

# Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography

Yoshikazu Uchiyama,<sup>a)</sup> Shigehiko Katsuragawa, Hiroyuki Abe, Junji Shiraishi, Feng Li, Qiang Li, Chao-Tong Zhang, Kenji Suzuki, and Kunio Doi

*Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, The University of Chicago, 5841 S. Maryland Ave., Chicago, Illinois 60637*

(Received 16 March 2003; revised 8 June 2003; accepted for publication 12 June 2003; published 22 August 2003)

An automated computerized scheme has been developed for the detection and characterization of diffuse lung diseases on high-resolution computed tomography (HRCT) images. Our database consisted of 315 HRCT images selected from 105 patients, which included normal and abnormal slices related to six different patterns, i.e., ground-glass opacities, reticular and linear opacities, nodular opacities, honeycombing, emphysematous change, and consolidation. The areas that included specific diffuse patterns in 315 HRCT images were marked by three radiologists independently on the CRT monitor in the same manner as they commonly describe in their radiologic reports. The areas with a specific pattern, which three radiologists marked independently and consistently as the same patterns, were used as “gold standard” for specific abnormal opacities in this study. The lungs were first segmented from the background in each slice by use of a morphological filter and a thresholding technique, and then divided into many contiguous regions of interest (ROIs) with a  $32 \times 32$  matrix. Six physical measures which were determined in each ROI included the mean and the standard deviation of the CT value, air density components, nodular components, line components, and multilobular components. Artificial neural networks (ANNs) were employed for distinguishing between seven different patterns which included normals and six patterns associated with diffuse lung disease. The sensitivity of this computerized method for a detection of the six abnormal patterns in each ROI was 99.2% (122/123) for ground-glass opacities, 100% (15/15) for reticular and linear opacities, 88.0% (132/150) for nodular opacities, 100% (98/98) for honeycombing, 95.8% (369/385) for emphysematous change, and 100% (43/43) for consolidation. The specificity in detecting a normal ROI was 88.1% (940/1067). This computerized method may be useful in assisting radiologists in their assessment of diffuse lung disease in HRCT images. © 2003 American Association of Physicists in Medicine. [DOI: 10.1118/1.1597431]

Key words: automated computerized scheme, diffuse lung diseases, high-resolution computed tomography, artificial neural networks

## I. INTRODUCTION

The differential diagnosis of diffuse lung disease is a major subject in high-resolution computed tomography (HRCT). However, it is considered a difficult task for radiologists, partly because of the complexity and variation in diffuse disease patterns on HRCT images, and also because of the subjective terms used for describing diffuse lung diseases. Therefore, our goal is to develop a computer-aided diagnostic (CAD) scheme for diffuse lung diseases on HRCT in order to assist the radiologists' image interpretation as a “second opinion.”

For the computerized detection of interstitial lung diseases on chest radiographs, we have developed three different CAD schemes which were based on the Fourier transform,<sup>1–3</sup> a geometric-pattern feature analysis,<sup>4</sup> and an artificial neural network (ANN) analysis<sup>5</sup> of image data. The results of an observer performance study<sup>6</sup> indicated that radiologists' performance in distinguishing between normal lungs and abnormal lungs with interstitial infiltrates was improved when the computer results were available.

The task subsequent to detection is the differential diagnosis for the identification of interstitial disease among many

possible diseases. We have also developed an ANN scheme for the determination of the likelihood of each of 11 interstitial diseases by using 10 clinical parameters and 16 radiologic findings.<sup>7</sup> Experimental results obtained with receiver operating characteristic (ROC) analysis indicated that the radiologists' performance in differential diagnosis was improved significantly when they used the computer output.<sup>8</sup>

In previous studies for the detection of diffuse lung diseases on HRCT images, Heitmann *et al.*<sup>9</sup> developed a method for the automated detection of ground-glass opacities on 120 HRCT images from 20 patients by using a hybrid network of three single networks with an expert rule. This hybrid network correctly classified 91 (75.8%) of 120 images from 20 patients. Uppaluri *et al.*<sup>10</sup> developed an adaptive multiple-feature method in assessing 22 independent texture features in order to classify six tissue patterns: honeycombing, ground glass, bronchovascular, nodular, emphysemalike, and normals. Delorme *et al.*<sup>11</sup> developed a texture-based pattern recognition method to classify normals, emphysematous changes, ground-glass lesion, intralobular fibrosis, and vessels by use of a multivariable discrimination analysis. They reported that 1,336 (70.7%) of 1,889 regions of interest

(ROIs) with  $5 \times 5$  matrix size obtained from five patients were classified correctly. These studies indicated that some diffuse lung diseases on HRCT can be detected by use of the computerized scheme. However, the number of cases used was relatively small, and the level of performance was rather low for clinical applications.

In this study, we attempted to determine physical measures on HRCT images in order to detect and characterize diffuse lung diseases, which will be a basis for application to the differential diagnosis of diffuse lung disease in the future. We compared the physical measures of normal slices with those of abnormal slices which included six typical patterns of diffuse lung diseases. In addition, we investigated the classification performance for distinction between normal and abnormal slices.

## II. MATERIALS

### A. Clinical cases

Clinical cases were selected from HRCT images in the Department of Radiology at the University of Chicago Hospitals based on normal cases and abnormal cases reported in clinical examinations from 1998 to 2002. Our image database included 315 HRCT images, which consisted of three images selected from each of 105 patients. Three images were selected by an expert radiologist based on the following criteria; an upper image at the aortic arch level or above, a middle image at the level of main bronchi, and a lower image at the level of lower lobar bronchi or below. These images were obtained with three CT systems (a HiSpeed CT/i, a LightSpeed QX/i, and a GENESIS HISPEED/RP; GE Medical Systems, Milwaukee, WI). The image matrix size was  $512 \times 512$  pixels. The original CT value ranged from  $-1000$  HU to  $1000$  HU. The slice thickness ranged from  $1.0$  mm to  $3.0$  mm. The x-ray tube voltages were  $120$  kVp and  $140$  kVp. The field of view was optimized for each patient during the examination, so that the pixel size in the database ranged from  $0.494$  mm to  $0.781$  mm.

### B. "Gold Standard" for normal and abnormal opacities in HRCT images

In this study, it is important to establish reliable cases with typical normal and abnormal patterns, which will be used as "gold standard," because the subjective terms and judgments by radiologists have generally been used to describe diffuse lung diseases. Therefore, we have carried out an observer test carefully to select areas with specific abnormal patterns as the "gold standard." Three chest radiologists independently marked areas by drawing boundaries that included specific diffuse patterns and other abnormalities in 315 HRCT images on a CRT monitor by use of a mouse. They were asked to mark abnormal areas in the same manner as they commonly describe in their radiologic reports. Therefore, the areas with very subtle and suspicious abnormal opacities, which might be considered unimportant clinically by the three radiologists, were not included as a "gold standard" in this study. The diffuse abnormal patterns marked

were the following: (1) ground-glass opacities, (2) reticular and linear opacities, (3) nodular opacities, (4) honeycombing, (5) emphysematous change, (6) consolidation, (7) non-specific diffuse opacities or indeterminate for classification, and (8) other abnormalities such as atelectasis, pleural thickening, bronchiectasis, pleural effusion, bulla, focal lung lesion, and artifacts.

Each abnormal pattern was marked independently, and areas that included two or more different patterns were also marked with different colors. The areas with a specific pattern, which three radiologists marked independently and consistently as the same pattern, were used as "gold standard" for the specific abnormal opacities in this study. Since the size and the shape of abnormal areas marked by radiologists tended to be different each other, the area with a specific abnormal pattern as "gold standard" was determined by a logical AND operation for areas marked by three radiologists. The areas identified by the three radiologists as abnormal, even if they were considered as different abnormal patterns involved in the same area, were determined as "abnormal areas," which were used for testing. A slice was determined as an "abnormal slice" when there was at least one "abnormal area," whereas a slice was determined as a "normal slice" when there was no area identified by any of the three radiologists as abnormal. The areas obtained from the "normal slice" were determined as "normal areas."

Figure 1 shows portions of enlarged HRCT images of one normal slice and six abnormal slices with ground-glass opacities, reticular and linear opacities, nodular opacities, honeycombing, emphysematous change, and consolidation. The white lines indicate the abnormal area of the "gold standard" for each of the specific opacities. In the following sections, these HRCT images were employed to demonstrate the effects of a number of physical measures on the detection and classification of abnormal areas due to diffuse lung disease.

## III. METHODS

### A. Overall scheme for classification of diffuse lung opacities

Figure 2 shows the schematic diagram of our method. The lungs in HRCT images were first segmented from the background in each slice by using a gray-level morphological opening and a thresholding technique. The gray-level morphological opening<sup>12,13</sup> was applied to removal of small light structures such as vessels, while maintaining the overall gray levels and larger light structures. A gray-level histogram indicating the distribution of pixel values was constructed from pixels within the smoothed thorax, and the gray level that maximize the separation between the two main peaks of the histogram was used as a threshold to segment the lungs. The majority of the lungs in HRCT images were segmented by use of this automated method. However, because eleven lungs with consolidation were not segmented correctly, we employed a manual method for segmentation of the lung regions in these cases.

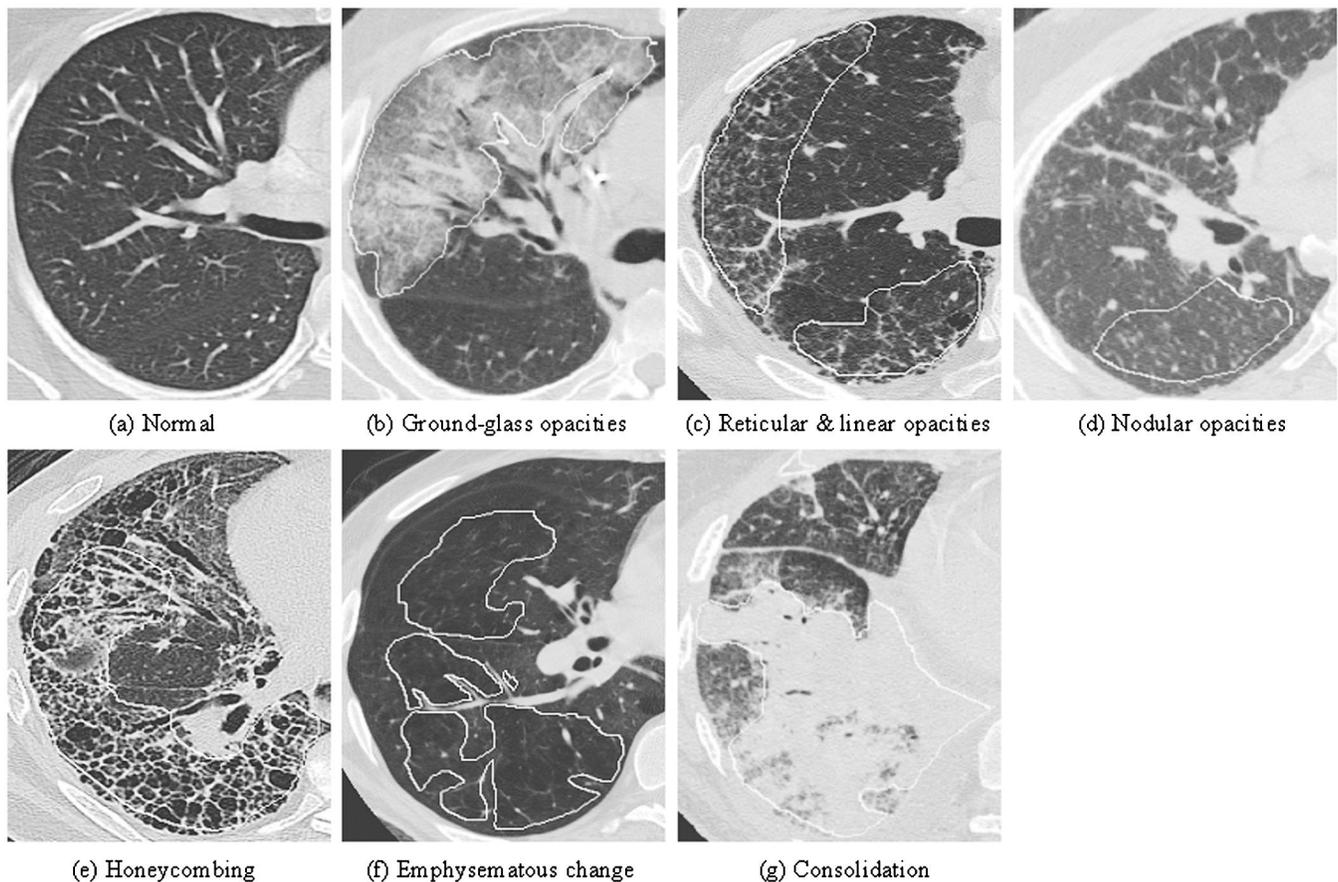


FIG. 1. An illustration of HRCT images with (a) normal slice and abnormal slices with (b) ground-glass opacities, (c) reticular and linear opacities, (d) nodular opacities, (e) honeycombing, (f) emphysematous change, and (g) consolidation. The white lines indicate abnormal areas of the “gold standard,” which were identified by three radiologists independently and consistently as the same type of abnormal patterns.

Many contiguous ROIs with a  $32 \times 32$  matrix size were selected automatically over the segmented lung region. The matrix size of the small ROI was chosen empirically. For abnormal ROIs used as a “gold standard,” at least 50% of the area in a  $32 \times 32$  ROI included specific abnormal opacities identified by the three radiologists. A large ROI with

$96 \times 96$  matrix size was also used in this study, which was placed centered over the small ( $32 \times 32$ ) ROI, i.e., one large ROI included nine contiguous small ROIs, and thus two adjacent large ROIs overlapped considerably. When some areas outside the lung regions were included in a ROI, the corresponding outside area was ignored for evaluating the diffuse lung disease. The physical measures determined in each ROI were used as the input data to an ANN. The ANN was employed for distinguishing between seven different patterns, which included normals and six patterns associated with diffuse lung disease, i.e., ground-glass opacities, reticular and linear opacities, nodular opacities, honeycombing, emphysematous change, and consolidation. In addition, we examined the usefulness of the Bayesian classifier as an alternative to the ANN.

## B. Determination of six physical measures

In order to detect and characterize the diffuse lung diseases on HRCT, we determined six different physical measures in this study. These included three measures related to the gray-level distribution and three measures for geometric patterns. The gray-level distribution measures were the mean and the standard deviation of CT values in a ROI, and also the fraction of the area with air density components in a ROI.

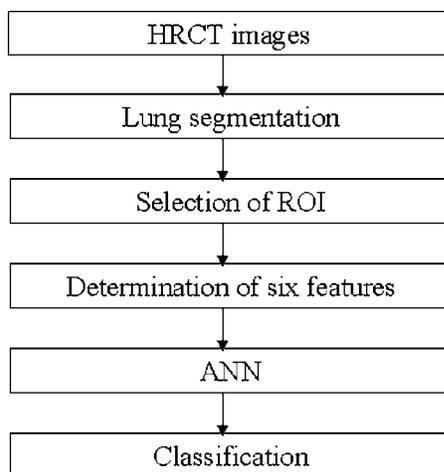


FIG. 2. Schematic diagram of the computerized classification method by use of ANN.

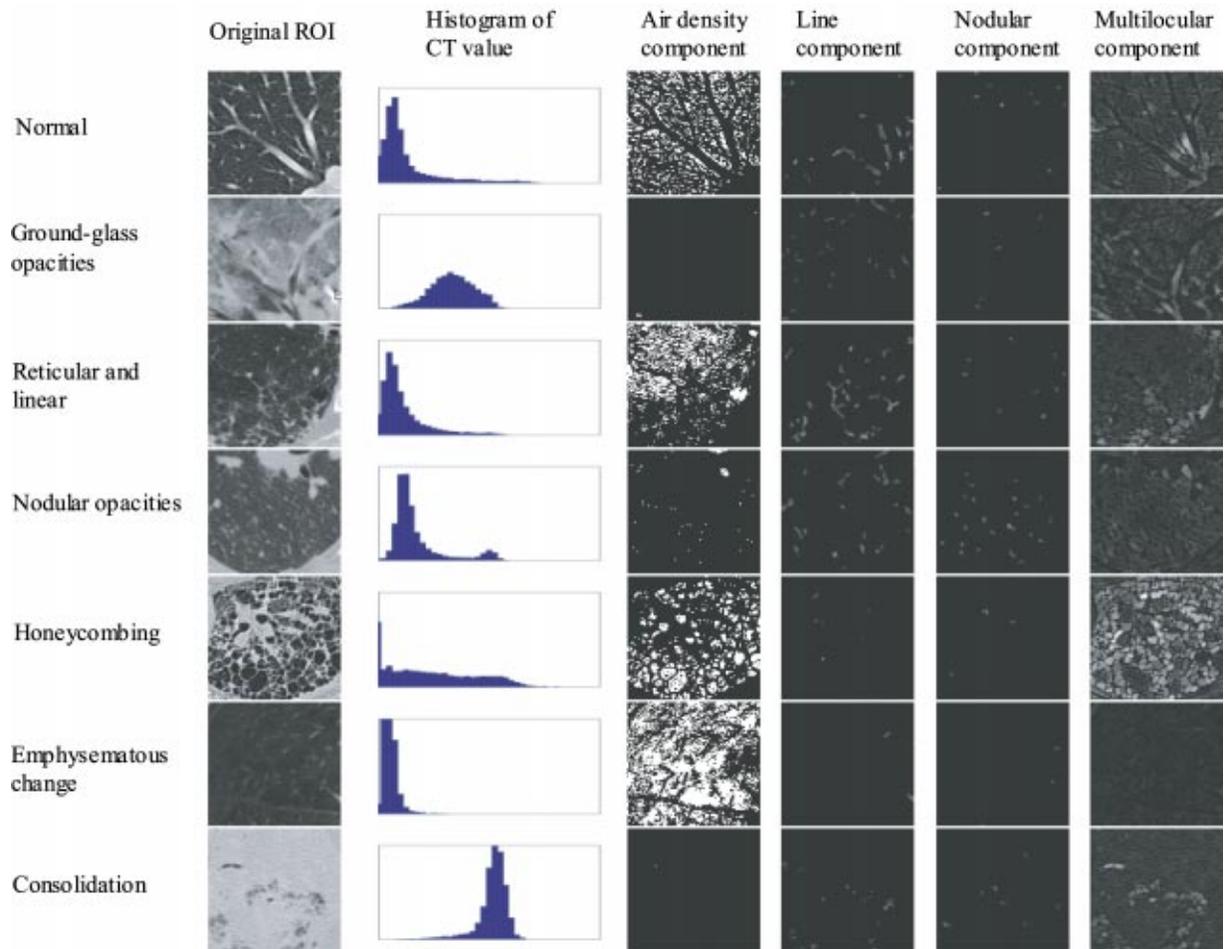


FIG. 3. An illustration of images (96×96) selected from the seven slices in Fig. 1, histograms of ROI images, and output images for air density components, line components, nodular components, and multilocular components.

The air density component was defined by the area having CT values between  $-910$  HU and  $-1000$  HU. The mean of CT values was employed for distinguishing some opacities which included very light areas or very dark areas in HRCT images such as consolidation and emphysematous change, as shown in Fig. 3. We used the standard deviation of CT values to characterize some opacities, which included a large variation in CT values due to the mixture of light and dark areas such as honeycombing. The air density component was quantified for the detection of some opacities, including air in the lungs. Although the measures obtained from the gray-level distribution are useful for the characterization of some diffuse lung diseases, it is difficult to detect nodular and reticular opacities because the gray-level distribution does not include information on the shape of opacities.

The geometric measures were then employed for the characterization of some aspects of the nodular components, line components, and multilocular components. In order to detect nodular components, we applied the morphological “white” top-hat transform<sup>12</sup> to an original CT image. The morphological white top-hat transform is defined by the subtraction of the opening of an original image from the original image. This operation corresponds to extracting “white” patterns smaller than the structure element used. The structure ele-

ment was a  $7 \times 7$  square, so that nodular opacities could be extracted while large vessels were removed. In order to remove small noise components in the background, we then applied the gray-level thresholding technique to the morphological white top-hat-transformed image by use of a threshold level of 175 pixel values which was selected empirically. The degree of circularity was defined by the fraction of the overlap area of the candidate with the circle having the same area as the candidate, and was determined for all of detected components to distinguish between nodular components and other linear components. The contrast was defined by the mean value of the five largest pixel values for each candidate in the white top-hat-transformed image. All detected components with a degree of circularity greater than 0.70 were considered to be initial candidates for nodular components. Because small blood vessels in a direction perpendicular to the slice tend to produce high contrast and circular patterns, these vessels were removed from candidates of nodular components which had contrast greater than 700, as illustrated in Fig. 4. Finally, the average pixel value of the corresponding image was defined as a measure indicating the fraction of nodular components in each ROI.

The line components were determined for extraction of the reticular and linear opacities. The gray-level thresholding

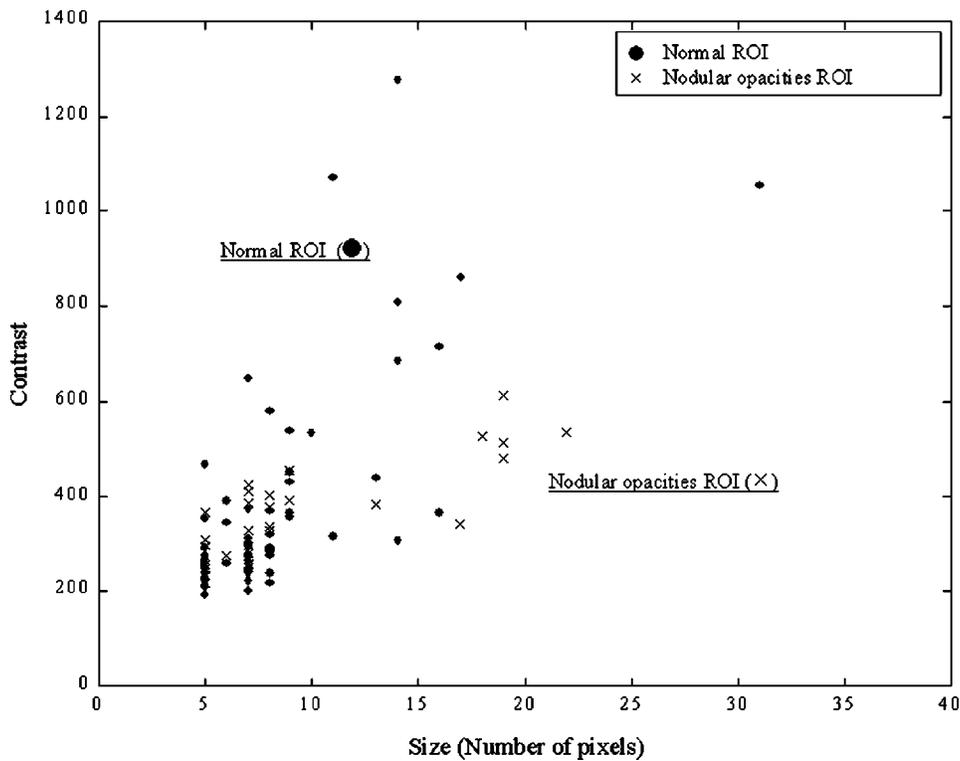


FIG. 4. The relationship between the size and contrast of initial candidates for nodular components in two ROIs selected for normal and nodular opacities. The initial candidates for the nodular component with the contrast above 700 were removed as small circular blood vessels.

technique was applied to the morphological white top-hat-transformed image of an original image by use of a threshold level of 175 pixel values which was also determined empirically. The degree of circularity was determined for all of the detected components for distinguishing between line components and nodular components. All detected components with a degree of circularity smaller than 0.70 were considered to be initial candidates for line components. The contrast was then calculated for all of the candidates for distinguishing between line components and medium-size vessels. All detected components with contrast smaller than 700 were considered to be line components, because medium-size blood vessels in a direction parallel to the slice tend to provide contrasts larger than 700, as illustrated in Fig. 5. The average pixel value of the resulting image was used as a measure for the fraction of line components in each ROI.

A measure for multilocular patterns was defined by the average pixel value of the image derived from the morphological “black” top-hat transform. The morphological black top-hat transform is given by subtraction of the original image from the opening of the original image. This operation corresponds to extracting “black” patterns that can fit into the area of the structure element used. The structure element was a  $7 \times 7$  square, so that honeycombing could be detected. Although the standard deviation of the CT value was useful for detecting honeycombing, it was difficult to distinguish between honeycombing and very large vessels. Therefore, we adopted the measure for multilocular patterns as another feature.

## C. Classification scheme

### 1. Artificial neural network

A three-layered ANN with a back-propagation algorithm<sup>14,15</sup> was employed as a classifier in this study. This ANN was applied for classifying all ROIs in the segmented lung regions into seven different categories, which included normals and six different abnormal patterns. The numbers of input, hidden, and output units were 12, 10, and 7, respectively. The number of hidden units was determined empirically. The input data for the ANN consisted of six features obtained from a small ROI with a  $32 \times 32$  matrix and another six features from a large ROI with a  $96 \times 96$  matrix. The six features for large ROIs were employed to take into account the information adjacent to the small ROIs with a  $32 \times 32$  matrix. The features were normalized by use of the average value and the standard deviation of each feature obtained from all normal ROIs in our database. In the feature space, therefore, the distribution of all features for normal ROIs was centered around the origin, whereas the distribution of features for abnormal ROIs was generally shifted from the origin. The output values for each of seven output units obtained with the ANN indicated the likelihood of each of the normal patterns and six abnormal patterns. The output unit yielding the largest value was considered to be the result of classification.

In order to investigate the performance of the classification between normal ROIs and abnormal ROIs in our database, we used the ROIs obtained from the “gold standard” for the training of the ANN. For each ROI with normals and

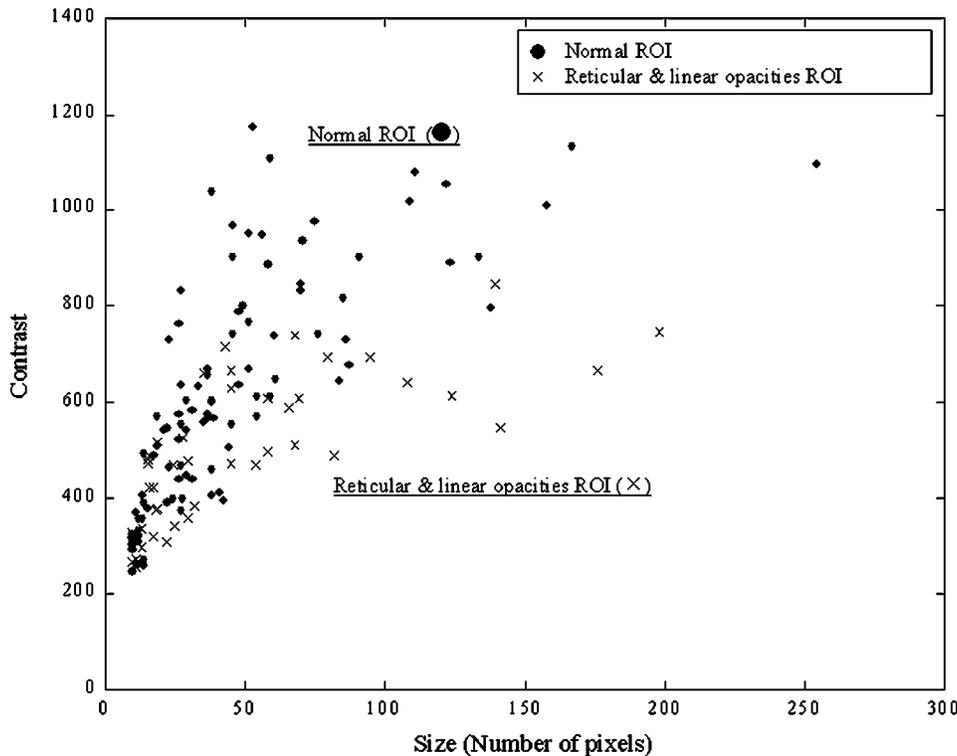


FIG. 5. The relationship between the size and contrast of initial candidates for line components in two ROIs selected for normals and reticular and linear opacities. The initial candidates for the line component with contrast above 700 were removed as medium-size horizontal blood vessels.

six abnormal patterns, twelve features were entered to each of 12 input units, whereas the teacher data were given to the output units, with 0.9 for the corresponding correct category and 0.1 for the remaining incorrect categories. For testing of the ANN, we used the normal ROIs obtained from the “normal area” and the abnormal ROIs obtained from the “abnormal area.” The abnormal ROIs consisted of the abnormal ROIs marked by the three radiologists as the same abnormal pattern and also as different abnormal patterns. Note that the normal ROIs and the abnormal ROIs identified as the same pattern were the same as those used for the training data set. In the testing, the ROI was considered abnormal when the largest output value of the ANN was given in the output unit corresponding to one of the six abnormal patterns, whereas the ROI was considered normal when the largest output value was given in the output unit for normals.

**2. Distinction between normal slice and abnormal slice**

In order to investigate the performance of the classification between normal slices and abnormal slices, we used all normal slices and all abnormal slices in our database as the testing data set. A normal slice was defined as a slice in which none of the three radiologists identified any abnormal ROI, whereas an abnormal slice corresponded to a slice in which the three radiologists identified at least one identical ROI as abnormal. Figure 6 shows the distributions of the number of ROIs that were classified as abnormal by our ROI-based classification scheme for both normal and abnormal slices. For normal slices, most slices included a relatively small number of incorrectly identified “abnormal ROIs,” which correspond to false-positive ROIs. Therefore,

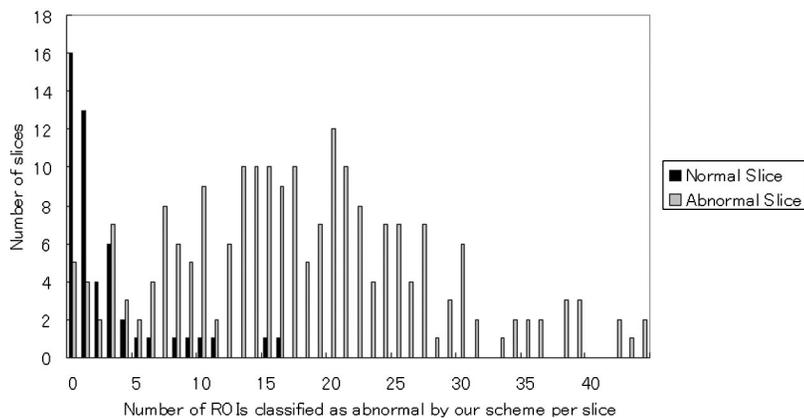
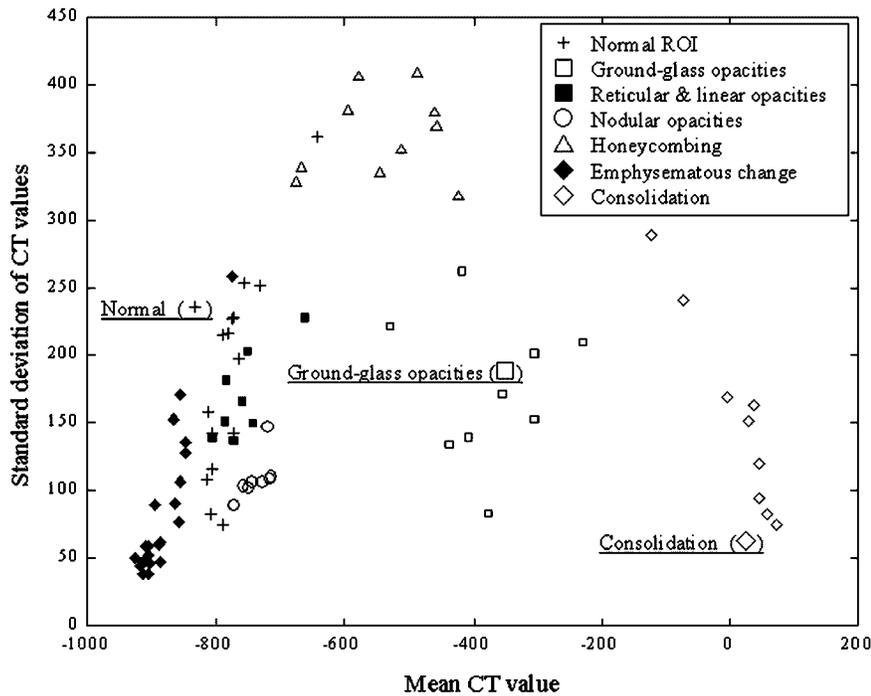
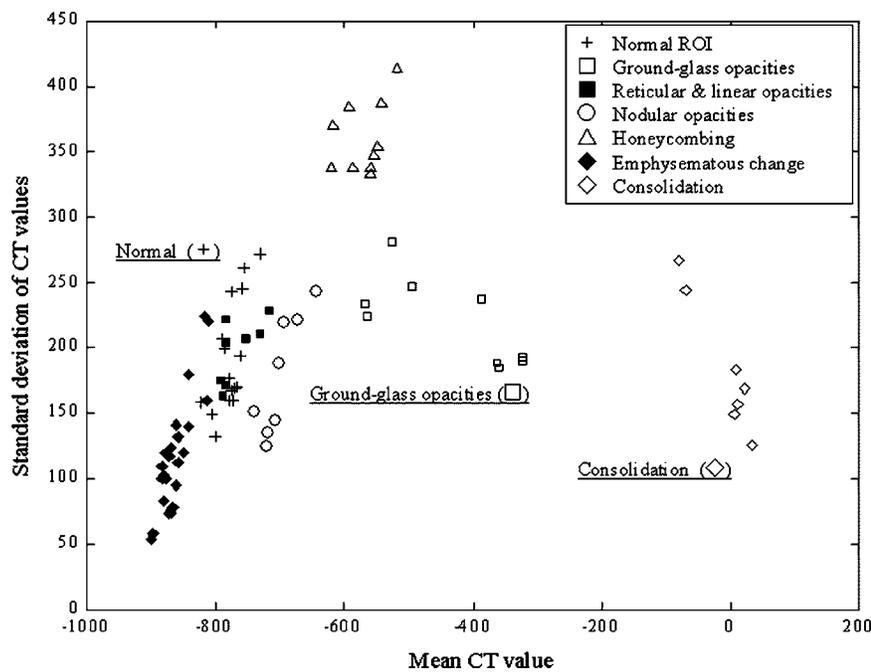


FIG. 6. Distribution of the numbers of ROIs which were classified as abnormal by our ROI-based classification scheme, obtained from normal slices and abnormal slices.



(a)



(b)

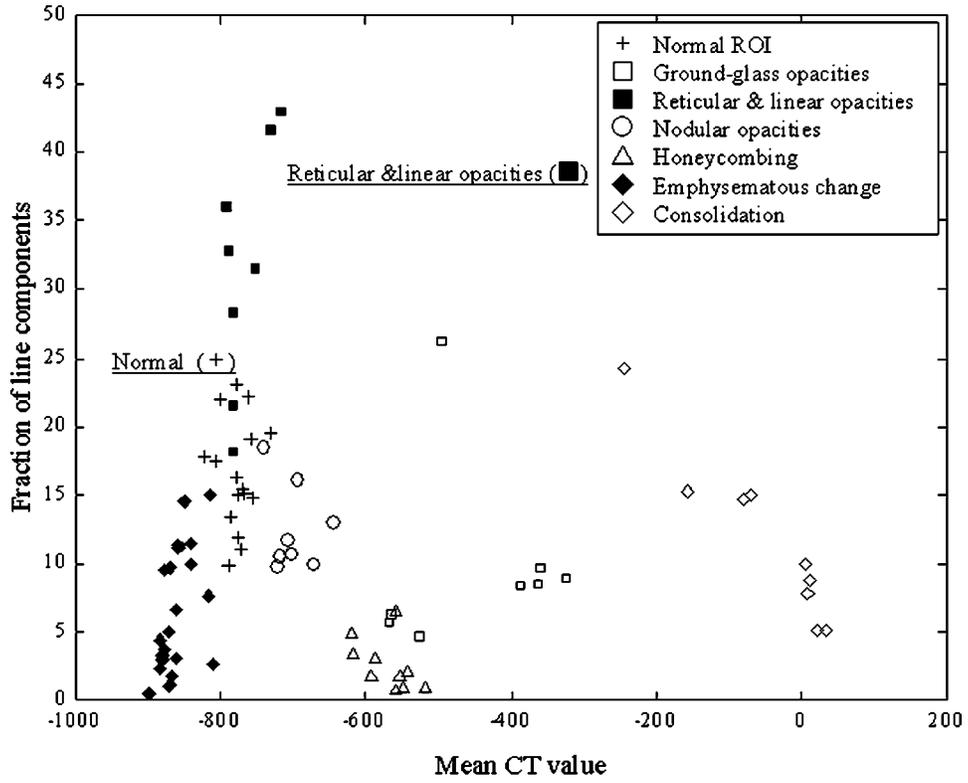
FIG. 7. A comparison of two distributions of physical measures obtained from (a) small ROIs ( $32 \times 32$ ) and (b) large ROIs ( $96 \times 96$ ).

we employed the first rule that a slice with more than five computer-abnormal ROIs would be considered as an abnormal slice. If the slice which was considered as an abnormal slice contained seven or fewer computer-abnormal ROIs, then the second rule was applied, i.e., a slice with four or more contiguous ROIs in one lung would be considered as an abnormal slice and otherwise considered as a normal slice.

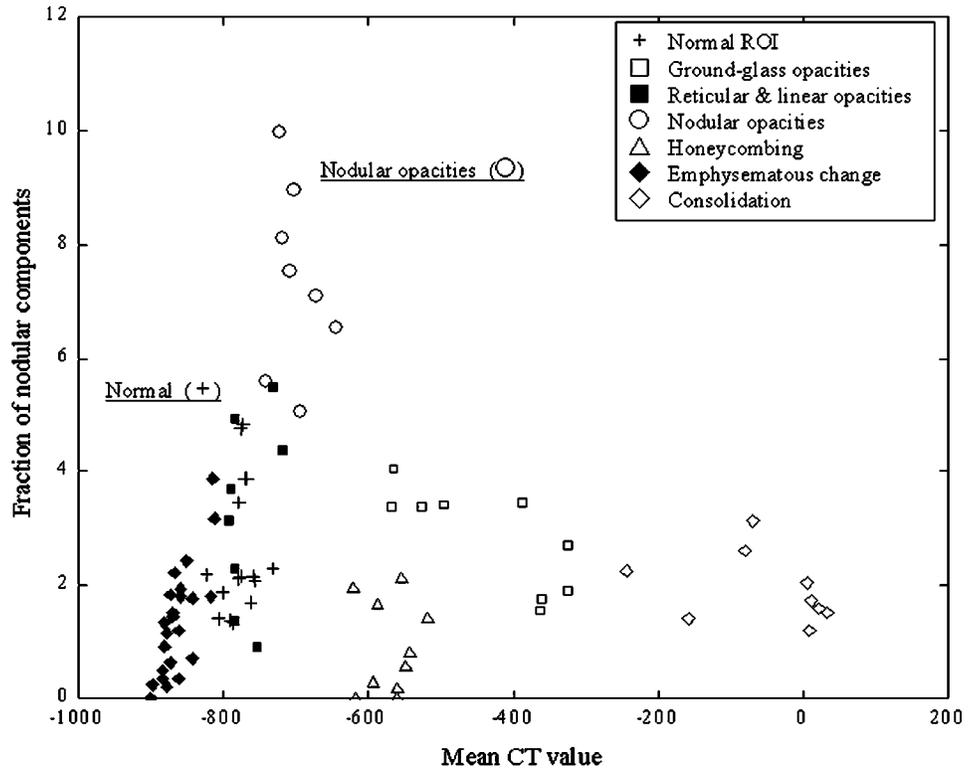
## IV. RESULTS AND DISCUSSIONS

### A. Effect of ROI size

In order to investigate the effect of ROI size, we selected several small ROIs ( $32 \times 32$ ) from seven slices, as illustrated in Fig. 1. The small ROIs were selected from the area of "gold standard." The physical measures for small ROI were

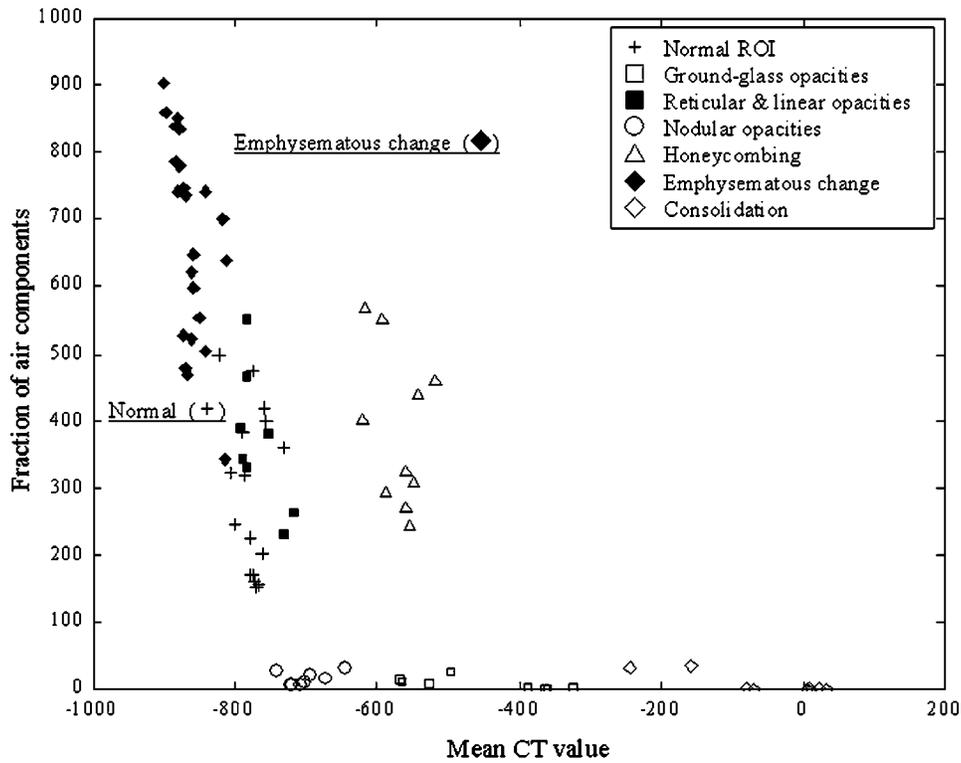


(a)

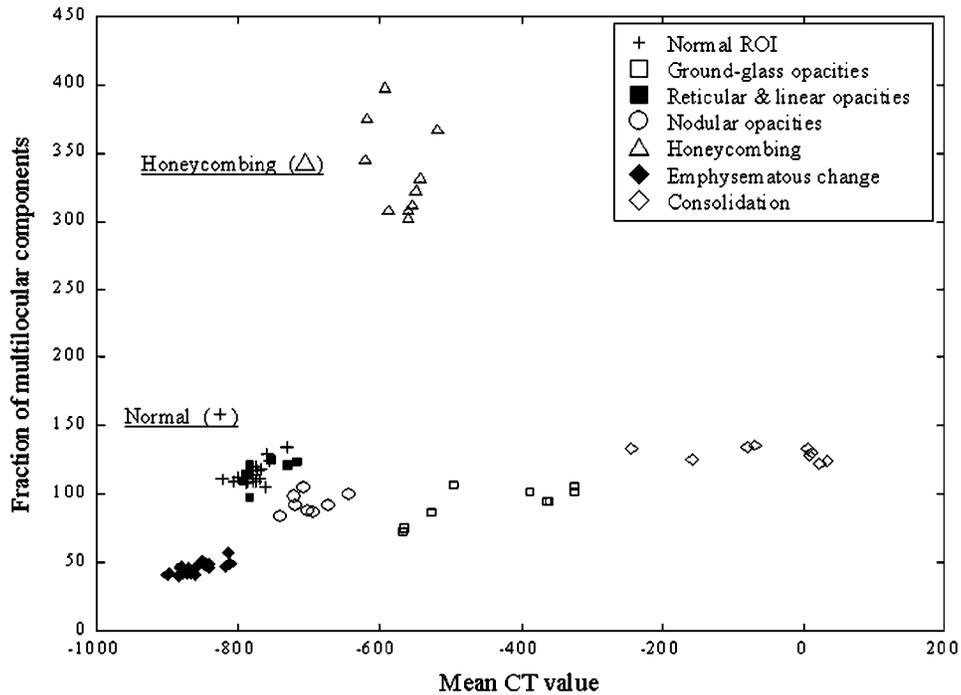


(b)

FIG. 8. Physical measures for large ROIs selected from the seven slices illustrated in Fig. 1. Features for the distinction between (a) normals and reticular and linear opacities, (b) normals and nodular opacities, (c) normals and emphysematous change, and (d) normals and honeycombing.



(c)

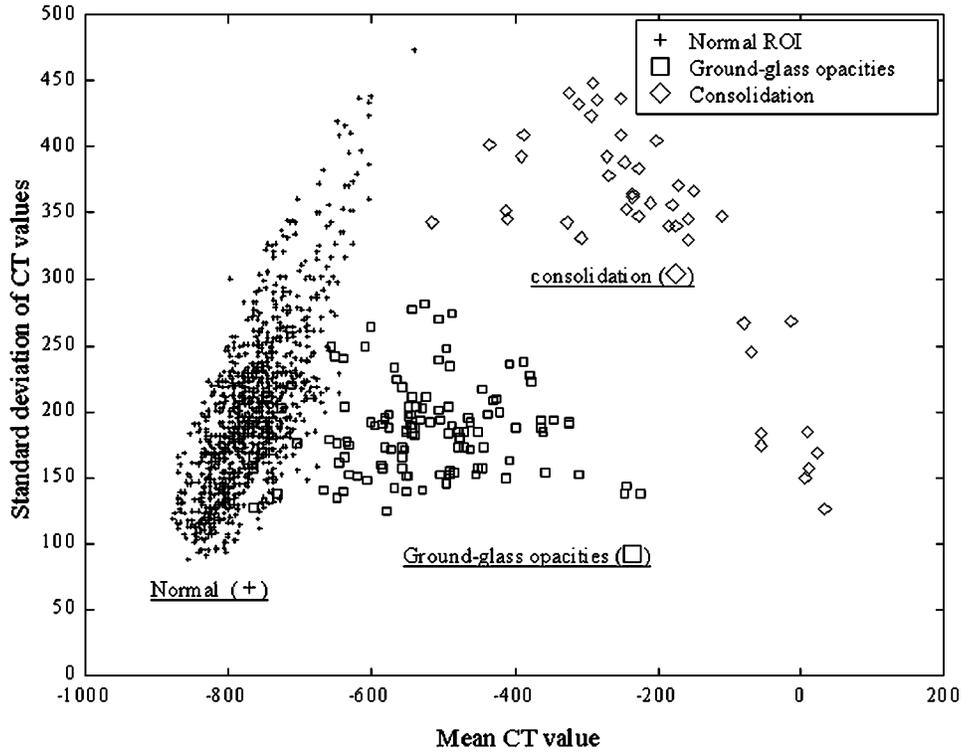


(d)

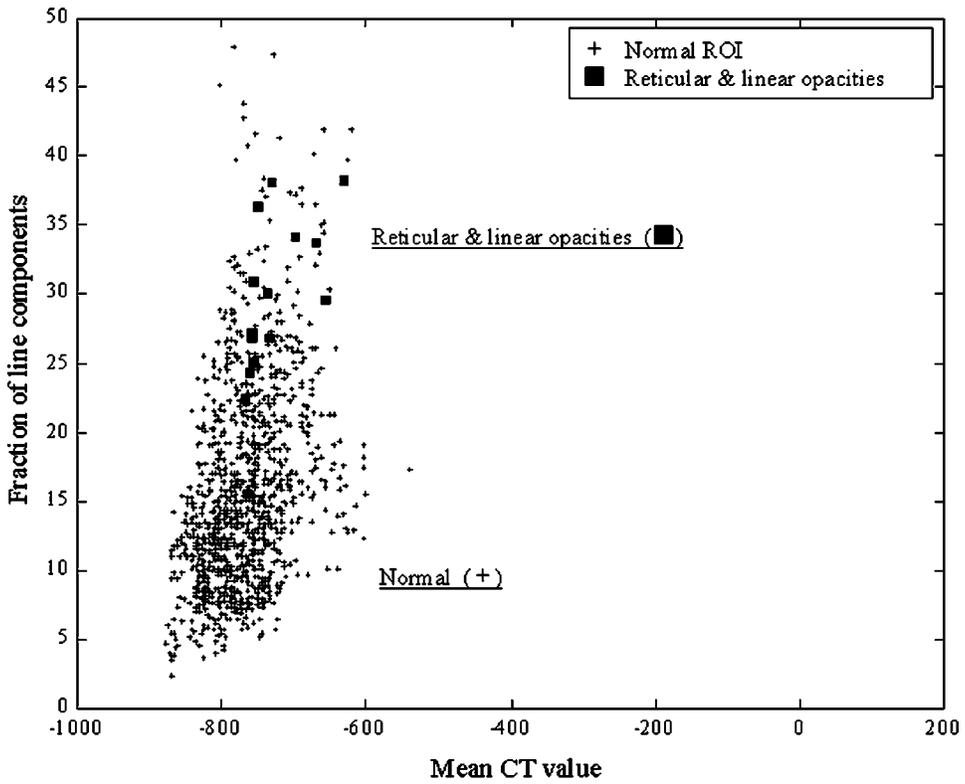
FIG. 8 (Continued.)

determined in each small ROI. The physical measures for large ROI (96×96) were determined in the 96×96 area located in the center of each selected small ROI. Figure 7 shows the effect of ROI size on the physical measures. As the ROI size increased from 32×32 in Fig. 7(a) to 96×96 in

Fig. 7(b), the distribution of data points for each of the different abnormal patterns was separated slightly. This is probably because large ROIs tended to give a better estimate of the statistical properties of physical measures than did a small ROI.

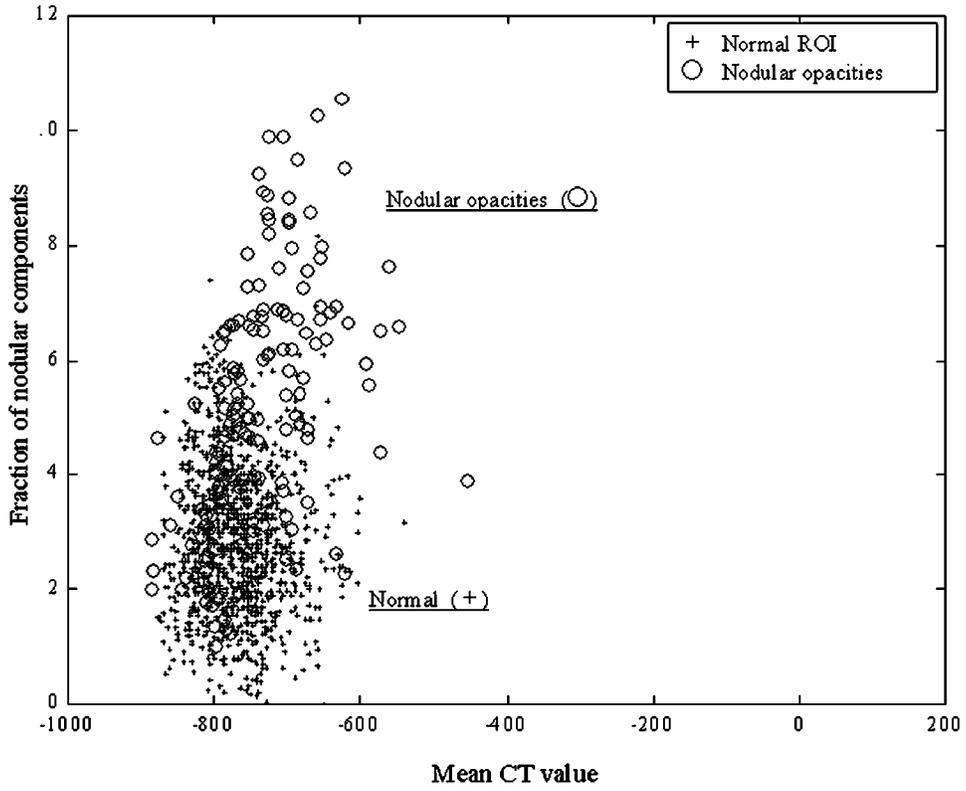


(a)

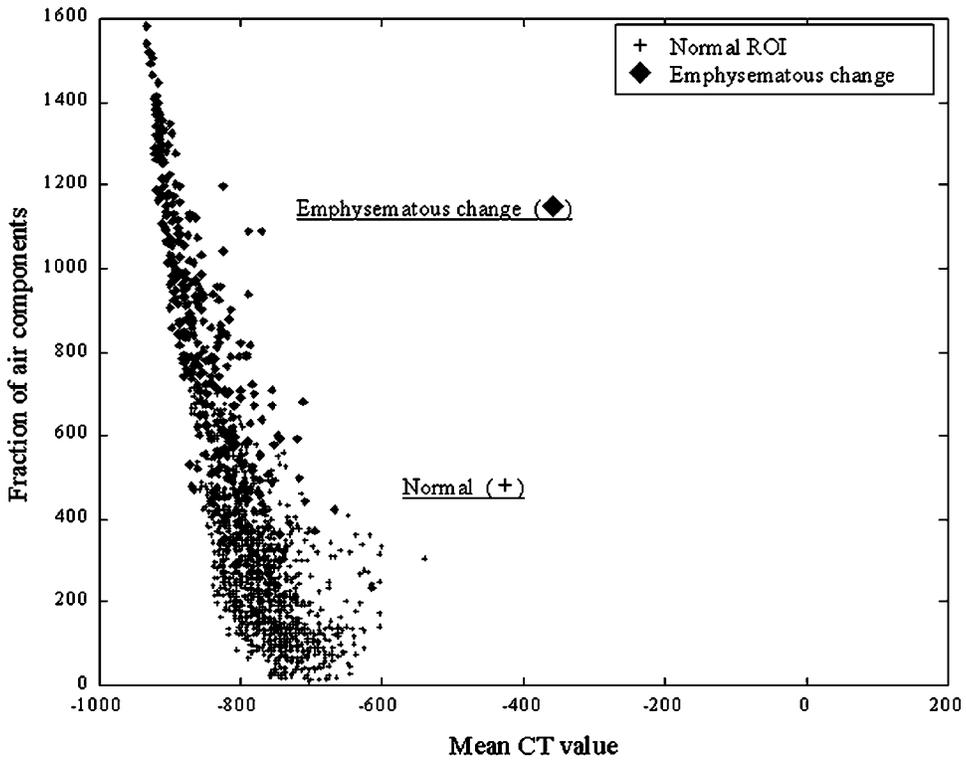


(b)

FIG. 9. Physical measures for large ROIs selected from the “gold standard” of all images in our database. Features for distinction between (a) normals, ground-glass opacities, and consolidation, (b) normals and reticular and linear opacities, (c) normals and nodular opacities, (d) normals and emphysematous change, and (e) normals and honeycombings.



(c)



(d)

FIG. 9 (Continued.)

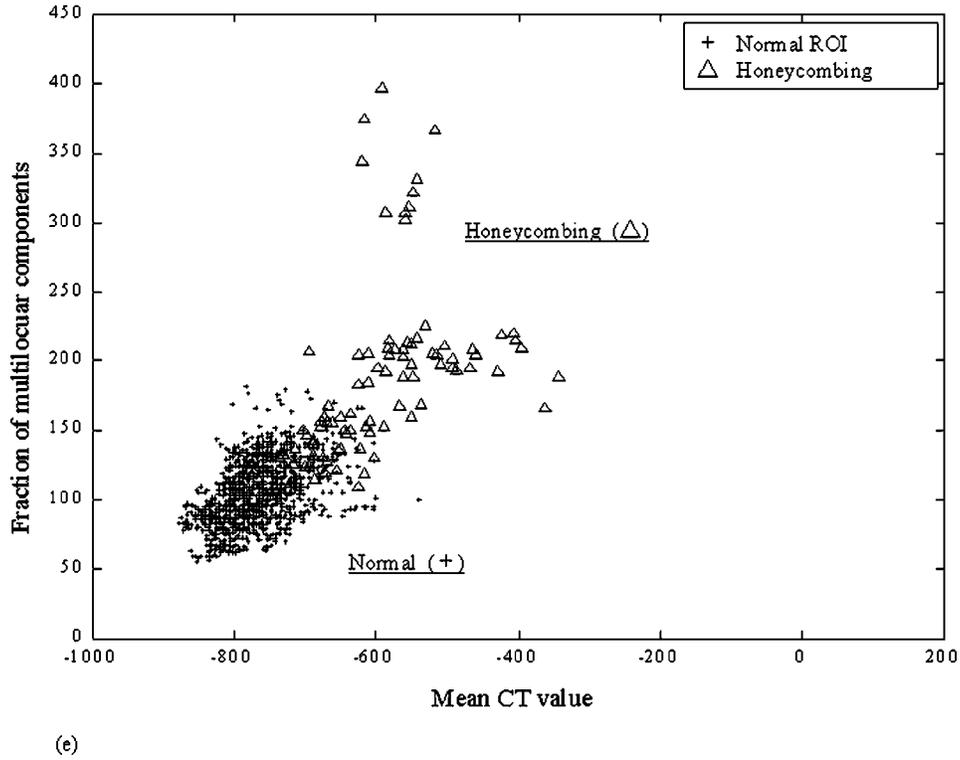


FIG. 9 (Continued.)

**B. Usefulness of six features**

Figure 7(b) shows the relationship between the mean and the standard deviation of CT values obtained with large ROIs (96×96). Note that all of physical measures in Figs. 8 and 9 discussed later were obtained with large ROIs. It appears in Fig. 7(b) that some of the abnormal patterns have distinctive features and thus can be distinguished from other opacities, even if only two features are employed. For example, the mean CT values for consolidation are larger than any other opacities, whereas the mean CT values for emphysematous change are smaller than any other opacities. The standard deviation of CT values for honeycombing is much larger than that for all of the other categories. The ground-glass pattern has relatively large CT values comparable to those of honeycombing, but its standard deviation is less than that of honeycombing. Figure 8(a) shows the relationship between the fraction of line components and the mean CT value. The

fractions of line components for reticular and linear patterns are greater than those of normals and other abnormal patterns. Figure 8(b) shows the relationship between the fraction of nodular components and the mean CT value. The fractions of nodular components for the nodular pattern are greater than those of normals and other abnormal patterns. Figure 8(c) shows the relationship between the fraction of air density components and the mean CT value. The ROIs with an emphysema pattern have more pixels with CT values between -910 HU and -1000 HU than do the other patterns. Figure 8(d) shows the relationship between the fraction of multilocular components and the mean CT value. The ROIs with a honeycomb pattern contain more multilocular components than do any other patterns.

We investigated the usefulness of the six features determined for all ROIs of normal and abnormal opacities from the “gold standard” of all slices in our database. Figure 9(a)

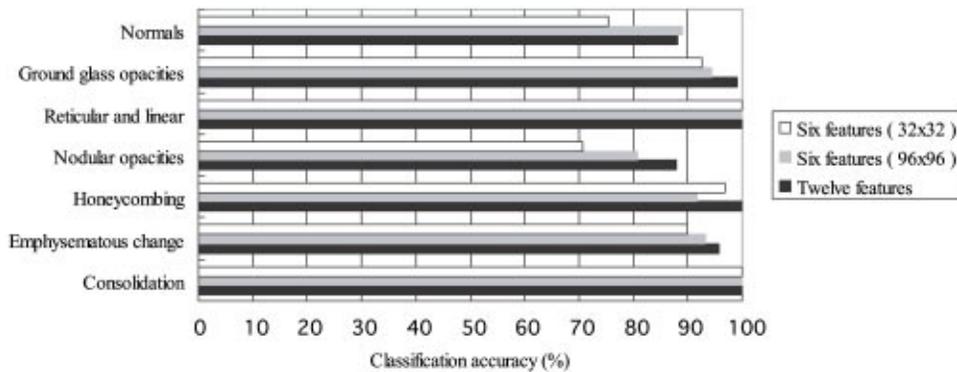


FIG. 10. Classification accuracy of the ANN obtained with six features and twelve features.

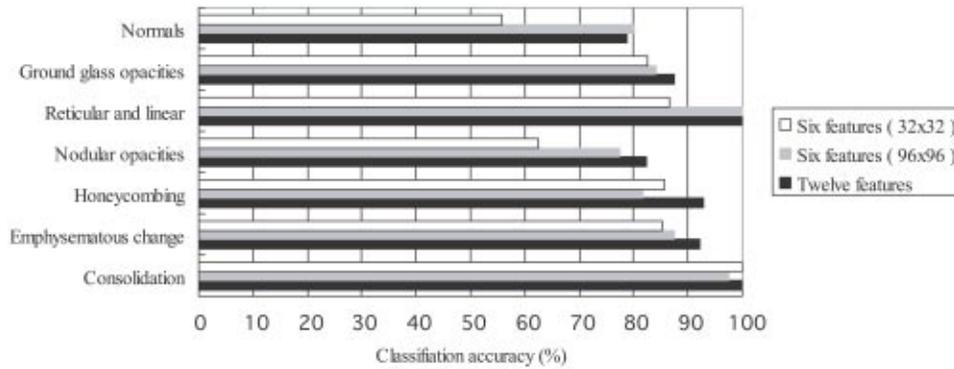


FIG. 11. Classification accuracy of the Bayesian classifier obtained with six features and twelve features.

shows the distribution of data points for normals, ground-glass opacities, and consolidations in terms of the mean and the standard deviation of CT values. All consolidations can be distinguished from normals and ground-glass opacities, whereas ground-glass opacities are slightly overlapped with normals. Figure 9(b) shows the relationship between the line components and the mean CT value for normals and reticular and linear opacities. Although some of the normals have large line components, the distribution for reticular and linear opacities tends to be shifted from that for normals. Figure 9(c) shows the relationship between the nodular components and the mean CT value for normals and nodular opacities. Although the two distributions for nodular opacities and normals are largely overlapped, some of the nodular opacities can be distinguished from normals by use of the fraction of nodular components. Figure 9(d) shows the relationship between air density components and the mean CT value for normals and emphysematous change. Approximately 80% of the emphysematous changes can be distinguished from normals. Figure 9(e) shows the relationship between the multilocular components and the mean CT value for normals and honeycombing. Approximately 75% of the honeycombings can be distinguished from normals.

**C. Comparison of classification accuracy using six features and twelve features**

The physical measures obtained with large ROI were useful to distinguish between some of abnormal patterns and normals. However, small ROIs may also be useful for capturing localized distinctive features of diffuse lung disease in small areas. In fact, it appears to be very useful to employ the combination of small and large ROIs to distinguish between some of the abnormal patterns and normals, as demonstrated below. Figure 10 shows the classification performance of the ANN with six features from small ROIs, and with six features from large ROIs, and with twelve features from small and large ROIs. As shown in Fig. 10, the ANN with twelve features provided a better classification performance for all of the abnormal categories than did the ANN with six features. Figure 11 shows the classification performance of the Bayesian classifier<sup>16,17</sup> with six features from small ROIs, with six features from large ROIs, and with twelve features from small and large ROIs. The Bayesian classifier with twelve features also provided a better overall classification performance than did the Bayesian classifier with six features, although the overall classification performance by the

TABLE I. Computerized classification results of the ROIs obtained from the “gold standard,” based on the use of the ANN and twelve features.

“Gold standard”	Computer output						
	Normal	GGO	Reticular	Nodular	Honeycombing	Emphysema	Consolidation
Normal (1067)	<b>940</b> (88.1%)	17 (1.6%)	14 (1.3%)	72 (6.7%)	6 (0.6%)	18 (1.7%)	0 (0.0%)
GGO (123)	1 (0.8%)	<b>122</b> (99.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Reticular (15)	0 (0.0%)	0 (0.0%)	<b>15</b> (100%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Nodular (150)	10 (6.7%)	0 (0.0%)	1 (0.7%)	<b>132</b> (88.8%)	0 (0.0%)	7 (4.7%)	0 (0.0%)
Honeycombing (98)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	<b>98</b> (100%)	0 (0.0%)	0 (0.0%)
Emphysema (385)	10 (2.6%)	2 (0.5%)	2 (0.5%)	0 (0.0%)	1 (0.3%)	<b>369</b> (95.8%)	1 (0.3%)
Consolidation (43)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	<b>43</b> (100%)

TABLE II. Computerized classification for distinguishing between normal ROIs and abnormal ROIs.

Cases (number of ROIs)	Computer output	
	Normal	Abnormal
Normal ROI (1067)	940 (88.0%)	127 (11.9%)
Abnormal ROI due to identical pattern (814)	21 (2.6%)	793 (97.4%)
Abnormal ROI due to different patterns (801)	119 (14.9%)	682 (85.1%)

Bayesian classifier was lower than that by the ANN. This result seems to indicate that it would be useful to take into account the information for both small and large ROIs. For the size of large ROIs, we confirmed that the matrix size of  $64 \times 64$  did not provide a better performance than did the  $96 \times 96$  matrix.

#### D. Performance of classification

We investigated the performance of the classification between normal ROIs and abnormal ROIs. Table I shows the results of a consistency test for the training cases. The sensitivity of this computerized method for the detection of the six abnormal patterns in each ROI was 99.2% (122/123) for ground-glass opacities, 100% (15/15) for reticular and linear opacities, 88.0% (132/150) for nodular opacities, 100% (98/98) for honeycombing, 95.8% (369/385) for emphysematous change, and 100% (43/43) for consolidation. The specificity in detecting a normal ROI was 88.1% (940/1067). Although the average performance was high, the efficiency is for discriminating between normals and nodular patterns was relatively low, i.e., 6.7% (72/1067) of normal ROIs were classified incorrectly as nodular opacities. This result is probably caused by the fact that there were many nodular-like patterns in normal slices. Table II shows the results of the classification performance between normal and abnormal ROIs. The sensitivity and specificity for the detection of abnormal ROIs were 97.4% (793/814) and 88.0% (940/1067), respectively. However, the sensitivity for abnormal ROIs which were identified by the three radiologists as abnormal, but with different opacities, were 85.1% (682/801), which is lower than that (97.4%) for abnormal ROIs that were identified by all radiologists as the same type of opacities. These results seem to indicate that the two distributions for normals and abnormal can be separated in the feature space obtained by the use of six physical measures. Table III shows the classification performance for normal slices, abnormal slices, and suspicious normal/abnormal slices. The normal slices were

TABLE III. Computerized classification for distinction among normal slices, abnormal slices, and suspicious normal/abnormal slices.

Cases (Number of slices)	Computer output	
	Normal	Abnormal
Normal slice (49)	41 (83.7%)	8 (16.3%)
Abnormal slice (213)	21 (9.9%)	192 (90.1%)
Suspicious normal/abnormal slice (53)	28 (52.8%)	25 (47.2%)

determined when there was no area identified by any of the three radiologists as abnormal. The abnormal slices were determined when there was an area identified by the three radiologists as abnormal even if they considered them as different abnormal patterns. The slices that did not belong to normal slices and abnormal slices were determined as a "suspicious normal/abnormal slice," i.e., they were determined when there was an area identified by one or two of the radiologists as abnormal. The sensitivity and specificity for the detection of abnormal slices were 90.1% (192/213) and 83.7% (41/49), respectively. However, 52.8% (28/53) of the suspicious normal/abnormal slices were classified as normal slices, whereas 47.2% (25/53) of the suspicious normal/abnormal slices were classified as abnormal slices.

#### V. CONCLUSION

Six physical measures were determined for the detection and characterization of diffuse lung diseases in HRCT images. The results indicated the usefulness of the six physical measures for the distinction between normals and six different types of diffuse lung diseases. This computerized method may be useful in assisting radiologists in their assessment of diffuse lung disease in HRCT images.

#### ACKNOWLEDGMENTS

The authors are grateful to Roger Engelmann, M.S., for his technical assistance, and to Elisabeth Lanzl for editing the manuscript. This study was supported by USPHS Grant No. CA 62625. K. D. and S. K. are shareholders in R2 Technology, Inc., Los Altos, CA. K.D. is a shareholder of Deus Technologies, Inc., Rockville, MD.

<sup>a)</sup>Corresponding author: Yoshikazu Uchiyama, Ph.D., Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, The University of Chicago, 5841 S. Maryland Ave., Chicago, Illinois 60637. Telephone: (773) 834-5096; fax: (773) 834-5091; electronic mail: yoshi@apex.bsd.uchicago.edu

<sup>1</sup>S. Katsuragawa, K. Doi, and H. MacMahon, "Image feature analysis and computer-aided diagnosis in digital radiography: Detection and characterization of interstitial lung disease in digital chest radiographs," *Med. Phys.* **15**, 311–319 (1988).

<sup>2</sup>S. Katsuragawa, K. Doi, and H. MacMahon, "Image feature analysis and computer-aided diagnosis in digital radiography: Classification of normal and abnormal lung with interstitial disease in chest images," *Med. Phys.* **16**, 38–44 (1989).

<sup>3</sup>S. Katsuragawa, K. Doi, N. Nakamori, and H. MacMahon, "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.* **17**, 72–78 (1990).

<sup>4</sup>T. Ishida, S. Katsuragawa, T. Kobayashi, H. MacMahon, and K. Doi, "Computerized analysis of interstitial disease in chest radiographs: Improvement of geometric-pattern feature analysis," *Med. Phys.* **24**, 915–924 (1997).

<sup>5</sup>T. Ishida, S. Katsuragawa, K. Ashizawa, H. MacMahon, and K. Doi, "Application of artificial neural networks for quantitative analysis of image data in chest radiographs for detection of interstitial lung disease," *J. Digit. Imaging* **11**, 182–192 (1998).

<sup>6</sup>L. Monnier-Cholley, H. MacMahon, S. Katsuragawa, J. Morishita, T. Ishida, and K. Doi, "Computer aided diagnosis for detection of interstitial opacities on chest radiographs," *Am. J. Radiol.* **171**, 1651–1656 (1998).

- <sup>7</sup>K. Ashizawa, T. Ishida, H. MacMahon, C. J. Vyborny, S. Katsuragawa, and K. Doi, "Artificial neural networks in chest radiographs: Application to differential diagnosis of interstitial lung disease," *Acad. Radiol.* **6**, 2–9 (1999).
- <sup>8</sup>K. Ashizawa, H. MacMahon, T. Ishida, K. Nakamura, C. J. Vyborny, S. Katsuragawa, and K. Doi, "Effect of artificial networks on radiologists' performance for differential diagnosis of interstitial lung disease on chest radiographs," *Am. J. Radiol.* **172**, 1311–1315 (1999).
- <sup>9</sup>K. R. Heitmann, H. U. Kauczor, P. Mildemberger, T. Uthmann, J. Perl, and M. Thelen, "Automatic detection of ground glass opacities on lung HRCT using multiple neural networks," *Eur. Radiol.* **7**, 1463–1472 (1997).
- <sup>10</sup>R. Uppaluri, E. A. Hoffman, M. Sonka, P. G. Hartley, G. W. Hunninghake, and G. McLennan, "Computer recognition of regional lung disease patterns," *Am. J. Respir. Crit. Care Med.* **160**, 648–654 (1999).
- <sup>11</sup>S. Delorme, M. K. Reichenbecher, I. Zuna, W. Schlegel, and G. V. Kaick, "Usual interstitial pneumonia: quantitative assessment of high-resolution computed tomography findings by computer-assisted texture-based image analysis," *Invest. Radiol.* **32**, 566–574 (1997).
- <sup>12</sup>J. Serra, *Image Analysis and Mathematical Morphology* (Academic, London, 1982).
- <sup>13</sup>R. C. Gonzalez and R. E. Woods, *Digital Image Processing* (Addison-Wesley, Cambridge, MA, 1992), pp. 518–560.
- <sup>14</sup>D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning Internal Representations by Error Propagation, Vol. 1: Parallel Distribution Processing* (MIT, Cambridge, MA, 1986), Chap. 8, pp. 318–362.
- <sup>15</sup>S. Haykin, *Neural Networks: a Comprehensive Foundation* (Prentice-Hall, Englewood Cliffs, NJ, 1999), pp. 156–251.
- <sup>16</sup>R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification* (Wiley, New York, 2001), pp. 20–64.
- <sup>17</sup>K. Fukunaga, *Statistical Pattern Recognition* (Academic, New York, 1990), pp. 124–177.