SYSTEMS-LEVEL QUALITY IMPROVEMENT

# Automatic Lung Segmentation Using Control Feedback System: Morphology and Texture Paradigm

**Norliza M. Noor · Joel C. M. Than · Omar M. Rijal ·
Rosminah M. Kassim · Ashari Yunus · Amir A. Zeki ·
Michele Anzidei · Luca Saba · Jasjit S. Suri**

**Abstract** Interstitial Lung Disease (ILD) encompasses a wide array of diseases that share some common radiologic characteristics. When diagnosing such diseases, radiologists can be affected by heavy workload and fatigue thus decreasing diagnostic accuracy. Automatic segmentation is the first step in implementing a Computer Aided Diagnosis (CAD) that will help radiologists to improve diagnostic accuracy thereby reducing manual interpretation. Automatic segmentation proposed uses an initial thresholding and morphology based segmentation coupled with feedback that detects large deviations with a corrective segmentation. This feedback is analogous to a control system which allows detection of abnormal or severe lung disease and provides a feedback to an online segmentation improving the overall performance of the system. This feedback system encompasses a texture paradigm. In this study we studied 48 males and 48 female patients consisting of 15 normal and 81 abnormal patients. A senior radiologist chose the five levels needed for ILD diagnosis. The results of segmentation were displayed by showing the comparison of the automated and ground truth boundaries (courtesy of

---

This article is part of the Topical Collection on *Systems-Level Quality Improvement*

N. M. Noor (✉)
Department of Engineering, UTM Razak School of Engineering and Advanced Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
e-mail: norliza@utm.my

J. C. M. Than
UTM Razak School of Engineering and Advanced Technology, UniversitiTeknologi Malaysia, Kuala Lumpur, Malaysia
e-mail: tcmjoel2@live.utm.my

O. M. Rijal
Institute of Mathematical Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia
e-mail: omarrija@um.edu.my

R. M. Kassim
Department of Diagnostic Imaging, Kuala Lumpur Hospital, Kuala Lumpur, Malaysia
e-mail: rosminahmk@gmail.com

A. Yunus
Institute of Respiratory Medicine, Kuala Lumpur, Malaysia
e-mail: ashdr64@yahoo.com.au

A. A. Zeki
UC Davis School of Medicine, Sacramento, CA, USA
e-mail: amir.zeki@ucdmc.ucdavis.edu

M. Anzidei
Department of Radiological Sciences, Policlinico Umberto I, University of Rome, Rome, Italy
e-mail: micheleanzidei@gmail.com

L. Saba
Azienda Ospedaliero Universitaria (A.O.U.) di Cagliari – Polo di Monserrato; Università di Cagliari, s.s. 554, Monserrato, Cagliari 09045, Italy
e-mail: lucasabamd@gmail.com

J. S. Suri
Global Biomedical Technologies, Roseville, CA, USA
e-mail: jsuri@comcast.net

J. S. Suri
AtheroPoint™ LLC, Roseville, CA, USA

J. S. Suri
Electrical Engineering Department (Affl.), Idaho State University, Pocatello, ID, USA

ImgTracer™ 1.0, AtheroPoint™ LLC, Roseville, CA, USA). The left lung's performance of segmentation was 96.52 % for Jaccard Index and 98.21 % for Dice Similarity, 0.61 mm for Polyline Distance Metric (PDM), −1.15 % for Relative Area Error and 4.09 % Area Overlap Error. The right lung's performance of segmentation was 97.24 % for Jaccard Index, 98.58 % for Dice Similarity, 0.61 mm for PDM, −0.03 % for Relative Area Error and 3.53 % for Area Overlap Error. The segmentation overall has an overall similarity of 98.4 %. The segmentation proposed is an accurate and fully automated system.

**Keyword** Lung segmentation · High resolution computed tomography thorax (HRCT) · Interstitial lung disease

## Introduction

Interstitial lung disease (ILD) encompasses a wide group of disorders [1]. Although ILD has a large array of diseases, these diseases share some common features in terms of clinical, radiologic and physiologic features. These disorders mostly have a common trait of progressive scarring or fibrosis of the lung tissue. These scaring alter oxygen gas exchange at the lungs thus affecting a person's ability to breathe normally. ILD affects the alveolar structures, pulmonary interstitum, and small terminal airways [2]. ILD also has many possible causes which will not be discussed in detail [3–5]. In ILD, lung volumes are reduced resulting in restrictive physiology. In addition to reductions in lung volumes diffusing capacity is also reduced [2].

High Resolution Computed Tomography (HRCT) of the lung is used in the diagnosis of Interstitial Lung Diseases (ILD). HRCTs are often interpreted by qualified radiologists and pulmonologists. The calculation of lung mass from CT scans is an accepted non-invasive method for determining lung tissue mass [6]. ILD can be analyzed using limited thin sections of HRCT slices called levels which are located based on anatomical landmarks, hence all our processing is based on these five levels as marked by radiologists.

Due to large number of CT images, radiologists have an interest in adapting Computer Aided Diagnosis (CAD)-based systems which can assist them in diagnostic evaluations [7–9]. There is a need for computer assistance for clinicians and radiologists arise. In a study by Beyer et al. [10], concluded that a CAD system could expedite the reading of chest CT cases for pulmonary nodules without relevant loss of sensitivity when used as a Concurrent Reader (CR) with a radiologist. Segmentation is one of the preliminary and crucial steps in the development of a CAD system to help radiologists [11].

Segmentation of structures in medical images is challenging due to several factors which include anatomical differences, abnormalities in lung tissue, image noise, and differences in acquisition parameters [12]. In abnormal lung images the inconsistencies can be in the form of ground glass opacities where there is increased attenuation of signal in the lung, which is caused by partial filling of lung parenchyma. In some advanced stages of the disease there can also be "honeycombing" where lung tissue is destroyed, fibrotic and contains multiple cystic airspaces with thick fibrous walls. These inconsistencies may affect the outcome of a segmentation algorithm causing low performance. Other efforts combat these problems were developed with specific diseases in mind and may not actually be effective on a wider database [13]. Thus lung segmentation specifically designed for abnormal lungs seems to be the solution that others have approached [14].

Thus our study aims to propose a system to segment lungs accurately and consistently over five levels of the lung for both healthy and diseased lungs. The novelty of this study lies in the development of the feedback system which is analogous to a control system which allows detecting abnormal or severe lung disease and provides feedback that encompasses texture paradigm to an online segmentation improving the overall performance of the system. The segmentation system uses an initial segmentation based on statistical threshold and mathematical morphology. The segmentation was compared with tracings obtained from a trained individual with knowledge of lungs. This comparison or feedback and corrective segmentation will help deal with the challenges of segmentation mentioned earlier. The large deviations obtained from the feedback would indicate severe cases of lung disease. These deviations would then be corrected using segmentation based on texture where the entropy of each pixel is used. The robustness and effectiveness of the segmentation is seen when it is applied to both normal and abnormal lungs and across five different levels representing the entire lung.

## Data acquisition

Ninety-six patients' HRCT Thorax images were obtained retrospectively from the Department of Diagnostic Imaging of Kuala Lumpur Hospital with ethical consent. These images consist of 15 healthy individuals (normal cases), 28 ILD cases and 53 other lung related diseases cases termed as non-ILD. In this study, the 28 ILD cases and 53 other lung related diseases were also combined together as one category termed as diseased cases. There were 48 males and 48 female samples. Images from all 96 patients were studied and evaluated for automatic lung segmentation. The HRCT scanner used was the Siemens SomatomPlus4 CT scanner. Each slice was obtained at 10 mm intervals in supine position will full suspended inspiration of the lung. This resulted approximately 30 HRCT Thorax image slices per patient. All the images are in DICOM format. A radiologist was assigned to view the

slices using SyngoFastView version VX57G27. The senior radiologist then individually determined the five slices of the HRCT Thorax image at five predetermined levels for each patient by viewing all the slices available and filling up a survey to determine the fulfillment of the criteria of five slices. The criteria of choosing these five predetermined levels of the HRCT are based on anatomic landmarks and represent the entire lung area from top area to the bottom. The levels and their corresponding landmarks are; level 1: aortic arch, level 2: trachea carina, level 3: pulmonary hilar, level 4: pulmonary venous confluence and level 5: 1 to 2 cm above the dome of right hemi-diaphragm. The five predetermined slices per patient will decreases the amount of slices being analyzed because it is able to represent the entire lung from top to bottom with five predetermined slices by the radiologist as shown in Figs. 1 and 2. The black arrows show the lungs.

## Methodology

Our primary goal is to establish an automated system for left and right lung segmentation in HRCT images. Automated systems are always susceptible to errors due to large variations in image data sets such as contrast, resolution, disease type, voluminous data size, different tissue types and variations in scanning protocols are some to name [15]. In such challenges, interactive paradigms play a vital role. Commercial regulatory systems these days require that the system be semi-interactive or semi-automated [16]. Interactive systems can correct, but cannot spot automatically cases which have high errors. We have designed an automated online system which can not only just automatically correct, but automatically spot large deviations. Such deviations are like a feedback control system where the online outputs are automatically compared against the ideal scenario which is the manual tracing by an expert and then undergo correction thereby improving the overall performance of the system. Our automated system is crude and fast based on global class separation infrastructure, while feedback control analogous system is more refined and follows local processing such as texture paradigm.

### Overall system

Figure 3 shows the overall system in a flowchart. The HRCT input undergoes global processing which involves the automatic segmentation based on tissue class segregation using regional statistics in global space. Using the analogy of control system theory, the comparator allows identification of the deviation between the intermediate output and trained system which already has information about the goal state. The comparator allows spotting of the large deviations in the intermediate output which is then fed to the feedback loop for local processing system to correct them. Thus our overall system has unique characteristics that not only spots large deviations but also corrects them. Another advantage of such a system is ensuring that the automated system does not reach the state of failures in real life and has a trained system ready for mitigation. Due to the iterative nature of the feedback system, the comparator has criteria to establish the refined output if the error control is below the defined threshold $\Delta_{Th}$. The overall novelty of such a system is a combination of fast global processing utilizing the global parameters and refined local processing in local space using a feedback control system. The complexity of system is low since the local processing is limited to only cases which have large deviations. The system is finally robust due to the knowledge derived from the trained system.

### Global system

The idea behind the global system is to capture the shape of the lung by differentiating the lung from its background. The regional information has statistics associated with it which is based on Hounsfield units (HU) generated from the X-ray attenuation. These HU values have specific ranges for the lung region on a global scale. These HU can be well captured by considering a two-class paradigm. Thus, our objective is to develop a class segregation system which can pick up regional statistics considering the two-class problem. A simple approach can be a threshold criteria embedded with statistical means and standard deviations. A fast and robust method such as classical Otsu threshold paradigm can be adapted for our model development [17].

Even though, such a threshold scheme has not be a novel contribution in the proposed work, but it adds as a component to fetch the global lung shape incorporating deviations which are corrected by our novel feedback system leading to final estimated accurate borders. Using such a framework, if $\omega_i$ represents the probability that the two classes are separated by threshold ($t$), $\sigma_i^2$ represents the variance of the classes,

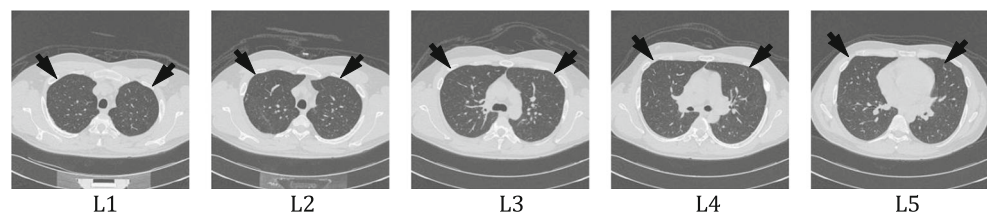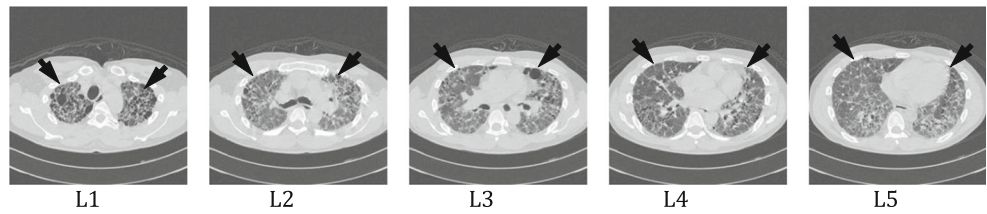**Fig. 1** Five levels of HRCT with *left lung* and *right lung* (normal)



| L1 | L2 | L3 | L4 | L5 |

**Fig. 2** Five levels of HRCT with *left lung* and *right lung* (abnormal)



| L1 | L2 | L3 | L4 | L5 |

the Otsu paradigm can lead to the formation of equation for optimal threshold computation and mathematically given as:

$$\sigma_\omega{}^2 = \omega_1(t) \ \sigma_1{}^2(t) + \omega_2(t) \ \sigma_2{}^2(t) \tag{1}$$

Thus we can get the optimum threshold given by $T_{opt}$ when it fits the criteria of the minimum variance of the classes in the Eq. 2. With a given input image as Fig. 4a, using $T_{opt}$, it is possible to the separate the non-body represented by black pixels from the body represented by white pixels in Fig. 4b.

$$\sigma_\omega^2 = \omega_1(T_{opt})\sigma_1^2(T_{opt}) + \omega_2(T_{opt})\sigma_2^2(T_{opt}) \tag{2}$$

To completely remove the background tissues, the global system requires an iterative process to establish a complete isolation of lung region in the CT lung image. This iterative threshold is empirically determined based on bias of the global Otsu threshold. Such a refinement will ensure that all regions not relevant to the lung region are eradicated for quantification. We call this threshold as $T_{emp}$ and it was defined for our database as 324 for our dynamic range. $T_{emp}$ is applied and the lung region is represented in the body as shown in Fig. 4c. The last stage of the global shape extraction system consists of smoothing and cleaning using binary morphology which consists of dilation followed by erosion. This morphological cleaning results in Fig. 4d. This consists of fundamental dilation and erosion equation expressed as Eqs. 3 and 4 respectively:

$$I \oplus H = \ \left\{ z \in E \middle| (H^S)_Z \ \cap \ I \neq \varnothing \right\} \tag{3}$$

$$I \ominus H = \ \left\{ z \in E \middle| H_Z \ \subseteq \ I \ \right\} \tag{4}$$

where $E$ is a Euclidean space or an integer grid, and I a binary image in $E$. $H^S$ is the symmetric of $H$. $(H^S)_Z$ is the translation of $H^S$ by the vector $z$. $H$ is the square structuring element $3 \times 3$ size. $H_Z$ is the translation of $H$ by the vector $z$.

Connected component analysis is used to detect the lung region in the binary image shown in Fig. 4e. There should be two large regions, one larger than the other. In some cases where the lungs are in close proximity, they will be grouped as one. The region where the two lungs connect is usually the lowest pixel width of lung region. To solve this, dynamic programming is done by calculating the lowest pixel width of each column to locate the region of separation. Pixels with the highest contrast are selected to be region of splitting between two lungs. Once two lungs are detected, erosion and dilation with the same structure element are again done to smoothen the boundary of the lungs and the boundaries are labelled in green boundaries for right and left lung in Fig. 4f.

## Local system

As discussed in introduction, segmentation of structures in medical images is challenging due to several factors which include anatomical differences, abnormalities, image noise, and differences in acquisition parameters [12]. As a result of the above challenges, it is always advisable to have a human trained system which can act as a tool to provide the correction to the global system challenges. Such a system was presented in Fig. 3 using the control system analogy by providing the feedback to correct the global challenges using the local system. The local system we provided uses the local
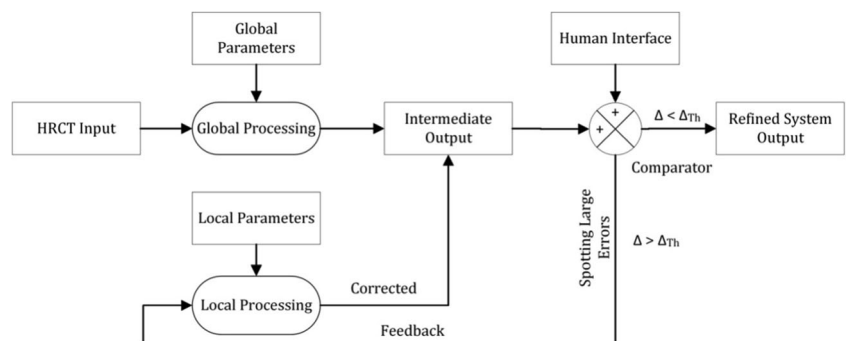
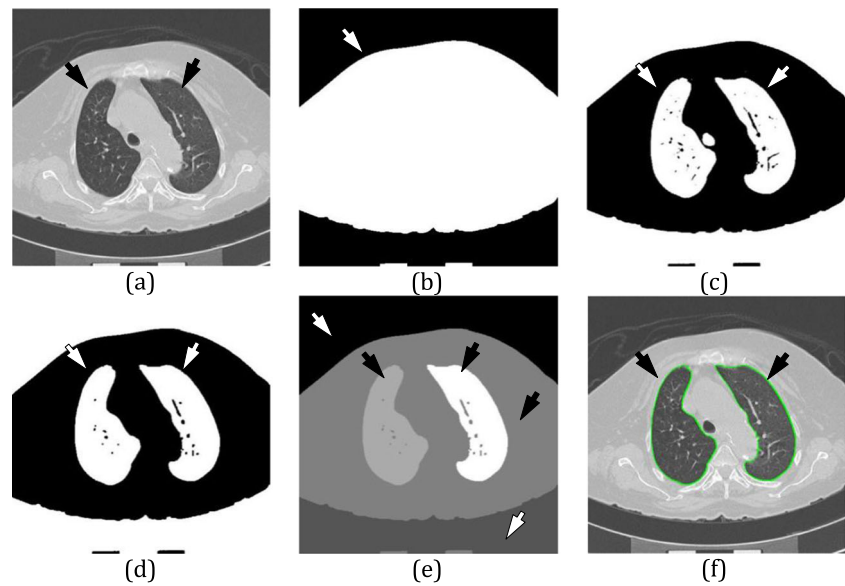**Fig. 3** System design using control system analogy

**Fig. 4** Overview of global processing: (**a**) Input image (**b**) Mask image (**c**) Iterative threshold output (**d**) Morphological cleaning (**e**) Connected component analysis (**f**) Labelled image



characteristics of the lung region. Such characteristics had two motives: (a) to classify the normal vs. diseased lungs and (b) able to automatically track the borders of the left and right lungs so deviations can be traced against the human trained system. Since the nature of the tissues in the diseased lung and normal lung could better be represented by aggressiveness of the tissue, we used the fundamental property of surface randomness to segregate the normal vs. disease lung. Thus we adapted a texture paradigm which had the property of tracing the tissues of diseased lung compared to normal lung. This texture was best adapted using the textured or entropy of the pixels in lung regions. We thus modelled this using entropy of the image as and defined as:

$$Entropy = -\sum_i P_i \log_2 P_i \qquad (5)$$

where $P_i$ is the probability that the difference between two adjacent pixels is equal to $i$, and $Log_2$ is the base 2 logarithm.

This was the distinguishing feature which leads to the correction of global weakness. To bring back in the classical framework of morphology, we have to compute the regional characteristics of the lung region followed by binarization. It was empirically computed threshold on the texture image leading to the clinical binary paradigm ($\vartheta_{Under}$ - Binary threshold for lung region texture under segmentation) and ($\vartheta_{Over}$ - Binary threshold for lung region texture over segmentation) followed by same clinical noise reduction and CCA. Though the feedback system offered the advantage of correction, it did require the automated spotting of the global system weakness. This can be done by the comparator system which provides the human trained system, in which consisted of the database of human trained borders from the human intervention system.

The effect of the local system is showed in the shift of lung region symbolised by the green borders and arrow in Fig. 5, when compared to the global system.
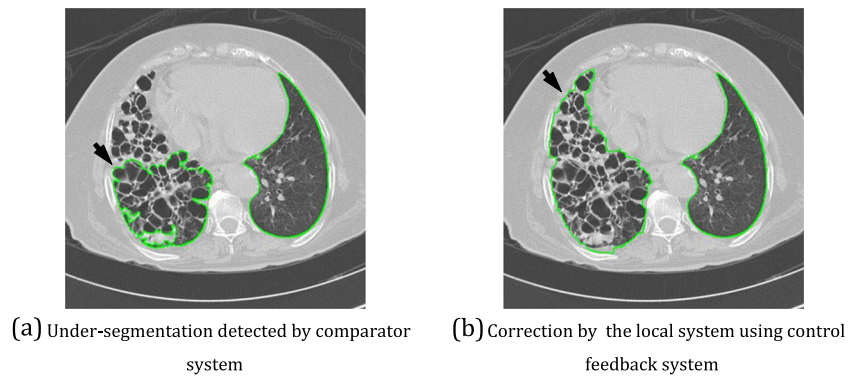
## Results

The process of segmentation involves computing the borders of left and right lung during global processing, local processing using feedback system and combined effect of global and local processing. Such a paradigm can be depicted visually and quantitatively evaluated. The main benefit for visually examination is to share the accuracy of automated method vs. manual tracings. Further, the clinical value of the quantification is to primarily compare normal vs. diseased subjects and understand the distribution. We use the variable "area" for our quantitative evaluation. We demonstrate the relationship between the automated lung area computed by the algorithm against the area computed by manually tracing the lungs using ImgTacer™, Global Biomedical Technologies, Inc., Roseville, CA, USA.

System's segmentation results

Figure 6 demonstrates the segmentation results for the right lung, where automated output (green) can be displayed with the manual lung borders (red). It shows 4x5 matrix representation, where the column represents five levels and four rows corresponds to four different subjects. Qualitatively, Fig. 6 shows encouraging results representing the close precision and accuracy between automated method and manual tracings. A similar display can be seen in Fig. 7 for the left lung. It is interesting to note that the automated segmentation follows the groves (like a jaw) shown by black arrows. On the

**Fig. 5** Feedback control system showing the segmentation correction using local system



(a) Under-segmentation detected by comparator system

(b) Correction by the local system using control feedback system

other side, the outer borders (like a fat belly) are also well followed between the automated system and the manual tracings.

An important observation of the results is that the segmentation method was able to yield high accuracy segmentation for all five levels showing the consistency as well as the accuracy of the method. In the expanded view in Fig. 8, a closer look is taken to see the high performance of the segmentation method proposed. The green arrows point to the segmentation boundary (green). The red arrows point to the ground truth

**Fig. 6** Overlays of *right lung* with automatic segmentation border (*green color*) and ground truth border (*red color*)
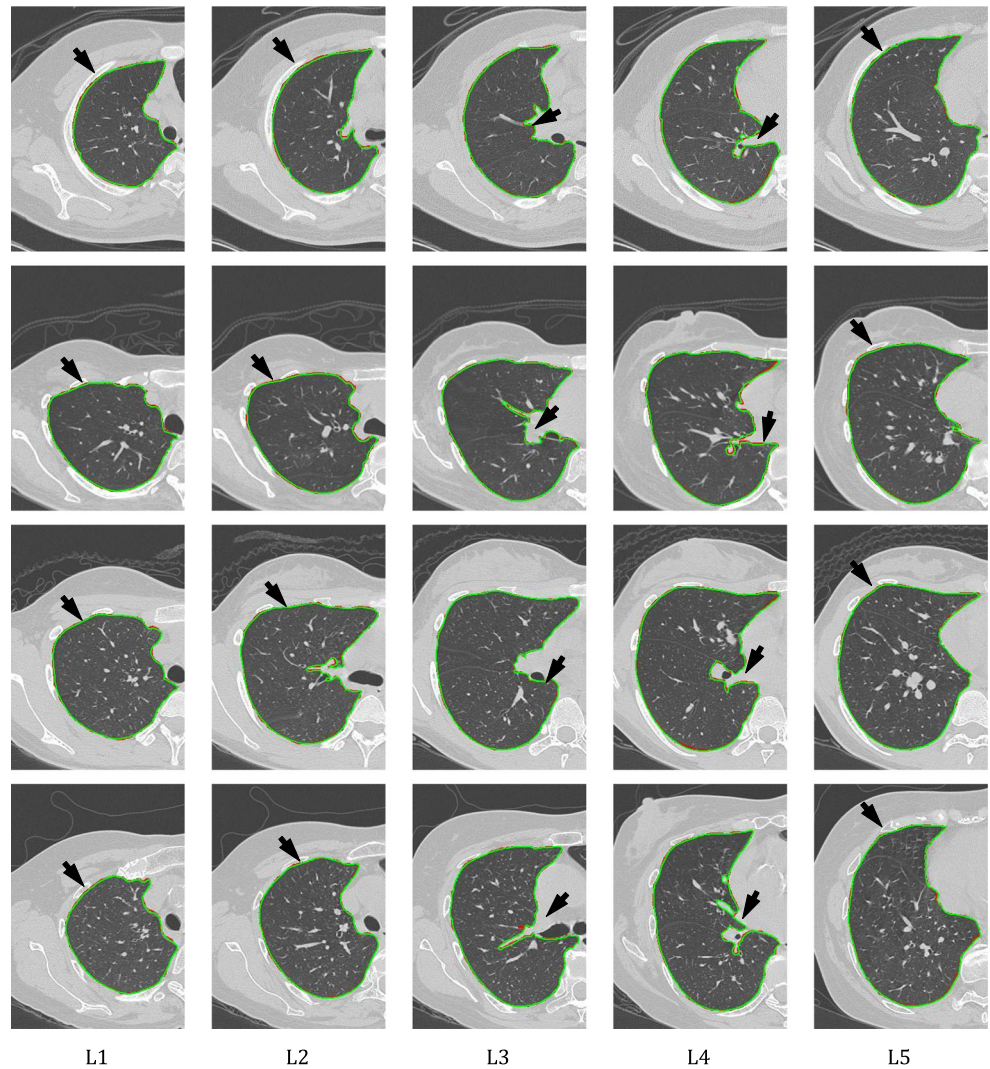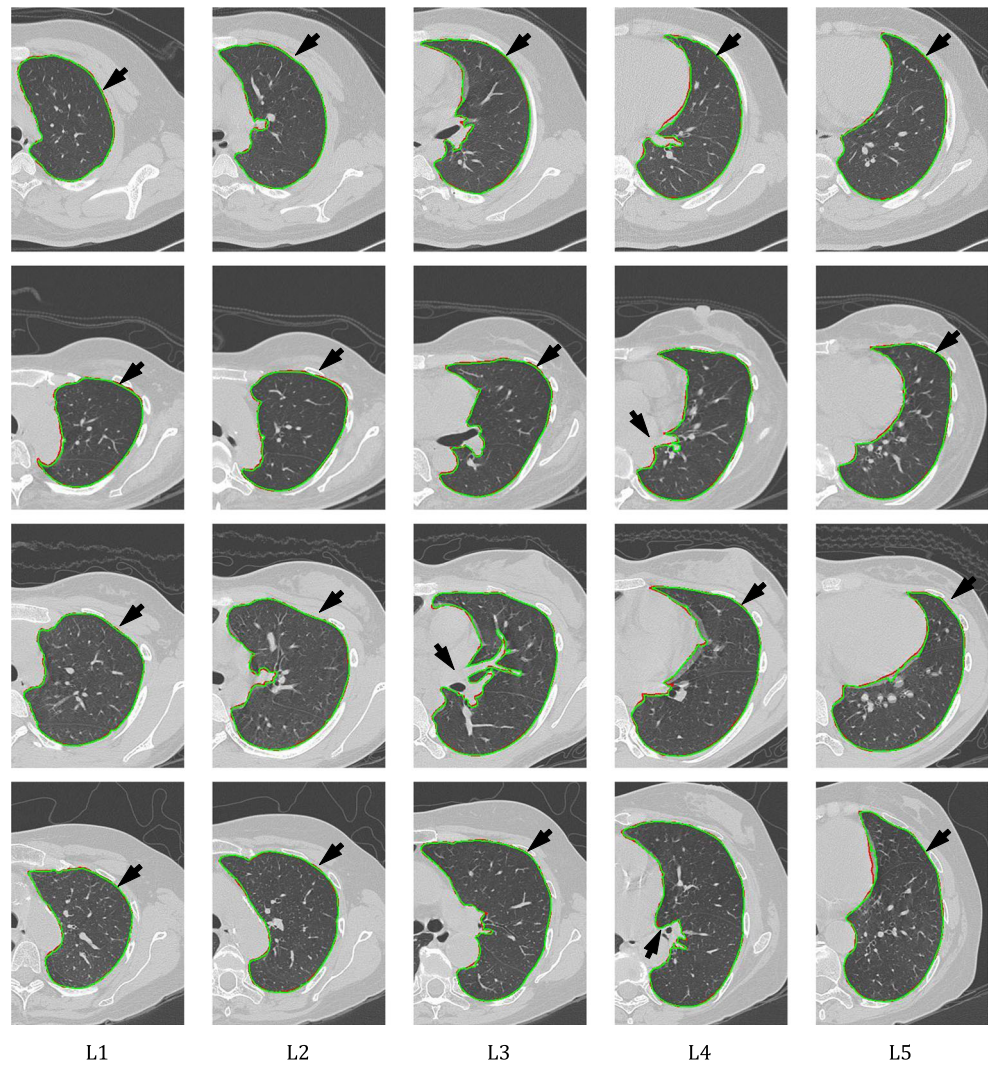


L1  L2  L3  L4  L5

**Fig. 7** Overlays of *left lung* with automatic segmentation border (*green color*) and ground truth border (*red color*)



boundary (red). The arrows show that the green and red borders are closely overlapping suggesting the high performance of the global and local segmentation system.

Global processing vs. local processing

The global processing or the initial segmentation which is based on a morphological approach can be inadequate especially when the lung region contains various or multiple tissues that are healthy and scarred which causes the information of each pixel to vary. The inadequacy of the initial segmentation as seen in Fig. 9 (under-segmentation) and Fig. 11 (over-segmentation) is due to the tissue's contrast inconsistency in the lung. To overcome this shortcoming of the global system, the local system utilizes the texture-based paradigm which is used for correction of large deviations using the feedback system. The results of this strategy are shown in Fig. 10 (for under-segmentation) and Fig. 12 (for over-segmentation). This approach was successfully implemented for both over

and under-segmentation cases. The local system managed to bring the green border much closer to the red boundary as seen by the arrows in Figs. 10 and 12. This is because the local system uses texture filter which probes locally into the pixel value to enable proper segregation between lung and non-lung regions especially when the lung pixels are inconsistent. The arrows in Figs. 9, 10, 11 and 12 show the borders of the lung.

Segmentation results: Normal vs. diseased lungs

One of the objectives is to understand how feedback dynamically behaves if the lung is normal vs. diseased. At the same time, it is important to understand the behavior of the system for the left and right lungs. In the diseased lungs, the lung regional intensities have large variability and inconsistencies. This poses a threat to global segmentation system causing under-segmentation.

Figure 14a and b represent such examples, where green color shows under-segmentation while red color shows the

**Fig. 8** Splitting the lung into four quadrants to understand the level of accuracy
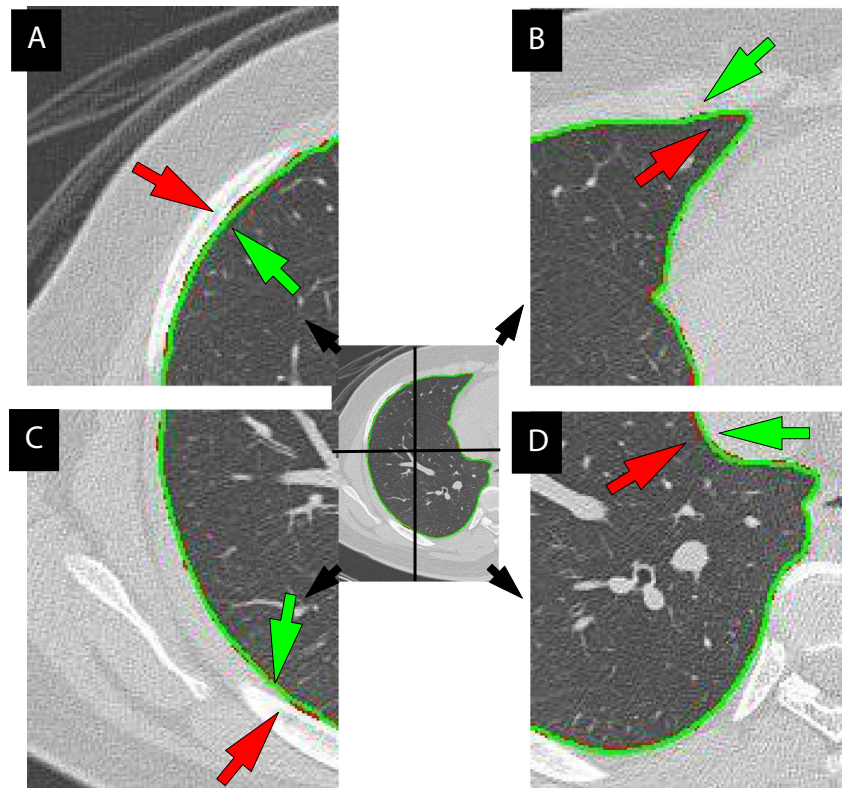


manual tracings. Due to this under-segmentation, the lung area is likely to be smaller, unlike in normal healthy subjects (Fig. 13a and b). Though, this behavior is spotted in certain levels, but is not always true at all levels of the lung due to the nature of the cancerous growth of cells. For example in Level 1 of Fig. 14, the lung region segmented using automated method (green) is still very close to that of the ground truth (red) for right lung and left lung but on moving to higher levels such as 2 to 5, the under-segmentation feature is more apparent using the output of the global system. The arrows in the figures show the lung borders.

Quantitative evaluation of normal vs. disease lungs

Table 1 shows the areas of the left and right lungs for all the five levels. The corresponding bar charts are shown in Fig. 15a and b for the right and left lungs. The difference in the overall areas for the healthy and diseased lung is not very

pronounced for the right lung, while more pronounced for the left lung indicated by the black arrows. Also note carefully that the bar chart shows the comparison between the automated method (grey color) and manual methods (black color). They show very close resemblance. Area is counted using information obtained from the pixel spacing in DICOM header and is counted as below:

$$A = h \times l \qquad (6)$$

where $h$=height of the pixel (mm) and $l$=length in (mm). Area ($A$) is in mm$^2$.

As seen in Table 1, following observations were seen:

(i)    The average normal lung area of the right lung (RL) (as shown in Column 1, labeled as Col 1) totaling 10098.85 mm$^2$ is higher than the average area of the left lung (LL) (as shown in Column 4, labeled as Col 4)

**Fig. 9** Under-segmentation error by global system (*green border*) and ground truth border (*red border*) for *right lung*
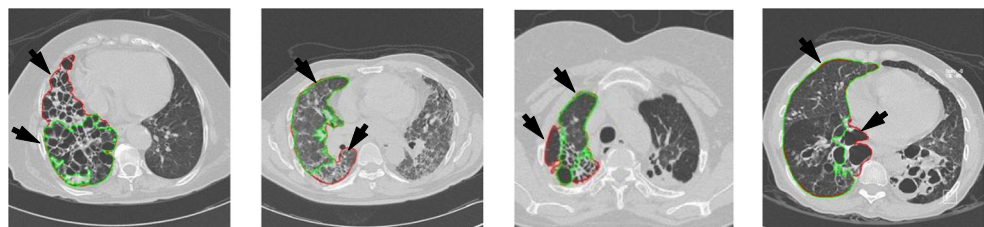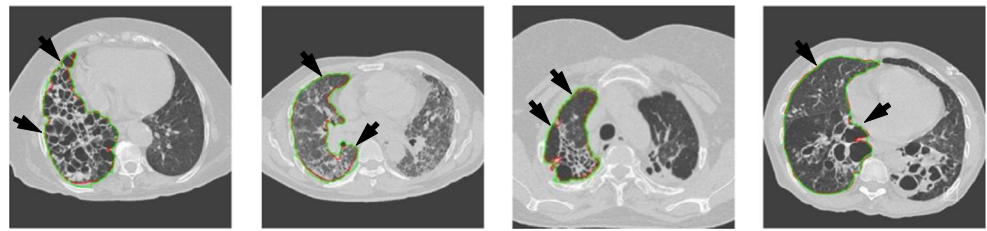
**Fig. 10** Corrected cases of under-segmentation by local system (*green border*) and ground truth border (*red border*) for *right lung*



totaling to 8546.21 mm$^2$. The area of the left lung is slightly smaller than the right lung because left lung has to accommodate the heart [18].

(ii)   The same behavior is observed for the diseased lungs. This can be seen in the Column 2 (labeled as Col 2) and Column 5 (labeled as Col 5). The average area for diseased right lung (RL) is 10462.46 mm$^2$ while the diseased left lung (LL) is 7765.02 mm$^2$.

(iii)   For the right lung (RL), the difference between the normal and abnormal was larger for Level 1 and Level 2 (12.04 and 11.83 %) compared to Level 3, 4 and 5 as (2.29, 0.38 and 3.42 %). The clinical interpretation and justification is the more aggressive disease in Level 1 and Level 2 compared to Level 3, Level 4 and Level 5.

(iv)   For the left lung (LL), the difference in the normal and abnormal was larger for Level 1, 3, 4 and 5 showing the values of 10.37, 10.68, 24.29 and 28.54 %. Only Level 2 showed a small difference between normal and abnormal patients (3.51 %). This also means that the left lung is more diseased compared to the right lung for these 81 patients.

(v)   Overall statistics shows that the average difference between the normal and abnormal for the left lung (LL) (10.06 %) is higher compared to right lung (RL) (3.48 %). The clinical interpretation of such statistics shows that the left lung is more diseased compared to the right lung. We also validated the above statistics using our manual tracings and our observations show the same general trend. This validates our clinical inference and interpretations.

Qualitative classification of normal vs. diseased lungs

The difference of normal and diseased lungs is seen also in the scatter plot in Fig. 16. The deviations from the trend
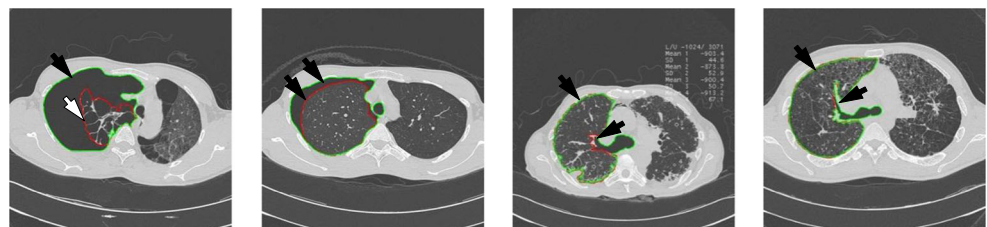
contributed by abnormal cases labeled by 'x' are mostly above the trend. This signifies that the segmented region is mostly smaller than the ground truth region for abnormal cases in right and left lung which is detected by the local system. This is a feature of detecting the lung is abnormal or not which is determined by the drop in area due to the ineffectiveness to segment based on the lung irregularities. There are also diseased lung areas closer to the trendline because the disease as shown before in Table 1 is not evident in all slices. Thus the results in this section show the ability of the feedback and corrective segmentation ability to handle high deviations of the segmentation where under and over-segmentation was present.

**Performance evaluation**

Performance evaluation was done to evaluate the quality of the segmentation and the performance of the segmentation system proposed. Several similarity coefficients are used including the Dice Similarity, Jaccard Index, relative area error, area overlap error and polyline distance metric. The consistency and accuracy of the segmentation is shown with the scatter plot and Bland Altman plot. Lastly, the precision of merit is presented.

Validation of segmentation system

Dice Similarity Coefficient (DSC) also known as Soren-Dice similarity coefficient gives an indication of similarity between two regions. Region *A* represents the area of the automated segmentation and region *B* represents the area enveloped using manual tracings. DSC is the ratio of area in common to both region A and region B to the average size of region A and region B.

**Fig. 11** Over-segmentation error by global system (*green border*) and ground truth border (*red border*) for *right lung*

**Fig. 12** Corrected cases of over-segmentation by local system (*green border*) and ground truth border (*red border*) for *right lung*
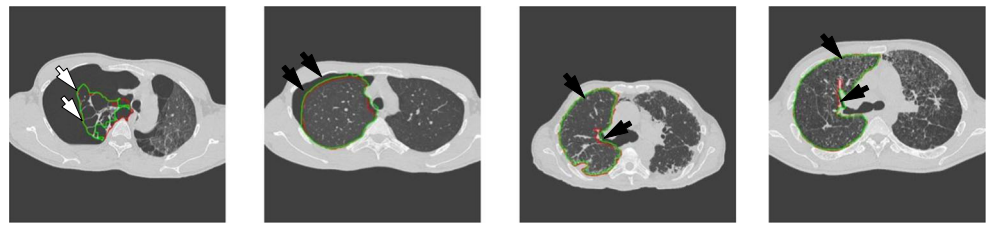


$$DSC = 200 \times \left( \frac{A \cap B}{A + B} \right) \tag{7}$$

Jaccard Similarity is shown below where region $A$ represents the area of the automated segmentation and region $B$ represents the area of the ground truth. Although very similar to Dice, Jaccard in the case of this study shows the ratio of area in common to both region A and B to the total size of region A and B available.

$$Jaccard = \left( \frac{A \cap B}{(A + B) - (A \cap B)} \right) \times 100 \tag{8}$$

Tables 2 and 3 show the DSC and Jaccard Index percentages for global and combined global and local systems. By adding the local feedback system, the DSC and Jaccard index increases the system performance. It is interesting to note that for all levels the DSC increased using feedback system (with local processing) compared to without feedback system (global system alone). The similar pattern was obtained for the Jaccard index. There was 4.89 % increase in DSC by adding the local feedback system. Jaccard increased by 0.79 % by adding the local feedback system. Corresponding behavior

was observed for the left lung also. DCS increased by 0.37 %, while Jaccard increased by 0.05 %. Thus, the right lung had more pronounced effect due to the local feedback system.

**Automated system against manual**

The performance of the automated system was evaluated by comparing against the manual tracings. The manual tracings were accomplished using ImgTracer™ 1.0 (AtheroPoint™ LLC, Roseville, CA, USA). Figure 17 shows the performance of the automated system for the left and right lung. This performance curve is shown to have area comparison of the automated system against the manual area computed from the manual tracings. As can be seen, the trend of the regression curve is nearly linear. This shows that the automated segmentation shows very promising results. It is however interesting to note that right lung has a longer regression line compared to left lung.

As part of the performance evaluation of the automated system, we compute the Bland-Altman Plot. The observation from the scatter plots is echoed in the Bland-Altman (BA) plots in Fig. 18. The vertical axis represents the difference

**Fig. 13** Five levels for *right lung* and *left lung* after segmentation using combined global and local processing for **normal cases**



L1          L2          L3          L4          L5

(a) Right lung



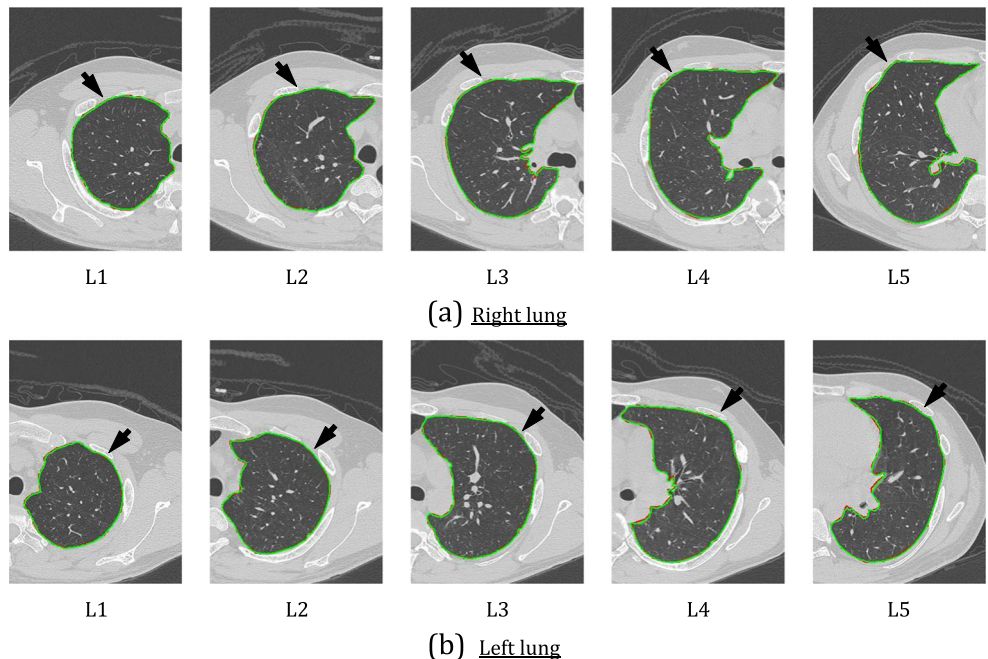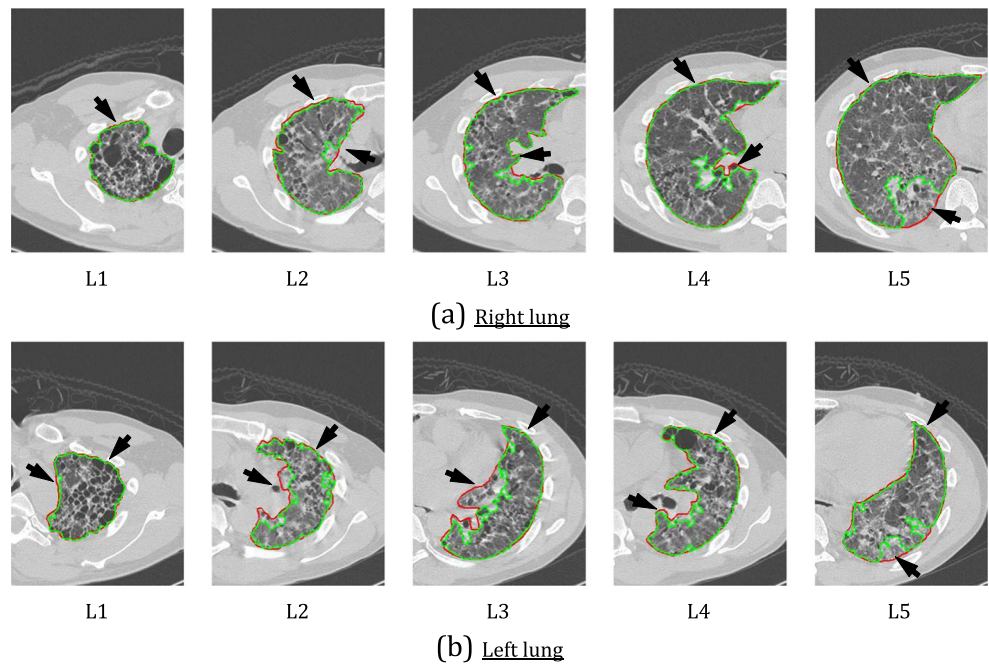L1          L2          L3          L4          L5

(b) Left lung

**Fig. 14** Five levels for *right lung* and *left lung* after segmentation using combined global and local processing for **abnormal cases**



(a) Right lung

(b) Left lung

between the ground truth area and the segmented area. There are three horizontal lines present in the plot, first the mean difference between the two regions and then the positive and negative 2 Standard Deviation (2SD) also known as the product of the mean difference by ±1.96.

Both BA plots for right and left lung show that large majority of the samples have a level of agreement between the automatic segmentation and manual segmentation. This is observed from the closely packed samples below the positive and negative 1.96 mean differences. This character of being closely packed is seen even along the horizontal axis which signifies the increase in mean area of lung regions. The observation of most samples within the ranges of 2 SD also signifies agreement between the segmented lung area and manual reading suggesting accuracy of segmentation. This further strengthens the observation of the robustness of the method when dealing with varying sizes of lungs.

**Table 1**  Area of normal and abnormal for abnormal and normal cases for *right lung* (RL) and *left lung* (LL)

| Level | Area (mm²) | | | | | |
|---|---|---|---|---|---|---|
| | Col 1 Normal (RL) | Col 2 Abnormal (RL) | Col 3 % Diff (RL) | Col 4 Normal (LL) | Col 5 Abnormal (LL) | Col 6 % Diff (LL) |
| L1 | 7152.30 | 8130.88 | 12.04 | 6452.36 | 7198.70 | 10.37 |
| L2 | 8610.58 | 9766.06 | 11.83 | 7887.46 | 8174.07 | 3.51 |
| L3 | 10406.22 | 10649.57 | 2.29 | 9333.13 | 8432.83 | 10.68 |
| L4 | 11566.03 | 11522.22 | 0.38 | 9654.61 | 7767.53 | 24.29 |
| L5 | 12574.97 | 12158.58 | 3.42 | 9272.64 | 7214.07 | 28.54 |
| Average | 10098.85 | 10462.46 | 3.48 | 8546.21 | 7765.02 | 10.06 |

Precision-of-merit

*Relative area and overlap area error*

Area plays an important role in evaluating the overall performance metric of the system. We define two different categories of overall performance evaluation precision-of-merit. They are information depending upon the region and boundary. The information using the region is defined in terms of relative area error and overlap error, while the boundary information is characterized using the polyline distance metric. Using the definition of area as region and defined A and B for automated and manual, the relative area error is given as:

$$\text{Relative Area Error } (\%) \ = \left(\frac{A - B}{B}\right) \times 100 \qquad (9)$$

Correspondingly, the Overlap error in terms of area is given as:

$$\text{Area Overlap Error } (\%) \ = \left(1 - \left(\frac{A \cap B}{(A + B) - (A \cap B)}\right)\right) \times 100 \qquad (10)$$

From levels one to five, the area of segmented and ground truth increases gradually for the right lung. The right lung is also larger than the left lung for all five levels from Tables 4 and 5. It is noticeable that the Relative Area in right lung is lower than that of the left lung for overall and most levels except Level 1 where large deviations are detected. However the low overall value of relative area error which −0.03 % suggests that the segmentation is accurate. The left lung has

**Fig. 15** Bar chart of mean area of *right lung* and *left lung* of normal and abnormal cases for segmented (*grey*) and ground truth (*black*)



(a) Right lung     (b) Left lung
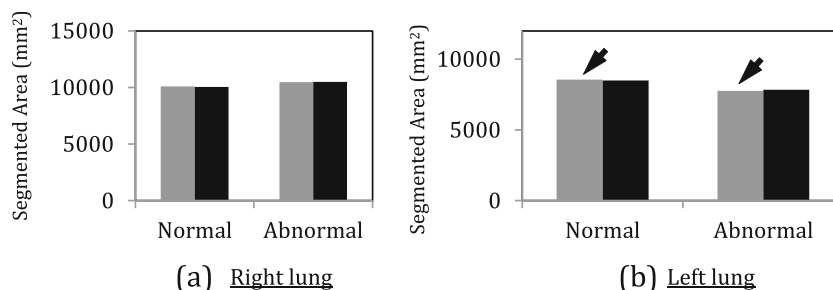
an overall of −1.15 % error also suggesting that the segmentation is accurate. Positive values of the Relative Area Error indicate that the area of segmentation is larger than manual area. Negative values of the Relative Area Error indicate the area of segmentation is smaller than manual area as seen in Levels 2, 4 and 5 for the right lung in Table 4 and all levels for the left lung in Table 5. For both right and left lung the segmentation is smaller than the manual area. The low errors for consistently over five levels also suggest the consistency and accuracy of the segmentation over all the samples of increasing area.

The Area Overlap Error is more sensitive to errors that are outside the region of intersection between the segmented area and the ground truth. Therefore there is an increase of error for both right and left lung. Both lungs show low error of 3.53 and 4.09 % for right and left lung respectively. Again, the low errors over five levels suggest the consistency and accuracy of the segmentation regardless of the level. The level can be an indicator of area or slice. Thus the segmentation errors suggest that the segmentation method is effective and consistent regardless of area or increased height of the slice taken.

*Polyline distance metric*

Polyline Distance Metric (PDM) used in this study measures the changes of the contours of the two regions. The reference

contour used is the ground truth and is denoted by $B_1$. A point on the reference contour $B_1$ is chosen as the reference point $(x_0, y_0)$. Next, the nearest point at the automated segmentation contour, $B_2$ was found using Euclidian distance. This is the 1st point $(x_1, y_1)$ to be evaluated. Then the 2nd point $(x_2, y_2)$ is established as the point next to the 1st point on the automated contour. The two points actually form a line segment, s. Next $d(v,s)$ was obtained which is the distance between the reference point, v $(x_0, y_0)$ and the line segment formed by 1st point and 2nd point.

The distance between the 1st point to the reference point is called $d_1$ whereas the 2nd point to the reference point is called $d_2$. Another term used in the process towards finding $d(v,s)$ is Lambda, $\lambda$ which is the distance of the reference point, v towards the line segment, s. The perpendicular distance between the line segment and the reference point, v, is given by $d^\perp$. The formulas to calculate $\lambda$ and $d^\perp$ are below;

$$\lambda = \frac{(y_2-y_1)(y_0-y_1) + (x_2-x_1)(x_0-x_1)}{(x_2-x_1)^2 + (y_2-y_1)^2} \tag{11}$$

$$d^\perp = \frac{(y_2-y_1)(y_0-y_1) + (x_2-x_1)(x_0-x_1)}{\sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}} \tag{12}$$

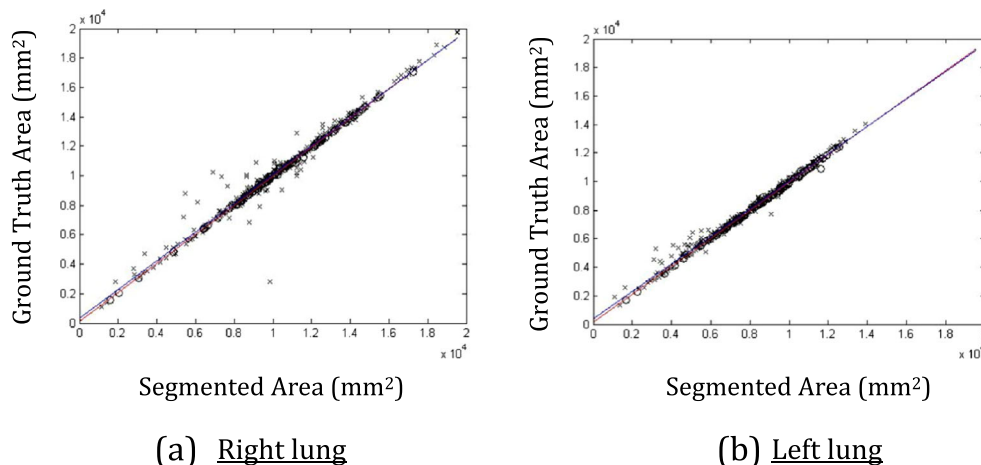**Fig. 16** Scatter plot of abnormal case (x) vs. normal case (o) for *right lung* and *left lung* area



(a) Right lung     (b) Left lung

**Table 2**    Dice similarity and Jaccard index for the *right lung*

| Level | DSC | | Jaccard | |
|---|---|---|---|---|
| | Mean | ±SD | Mean | ±SD |
| (a) Global System Alone | | | | |
| L1 | 97.73 | 6.53 | 96.11 | 8.95 |
| L2 | 98.28 | 2.05 | 96.70 | 3.69 |
| L3 | 98.27 | 1.64 | 96.65 | 3.02 |
| L4 | 98.06 | 2.38 | 96.28 | 4.14 |
| L5 | 98.16 | 3.71 | 96.60 | 6.17 |
| Average | 98.10 | 3.66 | 96.47 | 5.56 |
| (b) Combined Global+Local System | | | | |
| L1 | 98.55 | 1.90 | 97.20 | 3.44 |
| L2 | 98.66 | 0.87 | 97.37 | 1.67 |
| L3 | 98.43 | 1.18 | 97.02 | 2.07 |
| L4 | 98.39 | 1.21 | 96.86 | 2.25 |
| L5 | 98.87 | 0.95 | 97.78 | 1.78 |
| Average | 98.58 | 1.28 | 97.24 | 2.33 |

Therefore $d(v,s)$ is obtained using the following equation.

$$d(v,s) = \begin{cases} \min\{d1, d2\}; & if \ \lambda < 0, \lambda > 1 \\ |d\perp| & if \ 0 \le \lambda \le 1 \end{cases} \quad (13)$$

The process to obtain $d(v,s)$ is then repeated for the rest of the points of the contour $B_1$ and this is given by;

$$d(B_1, B_2) = \sum_{i=1}^{n} d(v_i, s_{B2}) \quad (14)$$

**Table 3**    Dice similarity and Jaccard index for the *left lung*

| Level | DSC | | Jaccard | |
|---|---|---|---|---|
| | Mean | ±SD | Mean | ±SD |
| (a) Global System Alone | | | | |
| 1 | 98.64 | 1.42 | 96.11 | 8.95 |
| 2 | 98.33 | 1.53 | 96.70 | 3.69 |
| 3 | 98.16 | 1.36 | 96.64 | 3.01 |
| 4 | 97.20 | 2.78 | 96.28 | 4.14 |
| 5 | 96.89 | 5.24 | 96.60 | 6.17 |
| Average | 97.84 | 2.94 | 96.47 | 5.56 |
| (b) Combined Global+Local System | | | | |
| L1 | 98.71 | 1.06 | 97.47 | 1.99 |
| L2 | 98.48 | 0.83 | 97.01 | 1.60 |
| L3 | 98.30 | 0.89 | 96.67 | 1.70 |
| L4 | 97.64 | 2.84 | 95.44 | 3.25 |
| L5 | 97.96 | 1.63 | 96.06 | 2.96 |
| Average | 98.21 | 1.35 | 96.52 | 2.49 |

where $n$, is the number of points in contour, $B_1$ and $s_{B2}$ is the segment on contour $B_2$. Secondly the algorithm above is repeated where $B_2$ now becomes the reference contour and $B_1$ becomes the segment contour $s_{B1}$ The reverse can be represented by $d(B_2,B_1)$. Lastly combining both $d(B_1,B_2)$ and $d(B_2, B_1)$ will yield the equation below which is the polyline distance metric;

$$D_S(B_1 : B_2) = \frac{d(B_1, B_2) + d(B_2, B_1)}{(\#vertices \in B_1 + \#vertices \in B_2)} \quad (15)$$

Polyline Distance Metric (PDM) yields a low average difference of less than 1 mm difference. The right lung has an average of 0.68 mm difference and the left lung has 0.61 mm difference. Again the PDM are consistently low for all five levels for both lungs. This supports the observation that the segmentation seems to have a high precision and consistent over a large varying area and height at which the slice was taken. Overall, the performance evaluation suggests that the segmentation method provides results which are consistent and have acceptable error ranges.

## Discussion

The objective of this research was to develop a pilot study which can automatically detect and quantify the normal versus diseased lung using a combination of a global system embedded with a feedback system characterized by the local system. Using the analogy of control system theory, the comparator was developed that allowed to identify the deviation between the intermediate output (global system) and trained system which already had information about the goal state which is the tracing from the lung expert. The comparator allowed spotting the large deviations in the intermediate output (global state) which was then fed to the feedback loop for local processing system to correct them. Thus our overall system has unique characteristics which permitted not only spotting large deviations but also corrected them using a texture-based paradigm. The system was demonstrated in Fig. 3. We took 96 patients consisting of classified patient population of 15 healthy individuals (normal cases), 28 ILD cases and 53 other lung related diseases cases termed as non-ILD. There were 48 males and 48 female samples. The HRCT scanner used was the Siemens SomatomPlus4 CT scanner. For left lung, the performance of segmentation was 96.52 % using Jaccard Index, and 98.21 % using Dice Similarity, 0.61 mm using Polyline Distance Metric (PDM), −1.15 % using Relative Area Error, and 4.09 % using Area Overlap Error. The right lung's performance of segmentation was 97.24 %, using Jaccard Index, 98.58 % using Dice Similarity, 0.61 using PDM, −0.03 % using Relative Area Error, and 3.53 % using

Fig. 17 Scatter plots of *left lung* and *right lung* area
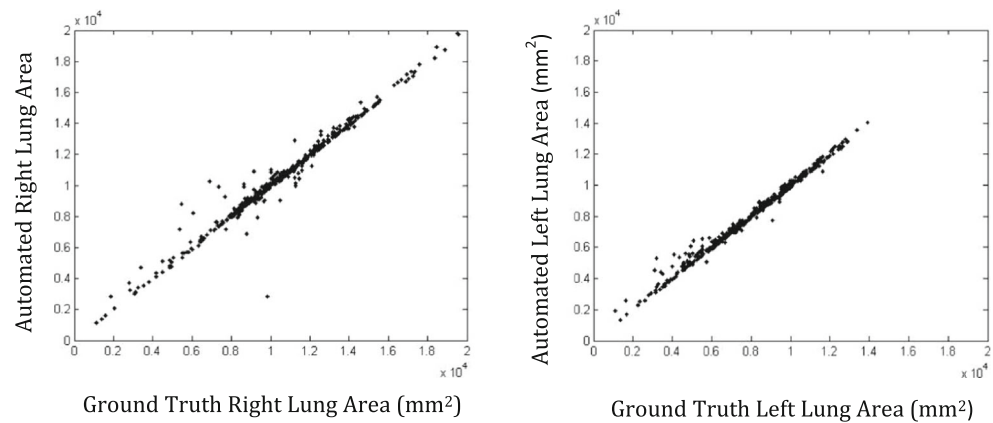
Area Overlap Error. The segmentation overall has an overall similarity of 98.4 %. This feedback control system was fully automated keep regulatory constraints and meeting the objective and goals for precision and stability.

The attempt of classification of abnormal lungs compared to normal lungs was shown in Table 1 and Fig. 16. The attempt showed that there was difference between the normal and abnormal lungs in some levels (L1 to L5). The difference of area feature of abnormal lungs to normal lungs may be one of the indicators but cannot be used alone for classification. Classification needs higher level efforts such as machine learning [19], expert systems and neural networks[20, 21], and texture based strategies for classification into abnormal lungs and normal lungs. Higher level classification methods offer longer and wider assessment of the decisions, higher consistency in decision making, and a shorter decision-making process.

Generally, there are two main types of segmentation. The first type is region-based segmentation. Region-based methods utilize on information within a region of interest rather than the boundaries noticeable. This method focuses on initializing a seed points within the region of interest to help the segmentation to progress. Neighboring pixels or voxels are evaluated and compared to the region or seed point

that has been initialized. This method's downside is that it tends to over-segment and gives a bigger region than supposed. The boundaries become blur and usually inaccurate. Region-based methods include but are not limited to texture analysis, threshold-based, region growing and deformable models. A study by Hu et al. was among the first to present a fully automatic method for identifying lungs in 3-D pulmonary X-Ray CT images [22]. The method is divided into three main steps. First, the lung region is extracted from CT scan image by gray level threshold. Secondly the left and right lungs are separated by identifying the anterior and posterior junctions by dynamic programming. Lastly morphological operations are utilized to smooth the irregular boundary along the mediastinum in order to obtain results consistent with those obtained by manual analysis, in which only most central pulmonary arteries are excluded from the lung region.

The second category is boundary-based techniques. This category utilizes the contours or boundaries of a certain region to segment out an image. It is usually implemented in 2-D images but restrictions are present especially for 3-D images. It is generally faster than region-based techniques because no seed points are required. Contour-based techniques include but are not limited to active contours or snakes. Tobata and Hospital in 2007 showed the application of snakes for

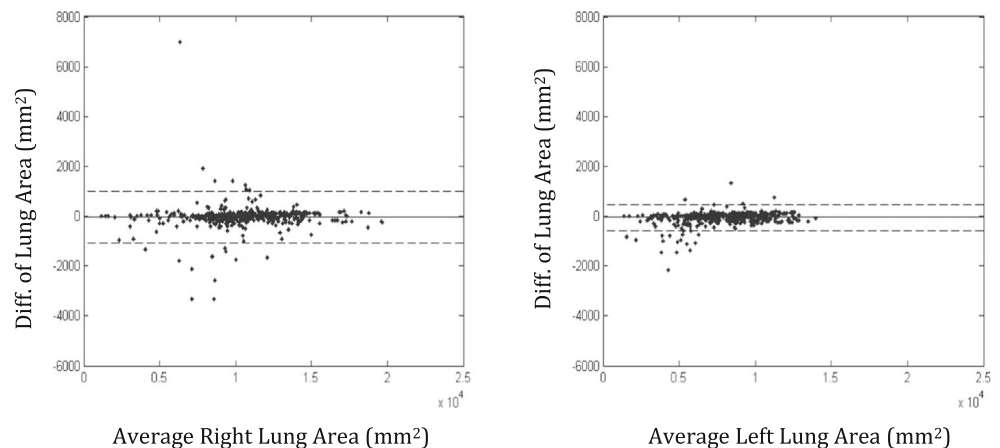Fig. 18 Bland Altman plots of *left lung* and *right lung* area

**Table 4**   Relative area error, overlap area error and boundary error for *right lung*

| Level | Area Segmented (mm$^2$) | | Area Ground Truth (mm2) | | Relative Area Error (%) | | Area Overlap Error (%) | | Boundary Error (PDM, mm) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | ±SD | Mean | ±SD | Mean | ±SD | Mean | ±SD | Mean | ±SD |
| 1 | 7965.95 | 2719.70 | 7890.10 | 2709.32 | 2.34 | 26.92 | 3.89 | 8.95 | 0.72 | 2.04 |
| 2 | 9567.27 | 2584.53 | 9563.67 | 2561.04 | −0.09 | 4.07 | 3.30 | 3.69 | 0.62 | 0.64 |
| 3 | 10608.15 | 2438.29 | 10575.67 | 2407.39 | 0.26 | 3.79 | 3.35 | 3.02 | 0.62 | 0.58 |
| 4 | 11529.68 | 2533.59 | 11627.82 | 2435.93 | −1.03 | 4.43 | 3.72 | 4.14 | 0.74 | 0.78 |
| 5 | 12230.99 | 3029.90 | 12374.57 | 2836.69 | −1.52 | 6.40 | 3.40 | 6.17 | 0.72 | 1.34 |
| Mean | 10400.29 | 3045.21 | 10426.90 | 3018.80 | −0.03 | 12.58 | 3.53 | 5.56 | 0.68 | 1.19 |

segmentation. Snake method is part of edge-based segmentation. The study was among the first to present an automatic approach using snakes or active contour model without manual input. It was able to deal with abnormal contrasting areas of the image which included the ground-glass opacity of the lung image [23].

There is a hybrid category that involves both region- and contour-based techniques. This category arose because of the limitations of both categories. Graph cut is an example that is a hybrid. This method successfully utilized by Boykov and Jolly for segmenting organs. The study utilized graphs by forming lines connecting all pairs of neighboring image pixels which can be called voxels by weighted edges. The study's objective was to separate an object of interest from the background based on graph cuts. The study offers a globally optimal efficient solution in a general N-dimensional setting. This allows images 2-D, 3-D and even 4-D images to be processed. It utilises both boundary and region information to form the segmentation. The region-based technique used in graph cut allows natural propagation of information throughout the volume of an N-dimensional image whereas contour-based technique help deals with the over-segmentation problem [24]. Limitations of vague lung borders, acquisition artifacts, low contrast and variability of objects add complexity to the segmentation task and require more complex methods. Osareh and Shadgar did a study to combat these problems and

proposed a method combining fuzzy c means segmentation and region aided geometry snakes segmentation method [25]. Again this method is a hybrid method that combines both local and global methods which actually was found to have higher accuracy in segmentation compared to the conventional region aided geometry snake segmentation method.

Table 6 shows the latest methods and their accuracies as part of the benchmarking protocol. The comparable error measures used were Dice Similarity Coefficient (DSC) and Area Overlap Error [26–32]. Ideally for image segmentation evaluation to be meaningful for comparison it should display two characters. The first character is objectivity which means that all the ground truths tracings are clear and unambiguous. Secondly is generality which means large variability in the number of images used [33]. The images used in this study fulfil both criterion with carefully traced manual borders and images consisting of normal and abnormal (both ILD and Non-ILD). The usage of five slices ensures generality as well when all these five levels present distinct properties compared to one another. In summary the proposed method of this study with local processing yielded the second best similarity of 98.4 % average for both left and right lungs. Comparing a high amount of cases segmentation to a low amount cases may not be fair to the large database segmentation. However it is important to take note if there are large variations, for example surface overlap difference of 2 or 3 % is quite significant [31].

**Table 5**   Relative area error, overlap area error and boundary error for the *left lung*

| Level | Area Segmented (mm$^2$) | | Area Ground Truth (mm2) | | Relative Area Error (%) | | Area Overlap Error (%) | | Boundary Error (PDM, mm) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | ±SD | Mean | ±SD | Mean | ±SD | Mean | ±SD | Mean | ±SD |
| 1 | 7072.91 | 2382.24 | 7066.06 | 2354.94 | −0.03 | 2.60 | 2.65 | 2.60 | 0.49 | 0.81 |
| 2 | 8124.76 | 2260.27 | 8196.53 | 2218.08 | −1.07 | 2.79 | 3.25 | 2.79 | 0.50 | 0.33 |
| 3 | 8586.07 | 2133.26 | 8581.31 | 2106.02 | −0.04 | 2.52 | 3.58 | 2.52 | 0.55 | 0.43 |
| 4 | 8088.73 | 2498.35 | 8183.89 | 2376.86 | −1.76 | 4.93 | 5.31 | 4.93 | 0.75 | 0.65 |
| 5 | 7572.08 | 2595.34 | 7693.77 | 2436.90 | −2.80 | 8.28 | 5.62 | 8.28 | 0.75 | 0.97 |
| Mean | 7898.60 | 2424.19 | 7954.35 | 2348.57 | −1.15 | 4.89 | 4.09 | 4.89 | 0.61 | 0.68 |

**Table 6**   Benchmarking our proposed method against previous methods

| First Author | Year | Method | No. of Cases | Dice Similarity Coefficient | Overlap Area Error |
| --- | --- | --- | --- | --- | --- |
| Korflatiiset al.[26] | 2008 | Texture classification-based | 30 | 94.60 % | – |
| Rikxoortet al.[27] | 2009 | Active Contours | 100 | 95.00 % | – |
| Kakaret al.[28] | 2009 | Fuzzy C Means clustering | 41 | 94.19 % | – |
| Zhou et al.[29] | 2006 | Threshold Based | 7 | 98.26 % | – |
| Wang et al.[30] | 2009 | Threshold Based | 76 | 98.40 % | – |
| Massoptieret al.[31] | 2009 | Graph Cut Based | 11 | 97.42 % | – |
| Abbas et al. [32](RL) | 2013 | Particle Swarm Optimizer | 48 | – | 8.30 % |
| Abbas et al. [32](LL) | 2013 | Particle Swarm Optimizer | 48 | – | 9.12 % |
| Proposed Method (RL) - Morphology | | | 96 | 98.10 | 3.53 % |
| Proposed Method (LL) - Morphology | | | 96 | 97.84 | 4.10 % |
| Proposed Method (RL & LL Combined) - Morphology | | | 96 | 97.97 | 3.82 % |
| Proposed Method (RL) -Rejection Analysis with Texture Filter | | | 96 | 98.58 | |
| Proposed Method (LL) - Rejection Analysis with Texture Filter | | | 96 | 98.21 | |
| Proposed Method (RL & LL Combined) Rejection Analysis with Texture Filter | | | 96 | 98.40 | |

Zhou et al. yielded the third best DSC of these works which is 98.26 % was achieved by using threshold-based segmentation technique implemented with seed points [29]. However Zhou only utilized 7 patients in their study. This high percentage of similarity is debatable due to the limited database used. Massoptier et al. ranked fourth with 97.42 % similarity using graph cut method, with 11 patients [31]. The other have increasing amount of patients and the most comparable was Van Rikxoort using active contours with 100 patients yielding similarity percentage of 95 % [27]. Though the proposed method uses the basic method of thresholding, morphology, it is coupled with texture filter to increase segmentation to provide accurate segmentation that outperforms other higher level methods. Abbas et al. proposed a method based on particle swarm optimization with overlap area errors of 8.3 % for right lung and 9.12 % for left lung [27]. There is a large significant difference where the proposed method yields an average of 3.82 % for both lungs. This supports the use of the basic method coupled with texture filter for local processing.

A study done by Wang et al. utilises a similar approach of thresholding and texture which yields a performance accuracy of 98.5 % agreement which is similar to 98.4 % of the proposed method [30]. Wang's study uses 31 normal cases and 45 abnormal cases with moderate or severe Interstitial Lung Disease (ILD). Wang's study utilized three levels for one patient which were first manually selected by a medical physicist-based on three criteria; L1: aortic arch, L2: main bronchi and L3: lower lobar bronchi. Manual tracings are done by the medical physicist and then confirmed by an expert chest radiologist. Wang's study segmentation is divided into three sections, removal of airways, initial segmentation using threshold, repairing of severe ILD cases using calculation co-occurrence matrix and threshold. For the removal of airways the study used seed points to remove airways that are not part of the lung lobe. The second stage of segmentation in the study involves the empirically selected a threshold of −300 HU to estimate initial lungs and used connected component analysis to separate the lung from outer regions. The third stage of the segmentation involves the usage of analysis texture characteristics of several components energy, entropy, maximum probability, and inverse difference moment to make up a co-occurrence matrix. The matrix is evaluated for severe ILD cases. With the matrix made enhancement on the original image to exhibit the high entropy. A fixed threshold of 600 was applied on the enhanced image and missing components that were left out from first segmentation is added in.

Unlike Wang, the proposed method uses five levels selected by a radiologist based on five levels based on five anatomy landmark which are level L1: aortic arch, L2: trachea carina, L3: pulmonary hilar, L4: pulmonary venous confluence and L5: 1 to 2 cm above the dome of right hemi-diaphragm. The result representation by Wang was a general mean of left and right and three levels. Wang's study used 31 normal and 45 abnormal patients. This ratio of normal to abnormal is relatively higher compared to the study of 15 normal cases and 81 abnormal cases. The higher ratio by Wang of total normal patients to total abnormal patients will give a more positive bias to the results. Our study, on the other hand, displays the results and performance of all five levels in its entirety and is split to left and right to show specificity and consistency of the segmentation method. Thus this study has more in depth representation of the lung with five levels compared to three levels. The threshold used in this study is also empirically selected at −324 HU. Unlike Wang who uses a co-occurrence for severe ILD, this study uses a feedback based on the manual tracings which detects deviations not just based

on the case type. The major disadvantage of Wang's method is that it requires the analysis of the all 30 slices of a patient before segmentation on 3 slices is done for one patient.

The parameters used in this study are listed in Table 7 to ensure repeatability. The first dynamic threshold $\vartheta_{Body}$, was Otsu Threshold to separate the body from the background of HRCT image. The second threshold used was an empirical threshold $\vartheta_{Lung,}$ of 700 pixel intensity or −324 HU for binary conversion. Erosion and Dilation was done using a structure element of square sized $3 \times 3$ pixels. The texture filter used was that of entropy. The empirical threshold used for local processing segmentation in under segmentation $\vartheta_{Under,}$ was 0.5 and for empirical over segmentation $\vartheta_{Over}$ was 0.2.

The proposed study has the following strengths: (i) this study utilizes a trained human interaction help or an expert which makes it a complete study. This interaction comes from the comparator of the system that allows spotting large deviations and correcting via feedback system leading to an accuracy of 99 %. Further, this meets the regulatory requirements a market needs for CAD systems. Secondly, (ii) the system consists of a global and local system that compliments each other offering the best of both worlds using local and general parameters. Next, (iii) the segmentation system has the ability to spot large deviations using a comparator. Thus (iv) the system proposed is also able to correct large deviations from the ground truth decreasing error of segmentation based on the local system. The system proposed also has (v) the ability to classify diseased and normal lungs. Lastly (vi) the system uses five levels compared to three levels other studies utilized for ILD patients giving a more complete and larger coverage of the lung increasing diversity on lung shape. The system has no restrictions on the type of lung disease such as obstructive diseases, as long as the segmentation paradigm is adapted.

The limitation of our study is that since this is a pilot study with limited resources, (i) it lacks validation from multiple observers even though the segmentation proposed is accurate. This can yield a positive or even negative bias of the ground truth if the tracer enlisted in not consistent. Therefore, having multiple tracers will provide a relatively non-biased segmentation. ii) The study also lacks intra-observer variability, where the same observer does the tracing more than once to validate the accuracy of the segmentation. This will give a more complete performance evaluation. With the current time constraints, we did evaluate the performance of all five levels of the lung when compared to an observer. (iii) Is that the study did not evaluate the segmentation on other images than the one obtain. However, given set of the parameters in the Table 7, we believe that with the exact same acquisition protocol used that the system is reproducible since the system is fully automatic. Lastly, (iv) the study only uses five levels per patient for evaluation which is a small number. Having more slices to be evaluated will further validate the segmentations accuracy.

As part of cost-benefit analysis, we highlight that such a segmentation system has benefits towards stability and reliability due to the role of the human observer, but it introduces an extra cost which is needed during the initial setup. Even though our system requires this cost of training of the human observer, but in normal circumstances, such a trained radiologist is normally present in the CT laboratory, over-looking the CT readings. Thus, the cost is not truly over-burden to the radiological CT lung laboratory. In future works with larger resources available, the authors would like to evaluate the significance of using more than one tracer to give inter-observer analysis to the accuracy of the segmentation. The variations from other tracers would give a more complete analysis. Secondly the intra-observer analysis can be done when a trained person to do lung tracing does the tracing repeatedly on the same image. The difference can be studied and documented and give further validation of the segmentation. Third, in the future, the study will move towards the extension from five levels in two dimensional to complete slices in three dimensional to compare normal and diseased which offers a more complete and also real time medical application since this is just a pilot study.

## Conclusion

As a conclusion the study is a pilot study that has fulfilled the aim to propose an automatic segmentation based on global and local system. The global system is based on morphology and the local system is based on texture with an embedded control feedback that detects and corrects large deviations and failures of segmentation. It was able to segment the lung accurately and highly similar to the ground truth for all five levels of the lung. Exhaustive data analysis was performed demonstrating three kinds of accuracy measures such as Jaccard index, Dice similarity and Polyline distance metrics. The results were consistent and show a promising accuracy measure. The system is

**Table 7** System parameters adapted for global and local processing

| Global parameters for global processing | Values |
| --- | --- |
| $\vartheta_{Body}$(Global system, threshold type) | Otsu |
| $\vartheta_{Lung}$( Binary threshold for lung region (morphology)) | −324 HU |
| S (Structure element) | square |
| $nxn$(size of $S$ in erosion) | $3 \times 3$ |
| $mxm$(size of $S$ in dilation) | $3 \times 3$ |
| Local parameters for local processing | Values |
| Type of texture filter (Local system) | Entropy |
| $\vartheta_{Under}$(Binary threshold for lung region texture under-segmentation) | 0.5 |
| $\vartheta_{Over}$(Binary threshold for lung region texture over-segmentation) | 0.2 |

able to segregate normal and diseased lungs for the left and right sides. Since this was a pilot study, there is a potential and scope of improving data size and exhaustive inter- and intra-observer variability analysis. Inspite of the above challenges, our system shows comparable accuracy measures with attempt to model towards the respiratory standard of care.

# References

1. Schwarz, M. I., Matthay, R. A., Sahn, S. A., Stanford, R. E., Marmorstein, B. L., and Scheinhorn, D. J., Interstitial lung disease in polymyositis and dermatomyositis: analysis of six cases and review of the literature. *Medicine* 55(1):89–104, 1976.

2. Peroš-Golubičić, T., and Sharma, O., *Clinical atlas of interstitial lung disease*. Springer, London, 2006.

3. Sharman, P., and Wood-Baker, R., Interstitial lung disease due to fumes from heat-cutting polymer rope. *Occup. Med.* 63(6):451–453, 2013.

4. O'Dwyer, D. N., Armstrong, M. E., Cooke, G., Dodd, J. D., Veale, D. J., and Donnelly, S. C., Rheumatoid Arthritis (RA) associated interstitial lung disease (ILD). *Eur. J. Intern Med.* 24(7):597–603, 2013.

5. Washko, G. R., Hunninghake, G. M., Fernandez, I. E., Nishino, M., Okajima, Y., Yamashiro, T., et al., Lung volumes and emphysema in smokers with interstitial lung abnormalities. *N. Engl. J. Med.* 364(10):897–906, 2011.

6. Henne, E., Anderson, J. C., Lowe, N., and Kesten, S., Comparison of human lung tissue mass measurements from ex vivo lungs and high resolution CT software analysis. *BMC Pulm. Med.* 12(1):18, 2012.

7. Krupinski, E. A., Berbaum, K. S., Does reader visual fatigue impact interpretation accuracy? *Proc. SPIE. Med. Imaging.* 7627, 76270M-1-6, 2010.

8. Jiang, Y., Nishikawa, R. M., Schmidt, R. A., Metz, C. E., Giger, M. L., and Doi, K., Improving breast cancer diagnosis with computer-aided diagnosis. *Acad. Radiol.* 6(1):22–33, 1999.

9. Doi, K., Computer-aided diagnosis in medical imaging : Historical review, current status and future potential. *Comput. Med. Imaging Graph.* 31(4):198–211, 2007. doi:10.1016/j.compmedimag.2007.02.002.

10. Kobayashi, T., Xu, X. W., MacMahon, H., Metz, C. E., and Doi, K., Effect of a computer-aided diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs. *Radiology* 199(3):843–848, 1996.

11. Nagaraj, S., Rao, G. N., and Koteswararao, K., The role of pattern recognition in computer-aided diagnosis and computer-aided detection in medical imaging: A clinical validation. *Int. J. Comput. Appl.* 8(5):18–22, 2010.

12. van Rikxoort, E. M., de Hoop, B., Viergever, M. A., Prokop, M., and van Ginneken, B., Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. *Med. Phys.* 36(7):2934–2947, 2009. doi:10.1118/1.3147146.

13. van Rikxoort, E. M., and van Ginneken, B., Automated segmentation of pulmonary structures in thoracic computed tomography scans : A review. *Phys. Med. Biol.* 58(17):R187, 2013. doi:10.1088/0031-9155/58/17/R187.

14. Sluimer, I., Prokop, M., and van Ginneken, B., Toward automated segmentation of the pathological lung in CT. *IEEE Trans. Med. Imaging* 24(8):1025–1038, 2005. doi:10.1109/TMI.2005.851757.

15. El-Baz, A., Suri, J. S., (Eds) *Lung Imaging and Computer Aided Diagnosis*: CRC Press, Boca Raton, 2011.

16. Nandy, K., Interactive segmentation and tracking in optical microscopic images. *Cytom. Part A* 81(5):357–359, 2012.

17. Otsu, N., A threshold selection method from gray-level histograms. *Automatica* 11:23–27, 1975.

18. Churg, A., *Thurlbeck's Pathology of the Lung*. Thieme, New York, 2005.

19. Li, L.-N., Ouyang, J.-H., Chen, H.-L., and Liu, D.-Y., A computer aided diagnosis system for thyroid disease using extreme learning machine. *J. Med. Syst.* 36(5):3327–3337, 2012. doi:10.1007/s10916-012-9825-3.

20. Sheikhtaheri, A., Sadoughi, F., and Hashemi, D. Z., Developing and using expert systems and neural networks in medicine: A review on benefits and challenges. *J. Med. Syst.* 38(9):1–6, 2014. doi:10.1007/s10916-014-0110-5.

21. Wang, Q., Zhu, W., and Wang, B., Three-Dimensional SVM with Latent Variable: Application for detection of lung lesions in CT images. *J. Med. Syst.* 39(1):171, 2015. doi:10.1007/s10916-014-0171-5.

22. Hu, S., Hoffman, E. A., and Reinhardt, J. M., Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. *IEEE Trans. Med. Imaging* 20(6):490–498, 2001. doi:10.1109/42.929615.

23. Itai, Y., Kim, H., Ishikawa, S., Yamamoto, A., and Nakamura, K., A segmentation method of lung areas by using snakes. *Int. J. Innov. Comput. Inf. Control* 3(2):277–284, 2007.

24. Boykov, Y., Jolly, M. P., Interactive organ segmentation using graph cuts. In Medical Image Computing and Computer-Assisted Intervention–MICCAI2000, 276–286, 2000.

25. Osareh, A., and Shadgar, B., A segmentation method of lung cavities using region aided geometric snakes. *J. Med. Syst.* 34(4):419–433, 2010.

26. Korfiatis, P., Kalogeropoulou, C., Karahaliou, A., Kazantzi, A., Skiadopoulos, S., and Costaridou, L., Texture classification-based segmentation of lung affected by interstitial pneumonia in high-resolution CT. *Med. Phys.* 35(12):5290–5302, 2008.

27. van Rikxoort, E. M., de Hoop, B., van de Vorst, S., Prokop, M., and van Ginneken, B., Automatic segmentation of pulmonary segments from volumetric chest CT scans. *IEEE Trans. Med. Imaging* 28(4):621–630, 2009. doi:10.1109/TMI.2008.2008968.

28. Kakar, M., and Olsen, D. R., Automatic segmentation and recognition of lungs and lesion from CT scans of thorax. *Comput. Med. Imaging Graph.* 33(1):72–82, 2009. doi:10.1016/j.compmedimag.2008.10.009.

29. Zhou, X., Hayashi, T., Hara, T., Fujita, H., Yokoyama, R., Kiryu, T., and Hoshi, H., Automatic segmentation and recognition of anatomical lung structures from high-resolution chest CT images. *Comput. Med. Imaging Graph.* 30(5):299–313, 2006. doi:10.1016/j.compmedimag.2006.06.002.

30. Wang, J., Li, F., and Li, Q., Automated segmentation of lungs with severe interstitial lung disease in CT. *Med. Phys.* 36(10):4592–9, 2009. doi:10.1118/1.3222872.

31. Massoptier, L., Misra, A., Sowmya, A., Automatic lung segmentation in HRCT images with diffuse parenchymal lung disease using graph-cut. 24th Inter Conf In Image and Vision Computing New Zealand, 2009. IVCNZ'09. 266–270, 2009.

32. Abbas, Q., Khan, M. T. A., Farooq, A., and Celebi, M. E., Segmentation of lungs in HRCT scan images using particle swarm optimization. *Int. J. Innov. Comput. Inf. Control* 9(5):2155–2165, 2013.

33. Boykov, Y., and Kolmogorov, V., An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Anal. Mach. Intell.* 26(9):1124–1137, 2004. doi:10.1109/TPAMI.2004.60.