RESEARCH ARTICLE

Fuzzy-C-Means Clustering Based Segmentation and CNN-Classification for Accurate Segmentation of Lung Nodules

Jalal Deen K[1]*, Ganesan R[2], Merline A[3]

Abstract

Objective: Accurate segmentation of abnormal and healthy lungs is very crucial for a steadfast computer-aided disease diagnostics. **Methods:** For this purpose a stack of chest CT scans are processed. In this paper, novel methods are proposed for segmentation of the multimodal grayscale lung CT scan. In the conventional methods using Markov–Gibbs Random Field (MGRF) model the required regions of interest (ROI) are identified. **Result:** The results of proposed FCM and CNN based process are compared with the results obtained from the conventional method using MGRF model. The results illustrate that the proposed method can able to segment the various kinds of complex multimodal medical images precisely. **Conclusion:** However, in this paper, to obtain an exact boundary of the regions, every empirical dispersion of the image is computed by Fuzzy C-Means Clustering segmentation. A classification process based on the Convolutional Neural Network (CNN) classifier is accomplished to distinguish the normal tissue and the abnormal tissue. The experimental evaluation is done using the Interstitial Lung Disease (ILD) database.

Keywords: Multimodal image- lung segmentation- Fuzzy-C-Means- CNN classifier- feature extraction

Introduction

Pulmonary diseases and disorders affect the human severely and lead to death and hospitalization. Radiological imaging methods like computed tomography (CT) is used in the standard hospitals to detect and diagnose the lung disease. It is also helpful to measure the rigorousness of the disease and to intend the treatment or surgical procedure (Mansoor et al., 2014). As a result of the technical developments in imaging methods, computerized investigation becomes automatic and does not require the effort of human. Several decision support modules are frequently aspired by physicians and radiologists to help out the diagnostic processes. For radiological based analysis of lung segmentation process, vigorous image processing tools are essential in order to extract features related to lung disease and lung structure in an effective manner. As soon as detecting the lung from the chest CT image, the segment with the disease is identified further using classification process to get accurate results.

Since the diagnosis of the lung diseases requires a huge range of measurements, a precise and hasty analysis tool is required an it should be economical too. The attainment of such a tool becomes a great challenge. As studied in paper (Armato and Sensakovic, 2004) the precision degree of lung segmentation process can disturb lung nodules recognition by 17%. The studies from (Hu et al.,

2001; Ganesan and Radhakrishnan, 2009) illustrated that the lungs are more black than the other chest parts and hence the accuracy of detection would become less in the presence of acute lung diseases. The latter segmentation processes (Wang et al., 2009; Kockelkorn et al., 2014) presented methods in order to circumvent these issues. These methods take into account several factors like graphical appearances, contours.

The conventional models, like Markov–Gibbs random fields (MGRF), demonstrate substantial possibilities in segmentation or noise removal process in multimodal images. The process stipulates a joint probability distribution of images and region of interest to approximate a preferred map for a given image. The image pattern commonly depends on the hypothesis of statistically autonomous image signals with dissimilar marginal probability distributions in every region. After recuperating the distributions from the diverse empirical signal distribution over the image, the preliminary segmentation process is accomplished by low-level pixel-wise classification. Typically, accurate segmentation results can be attained with the employment of peripheral probability distributions. A region map obtained after the preliminary pixel-wise classification is then distinguished by optimal statistical estimation of a concealed Markov model of regions. Though this process is seedy, it performs the segmentation process hastily. This Markov model

[1]*Department of Electronics and Instrumentation Engineering, Sethu Institute of Technology, Virudhunagar, [2]Department of Electronics and Instrumentation Engineering, Saveetha Engineering College, Chennai, [3]Department of Electronics and Communication Engineering, Sethu Institute of Technology, Madurai Tamilnadu, India. *For Correspondence: jalaldeensit@gmail.com*

based analysis is useful in a huge range of real-world applications for instance real-time robotic vision or automated analysis of medical images from computer tomography (CT), magnetic resonance imaging (MRI), and magnetic resonance angiography (MRA) (Gimel'farb et al., 2004). However, it bumps in to complications in perceiving exact region boundaries.

Cluster investigation is a process of clustering data set into many classes of analogous entities. In clustering methods the hard c-means process is the eminent and traditional clustering method which controls each map of the data set to one cluster (Wu and Yang, 2002).The hard c-means process is known as k-means method. Fuzzy c-means (FCM) are obtained from hard c-means (HCM). FCM has improved performance than HCM or k-means. The FCM is vigorous and provide accuracy. Another issue in the diagnosis of lung pathology is the false positive rate. This rate should be very low. Hence the proposed methodology must lessen the FP classifications without affecting the true positive classifications or sensitivity. Sophisticated neural network methods are used to categorize those scheduled features. Since it is a significant job to stipulate a good feature area, appropriate feature description, and respective extraction algorithms become very hard. The feature specification is very trivial for distinguishing the nodules and non-nodules of the lung image. In this paper, CNN is presented to specify the essential features of the multimodal chest image.

In this paper, a novel technique is presented based on FCM segmentation and CNN classifier. This is the fully automatic method covering the complete range of normally confronted diseases in chest CT scans. Additionally, the process is evaluated on Interstitial Lung Disease (ILD) database entailing various degrees and kinds of abnormalities. Section 2 gives the previous works done on the lung image segmentation. Section 3 illustrates the methods used in this paper. In section 4, experimental results are evaluated.

Materials and Methods

Appearance-based segmentation method employs texture features to extricate between elements which do not have perfect boundaries. In paper (Mansoor et al., 2014) an extensive range of lung images with disease are segmented in two stages. Primarily, the lung parenchyma was taken from fuzzy connectedness process, and the variances of the rib-cage from the lung parenchyma volumes were investigated to detect the disease. Wang et al., (2009) presented traditional Haralick's texture features to distinguish the normal and abnormal tissues in chest CT scans and analyzed indistinct boundaries for Inter stage Lung Disease (ILD). Initially, the simple thresholding method was employed to segment normal tissues and moderate ILD pathological tissue from the voxel-wise signals. From the developed segments, the severe ILD tissues are detected and merged with the primarily segmented areas. This method produced an average overlap of 96.7% with the conventional manual based segmentation method on a database consisting of 76 CT scans (31 normal and 45 ILD lungs). In paper

(Kockelkorn et al., 2014) the authors separated the chest CT scan into several 3D volumes containing voxels of equal intensities and classified each volume as the lung or local environment. Then, the missed voxels were adjusted through either a cooperative process or a slice-wise supervised classification method. Sluimer et al., (2005) utilized15 chest CT images to develop a probabilistic atlas of normal lung regions and registered a lung scan to the atlas with the intention of segmenting the lungs having severe ILD. In Zhou et al., (2014) an atlas-based segmentation model was implemented to segment lungs with tumors. In Nakagomi et al., (2013) employed a graph-cut segmentation technique which as simulated shape and other details of the adjacent lung regions. By conjoining more than one segmentation methods, hybrid segmentation method is attained to get high accuracies. Korfiatis et al., (2008) segmented lung image, with interstitial pneumonia based on the voxel-wise gray levels. Afterwards the original segmentation was sophisticated by classifying the voxels with the help of a support vector machine classifier. Lassen et al., (2010) employed a series of morphological functions to improve the threshold-based segmentation of the pulmonary spaces. Kockelkorn et al., (2010) performed segmentation of the lung CT images with a k-nearest neighbor classifier, trained on existing erstwhile data.

Shape-based segmentation method used precise information regarding lung shape with certain inconsistency in the scans to get more accuracy in the segmentation results. Sun et al., (2012) coordinated a 3D vigorous shape prototype to a CT chest image to coarsely describe the original lung borders. Then they distinguished the segmented regions with a global surface optimization technique presented by (Li et al., 2006). In (Rikxoort et al., 2009) the lungs are segmented using region growing and morphological operations techniques. Birkbeck et al., (2014) introduced statistical learning to functional constrictions originated from adjacent parts (e.g: heart, liver, and ribs) so as to segment lungs. These parts were first spotted by statistical classifiers and expended as geometrical restrictions for collapsing the lung region. Hua et al., (2011) proposed lung segmentation technique by augmenting a graph-supported cost function of voxel-wise intensities and their spatial gradients. They also considered the boundary flatness and rib controls. Nonetheless, these methods exhibit several shortcomings. Several techniques are affected by exclusive and user-reliant collaborations with a radiologist. In addition, they are affected by anatomical standards. These standards are problematic during the existence of pathological tissues. Furthermore, these methods cause segmentation process too subtle to demonstrate initialization and a number of control points. Many of the proposed detect only a certain kind of lung diseases like nodules and did not detect other kinds of lung tissues. The proposed method prevails over these problems and segment normal and abnormal lungs from chest CT scans with greater accuracy. The 3D CT chest images are first undergone for pre-processing step to detect their contextual voxels like air and bed depicted in Figure 1. The region growing methods are used for this purpose. In this paper, with the application of Gaussian Scale-Space

filter, the pre-processing stage is accomplished. In order to improve the accuracy of the segmentation process, a comprehensive former data of lung contours are attained from the training dataset of 3D CT images of normal and affected lungs. Then the features required for classifying the pathological lung nodules are extracted by using GLCM feature extraction method.

The aberration in the pathological lung CT image is detected by examining the visual outlines such as shape, texture, and attenuation information. These features offer the indispensable information so that they can assist enhancing diagnostic confidence and reliability. It is challenging issue to quantify the abnormal tissue area and to detect the exact type of disease when there are many abnormalities in the same voxel region. The GSS smoothing filter introduces long-range attributes to the actual voxel-wise intensities and the pair wise co-occurrences in the adjacent voxels of each voxel. GSS Filter treats the image assemblies at several scales, by demonstrating an image as a one-bound group of smoothed images. Gaussian scale space is a kind of linear scale space. The interesting scale of an image is the scale where the significant constitution is perceptible. An explicit scale factor called sigma is used to scale Gaussian derivatives. Scale-space illustration is parameterized by the dimensions of the smoothing kernel manipulated for quashing fine-scale assemblies.

Feature extraction using GLCM

In this paper, gray-level run length matrix (GLRLM), gray-level co-occurrence matrix (GLCM), and histogram methods are used to extract features. Previous works illustrated that texture, intensity, and gradient are the key features for automatic detection of abnormal image segments. In general, GLCM and GLRLM were computed using four different orientations (Mansoor et al., 2014). For every voxel in the region of interest, the features are extracted by taking a patch on that voxel with 7x7x7 neighborhood. The 24 distinct features extracted using GLCM, GLRLM and histograms are depicted in Table 1. The left and right lungs are examined separately to get the features of each individual lung.

Fuzzy-C-Means Clustering Algorithm

Clustering is the method of separating the data into homogenous units by considering the relationship of objects. The clustering method is the allocation of the feature vectors into N clusters. Every nthcluster has Cn as its center. Fuzzy Clustering is employed in numerous areas such as pattern recognition and Fuzzy detection. Among various kinds of fuzzy clustering methods, Fuzzy C-Mean clustering (FCM) is the extensively used one. FCM utilizes reciprocal distance to determine fuzzy weights. The input of this process is a pre known number of clusters, N. The mean position of every the members of a cluster is identified. The output is the segregating of N clusters on a class of objects. The goal of the FCM cluster is to reduce the total weighted mean square error, (MSE). The FCM consents each feature vector to match with several clusters of different fuzzy membership values.

The final segmentation is based on the optimum weight of the feature vector over all clusters. The steps involved in the FCM algorithm are given below.

Input: Feature vectors (image voxels)$v = \{v_1, v_2, \ldots, v_n\}, N=$ number of clusters.

Output: A group of Clusters' that lessening the sum of error of distance.

Steps:

1. Set random weight for every pixel using fuzzy weighting with positive weights $\{W_v^n\}$ ranging from 0-1.

2. Normalize the starting weights for each vth voxel on all N clusters by using the below equation

$$W_v^n \Big/ \sum_{i=1,N} W_v^i \tag{1}$$

3. Normalize the weights on n= 1,...,N for each v to get W_v^n as given below

$$W_v^n = \frac{W_v^n}{\sum_{i=1}^{V} W_i^n}, v = 1, 2, \ldots \ldots V \tag{2}$$

4. Estimate new centroids C_n, n = 1,...,N from

$$C_n = \sum_{v=1}^{V} W_v^n X(v), n = 1, 2, \ldots \ldots N \tag{3}$$

5. Update the weights $\{W_v^n\}$ by using (4)

$$W_v^n = \frac{(1/||x_v - c_n||^2)^{1/(p-1)}}{\sum_{i=1}^{N}(1/||x_v - c_i||^2)^{1/(p-1)}}, n = 1, 2, \ldots N, v = 1, 2, \ldots V \tag{4}$$

6. If the input is altered, do again from step 3, else stop the process.

7. Set each pixel to a cluster according to the maximum weight.

Convolution Neural Network

Architecture of one hidden layer is depicted in Figure 2. It is examined for its skill to classify the nodules. This network consists of three layers namely, one input layer, one hidden layer, and one output layer. The input layer has P^2 neurons that represent the $P x P$ pixel of the image obtained from segmentation process. The hidden layer contains groups of $N x N$ neurons organized as a sovereign $N x N$ feature map (where $N=P-r+1$) and the $r x r$ area is represented as the interested area. Each hidden neuron selects input from a $r x r$ adjacent section on the input image section. If the neurons in the similar feature map are one neuron distant, then their interested areas in the input layer are one pixel distant. Each neuron of the similar feature map is reserved to take the identical group of r^2 weights and accomplish the equal action on the resultant fragments of the input image.

The advantage of hindering the weights permits the network to achieve shift-invariant pattern recognition. Hence, the total action is represented as the $r x r$ convolution kernel. The feature map is the output obtained from the convolution of the input with the $r x r$ convolution kernel. Each hidden neuron yj creates its output by means of an activation function represented as in (5). The minimum and maximum activation functions are zero and one, correspondingly.

$$y_j(w, x, a_j) = f\left(\sum_{i=1}^{r^2}(w_{ji}x_i) + a_j\right) \tag{5}$$

w_{ji}- the weight between the hidden neuron, *j* and the pixel, *i* of the input image.

x_i- gray value of the input pixel i.

a_j- the bias of the hidden neuron j.

$x_1, x_2, \ldots x_r^2$ Pixels on input image and they are

connected to the neuron, j.

The output layer is entirely linked to the hidden layer. The sigmoid activation function, zo of the output neuron is represented by,

$$z_0(w,y,g_0) = \frac{1}{1 + exp\left\{-\left[\sum_{j=1}^{nN^2}\left(w_{0j}y_j\right) + g_0\right]\right\}} \quad (6)$$

w_{0j}- Weight between the output neuron and neuron, j in the hidden layer

nN^2- total number of neurons in the hidden layer

g_0- bias of the output neuron.

Hence, the network contains $(O+P^2+nN^2)$ number of neurons and $(nN^2 (r^2+O+ 1) +O)$ number of links. These numbers include the input neurons and bias links also. The number of independent links is given by $nN^2 (O+ 1) + nk^2+ O$. O represents the number of output neurons.

The network weights as well as the bias weights are altered by the application of the Back Propagation (BP) algorithm (Rumelhart et al., 1985). The BP algorithm iteratively alters the weights with the intention of reducing the total error of the actual output vector from the target vector. The error function to be reduced is called as the Sum-of-Squared Error (SSE).

During training, the interested areas within one hidden class are restricted to consume the equal form of weights. The weights between hidden and output layers and the weights of every interested area, are altered by means of stochastic mode. In this method, the weight difference for each training sample is obtained from each back-propagated error and are altered instantaneously for every neuron.

Results

Dataset used

The public database of ILD in the lung is used for evaluating the results of the proposed method experimentally. The considered database consists of 113
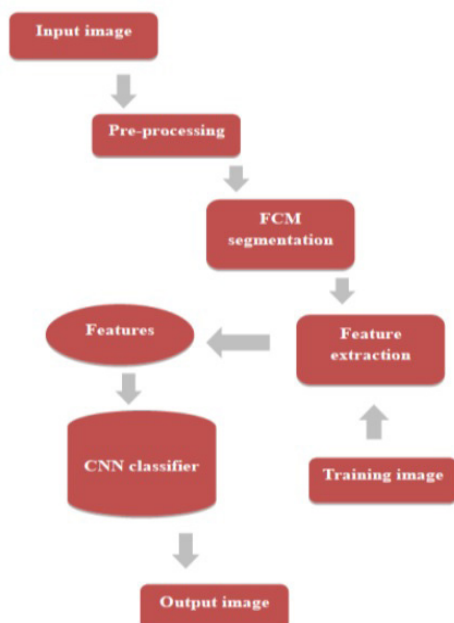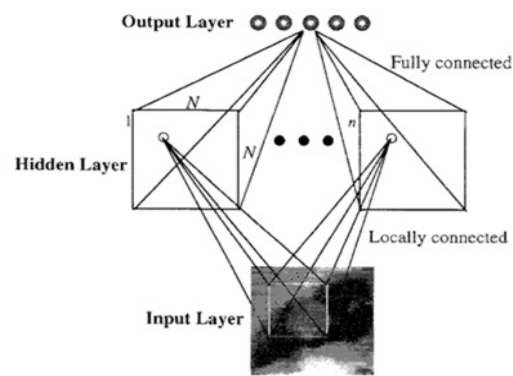

Figure 1. Block Diagram of the Process

Figure 2. Architecture of One Hidden Layer CNN (Lin, Lo, Hasegawa, Freedman, Mun, et al., 1996).

image sets of high-resolution CT (HRCT) images of dimension of 512 x512 pixels per slice. In this database, for the region of interest of every image set a tissue pattern annotation is given with 17 different tissue patterns. Five general tissue patterns are analyzed by the researchers in this database. These patterns are normal, emphysema, ground glass, fibrosis, and micro nodule are depicted in Table 2.

Implementation

By using the dataset mentioned above, the proposed method is evaluated in MATLAB software. Initially, the left and right lungs are segmented from the chest

Table 1. Feature Extraction Method and Features

S. No.	Features	Methods used for extraction
1	Energy	GLCM
2	Entropy	
3	Correlation	
4	Inverse Difference Momentum (IDM)	
5	Inertia	
6	Cluster shade	
7	Cluster prominence	
8	Short Run Emphasis	GLRLM
9	Long Run Emphasis	
10	Gray-level non uniformity	
11	Run length non uniformity	
12	Run percentage	
13	Low gray level Run Emphasis	
14	High gray level Run Emphasis	
15	Short Run Low gray level Run Emphasis	
16	Short Run High gray level Run Emphasis	
17	Long Run Low gray level Run Emphasis	
18	Long Run High gray level Run Emphasis	
19	Mean	Histogram
20	Variance	
21	Skewness	
22	Kurtosis	
23	Min.	
24	Max.	
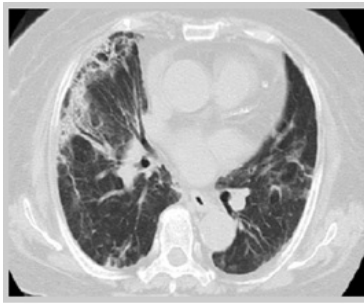
Figure 3. Input CT Image

Table 2. Summary of Dataset Used

S.No.	Tissue Category	Images	AROI	Patches
1	Normal(TN)	15	157	6,934
2	Emphysema (TE)	9	108	1,474
3	Ground glass (TG)	35	416	2,974
4	Fibrosis (TF)	35	479	4,456
5	Micronodule (TM)	18	298	7,893

Table 3. Comparison of Performance Parameters

S.No.	Parameters	MGRF [29]	FCM+CNN
1	MSE	0.3664	0.16
2	PSNR	52.4911	56.0902
3	Loss percentage	20	8
4	Accuracy	97	99

CT scan available in the set, using FCM segmentation process. Then GLCM, GLRLM and histogram methods are implemented to extract the features of the test image set as well as the training image set used to identify the lung nodules. The extracted features of the training image set and that of the test image set are applied to the CNN classifier. From the test and training samples, the classifier
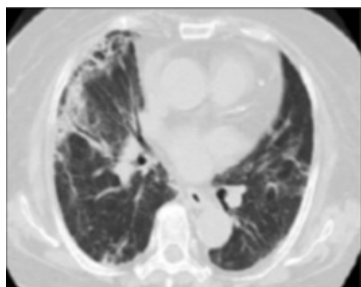

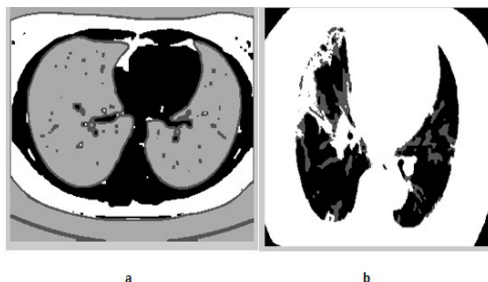Figure 4. Gaussian Scale Space Filtered Image

Figure 5. Existing Method (Soliman, Kalifa, Elnakip, et al., 2017) a) MGRF segmentation image b) classified image
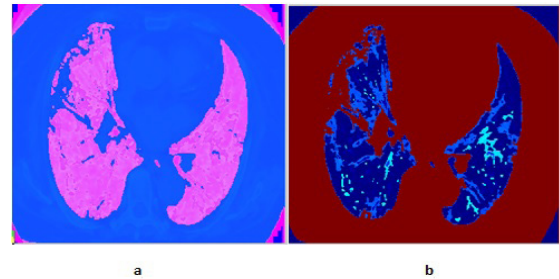

Figure 6. Existing Method a) FCM segmentation image b) CNN classified image

detects the lung nodules and label the pathological lung tissue and normal lung tissue. If inappropriately, there are only fewer training samples, then samples in various are produced from a single training input and then employed in the training phase. The oriented samples for a given training sample is generated by rotating the training sample at 90o, 180o and 270o.Moreover, by means of flipping the patterns and then rotating again at 0o, 90o, 180o and 270o, more number of training patterns are obtained. All these samples have the similar target output vector during training. Similarly, for every test pattern, eight rotated patterns are obtained by employing the same process mentioned above for training patterns. Then the output values are produced by the CNN for every rotated testing sample. Then true and false nodules are labeled.

Figure 3 represents the input image taken from the dataset for the implementation of the proposed method. After applying GSS filter the image is enhanced. The corresponding image after the application of GSS is depicted in Figure 4. The MGRF segmentation process proposed in the previous study (Soliman et al., 2017) is tested for the given input image. The segmentation result is shown in Figure 5a.The output of classification result is shown in Figure 5b. Similarly the segmentation output of the proposed FCM method is depicted in figure 6a and the classification image obtained from CNN classifier is shown in Figure 6b.

Performance comparison

The parameters used for comparing the performance of existing and proposed classification process are Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR) in Table 3. These parameters are the two error measures employed to compare image quality. The MSE signifies the cumulative squared error between the output image and the input image. PSNR denotes a measure of the peak error. For small value of MSE, the error is also less. The PSNR gives the peak signal-to-noise ratio (in decibels (dB)) between two images. This ratio is used as a factor for measuring the quality between the input and output image. For large PSNR value, the improved image quality is obtained.

Discussion

In computer aided diagnostics of lung image, detection, classification and quantification and segmentation are the vital steps. Existing methods of lung segmentation exploited the difference presenting in the image contrast

of the lung area and its adjacent tissues. These methods did not identify the abnormalities in the lung. In general, lung images have some pathologies and abnormalities happened due to experimental environment. In this paper, a novel method for fully automatic segmentation is proposed. This method contains machine-learning classification to manipulate pathologies. The performance of the proposed method evaluated with image sets available in the public database. The performance measures are also compared with that obtained from the existing method. To improve the accuracy of the image obtained from the segmentation process, GSS filter is used in the preprocessing step.

Statement conflict of Interest
Nil.

Acknowledgements

References

Armato SG, Sensakovic WF (2004). Automated lung segmentation for thoracic CT Impact on computer-aided diagnosis. *Acad Radiol*, **11**, 1011-21.

Birkbeck N (2014). Lung segmentation from CT with severe pathologies using anatomical constraints. *Med Image Comput Comput Assist Intery*, **17**, 804-11.

Campadelli P, Casiraghi E, Artioli D (2006). A fully automated method for lung nodule detection from postero-anterior chest radiographs. *IEEE Trans Med Imaging*, **25**, 1588-603.

Ganesan R, Radhakrishnan S (2009).Segmentation of computer tomography brain images using genetic algorithm. *Int J Soft Comput*, **4**, 157-61.

Gimel'farb G, Farag AA, Bag A, et al (2004). Expectation-Maximization for a linear combination of Gaussians. in Proc. IAPR Int. Conf. Pattern Recognition, **3**, 422–5.

Hu S, Hoffman EA, Reinhardt JM, et al (2001).Automatic lung segmentation for accurate quantization of volumetric X-ray CT images. *IEEE Trans Med Imaging*, 20, 490-8.

Hua S, Sonka H, Reinhard t, et al (2011).Segmentation of pathological and diseased lung tissue in CT images using a graph-search algorithm. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2072-5.

Korfiatis, Kalogeropoulou, Karahaliou, et al (2008).Texture classification-based segmentation of lung affected by interstitial pneumonia in high-resolution CT. *Med physics*, **35**, 5290-302.

Korfiatis P, kiadopoulos S, Sakellaropoulos P, et al (2014). Combining 2D wavelet edge highlighting and 3D thresholding for lung segmentation in thin-slice CT. *Br J Radiol*, **80**, 996-1004.

Kockelkorn TJP, Rikxoort M, Grutters C, et al (2010). Interactive lung segmentation in CT scans with severe abnormalities. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, **14**, 564-7.

Kockelkorn TJP, Prokop M, Bozovic, et al (2014). Interactive lung segmentation in abnormal human and animal chest CT scans. *Med physics*, **41**, 417-29.

Lassen K, Friman K, Peitgen H, et al (2010). Automatic segmentation of lung lobes in CT images based on fissures, vessels and bronchi. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, **14**, 560-3.

Lin JS, Lo S, Hasegawa A, Freedman MT, et al (1996).Reduction of false positives in lung nodule detection using a two-level neural classification. *IEEE Trans Med Imaging*, **15**, 206-17.

Li K, Wu X, Chen Z, Sonka, et al (2006). Optimal surface segmentation in volumetric images–A graph-theoretic approach. *IEEE Trans Pattern Anal Mach Intell*, **28**, 119-34.

Mansoor A, Bagci U, Xu, et al (2014). A generic approach to pathological lung segmentation. *IEEE Trans Med Imaging*, **33**, 2293-10.

Mendonça AM, Silva JA, Campilho A, et al (2004). Automatic delimitation of lung fields on chest radiographs. 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro, **2**, 1287-90.

Nakagomi K, Shimizu A, Kobatake H, et al (2013). Multi-shape graph cuts with neighbor prior constraints and its application to lung segmentation from a chest CT volume. *Med Image Anal*, **17**, 62-77.

Ross JC, Estepar, Diaz A, et al (2009).Lung extraction, lobe segmentation and hierarchical region assessment for quantitative analysis on high resolution computed tomography images. *Med Image Comput Comput Assist Intery*, **12**, 690-8.

Rumelhart DE, Hinton GE, Williams RJ, et al (1985). Learning internal representations by error propagation. *California Univ San Diego La Jolla Inst for Cognitive Science*, **8**, 317-62.

Silveira M, Nascimento J, Marques J, et al (2007). Automatic segmentation of the lungs using robust level sets. *Conf Proc IEEE Eng Med Biol Soc*, **10**, 4414-7.

Sluimer I, Prokop M, Ginneken B, et al (2005).Toward automated segmentation of the pathological lung in CT. *IEEE Trans Med Imaging*, **24**, 1025-38.

Soliman A, Kalifa, Elnakip A, et al (2017).Accurate Lungs Segmentation on CT Chest Images by Adaptive Appearance-Guided Shape Modeling. *IEEE Trans Med Imaging*, **36**, 263-76.

Sun S, Bauer C, Beichel R, et al (2012).Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach. *IEEE Trans Med Imaging*, **31**, 449-60.

Taher, Sammouda R (2011). Lung cancer detection by using artificial neural network and fuzzy clustering methods. IEEE GCC Conference and Exhibition, **10**, 295-8.

Ginneken B, Mikkel BS, Loog M, et al (2006). Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. *Med Image Anal*, **10**, 19-40.

Rikxoort EM, Hoop VA, Ptokop M, et al (2009). Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. *Med Phys*, **36**, 2934-47.

Wang J, Li F, Li Q, et al (2009). Automated segmentation of lungs with severe interstitial lung disease in CT. *Med phys*, **36**, 4592-9.

Wu KL, Yang MS (2002).Alternative c-means clustering algorithms. *Pattern Recognit*, **35**, 2267-78.

Zhou J, Yan G, Zhang S (2014).SU-F-BRF-02: Automated lung segmentation method using atlas-based sparse shape composition with a shape constrained deformable model. *Med Phys*, **41**, 399.