Published in final edited form as:

Phys Med Biol. 2011 February 21; 56(4): 1139–1153. doi:10.1088/0031-9155/56/4/016.

Computer-aided detection of early interstitial lung diseases using low-dose CT images

Sang Cheol Park¹, Jun Tan², Xingwei Wang², Dror Lederman², Joseph K Leader², Soo Hyung Kim¹, and Bin Zheng^{2,3}

¹School of Electronics and Computer Engineering, Chonnam National University, Gwangju 500-757, Korea

²Department of Radiology, University of Pittsburgh, Pittsburgh, PA 15213, USA

Abstract

This study aims to develop a new computer-aided detection (CAD) scheme to detect early interstitial lung disease (ILD) using low-dose computed tomography (CT) examinations. The CAD scheme classifies each pixel depicted on the segmented lung areas into positive or negative groups for ILD using a mesh-grid-based region growth method and a multi-feature-based artificial neural network (ANN). A genetic algorithm was applied to select optimal image features and the ANN structure. In testing each CT examination, only pixels selected by the mesh-grid region growth method were analyzed and classified by the ANN to improve computational efficiency. All unselected pixels were classified as negative for ILD. After classifying all pixels into the positive and negative groups, CAD computed a detection score based on the ratio of the number of positive pixels to all pixels in the segmented lung areas, which indicates the likelihood of the test case being positive for ILD. When applying to an independent testing dataset of 15 positive and 15 negative cases, the CAD scheme yielded the area under receiver operating characteristic curve (AUC = 0.884 ± 0.064) and 80.0% sensitivity at 85.7% specificity. The results demonstrated the feasibility of applying the CAD scheme to automatically detect early ILD using low-dose CT examinations.

1. Introduction

Interstitial lung disease (ILD) includes a heterogeneous group of abnormal conditions or diseases that involve the lung parenchyma and fibrosis (King 2005). These diseases typically have similar clinical, radiographic, physiologic, and/or pathologic features. ILD can be caused by infection, an exposure at the workplace, medications, and various disorders. Studies have demonstrated that ILD was an important symptom in the diagnosis and assessment of a large number of lung diseases and when ILD was present these diseases often resulted in a higher occurrence of lung cancer and/or mortality rate (Katabami *et al* 2000, Bourke 2006). ILD is chronic and progressive (Rosas *et al* 2007, Gochuico *et al* 2008, Song *et al* 2009). It also tends to be widespread and not to be confined to single lobes or segments (Marten *et al* 2009). As the disease develops, ILD gradually leads to reduced lung volume, decreased lung compliance and restrictive physiology (Song *et al* 2009), which eventually causes considerable morbidity and mortality rate of the patients (Arzhaeva *et al* 2007). For example, the patients diagnosed with usual interstitial pneumonia (the largest subgroup of ILD) have the median survival rate of 2.8 years or a 15–30% 5 year survival rate, which is quite comparable to non-squamous cell lung cancer and significantly worse

than many other cancers (Bjoraker *et al* 1998). Previously clinical researches supported the hypothesis that early detection and treatment of ILD provide a better chance of controlling or slowing disease progression, and hence, increasing the survival probability of the patients (Bouros *et al* 2002, Goh *et al* 2008).

X-ray chest radiography and computer tomography (CT) are two common anatomic imaging modalities that are routinely used in the detection and diagnosis of a variety of lung diseases. The study has demonstrated the high correlation between CT finding, pulmonary function tests and bronchoalveolar lavage cytology in diagnosis of ILD (Biederer *et al* 2004). Thus, using CT images, physicians and/or radiologists can detect and identify the presence of diseases (Uppaluri *et al* 1999), directly visualize the morphologic extents of diseases (Best *et al* 2003), characterize the patterns and severity of diseases (Uchiyama *et al* 2003, Sluimer *et al* 2006), and assess the clinical course of diseases and response to therapy (King 2005). As the advance of CT technology, the high-resolution CT examination has become the imaging modality of choice for the detection and diagnosis of lung diseases (Schaefer-Prokop *et al* 2001, Sluimer *et al* 2006). Although HRCT offers images of the lung with increasingly improved anatomic resolution, visual interpretation or assessment of a large number CT image slices (i.e. >250 at the section thickness of 1.25 mm) remains as a difficult and time-consuming task that results in substantial intra- and inter-reader variabilities (Leader *et al* 2005).

To help improve efficiency and accuracy in detecting and diagnosing lung diseases using CT images, developing computer-aided detection or diagnosis (CAD) schemes has been attracting research interests for the past two decades. It is believed that CAD can provide objective and reproducible measures to detect and/or classify between the normal and the diseased lung or lung areas (regions of interest (ROIs)). As a result, a number of CAD schemes for detecting or classifying ILD using CT images have been developed and reported. Some studies divided each CT image slice into ROIs with the fixed size using a simple grid structure and then focused on distinguishing or classifying the ROIs into several different patterns including normal and several abnormal patterns associated with ILD (i.e. fibrosis, ground glass, honeycombing, bronchovascular, and hyperlucency) (Uppaluri *et al* 1999, Uchiyama *et al* 2003, Sluimer *et al* 2006). Recently, several studies also proposed to segment lung areas depicting ILD in CT images (Korfiatis *et al* 2008, Wang *et al* 2009) or characterize the severity of ILD (Zheng *et al* 2004, Song *et al* 2009). In these studies, various image features including run-length matrix textures, co-occurrence matrix, moments of histogram, fractal dimensions, multi-scale Gaussian filter bank, and wavelet-based texture features have been investigated and applied to detect and classify between normal lung tissue and a variety of ILD-related tissue patterns. Previous studies have also preliminarily assessed radiologists' performance in distinguishing between normal and abnormal (ILD) lung tissue image patterns when using CAD as an assisting tool (Uchiyama *et al* 2003, Korfiatis *et al* 2008).

Although a number of CAD schemes for ILD have been previously developed and tested to detect and classify several ILD-related patterns depicted on manually selected ROIs from CT image slices (Uppaluri *et al* 1999, Sluimer *et al* 2006) or quantify ILD severity among the identified positive cases (Best *et al* 2003, Zheng *et al* 2004, Marten *et al* 2009), these schemes were typically applied to the severe ILD cases and the CT images acquired from the examinations with regular dose for diagnostic purpose. Due to the large inter-reader variability in detecting ILD regions and rating its severity, assessing the performance of these CAD schemes is often difficult and unreliable. Meanwhile, despite the great effort, none of any CAD schemes have been accepted in the clinical practice to effectively assist radiologists detecting and/or diagnosing ILD to date. However, in the past decade, the low-dose HRCT examinations have been widely investigated as a potential screening tool to

detect lung cancer at an early stage (Pastorino 2010). As the availability of increasing number of low-dose CT examinations, whether using these low-dose CT images could also affectively detect early ILD cases have not been fully investigated to date. Based on the fact that early detection of ILD from the asymptomatic cases is clinically important to effectively control and/or prevent the rapid disease progression, we adopted a different approach in this study. Unlike the previous studies that focused on detecting and classifying detailed ILD patterns or severity, this study aims to develop and test a new CAD scheme that is specifically designed to prescreen and detect the early asymptomatic ILD cases from a group of mixed lung cancer screening cases using low-dose CT examinations.

2. Materials and methods

2.1. Image dataset

From an established low-dose CT lung cancer screening image database in our medical center (Leader *et al* 2005), we queried and selected 19 cases that depict mild (or early) ILD. The CT images of these cases were retrospectively reviewed by an experienced chest radiologist and the mild ILD was re-confirmed in all of these selected cases. We then randomly selected another set of 19 negative cases for ILD from the same lung cancer screening database. Thus, the dataset used in this study includes a total of 38 cases with equal number of positive and negative cases for ILD. The ILD negative or positive status among each of these 38 cases was confirmed by the follow-up diagnostic reports. Since the purpose of this study is to develop a CAD scheme to detect mild ILD at an early stage by using the low-dose CT examinations, no severe ILD cases were included in the dataset. As a result, there are no honeycombing and/or consolidation patterns depicted on these confirmed ILD cases. The dataset was further divided into two independent training and testing subsets. The training subset includes four positive and four negative examinations for ILD. From these eight training cases, a large number of negative and positive pixels for ILD were extracted to train our pixel-based classifier (that will be described in the following section in detail). The testing subset includes the rest of 30 cases including 15 positive and 15 negative ILD examinations.

The detailed description of the CT image acquisition protocol for the cases selected in our dataset has been previously reported (Leader *et al* 2005). In brief, these CT examinations were acquired using the light-speed plus or light-speed ultra CT machines (GE Healthcare). The CT protocol varied depending on the subject size. Specifically, in these CT screening examinations the x-ray tube voltage was either 120 or 140 kVp, and the mean tube current per second scanning ranged from 22.8 to 51.9 mAs. The plane pixel dimensions in these CT examinations varied from 0.56 to 0.82 mm depending on the body size of a participant being scanned. The scanning CT acquisitions were contiguous and the images were reconstructed at slice section thickness of 2.5 mm with 512×512 in plane pixel matrices using the GE high-spatial frequency 'lung' convolution kernel.

2.2. CAD scheme

Our CAD scheme aims to detect and classify the cases of low-dose CT examinations into negative or positive group for ILD. The scheme includes multiple image processing and feature classification steps. We first provide a general description of the scheme and then discuss each step in detail in the following subsections. First, the scheme segments the lung areas depicted on each CT image slice and computes the total lung volume. Second, the scheme applies a mesh-grid region growth algorithm to scan through the entire segmented lung areas and select or identify the pixels that need to be analyzed and classified as positive or negative for ILD. All unselected pixels are automatically classified to the negative group for ILD. For each selected pixel, a ROI centered at this pixel is created with a predefined

size based on our experimental results. The scheme extracts a set of histogram and texture features from the ROI. Based on these features, an artificial neural network (ANN) was applied to determine the likelihood of the ROI being positive for ILD. If the ROI is classified depicting the ILD pattern, the pixel at the center of the ROI is classified to the group of positive for ILD. Otherwise, this pixel is classified to the group of negative for ILD (figure 1). Third, after all pixels identified on the segmented lung areas have been classified into one of the two groups, the CAD scheme counts all pixels classified to the positive group for ILD and all pixels depicted on the segmented lung areas in the complete CT examination. The scheme then computes a ratio of the computed ILD volume to the entire segmented lung volume, which is used as a case-based detection or a classification score indicating the likelihood of the case being positive for ILD. Following are the detailed descriptions of each steps of the CAD scheme.

2.2.1. Lung segmentation—A computerized scheme previously developed in our group (Park *et al* 2009a) was applied to segment the 3D lung volume depicted on the complete CT examination. In brief, the lung segmentation algorithm includes a 2D-based adaptive threshold (segmentation) step and a 3D-based region growing processing step. In the first step, the scheme determines a segmentation threshold value for each CT slice image. To automatically select this threshold, the scheme searches the histogram of the pixel values for the minimum valley between two peaks representing the center of the foreground and background of lung regions. The scheme typically generates multiple segmented and frequently isolated regions, which are named ‘2D segments’. Some of the segments actually represent lung areas and others are related to non-lung areas and/or the CT table. To identify lung areas in each CT image and reconstruct a 3D lung volume, the second step of the scheme applies a 3D region growing method based on grouping of 2D segments. This step starts from a CT image that depicts the two largest 2D segments representing the left and right lungs that are typically located in the middle section of the two lungs. The scheme then separately scans upward and downward to the top (apices) and bottom (diaphragm) of the segmented lung areas. During this scanning process, the scheme gradually builds a 3D lung volume dataset by connecting 2D lung areas segmented on all paired adjacent images and deleting all other isolated (unconnected) 2D segments initially segmented on CT images. Finally, a rolling ball algorithm is applied to generate the smoothed lung segmentation boundary.

2.2.2. Selection of training samples—After lung segmentation, the next step of the CAD scheme is to scan the entire lung areas and classify each pixel depicted on the segmented lung areas into one of the two groups (positive or negative group for ILD). For this purpose, we need to assemble a diverse training dataset to train and optimize CAD scheme to recognize the difference of pixels between these two groups. Since the optimal performance can be achieved when the system is trained with the negative samples acquired from both negative cases and normal regions inside the positive cases containing ILD lesions, and the positive samples from the pixels being associated with ‘obvious’ (typical) categories of ILD lesions (Arzhaeva *et al* 2007), we manually selected 1146 negative pixels from normal lung tissue in four negative and four ILD verified (positive) cases, as well as 674 positive pixels located inside the obvious ILD regions circled by the radiologist in four positive cases for ILD. When a pixel was pointed by the computer mouse and selected as a sample, the information related to the pixel was saved into a sample file, which includes CT image file name, *x* and *y* coordinate of the pixel on the CT image, class (negative or positive for ILD), and the folder name where the CT image was saved. After building this sample file, a computerized scheme for the training purpose reads the file and retrieves the recorded (selected) pixel one-by-one. For each selected pixel, the scheme creates a square ROI centered at this pixel. The scheme then computes a set of texture features from the ROI. The

features computed from all training samples (ROIs) were normalized and saved into the feature file with the associated class (negative or positive for ILD).

2.2.3. Feature description and extraction—To classify a pixel at the center of ROI as the positive or negative for ILD, we computed 4 histogram-based statistical features (mean, variance, skewness and kurtosis) and 22 run-length matrix (RLM) texture features in the 0° and 90° direction from the ROI (Galloway 1975, Chu *et al* 1990, Dasarathy and Holder 1991). The reason for selecting these features is primarily based on the fact that the first-order gray-level statistics to characterize gray-level distribution within a ROI have been previously used and reported to be effective in the classification and quantification of ILD (Uppaluri *et al* 1999, Uchiyama *et al* 2003, Sluimer *et al* 2006, Arzhaeva *et al* 2007, Song *et al* 2009), as well as in the segmentation of lung areas with relatively severe ILD cases (Korfiatis *et al* 2008, Wang *et al* 2009). In addition, these computed histogram-based features derived from the lung CT images have also been found to correlate with the results of the pulmonary function tests (PFT) in the patients diagnosed with ILD (Best *et al* 2003).

The concept of run length matrix (RLM) texture features was introduced by Galloway (1975) to extract information of an image from its gray-level runs. A run is composed of consecutive pixels of the same gray level in a given direction. Various runs with different gray levels and their lengths are arranged according to the lengths and gray levels and construct a two-dimensional matrix called a RLM. Two new features were introduced by Chu *et al* (1990) using the distribution of gray levels of runs to complement the five earlier features. Four additional features were also proposed by Dasarathy and Holder (1991) to emphasize the joint distribution properties of the run lengths and gray levels. Texture analysis based on run lengths of gray levels has become a popular approach in many areas of medical image processing including the interpretation of ILD image patterns (Uppaluri *et al* 1999, Sluimer *et al* 2006, Arzhaeva *et al* 2007, Song *et al* 2009).

Since the values of these texture features vary as the change of ROI sizes, one of the most difficult issues in using these texture features to build an optimal classifier is to define an optimal ROI size. In this study, we conducted a series of experiments to select (test) five different ROI sizes (ranging from 9×9 to 63×63 pixels) and compare the final CAD testing performance when applying each of these ROIs.

2.2.4. Classifier optimization—Once a set of features was computed from each of selected training ROIs, we built and trained a multi-feature-based machine learning classifier to categorize each ROI (or centered pixel). Although many types of machine learning classifiers, such as an ANN, a support vector machine (SVM), and a Bayesian belief network (BBN), can be used for this purpose, we selected to build an optimal ANN-based classifier in this study. The ANN is a popular machine learning tool used in CAD schemes applying to the different medical images due to its advantages in learning the function to optimally approximate the relationship between the input features and desired classes using the relatively noisy or partially available training data (Park *et al* 2009b). The ANN built in this study contains three feed-forward layers (one input layer, one hidden layer, and one output layer) and uses a sigmoid function as the activation function. The ANN was trained by the back-propagation method that is one of the most commonly used approaches to minimize the total squared error of the ANN output.

To select a set of optimal features from the original feature pool of 26 features and determine the optimal number of hidden neurons of the ANN, we used a genetic algorithm (GA)-based optimization method that has been previously established and tested in our group for optimizing a variety of machine learning classifiers (including ANN) used in CAD schemes for mammograms (Zheng *et al* 2001). In brief, a specifically designed binary

coding method was used to define the GA chromosomes. Each chromosome has 30 genes in which the first 26 represent ROI features and the last 4 genes represent the number of hidden neurons (i.e. $M = 0101$ indicating five hidden neurons). Each GA chromosome determines a unique ANN structure including the input features and the number of hidden neurons. In the first 26 genes of the GA chromosome, the binary code of '1' indicates that the feature represented by this gene is selected as an input feature of the ANN and '0' indicates that the represented feature is discarded. Thus, in each GA chromosome, N features ($N \leq 26$) were selected and the rest of $26 - N$ features were discarded. Once a GA chromosome was initially randomly selected by the program, an ANN using N input neurons and M hidden neurons was trained using the training ROIs with a fixed training iterations (i.e. 500 in this study) and a large ratio of the momentum (0.95) to the learning rate (0.01) to minimize the potentially negative impact of over-training and improve scheme robustness. The ANN-generated classification accuracy was measured by the percentage of ROIs being correctly classified (or hit rate). The GA chromosomes that produce higher classification accuracy have higher probability to be selected to create the GA chromosomes in the next optimization generation using the cross-over and mutation operation aiming to find the composition of genes that are able to further improve ANN classification accuracy to distinguish between the positive and negative pixels (or ROIs) for ILD. The GA optimization process continued until either there is no further improvement of classification accuracy in the new generation or the searching generation reaches the predetermined maximum number (i.e. 300 in our studies). The final GA-selected chromosome (or feature set) was selected and used to build the optimal ANN.

2.2.5. Region growing-based pixel classification—Although the optimized ANN is able to classify each pixel depicted on the segmented lung areas into the positive or negative groups for ILD, it is a very time-consuming process to scan such a large number of segmented lung pixels and compute the corresponding texture features. To reduce processing time and increase computational efficiency of our CAD scheme, we investigated and applied a region growing-based pixel classification method. For each CT examination, a virtual mesh grid was applied to cover the segmented lung area. Instead of processing all pixels depicted on the segmented lung areas, only the pixels located in the intersections of the grid are used as the growth seeds. The size of the grid is 8×8 pixels, which was empirically determined based on our observation and experimental data analysis in this study. The scheme, first, detects whether a seed point is classified as positive for ILD. If the seed pixel is determined as negative, the scheme jumps to detect and classify the next seed point (the next intersection of mesh grid). Otherwise, the scheme applied region growing at the seed point that means neighbor pixels adjacent to the seed point are investigated until no positive pixel is found. In this way, the scheme can skip processing and classifying the high percentage of pixels inside the grid (98.4% or 63/64) that are negative for ILD.

2.2.6. Case-based detection score—Since the ultimate goal of our CAD scheme is to detect whether the patients have early ILD based on the computerized detection and analysis of low-dose CT examinations, we proposed a new measurement index (or the detection score of the CAD scheme) to predict the likelihood of the patients being positive for ILD. This index or detection score is computed by the ratio of the computed suspicious ILD volume (Vol_{ILD}) to the computed total lung volume (Vol_{Lung}) ranging from 0 to 100 as shown in equation (1):

$$\text{score} = \frac{\text{Vol}_{\text{ILD}}}{\text{Vol}_{\text{Lung}}} \times 100. \quad (1)$$

In general, the larger the detection score, the higher the likelihood of the case being associated with ILD. Similar to using the density mask (i.e. Hounsfield unit < -950 (Aziz *et al* 2005)) to detect and quantitatively assess the status (or severity) of emphysema using CT images, negative cases for ILD may still have a small fraction of pixels that are identified as positive for ILD, a threshold should be defined. If the detection score is higher than the predefined threshold (T), the test case is classified positive for ILD (abnormal case). Otherwise, the case is classified as negative for ILD.

2.3. Experimental procedure and performance evaluation

To select the effective size of the ROI surrounding an examined pixel for computing histogram and texture-based features, we tested and evaluated five sizes of ROIs including 9×9 , 15×15 , 21×21 , 33×33 , and 63×63 pixels in this study. Specifically, our CAD scheme was trained and tested five times using five different ROIs. In each training phase, the scheme first retrieved 1146 negative pixels and 674 positive pixels saved in sample files. The scheme then extracted a ROI from the corresponding CT image (which is centered on the retrieved pixel and uses the defined ROI size) and computed 26 image features. After building two image feature files (one includes 1146 negative ROIs and one includes 674 positive ROIs), we applied genetic algorithm (GA) to search for an optimal feature set among the initial pool of 26 features. This GA optimization process was repeated five times using five different ROIs (from 9×9 to 63×63 pixels). We then compared the classification accuracy of five ‘best performed’ ANNs selected by the GA during the five independent optimization processes and finally selected one that achieved the highest classification accuracy.

After implementing the optimal ANN into our CAD scheme, we tested CAD final case-based classification performance using the testing dataset of 30 CT examinations including 15 negative and 15 positive cases for ILD. The CAD scheme generated one detection (or classification) score for each test case. We then applied a receiver operating characteristic (ROC) data fitting and analysis program (ROCKIT (Metz 1998)) to generate a ROC curve and compute the area under the ROC curve (AUC) including its standard deviation based on the CAD-generated detection scores of all 30 testing cases. From the ROC curve, we also examined the detection sensitivity and specificity of the CAD scheme applying to this testing dataset.

3. Results

The relationship between the selected ROI size and computing time for feature extraction, the number of selected image features, as well as the highest classification accuracy acquired during the GA optimization process are summarized in table 1. The results show that as the size of the ROI increases from 9×9 to 63×63 pixels, the computing processing time to extract and compute 26 texture features from a ROI also monotonically increases from 3.89 to 9.02 ms using a desktop personal computer (AMD Athlon™ 64 × 2 Dual Core Processor 5000 + 2.61 GHz with 3.5GB of RAM). During the training phase, GA selected 8 to 18 features from the initial 26 features using the five ROI sizes. The classification accuracy (hit rate) of the ANN initially monotonically increases from 90.1% to 95.4% as the ROI size increases from 9×9 to 33×33 pixels. Then, the classification accuracy starts to degrade back to 91.7% as the ROI size increases to 63×63 pixels. Therefore, the ANN using features computed from the ROIs of 33×33 pixels achieved the superior classification accuracy (95.4%) over the other ANNs based on ROIs with the other four different sizes. This ANN was selected and implemented in our CAD scheme. In summary, the implemented optimal ANN has 11 input neurons and 10 hidden neurons. These 11 features include one histogram-based statistical feature and 10 run-length matrix texture features (table 2).

Our experiment with this testing dataset showed that the average number of pixels depicted on the segmented lung areas was 4 317 211 for a CT examination with slice thickness of 2.5 mm. When using the same desktop personnel computer applied in the training process, it took approximately 7 h to process and classify ROIs generated from 4 317 211 pixels depicted on the whole lung area. However, by applying the mesh-grid-based region growing method implemented in the CAD scheme, the average number of processed pixels were reduced to 280 292 (a 93.5% reduction) in the positive cases and the corresponding processing time was also reduced to less than 27.0 min for processing and classifying ROIs generated from 280 592 pixels. The results also show when applying to the negative cases in our dataset, the CAD scheme requires to averagely process and classify 91 074 pixels further reducing the CAD processing time to approximately 9 min for one case.

Figure 2 shows an example of applying our CAD scheme to automatically detect and segment the suspicious regions for ILD lesions (including both true-positive and false-positive regions). When applying to the entire testing dataset, the CAD scheme achieved the performance of $AUC = 0.884 \pm 0.064$ (figure 3). From the ROC curve, CAD can achieve 80.0% sensitivity at 85.7% specificity. The detailed data analysis also showed that distribution of detection scores for negative cases was narrow from 0.91 to 4.60. On the other hand, the detection scores of positive cases were widely spread out from 1.07 to 15.40 (figure 4). The results indicate that since these are early ILD cases, the overall percentage of the lung areas depicting ILD-type texture patterns are relatively small (<15.4%) in this testing dataset. If using a threshold on the detection score of 3.5, the CAD scheme correctly classified 26 of 30 CT examinations (86.7%), while misclassified 4 others (13.3%) including 1 negative case and 3 positive cases.

4. Discussion

In this study, we developed and tested a new CAD scheme that focused on detecting early ILD from asymptomatic cases. The study has a number of unique characteristics. First, unlike many of previously developed CAD schemes that focused on classifying certain types of ILD-related patterns depicted on the ROIs with the fixed size, our scheme classifies each pixel depicted on the segmented lung areas into one of the two groups (positive or negative for ILD) and thus identifies (segments) the ILD lesions in varying size and shape. Second, this CAD scheme aims to detect ILD at a relatively early stage using images acquired from the low-dose CT examinations that are originally used for lung cancer screening purpose. The CAD scheme was trained using the CT images that depict relatively mild ILD and it does not classify specific ILD patterns depicted on ROIs of CT images. Third, similar to the quantitative index used to detect and evaluate emphysema (using a density mask), we proposed a new quantitative scoring index (namely the ratio of the number of pixels classified as positive for ILD to the number of total pixels depicted on the entire segmented lung areas) to detect and assess ILD. Thus, the CAD scheme is able to generate one classification score to indicate the likelihood of a test case being positive for ILD. Fourth, our testing dataset included both positive and negative cases randomly acquired from a large pool of low-dose CT examinations for screening lung cancer. The results demonstrated that the CAD scheme achieved relatively high performance with the area under ROC curve ($AUC = 0.884 \pm 0.064$) showing the feasibility of applying CAD to detect early stage ILD using low-dose CT examinations that do not specifically target on the detection or diagnosis of ILD. In summary, considering that a majority of ILD patients present with advanced disease stages in the diagnosis and treatment (Rosas *et al* 2007), it is important to detect ILD at an early stage. Since the asymptomatic (or preclinical) ILD is detectable by CT examinations (Gochuico *et al* 2008), this study developed a new CAD scheme and demonstrated that CAD was able to identify asymptomatic ILD patients who underwent low-dose CT examinations for other purposes. Last, we emphasized that due to the different

application purposes and classification criteria, our CAD scheme did not directly compete with the previously reported CAD schemes for ILD and the performance level of our scheme was also not directly comparable to the previously reported results.

To develop an optimal and efficient CAD scheme for detecting ILD cases, we investigated and solved several technical difficulties in this study. First, for the texture computation and analysis of ILD patterns, defining an optimal ROI size is an important issue because the small ROIs may not include enough texture information and the large ROIs may also fail to describe texture characteristics due to the unrelated normal lung tissues. Moreover, it also takes more computational time to extract texture information from the large ROIs. Thus, it is important to investigate and identify the optimal size of ROI that enables us to effectively represent ILD texture. In the experiments, we tested and evaluated five ROIs with different sizes and selected the most superior one (33×33 pixels) based on the performance and computational efficiency. Second, since an extremely large number of pixels is involved in the segmented lung areas (volume) and the majority of the pixels are negative (not associated with ILD) in the negative or early ILD cases, processing and classifying each pixel depicted on the segmented lung areas is very inefficient and unnecessary. To solve this problem, we developed and tested a special region growing method with seed points selected at the intersections of a virtual mesh-grid. Using this method, it substantially reduces the number of processing pixels and thus increases computational efficiency.

Since low-dose CT has been widely investigated and considered as a promising imaging modality in screening or early detecting lung nodules or cancer, the another potential benefit of developing this CAD scheme is to help assess the risk of patients in developing lung cancer. For example, one study showed that patients with diffuse ILD had an exceedingly high occurrence of lung cancer as compared with patients without ILD (53% versus 15%, $p < 0.001$ (Katabami *et al* 2000)). Studies also found that lung cancer patients with ILD showed markedly higher postoperative pulmonary morbidity and mortality than those without ILD. As a result, careful preoperative evaluation and postoperative management are required to achieve optimal surgical outcome in lung cancer patients who also have ILD (Kawasaki *et al* 2002, Fujimoto *et al* 2003). The fact that lung cancer is frequently detected in the lower lobes of the lung and often in the vicinity of pre-existing scars, where fibrosis is predominant, supports the association between ILD and lung cancer (Bouros *et al* 2002). Although a large number of CAD schemes have been developed to detect lung cancer (or lung nodules) depicted on low-dose CT images (Wiemker *et al* 2005), CAD schemes for detecting ILD depicted on the same low-dose CT images have not been previously developed and reported. Thus, our CAD scheme can be a potentially valuable tool for the early detection of ILD when the patients go to screen for lung cancer. Although the patients may not have lung cancer at the current stage of the low-dose CT screening examination, if the ILDs are detected, these patients may be classified into a group with high risk of developing lung cancer. Thus, detecting mild ILD results in early treatment or monitoring of the diseases, which may significantly improve lung function as measured by the pulmonary function tests (Schnabel *et al* 2003) and reduce the risk of developing lung cancer (Bouros *et al* 2002).

In summary, we demonstrated in this study a new concept and approach in developing CAD schemes that aim to detect early ILD from asymptomatic participants using low-dose CT examinations. Although the results are encouraging, this is a very preliminary study with a number of limitations. First, since the performance of CAD schemes for detecting ILD considerably relies on the lung segmentation, our rolling ball algorithm may fail to segment (include) a fraction of diseased lung areas near the lung boundary (in particular for ILD at advanced disease stages). However, because our CAD scheme focused on the detection of mild (early stage) ILD, such impact is very minimal. Second, due to artificial noises, the

scheme may detect false positive regions on the top (apices) and bottom (diaphragm) slices of the chest CT examinations, as well as between the vessel and normal lung tissue (figure 2). How these false-positive regions affect the overall case-based classification results needs to be further investigated. Third, we only applied and tested the mesh-grid-based ROIs with the fixed size. Whether using adaptively adjusted or scaled ROI size can significantly improve computational efficiency while maintaining high classification accuracy needs to be further investigated. Fourth, the testing dataset used in this study is limited. The performance and robustness of our scheme needs to be further tested using larger and more diverse datasets in the future studies. Fifth, due to the focus of this study, our CAD scheme does not attempt to classify textures between different types of ILD patterns and to compute the score of ILD extent (severity). Last, a simple classification index (detection score) based on the size ratio of the ILD positive regions detected and segmented from one CT examination was used in our CAD scheme to classify between the positive and negative cases. Whether extracting the other image features from the detected ILD regions and developing a multi-feature-based machine learning classifier can further improve CAD performance have not been investigated in this study. Despite these limitations, this preliminary study demonstrated the feasibility of developing a CAD scheme to detect early ILD from the asymptomatic participants using low-dose CT examinations. More research work is still needed to build and test the highly performed and robust CAD schemes before CAD can be used in the clinical practice to help radiologists more accurately and consistently detect and quantitatively assess ILD at an early stage.

Acknowledgments

This work is supported in part by grants P50 HL084948 and R01 HL085096 from the National Heart, Lung, and Blood Institute, National Institutes of Health, USA, to the University of Pittsburgh.

References

- Arzhaeva Y, et al. Computer-aided detection of interstitial abnormalities in chest radiographs using a reference standard based on computed tomography. *Med. Phys.* 2007; 34:4798–4809. [PubMed: 18196808]
- Aziz ZA, et al. Functional impairment in emphysema: contribution of airway abnormalities and distribution of parenchymal disease. *Am. J. Roentgenol.* 2005; 185:1509–1515. [PubMed: 16304005]
- Best AC, et al. Quantitative CT indexes in idiopathic pulmonary fibrosis: relationship with physiologic impairment. *Radiology.* 2003; 228:407–414. [PubMed: 12802000]
- Biederer J, et al. Correlation between HRCT finding, pulmonary function tests and bronchoalveolar lavage cytology in interstitial lung disease associated with rheumatoid arthritis. *Eur. Radiol.* 2004; 14:272–280. [PubMed: 14557895]
- Bjoraker JA, et al. Prognostic significance of histopathologic subsets in idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* 1998; 157:199–203. [PubMed: 9445300]
- Bourke SJ. Interstitial lung disease: progress and problems. *Postgrad. Med. J.* 2006; 82:494–499. [PubMed: 16891438]
- Bouros D, Hatzakis K, Labrakis H, Zeibecoglou K. Association of malignancy with diseases causing interstitial pulmonary changes. *Chest.* 2002; 121:1278–1289. [PubMed: 11948064]
- Chu A, Sehgal CM, Greenleaf JF. Use of gray value distribution of run lengths for texture analysis. *Pattern Recognit. Lett.* 1990; 11:415–420.
- Dasarathy BR, Holder EB. Image characterizations based on joint gray-level run-length distributions. *Pattern Recognit. Lett.* 1991; 12:497–502.
- Fujimoto T, Okazaki T, Matsukura T, Hanawa T. Operation for lung cancer in patients with idiopathic pulmonary fibrosis: surgical contraindication. *Ann. Thorac. Surg.* 2003; 76:1674–1679. [PubMed: 14602310]

- Galloway MM. Texture analysis using gray level run lengths. *Comput. Graphics Image Process.* 1975; 4:172–179.
- Gochuico BR, et al. Progressive preclinical interstitial lung disease in rheumatoid arthritis. *Arch. Int. Med.* 2008; 168:159–166. [PubMed: 18227362]
- Goh NS, et al. Interstitial lung disease in systemic sclerosis: a simple staging system. *Am. J. Respir. Crit. Care Med.* 2008; 177:1248–1254. [PubMed: 18369202]
- Katabami M, et al. Pneumoconiosis-related lung cancers: preferential occurrence from diffuse interstitial fibrosis-type pneumoconiosis. *Am. J. Respir. Crit. Care Med.* 2000; 162:295–300. [PubMed: 10903257]
- Kawasaki H, et al. Postoperative morbidity, mortality, and survival in lung cancer associated with idiopathic pulmonary fibrosis. *J. Surg. Oncol.* 2002; 81:33–37. [PubMed: 12210025]
- King TE. Clinical advances in the diagnosis and therapy of the interstitial lung diseases. *Am J Respir Crit Care Med.* 2005; 172:268–279. [PubMed: 15879420]
- Korfatis P, et al. Texture classification-based segmentation of lung affected by interstitial pneumonia in high-resolution CT. *Med. Phys.* 2008; 35:5290–5302. [PubMed: 19175088]
- Leader JK, et al. Pulmonary nodule detection with low-dose CT of the lung: agreement among radiologists. *Am. J. Roentgenol.* 2005; 185:973–978. [PubMed: 16177418]
- Marten K, et al. Interstitial lung disease associated with collagen vascular disorders: disease quantification using a computer-aided diagnosis tool. *Eur. Radiol.* 2009; 19:324–332. [PubMed: 18726597]
- Metz, CE. ROCFIT 0.9B Beta version. Chicago, IL, USA: University of Chicago; 1998. <http://www-radiology.uchicago.edu/krl/>
- Park SC, et al. Pulmonary airway tree segmentation using adaptive volume of interest in CT images. *Proc. SPIE.* 2009a; 7259:72593U.
- Park SC, Pu J, Zheng B. Improving performance of computer-aided detection scheme by combining results from two machine learning classifiers. *Acad. Radiol.* 2009b; 16:266–274. [PubMed: 19201355]
- Pastorino U. Lung cancer screening. *Br. J. Cancer.* 2010; 102:1681–1686. [PubMed: 20424610]
- Rosas IO, et al. Early interstitial lung disease in familial pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* 2007; 176:698–705. [PubMed: 17641157]
- Schaefer-Prokop C, Prokop M, Fleischmann D, Herold C. Highresolution CT of diffuse interstitial lung disease: Key findings in common disorders. *Eur. Radiol.* 2001; 11:373–392. [PubMed: 11288840]
- Schnabel A, et al. Interstitial lung disease in polymyositis and dermatomyositis: clinical course and response to treatment. *Semin. Arthritis Rheum.* 2003; 32:273–284. [PubMed: 12701038]
- Sluimer IC, Prokop M, Hartmann I, van Ginneken B. Automated classification of hyperlucency, fibrosis, ground glass, solid, and focal lesions in high-resolution CT of the lung. *Med. Phys.* 2006; 33:2610–2620. [PubMed: 16898465]
- Song G, et al. A comparative study of HRCT image metrics and PFT values for characterization of ILD and COPD. 2nd Int. Workshop on Pulmonary Image Analysis (London, UK). 2009
- Uchiyama Y, et al. Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography. *Med. Phys.* 2003; 30:2440–2454. [PubMed: 14528966]
- Uppaluri R, et al. Computer recognition of regional lung disease patterns. *Am. J. Respir. Crit. Care Med.* 1999; 160:648–654. [PubMed: 10430742]
- Wang J, Li F, Li Q. Automated segmentation of lungs with severe interstitial lung disease in CT. *Med. Phys.* 2009; 36:4592–4599. [PubMed: 19928090]
- Wiemker R, Rogalla P, Blaffert T. Aspects of computer-aided detection (CAD) and volumetry of pulmonary nodules using multislice CT. *Br. J. Radiol.* 2005; 78:S46–S56. [PubMed: 15917446]
- Zheng B, Chang YH, Good WF, Gur D. Performance gain in computer-assisted detection schemes by averaging scores generated from artificial neural networks with adaptive filtering. *Med. Phys.* 2001; 28:2302–2308. [PubMed: 11764037]
- Zheng B, et al. Automated detection and classification of interstitial lung diseases from low-dose CT images. *Proc. SPIE.* 2004; 5370:849–856.

Figure 1.

The general processing steps of the CAD scheme to classify each pixel into either negative or positive group for ILD.

Figure 2.

Examples of partial detection results using the CAD scheme. Two images in the first column and images in the second column are the original images selected from different CT scans for demonstration and the images marked with CAD-generated results (including the boundary contour of the segmented lung areas and identified ILD lesions), respectively. The white circles in the result images indicate false positives because the tissue around vessel is similar to the ILD texture pattern.

Figure 3.
A ROC-type performance curve for applying the CAD scheme to 30 testing CT examinations.

Figure 4. Distributions of detection score calculated by the CAD scheme applied to negative (normal) and positive (abnormal) CT examinations.

Table 1

Summary of performances (accuracy rate and processing time) of ANN independently using various sizes of ROIs and an optimal feature sets selected by the genetic algorithm.

	Processing time (ms/ROI)	Optimal feature dimension	ANN accuracy rate (%)
9 × 9	3.79	8	90.1
15 × 15	4.20	18	92.9
21 × 21	4.81	14	94.0
33 × 33	5.78	11	95.4
63 × 63	9.02	11	91.7

Table 2

The list of the features extracted and computed from each ROI.

Histogram-based statistical features		4		Mean of the histogram
				Variance of the histogram [*]
				Skewness of the histogram
				Kurtosis of the histogram
Run-length matrix (RLM) texture features	0°	11	22	Short run emphasis
				Long run emphasis [*]
				Gray-level non-uniformity [*]
				Run length non-uniformity [*]
				Run percentage
				Low gray-level run emphasis
				High gray-level run emphasis [*]
				Short run low gray-level emphasis
				Short run high gray-level emphasis
				Long run low gray-level emphasis [*]
				Long run high gray-level emphasis
	90°	11		Short run emphasis
				Long run emphasis
				Gray-level non-uniformity [*]
				Run length non-uniformity
				Run percentage [*]
				Low gray-level run emphasis
				High gray-level run emphasis
				Short run low gray-level emphasis [*]
				Short run high gray-level emphasis [*]
				Long run low gray-level emphasis [*]
				Long run high gray-level emphasis

The optimal features selected by genetic algorithm for the optimal ANN using a ROI with 33×33 pixels are shown with (*) marks.