodules in CT scans: preliminary results. *Med Phys* 2001; 28:1552–1561.
- Ko JP, Betke M. Chest CT: automated nodule detection and assessment of change over time: preliminary experience. *Radiology* 2001; 218:267–273.
- Lee Y, Hara T, Fujita H, et al. Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. *IEEE Trans Med Imag* 2001; 20:595–604.
- Brown MS, McNitt-Gary MF, Goldin JG, et al. Patient specific models for lung nodule detection and surveillance in CT images. *IEEE Trans Med Imag* 2001; 20:1242–1250.
- Armato SG, Li F, Giger ML, et al. Lung cancer: performance of automated lung nodule detection applied to cancers missed in a CT screening program. *Radiology* 2002; 225:685–692.
- Gurcan MN, Sahiner B, Petrick N, et al. Lung nodule detection on thoracic computed tomography images: preliminary evaluation of a computer-aided diagnosis system. *Med Phys* 2002; 29:2552–2558.
- Suzuki K, Armato 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.
- Nagel RH, Nishikawa RM, Doi K. Analysis of methods for reducing false positives in the automated detection of clustered microcalcifications in mammograms. *Med Phys* 1998; 25:1502–1506.
- 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.
- Nakamura K, Yoshida H, Engelmann R, et al. Computerized analysis of the likelihood of malignancy in solitary pulmonary nodules by use of artificial neural networks. *Radiology* 2000; 214:823–830.
- Jiang Y, Nishikawa RM, Wolverson DE, et al. Malignant and benign clustered microcalcifications: automated feature analysis and classification. *Radiology* 1996; 198:671–678.
- Chan HP, Sahiner B, Petrick N, et al. Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network. *Phys Med Biol* 1997; 42:549–567.
- Kallergi M. Computer-aided diagnosis of mammographic microcalcification clusters. *Med Phys* 2004; 31:314–326.
- Ashizawa K, Ishida T, MacMahon H, et al. Artificial neural networks in chest radiographs: application to differential diagnosis of interstitial lung disease. *Acad Radiol* 1999; 6:2–9.
- Ishida T, Katsuragawa S, Ashizawa K, et al. Application of artificial neural networks for quantitative analysis of image data in chest radiographs for detection of interstitial lung disease. *J Dig Imag* 1998; 11:182–192.
- Chan HP, Sahiner B, Wagner RF, et al. Classifier design for computer-aided diagnosis: effects of finite sample size on the mean performance of classical and neural network classifiers. *Med Phys* 1999; 26:2654–2668.
- Sahiner B, Chan HP, Petrick N, et al. Feature selection and classifier performance in computer-aided diagnosis: the effect of finite sample size. *Med Phys* 2000; 27:1509–1522.
- Bishop CM. *Neural networks for pattern recognition*. New York: Oxford University Press, 1995.
- Suzuki K, Shiraishi J, Abe H, et al. False-positive reduction in computer-aided diagnostic scheme for detecting nodules in chest radiographs by means of massive training artificial neural network. *Acad Radiol* 2005; 12:191–201.
- Suzuki K, Armato 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 Med Imag (SPIE MI)* 2003; 5032:1355–1366.
- Arimura H, Katsuragawa S, Suzuki K, et al. Computerized scheme for automated detection of lung nodules in low-dose CT images for lung cancer screening. *Acad Radiol* 2004; 11:617–629.
- 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.
- Suzuki K, Horiba I, Ikegaya K, et al. Recognition of coronary arterial stenosis using neural network on DSA system. *Syst Computers Japan* 1995; 26:66–74.
- 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.
- 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 Information Sys* 2002; E85-D:1710–1718.
- 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.
- Rumelhart DE, Hinton GE, Williams RJ. Learning representations of back-propagation errors. *Nature* 1986; 323:533–536.
- 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; 318–362.
- Funahashi K. On the approximate realization of continuous mappings by neural networks. *Neural Networks* 1989; 2:183–192.
- Barron AR. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Trans Information Theory* 1993; 39:930–945.
- Suzuki K, Horiba I, Sugie N. A simple neural network pruning algorithm with application to filter synthesis. *Neural Proc Lett* 2001; 13:43–53.
- Suzuki K. Determining the receptive field of a neural filter. *J Neural Eng* 2004; 1:228–237.
- Metz CE. ROC methodology in radiologic imaging. *Invest Radiol* 1986; 21:720–733.

43. 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.
44. 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.
45. Suzuki K, Li Q, Li F, et al. Distinction between nodules and false positives in CAD scheme for lung nodule detection on multi-detector CT images by means of massive training artificial neural networks. *Radiology* 2004; 233:290-291.
46. Suzuki K, Li F, Sone S, Doi K. Computer-aided diagnostic scheme for distinction between benign and malignant nodules in thoracic low-dose CT by use of massive training artificial neural network. *IEEE Trans Med Imag* 2005; 24.

APPENDIX

The inputs of the MTANN are the pixel values in a subregion R_S on an input CT image $I(x,y)$. The output of the MTANN is an estimate $O(x,y)$ for a teaching value, represented by

$$O(x, y) = NN\{I(x - i, y - j) | i, j \in R_S\} \quad (1)$$

where x and y are the indices of coordinates, and $NN\{\bullet\}$ is the output of a linear-output multilayer ANN. The MTANN is trained with a large number of subregions selected from input CT images together with the corre-

sponding teaching pixels in teaching images $T(x,y)$. The error to be minimized by training of the MTANN is defined by

$$E = \frac{1}{P} \sum_{x,y \in R_T} \{T(x, y) - O(x, y)\}^2 \quad (2)$$

where P is the number of teaching pixels in the training region R_T . The output image is obtained by scanning of an input CT image with the trained MTANN. For distinction between a nodule and a non-nodule, a scoring method is applied to the output images of the trained MTANN. The score is defined by use of the output image and a 2D Gaussian weighting function, as described here:

$$S = \sum_{x,y \in R_E} f_G(\sigma; x, y) \times O(x, y) \quad (3)$$

where S is the output score for a given nodule candidate from the MTANN, R_E is the region for evaluation, $O(x,y)$ is the output image of the MTANN where its center corresponds to the center of R_E , and $f_G(\sigma; x, y)$ is a 2D Gaussian weighting function with standard deviation σ , with its center corresponding to the center of R_E .