Differentiation Between Nodules and End-On Vessels Using A Convolution Neural Network Architecture

Jyh-Shyan Lin, Akira Hasegawa, Matthew T. Freedman, and Seong K. Mun

In recent years, many computer-aided diagnosis schemes have been proposed to assist radiologists in detecting lung nodules. The research efforts have been aimed at increasing the sensitivity while decreasing the false-positive detections on digital chest radiographs. Among the problems of reducing the number of false positives, the differentiation between nodules and end-on vessels is one of the most challenging tasks performed by computer. Most investigators have used a conventional two-stage pattern recognition approach, ie, feature extraction followed by feature classification. The performance of this approach depends totally on good feature definition in the feature extraction stage. Unfortunately, suitable feature definition and corresponding extraction implementation algorithms proved to be very difficult to define and specify. A convolution neural network (CNN) architecture, trained by direct connection to the raw image is proposed to tackle the problem. The CNN, which uses locally responsive activation function, is directly and locally connected to the raw image. The performance of the CNN is evaluated in comparison to an expert radiologist. We used the receiver operating characteristics (ROC) method with area under the curve (A_z) as the performance index to evaluate all the simulation results. The CNN showed superior performance (A_z = 0.99) to the radiologist's (A_z = 0.83). The CNN approach can potentially be applied to other applications, such as the differentiation of film defects and microcalcifications in mammography, in which the image features are difficult to define or not known a priori.

KEY WORDS: Convolution neural network (CNN), computer-aided diagnosis (CADx), feature map, and receiver operating characteristics (ROC).

RECENTLY, various computer-aided diagnosis (CADx) systems,[1-8] which use both digital image-processing techniques and artificial neural networks, have been developed to automatically detect lung nodules on digital chest radiographs. Typically, CADx schemes successively perform two diagnostic functions: (1) location of suspected nodule areas, also known as a "prescan process," on the digitized chest radiograph and (2) differentiation of "true" nodules from "false" nodules. In recent literature, many effective digital processing algorithms have been developed to locate suspected nodules.[2-4,6] It has been determined that too many false-positive classifications per chest radiograph are made. These false-positives include rib crossings, rib-vessel crossings, vessel clusters, and end-on vessels.[5,9] The difficulty of identifying end-on vessels is among the problems of reducing false-positive detections.

The shadow of an end-on vessel is formed by the incident x-ray beam parallel to the vessel. The shape of its shadow may be characterized by a small, round pattern. Generally, these small, round spots contain high contrast with a high degree of circularity. In distinguishing these round nonnodule objects from nodules, the size versus contrast information is one of the most important criteria used by radiologists. End-on vessels tend to have higher contrast than nodules of the same size. To extract the image features of size and radius is not an easy task because a nodule and an end-on vessel are usually surrounded or overlapped with chest structures such as ribs, clavicles, and vessels. Accordingly, it is difficult to specify the contrast and size of the nodule or end-on vessel in such a noisy lung structure environment. Several morphology-based algorithms[3-5] have been proposed to extract specific features, such as circularity, size, and contrast. Matsumoto et al[6] used a region-growing method to find the transition points where the effective radius or circularity of the grown region increases abruptly. This transition indicates that the grown region has merged with the surrounding background adjacent to the suspected nodule. The contrast of the suspected nodule is defined as the gray-level interval between the average high pixel value of the region and the average pixel value of the transition points. Lin et al[10,11] proposed an

From the ISIS Center, Radiology Department, Georgetown University Medical Center, Washington, DC.

Supported in part by US Army Grant No. DAMD17-93-J-3007. The content of this manuscript does not necessarily reflect the position or the policy of the government.

Address reprint requests to Jyh-Shyan Lin, PhD, ISIS Center, Radiology Department, Georgetown University Medical Center, 2115 Wisconsin Ave, NW, Suite 603, Washington, DC 20007.

0897-1889/95/0803-0004$3.00/0

algorithmic approach to extract the contrast and radius information from the raw image of nodules and end-on vessels. Based on the defined features of contrast and radius, he investigated the feature distributions of the radius versus contrast/radius. The method achieved an area under the curve (A_z) value of 0.94. Though the method showed promise in the task, too many parameters, such as the number of high and low pixels and the minimum radius, needed to be adjusted.

Actually, the main concern of using algorithmic approach lies in the method of defining and quantifying the features that will discriminate nodules and end-on vessels. When nodules or end-on vessels are partially or fully superimposed on ribs, bronchi, and other anatomic chest structures, the size and contrast can not be clearly defined. Therefore, the algorithmic methodologies based on morphology alone and on a priori knowledge of known features have been problematic. In this study, to extract the suitable features from the two-dimensional (2D) images of chest radiographs, a convolution neural network (CNN) architecture,[12-15] trained by direct connection to the raw image, has been configured and tested to solve the problem of distinguishing nodules and end-on vessels. Performances of the CNN as well as that of an expert radiologist are evaluated.

MATERIALS AND METHODS

Database of Image Blocks

The generation of image blocks of nodules and end-on vessels involved three stages: acquisition of digitized chest radiographs, extraction of image blocks, and preprocessing.

Acquisition of digitized chest radiographs. For this study, we used posterior-anterior chest radiographs that were selected mainly from routine cases at Georgetown University Medical Center (Washington, DC). The chest radiographs (14 × 17 inches, actual size) were digitized to 2,048 × 2,500 × 10 bits by using a laser film scanner (model KDFR-S; Konika, Tokyo, Japan). For computational simplicity, the digitized chest films were averaged to 512 × 625 × 10 bits such that one pixel represents 0.7 × 0.7 mm, actual size.

Extraction of image blocks. Image blocks (each 32 × 32 pixels) of nodules and end-on vessels on the digitized chest images were extracted manually from the cancerous and normal chest images, respectively. The cancerous and normal chest images have been confirmed by computed tomography, biopsy, or follow-up films. Note that it is very difficult even for radiologists to identify a nodule that is close to the hilum area, the central portion of the lung where all the blood vessels enter the lung. To eliminate the possibility of mistakenly selecting a nodule as an end-on vessel, all image blocks were selected from the areas away from the hilum. The block size of 32 × 32 pixels (about 22 × 22 mm) is sufficient to encompass the various sizes of small nodules (diameter range, 3 to 15 mm) and end-on vessels in which we were interested. A minimal size of 3 mm is chosen because most radiographically detectable nodules are larger than 3 mm.[16]

Preprocessing. The quality of chest radiographs varies because of the noise introduced during film development, nonuniform thickness of chest walls, and the transmission properties of tissues, vessels, and ribs of different patients. In general, the perihilar area in a chest radiograph is more radiodense than the peripheral area because of differences in the amount of tissue traversed by the x-ray beam in the two regions. The variation in optical density observed in the lung field is caused by the gross anatomy of the lung and nonuniform chest wall (background trend). Because the image blocks were obtained from different regions of the lung, each image block contains nonuniform lung background. It is important to correct the background trend so that the overall background becomes uniform. Katsuragawa et al[17] used a 2D surface-fitting technique in which the lung background trend was estimated and corrected by fitting a second-order polynomial surface to the gradual change in the background density distribution in a selected region of a suspected nodule area. We applied the same method to rectify the local background discrepancy in each image block. There is no normalization or scaling process performed on the image blocks. The preprocessed image is directly connected to the input of the CNN. The potential use of CNN for extraction as well as classification of image features is investigated.

CNN Architecture

The CNN architecture shown in Fig 1 is investigated for its ability to distinguish nodules from end-on vessels. The network has one input layer, one hidden layer, and one

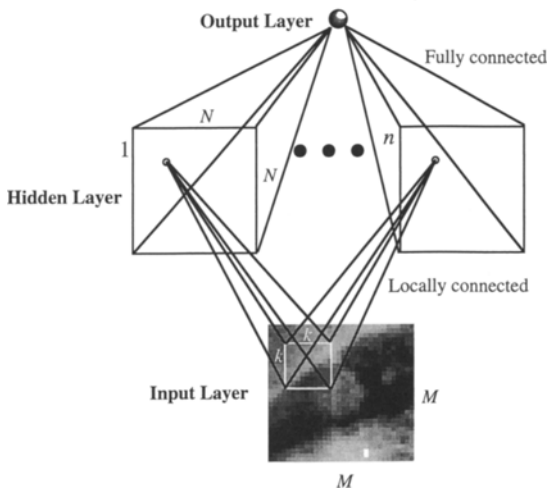

Fig 1. The CNN architecture with one output neuron is shown.

output neuron in the output layer. The input layer consists of M^2 neurons that correspond to the $M \times M$ pixel preprocessed input image.

The hidden layer is composed of n groups of $N \times N$ neurons arranged as n independent $N \times N$ feature maps, where N is equal to $M - k + 1$ and the $k \times k$ area is called the receptive field. Each hidden neuron takes input on a $k \times k$ neighborhood on the input image block. For neurons in the same feature map that are one neuron apart, their receptive fields in the input layer are one pixel apart. Moreover, each neuron in the same feature map is constrained to have the same set of k^2 weights and perform the same operation on the corresponding parts of the input image. Thus, the total effect of the operation can be expressed as a 2D discrete convolution with the $k \times k$ convolution kernel (ie, the receptive field). Namely, the feature map is the output of the input image convolution with the kernel. All neurons in other feature maps share other k^2 weights in the same way. Each hidden neuron y_j generates its output through a local responsive activation function which is given by

$$y_j(\underline{w}, \underline{x}, a_j) = A_j(\underline{w}, \underline{x}, a_j)(1 - A_j(\underline{w}, \underline{x}, a_j)), \quad (1)$$

where $A_j(\underline{w},\underline{x},a_j)$ is the sigmoid activation given by

$$A_j(\underline{w}, \underline{x}, a_j) = \frac{1}{1 + \exp\left\{-\left[\sum_{i=1}^{k^2} (w_{ji}x_i) + a_j\right]\right\}}, \quad (2)$$

where w_{ji} is the weigth between hidden neuron j and pixel i of the input image block, x_i is the gray value of input pixel j, and a_j is the bias of the hidden neuron j. Note that $x_1, \ldots, x_k{}^2$ are the pixel areas on the input image block that are connected to the neuron j. The input-to-output activations of the locally responsive and sigmoid functions (equations 1 and 2, respectively) are shown in Fig 2. Note that equations 1 and 2 are first order differentiable. Each locally responsive neuron becomes activated only for inputs in some small region of the input space. When the input has value greater than 10 or smaller than -10, the output activation is close to 0. The minimum and maximum activation are 0 and 1, respectively.

The output neuron is fully connected to the hidden layer.

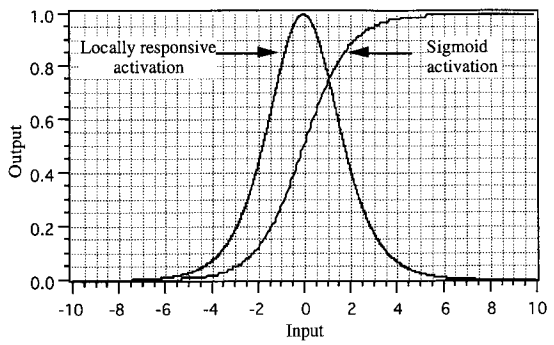

Fig 2. The sigmoid and locally responsive activation functions are shown.

The activation z_o of the output neuron is given by

$$z_o(\underline{w}, \underline{y}, g_o) = \frac{1}{1 + \exp\left\{-\left[\sum_{i=1}^{nN^2} (w_{oi}y_i) + g_o\right]\right\}}, \quad (3)$$

where w_{oi} is the weight between the output neuron and neuron i in the hidden layer, nN^2 is the total number of neurons in the hidden layer, and g_o is the bias of the output neuron.

In summary, the network consists of $1 + M^2 + n \times N^2$ neurons (including the input neurons) and $(k^2 + 2) \times n \times N^2 + 1$ links (including the bias links) in which $n \times [2 \times N^2 + k^2] + 1$ are independent links (including the independent bias links).

RESULTS AND DISCUSSION

In the following experiments, the CNN architecture (see Fig 1) has this setup: input image blocks of 32×32 pixels, ie, $M = 32$; kernel size of 5×5, ie, $k = 5$; 10 feature maps, ie, $n = 10$; 1 output neuron; and an additional bias input 1 to each neuron in the hidden and output layers.

Accordingly, each feature map in the hidden layer has 28×28 neurons. The total number of neurons is 8,865, the total number of links is 211,681, and the total number of independent links is 15,931. The number of neurons and links include, respectively, the input neurons and the bias links that have an input value of 1.

Network Training

All the network weights, including the bias weights, are updated using the back-propagation learning rule.[19] Initially, the network weights are preset with random numbers using a uniform distribution between $-1/F_i$ and $+1/F_i$ where F_i is the number of inputs (ie, fan-ins) from the previous layer that are connected to neuron i. The network weights (including bias weights) are updated after each presentation of a single training pattern by using the so-called stochastic gradient procedure.[20,21] The network weights are saved at every tenth iteration, where one iteration represents the completion of a learning sequence (ie, a feed-forward pass followed by an error back-propagation pass, through the entire set of training patterns). Sum of squared error (SSE)[19] over all the training patterns was used as the objective function that was to be minimized through network training. The desired output (teaching signal) at the output neuron has a highest activation value 1 and lowest value 0 when,

respectively, a nodule and an end-on vessel are presented at the input of the network. The initial learning rate is preset to 0.001 and is gradually decreased to 0.0001, 0.00001, and 0.000001. The training procedure is terminated whenever one of the following three conditions is satisfied: (1) the SSE reaches a value below 0.1, (2) there is less than 1% change of SSE for 100 consecutive iterations, or (3) there are more than 1,000 training iterations.

The training set consists of image blocks of 40 nodules (Fig 3A) and 53 end-on vessels (Fig 3B). The 40 nodules were selected from the image blocks extracted from four chest images containing multiple nodules. The 53 end-on vessels were randomly selected from the image blocks extracted from 10 normal chest images that have been confirmed to be "definitely no nodule" cases. To improve the situation where only a small and limited number of training patterns are available, patterns of different orientations are generated from a single training example and are used in network training. For each training pattern p, three rotated versions of p are generated by rotating p at 90, 180, and 270 degrees. In addition, four more patterns are obtained by left-right flipping of pattern p followed by rotations of 0, 90, 180, and 270 degrees. As a result, the number of training patterns is virtually increased eightfold. These patterns have the same desired output (either 1 or 0) during network training. Accordingly, the training set has virtually increased to 320 nodules and 424 end-on vessels. During training, all 744 training patterns of different orientations are presented in random order to the input of the CNN.

Learning Curve and Testing on the Training Set

The learning curve of the CNN is shown in Fig 4. The initial SSE is 99.1, which decreases to 1.92 after 500 iterations. Because there is no significant improvement of SSE (ie, less than a

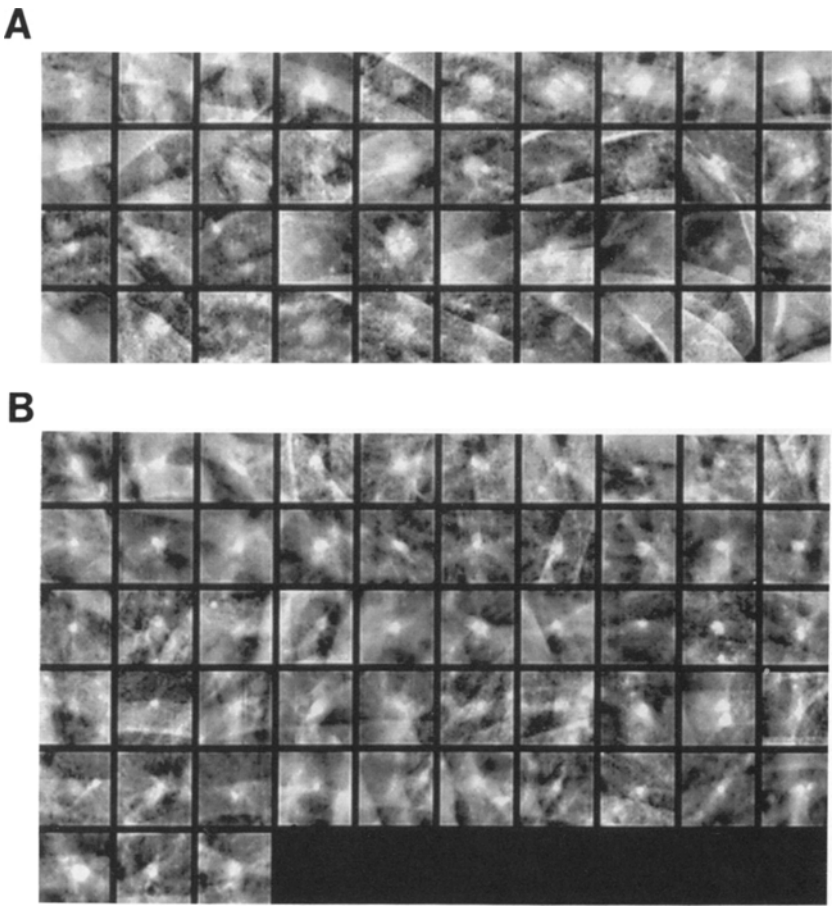

Fig 3. Image blocks for training are shown. (A) Nodules, (B) end-on vessels.

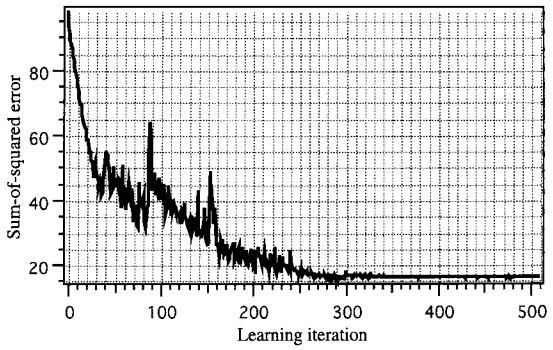

Fig 4. The learning curve of the CNN is shown.

1% change of SSE for 100 consecutive iterations), the training process is stopped. The total training time on a workstation (model 3000; Digital Equipment Corp, Maynard, MA) is approximately 3 hours.

The testing result of the CNN on the training set is shown in Fig 5. The white circles and black dots in Fig 5 indicate the desired outputs and the actual mean output values of the trained CNN, respectively, for the 93 training patterns (40 nodules and 53 end-on vessels). The white circles indicate the desired output values of ones for the first 40 image blocks of nodules and zeros for the following 53 image blocks of end-on vessels. The two well-separated clusters of black dots (ie, the learning results) in Fig 5 show that the trained CNN has achieved a 100% recognition rate on the training set.

Network Testing

The testing set consisting of 66 nodules (Fig 6A) and 46 end-on vessels (Fig 6B) was used to test the generalization performance[19,21] of the trained CNN. The 66 nodules were selected from the image blocks extracted from 17 chest images containing single or multiple nodules. The 46 end-on vessels were randomly selected from the image blocks extracted from nine normal chest images which have been confirmed to be "definitely no nodule" cases.

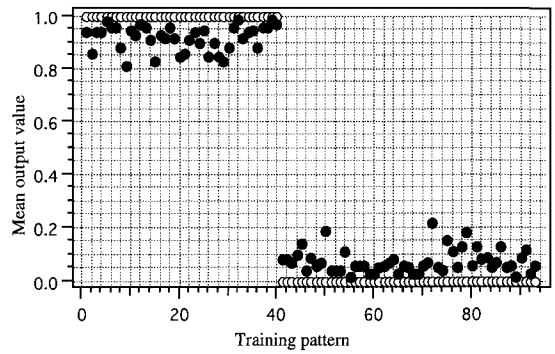

Fig 5. The testing results of the trained CNN on the training set is shown. (○), Desired outputs; (●), actual mean output values of the trained CNN.

Generalization Curve and Testing on the Test Set

For each test pattern of Fig 6, A and B, we obtain eight rotated versions by using the same method as in training the CNN. Then, a mean CNN output value is calculated as the average of the eight output values generated by the CNN for each rotated testing pattern. The mean CNN output values for the testing patterns were analyzed by the ROC method. To monitor the CNN's performance and avoid network over learning[20-22] during the training process, we have saved the learned weights and computed the A_z value at every tenth training iteration. Figure 7 shows the A_z performance of the CNN at every tenth learning iteration. The CNN achieves the consistent A_z value of 0.99 after 250 learning iterations.

The distribution of the mean output value for each testing pattern generated by the CNN is shown in Fig 8, in which white and black dots correspond to end-on vessels and nodules, respectively. The corresponding ROC curve with an A_z value of 0.99 is shown in Fig 9.

Analysis of False Negatives and False Positives

In Fig 8, when we use 0.57 as the decision boundary, no false positives are found. However, two false negatives are generated. The original image and corresponding mean CNN output value of the two false negatives are shown in Fig 10A. Image a in Fig 10A shows a large and radiodense nodule. Image b in Fig 10A shows a nodule partially overlapped by lung structure and resulting in a high-contrast object at the center of the image. If we use a mean output value of 0.50 as the decision boundary (Fig 8), then no false negatives are found. However, two false positives (Fig 10B) are generated. These images contain end-on vessels that are surrounded by lung structure and have contrast similar to a nodule.

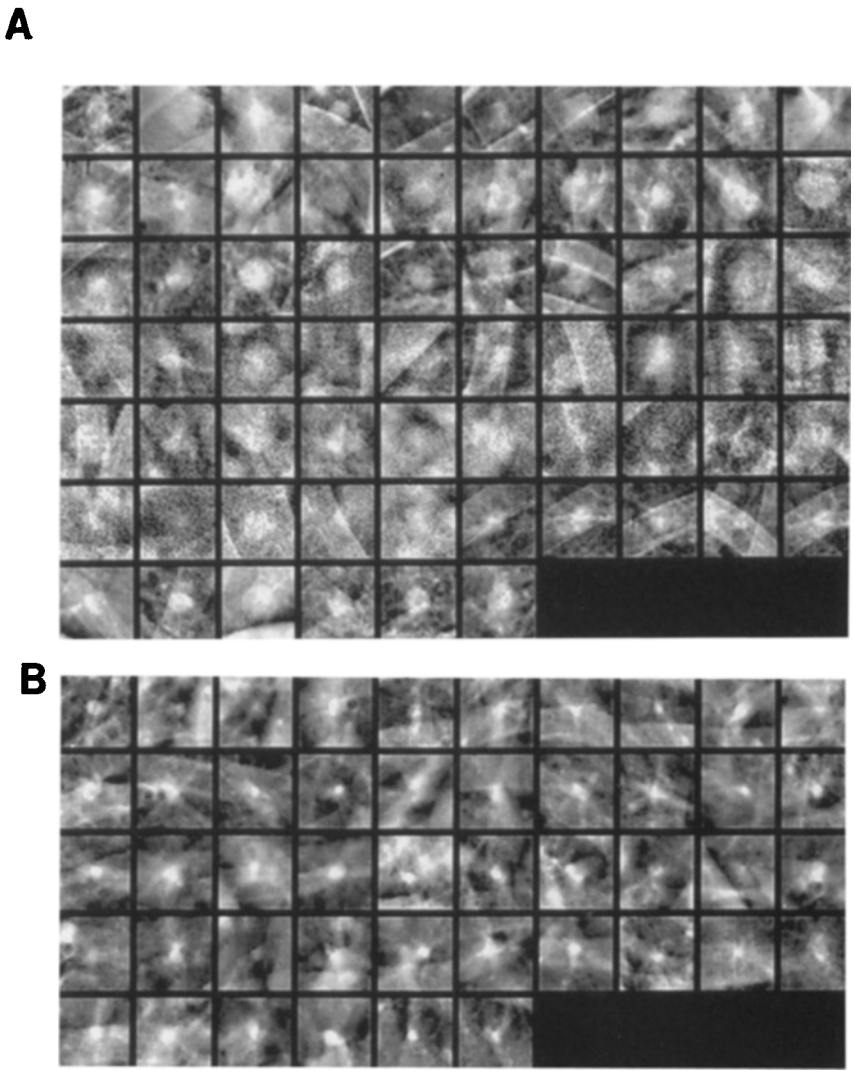

Fig 6. Image blocks for testing are shown. (A) Nodules, (B) end-on vessels.

Analysis of Feature Maps

We analyze the images of the 10 feature maps as shown in Fig 11, A and B, when a nodule (the fourth image from the left of the first row in Fig 6A) and an end-on vessel (the third image from the left of the first row in Fig 6B) are the trained CNN's inputs, respectively. The input image of Fig 11A contains a nodule that lies between two ribs at the lower left and upper right corners. The nodule is partially overlapped with the rib at the lower left corner. Note also that there is an end-on vessel at the upper left corner of the image. The input image of Fig 11B contains an end-on vessel that has a vessel line leading up to it from the bottom left corner. The second and third rows in Fig 11, A and B, show the images of the 10 feature maps before and after the process of locally responsive activation function (Fig 2), respectively. The CNN generated mean output values of 0.59 and 0.43 when the inputs are the nodule and the end-on vessel, respec-

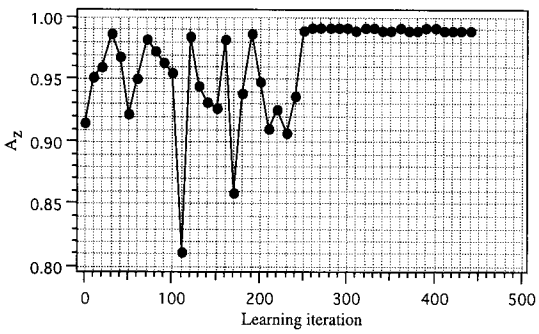

Fig 7. Testing results of the CNN at every tenth learning iteration are shown.

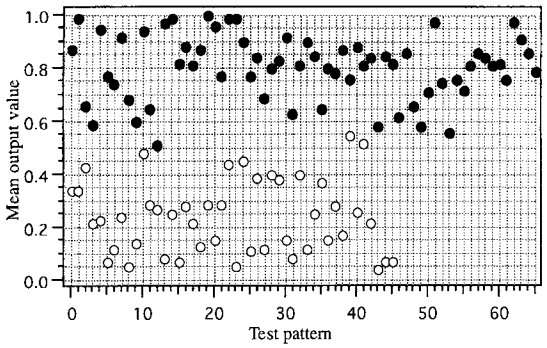

Fig 8. Distribution of the mean output value for each testing pattern generated by the trained CNN is shown. (●), 66 nodules; (○), 46 end-on vessels.

tively. It is interesting to see that the CNN generates a low mean output value for the input image (Fig 11A), which contains not only a nodule but also an end-on vessel. Namely, the CNN is uncertain about whether the image should be diagnosed as a nodule or an end-on vessel case. Note that we have similar types of training patterns (for example, the first and third images in the third row of Fig 3A) associated with desired output values of 1 for the supervised learning.

From the images of the 10 feature maps, we can see that the CNN performed various image processing techniques on the input image. For example, feature maps i and j in Fig 11A show that the CNN detected the rib edges while suppressing the round object signals. The feature map h in Fig 11B shows the result of enhancement of the vessel line leading up to the end-on vessel. The two feature maps indexed as a in Fig 11, A and B, show two smoothed (ie, low-passed filtered) image results. The feature maps f and g in Fig 11, A and B, show the effects of inverted gray level images. Some images (for example, the feature maps b, c, d, and e in Fig 11, A and B), show certain processed results; however, the appropriate names for those learned processing methods are not yet defined. The output images of the activation function (the bottom row in Fig 11, A and B) show a multi-thresholding phenomenon as viewed from left to right.

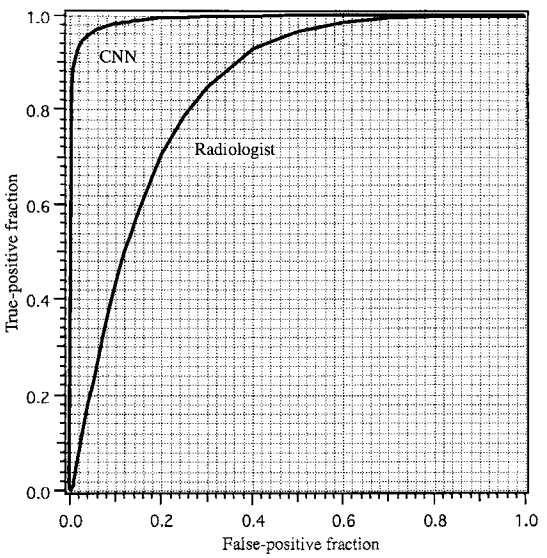

Fig 9. ROC curves with A_z = 0.99 for the trained one-hidden-layer CNN and A_z = 0.83 for the radiologist are shown.

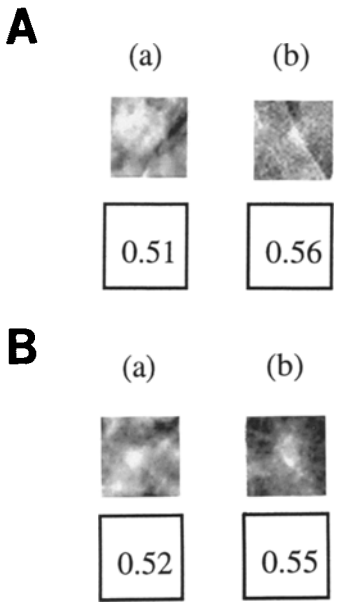

Fig 10. Shown are the original image blocks and mean output values of the (A) two false negatives and (B) two false positives when the mean output values of 0.57 and 0.50 are used as decision boundaries, respectively.

Performance of an Expert Radiologist

An experiment is set up to evaluate an experienced radiologist's performance in identifying end-on vessels. The radiologist has prior knowledge that the test is to differentiate nodules and end-on vessels. The image blocks of end-on vessels (Fig 6A) and nodules (Fig 6B) are randomly mixed and shown on the monitor. The radiologist is allowed to enlarge and adjust the intensity and contrast of the image blocks. An experienced radiologist applies mainly three criteria to identify an end-on vessel: (1) the size

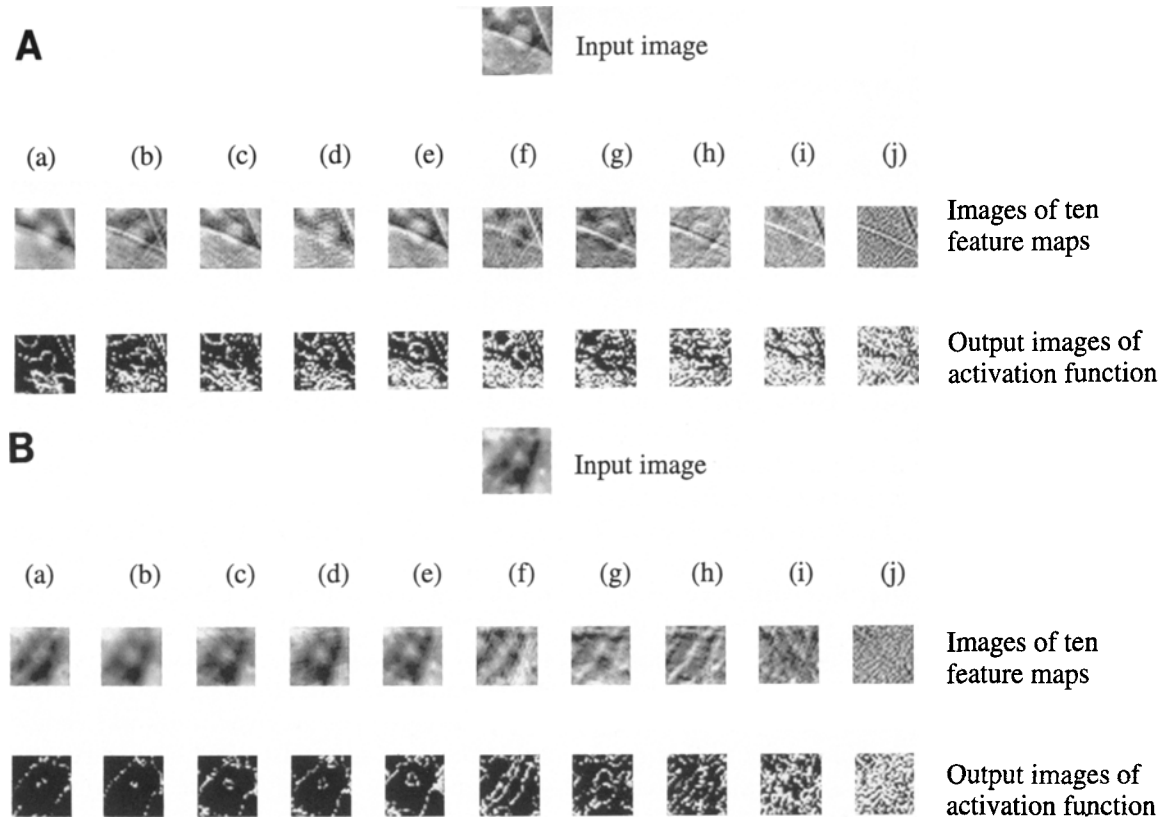

Fig 11. The images of ten feature maps in the hidden layer of the trained one-hidden-layer CNN when (A) a nodule and (B) an end-on vessel are the inputs to the network are shown.

of an end-on vessel is inversely proportional to its distance from the hilum; (2) some end-on vessels have vessel lines leading up to them; and (3) end-on vessels are more radiodense and higher in contrast than similarly sized nodules.

The radiologist does not necessarily apply these criteria in a fixed order. Because no location information is supplied, the radiologist can only apply criteria 2 and 3 to identify end-on vessels. For each image block, the radiologist assigns a real number (or rating) between 0 and 1; smaller and larger numbers represent lower and higher confidence levels that the testing image contains a nodule. The ratings for both testing image blocks are analyzed by the ROC method. Figure 12 shows the rating assigned by the radiologist for each testing pattern, in which white circles and cross bars correspond to nodules and end-on vessels, respectively. The corresponding ROC results show an A_z value of 0.83 for the radiologist (Fig 9). From Fig 12, the radiologist's ratings for the two false negatives identified by the CNN (Fig 10A) are 1 and 0.4 for images a and b, respectively. The radiologist has successfully identified the large and high radiodense object (image a) as a true nodule. The radiologist's ratings for the two false-positives as shown in images a and b in Fig 10B are 0.50 and 0.7, respectively. In other words,

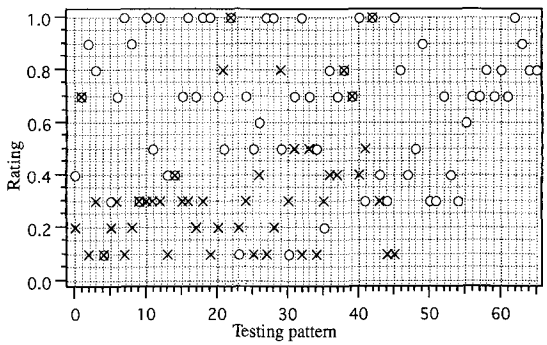

Fig 12. Distribution of the rating for each testing pattern given by the radiologist is shown. (○), nodules; (×), end-on vessels.

the radiologist is not sure about the diagnosis of image a. Moreover, the rating of 0.7 indicates that image b has characteristics that are visually similar to a nodule.

CONCLUSIONS

From the simulation results of differentiation between nodules and end-on vessels, the CNN has shown performance ($A_z = 0.99$) superior to the radiologist's ($A_z = 0.83$). The neural-network approach outperformed the methods based on the algorithmic design in extraction of known or defined new features. Based on the same testing data base, the CNN outperforms the algorithmic approach developed by Lin et al[10,11] which achieved an A_z value of 0.94. Without the laborious design and definition of feature extraction rules, the CNN was trained to learn the discrimination rules for the task. We have observed several interesting points in this study. Both the CNN and the radiologist directly dealt with the raw image blocks of nodules and end-on vessels. Without the information of size versus location from the hilum, the radiologist used mainly criteria 2 and 3 to differentiate between nodules and end-on vessels. The CNN dealt directly with the image block, rather than prespecified features, and was trained to extract and classify features, such as rib edges, vessel lines, and round object signals (Fig 11, A and B). To improve the CNN's diagnostic performance on testing images that contain both nodule and end-on vessel signals, more training patterns will be needed. Furthermore, incorporation of the radiologist's expertise, such as criterion 1, into the CNN will be useful in improving the network performance (for example, to eliminate the false-negative case of Fig 10A).

The success of the CNN method in this simulation indicates its potential in other applications, such as differentiation between film defects and microcalcifications in mammograms and classification of highly correlated and nonorthogonal image patterns. Future research will focus on the optimization of the number of feature maps and the size of the receptive field, on the analysis of trained convolution kernels, and on the minimization of the network architecture through weight elimination and pruning of least significant neurons or weights. The overfitting problem[20-23] of oversized networks should be further investigated to find a minimal network architecture with maximum generalization performance.

ACKNOWLEDGMENT

We thank Susan Kirby for her editorial assistance.

REFERENCES

1. Garg S, Floyd CE: Neural network localization of pulmonary nodules on digital chest radiographs. Radiology 185:157, 1992

2. Giger ML, 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

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

4. Giger ML, Ahn N, 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

5. Matsumoto T, Yoshimura H, Doi K, et al: Image feature analysis of false-positive diagnosis produced by automated detection of lung nodules. Invest Radiol 8:587-597, 1992

6. Matsumoto T, Yoshimura H, Giger ML: Potential usefulness of computerized nodule detection on screening programs for lung cancer. Invest Radiol 27:471-475, 1992

7. Wu Y, Doi K, Asada N, et al: Simulation studies of diagnostic decision making using artificial neural networks. Med Phys 17:525, 1990

8. Wu Y, Doi K, Metz CE, et al: Simulation studies of data classification by artificial neural networks: Potential applications in medical imaging and decision making. J Digit Imaging 6:117-125, 1993

9. Lin JS, Ligomenides PA, Freedman MT, et al: Application of artificial neural networks for reduction of false-positive detections on digital chest radiographs. Symposium on Computer Applications in Medical Care, Washington, DC, October 30-November 3, 1993

10. Lin JS, Lo SC, Freedman MT, et al: Application of artificial neural networks for reducing false-positives in lung nodule detection on digital chest radiographs. SPIE Medical Imaging, 1995 (in press)

11. Lin JS: Convolution Neural Network Architecture with Application for Lung Nodule Detection in Digital Chest Radiography. Doctoral dissertation, University of Maryland, College Park, MD, 1994

12. Lo SC, Lin JS, Freedman MT, et al: Application of artifical neural networks to medical image pattern recognition. World Congress on Neural Network. San Diego, CA, June 5-9, 1994

13. Zhang W, Hasegawa A, Itoh K, et al: Image processing of human corneal endothelium based on a learning neural network. Applied Optics 30:4211-4217, 1991

14. Hasegawa A, Lo SB, Freedman MT, et al: Convolution neural network based detect of lung structures. Proc SPIE 2167:654-662, 1994

15. LeCun Y, Boser B, Denker JS, et al: Handwritten digit recognition with a back-propagation network, in Touretsky DS (ed): Advances in Neural Information Processing Systems II. Denver, CO, Morgan Kaufmann, 1990 pp 396-404

16. Freedman MT: Clinical Imaging: An Introduction to the Role of Imaging in Clinical Practice. New York, NY, Churchill Livingstone, 1988

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

18. Giger ML, Doi K, MacMahon H: Computerized detection of lung nodules in digital chest radiographs. Proc SPIE 767:384-386, 1987

19. Rumelhart DE, Hinton GE, Williams RJ: Learning internal representations by error propagation, in Rumelhart DE, McClelland JL (eds): Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol 1. Cambridge, MA, Massachussetts Institute of Technology, 1986, pp 318-362

20. Anderson JA, Rosenfeld E, Neurocomputing: Foundations of Research. Cambridge, MA, Massachussetts Institute of Technology, 1988

21. Dayhoff JE: Neural Network Architecture: An Introduction. New York, NY, Van Norstrand Reinhold, 1990

22. Baum EB, Haussler D: What size net gives valid generalization? Neural Computation 1:151-160, 1989

23. LeCun Y, Denker JS, Solla SA: Optimal brain damage, in Toureszky DS (ed): Advances in Neural Information Processing Systems II. Denver, CO, Morgan Kaufmann, 1990, pp 598-605