International Journal of Advanced Information Science and Technology (IJAIST) ISSN: 2319:2682
Vol.4, No.2, February 2015 DOI:10.15693/ijaist/2015.v4i2.204-211

Region based CT Lung image analysis with neural Network Multi-Level Classifier

K.Jalal Deen
Department of EIE
Sethu institute of Technology,
Tamil Nadu, India

Dr.R.Ganesan
Prof.& Head, Department of EIE, SIT
Sethu institute of Technology,
Tamil Nadu, India

***Abstract*- An artificial neural network ensemble is a learning paradigm where several artificial neural networks are jointly used to solve a problem. The techniques for utilized feature extraction included statistical, morphological features, as well as features derived from texture analysis, Fourier-based features and wavelet-based features. A New Dynamic Multi-Level classifier Computer-Aided Diagnosis (CAD) system for enhanced automatic detection of CT Lung Images for abnormalities is proposed in this paper. In this paper we also proposed to enhance the CT Lung images using Selective Median Filter (SMF) for improving the quality of the image by reducing the noise. We present new neural network Multi-Level Classifier segmentation method to reduce the suspicious region from the enhanced CT scan image. In this paper we also design a New Neural Network based Multi-Level Classifier for classifying the CT scan images using the extracted textural features.**

Keywords - ***Neural networks, Segmentation, Lung cancer Diagnosis, Selective Median Filter (SMF), Genetic Algorithm, Feature Extraction***

I. INTRODUCTION

The mortality rate of lung cancer is the highest of all types of cancer, and it is one of the most serious cancers in the world, with the smallest survival rate after diagnosis and showing a gradual increase in the number of deaths every year. Lung cancer is one of the most common and deadly diseases in the world. Detection of lung cancer in its early stage is the key of its cure. In general, measures for early stage lung cancer diagnosis mainly includes those utilizing X-ray chest films, CT, MRI, isotope, bronchoscopy, etc., among which a very important measure is the so-called pathological diagnosis that analyses the specimens of needle biopsies obtained from the bodies of the subjects to be diagnosed. At present, the specimens of needle biopsies are usually analysed by experienced pathologists. Since senior pathologists are rare, reliable pathological diagnosis is not always available. During the last decades, along with the rapid developments of image processing and pattern recognition techniques, computer-aided lung cancer diagnosis attracts more and more attention. Many achievements have already been attained.

X-ray chest films are valuable in lung cancer diagnosis. However, there are cases where subsequent examination should be performed to increase the reliability of diagnoses. For example, if a patient has suffered pneumonia before, some tumors may appear in abnormal lung areas so that they are difficult to be distinguished from cicatrices. The lung consists of airways, vessels, and connective tissue called the interstitium: a framework of thin membranes that supports and defines the structure of the lung. The airways and vessels are laid out in an intricate tree-like network, dividing into increasingly finer branches starting at the hilum, the location where they enter the lungs.

With the demands of slice thickness and spatial resolution requirements continue to increase, the amount of data lung inspection obtained will be greater. Such a huge amount of data will bring a great challenge to doctor's diagnosis. Therefore, the computer-aided diagnosis emerged. The key issue of medical image segmentation is to extract the image objects or special interest organization is a indispensable means of diseased tissue such as qualitative and quantitative analysis, at the same time is the basic of three-dimensional reconstruction. The 3D reconstruction model accuracy is directly affected by the qualities of segmentation. As the characteristic difference between medical imaging principles and organizing its own, medical image always is affected such as noise / field offset effect / partial body effect and the impact of organized campaigns, medical imaging will be blur and non-uniformity and has other characteristics.

Image segmentation methods include thresholding method, region growing algorithm, based on edge detecting methods, method for segmenting image based on pattern classification, watershed segmentation method, method based on anatomical model[6] and so on. Thresholding method is easy and fast for us to carry out, this method is deemed as the pre-processing used in conjunction with other methods. Disadvantage of thresholding segmentation method is not applicable of multi-channel image and characters are not very different from the images, for gray-scale images do not exist

Significant differences or has a bigger overlap region of gray value images is difficult to get accurate results. Region growing algorithm is widely used in medical image segmentation; method first select a seed point, then turn the similar pixels around the seed pixel into the seed area. It is simple, especially for a small segmentation structure. The disadvantage is need selecting the growing seed points manually. Based on edge detecting methods is by detecting the characteristic values of adjacent pixels to obtain the edges between different regions.

Method for segmenting image based on pattern classification can achieve great results but it requires large amount of training samples and it needs to extract image feature classification. The processing result was highly dependent on the characteristics, and it could take us a lot of time that could not meet the real time requirement of clinical diagnosis. Watershed segmentation method requires several morphological operations and watershed transformation in the lung region segmentation process, speed deficits, and the segmentation is not accurate enough. Method based on anatomical model is a common pattern of medical image processing. The method improved the automation of medical image segmentation, but required to establish a detailed and complete knowledge model with the help of medical experts. In order to achieve fully automatic segmentation of thoracic CT images, this article put forward a new automatic segmentation of lung parenchyma, according to threshold processing and mathematical morphology operations.

II. RELATED WORK

The author analyse this work is the comparison process where the reconstructed images are compared to the original CT image in terms of how much deviation is incurred in the reconstructed images due to noise (removal/addition) after normalizing all extracted texture measures. Two nonparametric statistical distance measures were used for comparison. Although these distance measures are often used in determining accuracy of clusters separability, they are used here to indicate how non-separable (i.e. close) the reconstructed images are to the original. And also found the best non-separable texture measure between the original and reconstructed images which is less susceptible to noise. This work investigated the susceptibility of five different texture analysis measures to noise by using two distance measurement methods to compare the original CT images with their corresponding reconstructed clean and noisy versions. It was shown that the texture measures with few features such as the ACF and FD was the least affected by noise in both distance tests as compared to GLCM which had the highest number of features. Also adaptively filtered images can assist in reducing subtle noise, and hence offer better texture accuracy [1].

Problem identified to minimum size of image was possible to texture analysis measures to noise by using two distance measurement methods to compare the original CT images. Previous work in this area have involved training machine learning algorithm using statistical, textural and/or morphological features of common patterns extracted from CT scans and documenting the success of classifying these patterns correctly.

The ability of rough sets to handle images: Rough sets provide reasonable structures for the overlap boundary given domain knowledge. The case study for images of the heart on cardiovascular magnetic resonance (MR) images also extends to handling multiple types of knowledge including: myocardial motion, location and signal intensity. A study concerned with distinguishing different picture types of the central nervous system is introduced [2]. The basic idea of a histon is to build a histogram on top of the histograms of the primary colour components red, green, and blue.

The authors show that the base histogram correlates with the lower approximation, whereas the encrustation correlates with the upper approximation. The problem of a machine vision application where an object is imaged by a camera system is considered in [3]. The object space can be modeled as a finite subset of the Euclidean space when the objects image is captured via an imaging system. Rough sets can bound such sets and provide a mechanism for modelling the spatial uncertainty in the image of the object. This work introduced a rough sets approach for building pattern matching systems that can be applicable with a wide range of images in medical sciences [3]. In this paper region growing is discussed using automatic tools, where the region growing algorithm learns its homogeneity criterion automatically from characteristics of the region to be segmented, and it

allows a segmentation of individual structures. Segmentation of pulmonary X-ray computed tomography (CT) images is a precursor to most pulmonary image analysis applications.

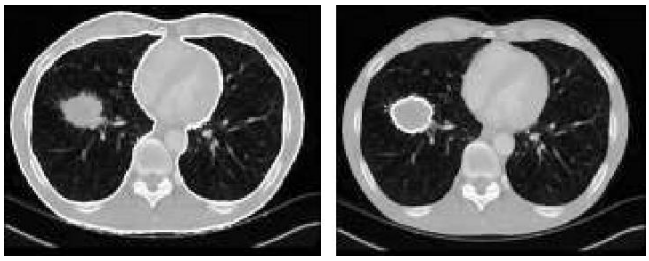
Manual segmentation of lung images is extremely time consuming for users, labor intensive and prone to human errors and hence an automated technique with proven algorithm is the only road to success in CAD applications. There is no generally applicable automatic segmentation technique that will work for all images as the images itself are quite complex and unique depending upon the domain application. The author presents an efficient region growing segmentation algorithm that can segment a region of interest into a more meaningful set of regions and objects. Some of them use a semi-automatic algorithm and still need some user interaction, while others are fully automatic and the user has only a verification role [4].

The author presents the potential for fractal analysis of time sequence contrast-enhanced (CE) computed tomography (CT) images to differentiate between aggressive and non-aggressive malignant lung tumors (i.e., high and low metabolic tumors). The aim is to enhance CT tumor staging prediction accuracy through Identifying malignant aggressiveness of lung tumors. As branching of blood vessels can be considered a fractal process, the research examines vascularized tumor regions that exhibit strong fractal characteristics. The analysis is performed after injecting 15 patients with a contrast agent and transforming at least 11 time sequence CE CT images from each patient to the fractal dimension and determining corresponding lacunarity.

The fractal texture features were averaged over the tumor region and quantitative classification showed up to 83.3% accuracy in distinction between advanced (aggressive) and early-stage (non-aggressive) malignant tumors. Also, it showed strong correlation with corresponding lung tumor stage and standardized tumor uptake value of fluorodeoxy glucose as determined by positron emission tomography. These results indicate that fractal analysis of time sequence CE CT images of malignant lung tumors could provide additional information about likely tumor aggression that could potentially impact on clinical management decisions in choosing the appropriate treatment procedure [5]. Review of existing literature has shown that studies have concentrated mostly on the analysis of CT-scan images to detect tumors and other anomalies of the lungs.

However, minimal work has been done in attempting to classify tumor classes based on these images for which new ground is broken here. The common practice to determine the tumor class is to perform a histopathological analysis on tissue samples obtained by invasive techniques such as a needle biopsy. As time and cost are crucial factors when it comes to the treatment of a lung tumor, an automated image based classifier could act as a precursor to histopathology-logical analysis, thus enabling the kick-starting of class specific treatment procedures. Based on CT-scan images of 74 patients with Non-Small Cell Lung Cancer (NSCLC) the first task was to develop an effective model to classify it into two subtypes, Adenocarcinoma and Squamous-cell Carcinoma [6].

The initial segmentation of the CT-scan images which segments out the lung region from the rest of the body was done using the built-in segmentation algorithm provided in the Lung Tumor Analysis (LuTA) software suite of Definiens [7]. On completion of the lung field segmentation, tumor identification was manually conducted by one of the radiologists at the H. Lee Moffitt Cancer Centre or a person with expertise in identifying lung tumors. The tumor, upon identification, was segmented out using LuTA's built-in region growing algorithm. The initial seed point for the algorithm was provided by the expert. The algorithm finds the tumor boundary across the image sequences. This boundary contains the tumor objects in each slice of the CT-image sequence. Figures 1.0 (a) and (b) shown the lung with tumor and with the tumor boundary outlined after region growing, respectively [7].



Figures 1 (a) and (b) shown the lung with tumor and with the tumor boundary outlined after region growing

Geometric Features

The first set of features to be looked at is geometric features. Geometric features, in both 2D and 3D provide vital structural information about the tumor object being analysed. The geometric features evaluated were deemed useful in analysing and quantifying biomedical images such as CT-scans [8]. Area (2D, 3D)

The number of pixels form an image object rescaled by using unit information. In scenes that provide no unit information, the area of a single pixel is 1 unit. Consequently, the area of an image object is the number of pixels forming it. If the image data provides unit information, the area of an image object is the true area covered by one pixel times the number of pixels form the image object.

The author shown that an automated texture-based system based on co-training is capable of achieving multiple levels of emphysema extraction in high-resolution computed tomography (HRCT) images. Co-training is a semi-supervised technique used to improve classifiers that are trained with very few labeled examples using a large pool of unseen examples over two disjoint feature sets called views. It is also shown that examples labeled by experts can be incorporated within the system in an incremental manner. The results are also compared against ‘‘density mask’’, currently a standard approach used for emphysema detection in medical image analysis and other computerized techniques used for classification of emphysema in the literature. The new system can classify diffuse regions of emphysema starting from a bullous setting.

The classifiers built at different iterations also appear to show an interesting correlation with different levels of emphysema, which deserves more exploration [9]. The author proposed an implementation of shape feature in the detection of obstructive lung diseases and the results shows improvement in classification sensitivity compared to feature based on texture only. However, their proposed system is dependent on the region size, e.g., 16x16, 32x32 and 64x64 pixels. Gathering region images from CT image is not an easy task especially with fixed size of region. To increase the efficiency while preserving the high sensitivity, a new feature is desired [10]. In this paper, a novel feature is called the continuous local histogram (CLH) is introduced. CLH integrates three basic types of features which are texture feature, shape power. The author described a fully automatic method for identifying the lungs in CT images.

The method has three main steps. First, the lung region is extracted from the CT images by gray-level thresholding. The left and right lungs are then separated by detecting the anterior and posterior junctions. Finally, we optionally smooth the lung boundary along the mediastinum. There are several distinctions between our method and previous work. First, instead of using a fixed threshold value, we use an optimal thresholding method to automatically choose a threshold value that reflects the gray-scale characteristics of a specific dataset. Second, we use an efficient method to find the anterior and posterior junction lines between the right and left lungs. Finally, to obtain more consistent results across time and to leave lung structures with the lung, we optionally smooth the irregular boundary along the mediastinum [11].

III. Description of Proposed Scheme

A New Dynamic Multi-Level Computer-Aided Diagnosis (CAD) system for enhanced automatic detection of CT Lung Images for abnormalities is proposed in this paper. In this paper we also proposed to enhance the CT Lung images using Selective Median Filter (SMF) for improving the quality of the image by reducing the noise. We present new neural network Multi-Level Classifier segmentation method to reduce the suspicious region from the enhanced CT scan image. In this paper we also design a New Neural Network based Multi-Level Classifier for classifying the CT scan images using the extracted textural features.

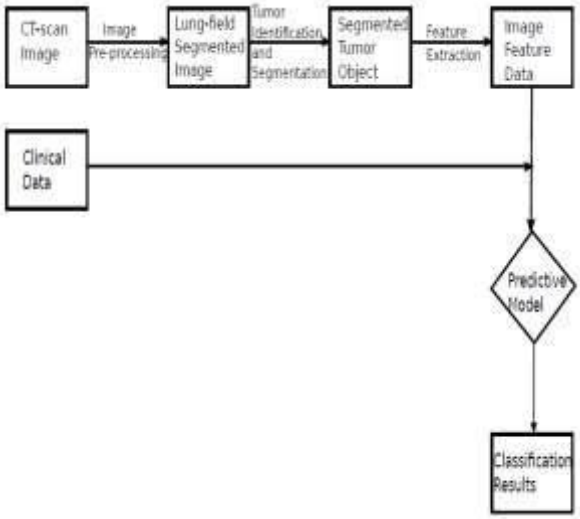


Figure 2.General workflow of the process of developing and using predictive models

A.Region based Lung region repair

Extracting the lung area is intended to find the tumor area in the lung area. While some tumor areas with tissue and the tumor area would not be included in the lung area after the binarization. For the characteristics of that expansion can expand the image morphology and corrosion can reduce the image morphology, this paper proposed a new multi-level boundary repair algorithm to enhance the CT Lung images using Selective Median Filter (SMF) for improving the quality of the image. The steps of the algorithms are as follows:

Steps: Using the fewer and more times of corrosion and expansion operations respectively to the original image.

We can only use the fewer times of corrosion and expansion operations once, while the times of the more we can obtain by experiments.

IV. Implementation Steps

The first stage involves a pipeline of image processing routines to separate the lung from other structures of original chest X-rays and identifying the region that is suspected of being a nodule. At this stage, the system extracts 32x32 square areas considering the suspicious point as the centre. Because it involves a multi-level region based pixel-based technique, all the pixels in the square region are considered as the inputs to the system. The intensity values of the pixels that fall within this region are extracted and are stored in a database which is used to train the system at the next stage.

In the second stage, new neural network Multi-Level segmentation is trained based on two types of inputs: pixel-based inputs in which we consider the intensity levels of pixels within the suspected region and dynamic feature-based inputs where we take into account first and second order dynamic features.

In the pre-processing stage selective Median filtering is required to remove the effect of poor contrast due to glare, noise and effects caused by poor lighting conditions during image capture. A low-frequency image was generated by replacing the pixel value with a median pixel value computed over a square area of 8x8 pixels centered at the pixel location. Sharpening and histogram equalization methods were used to enhance the contrast of the images.

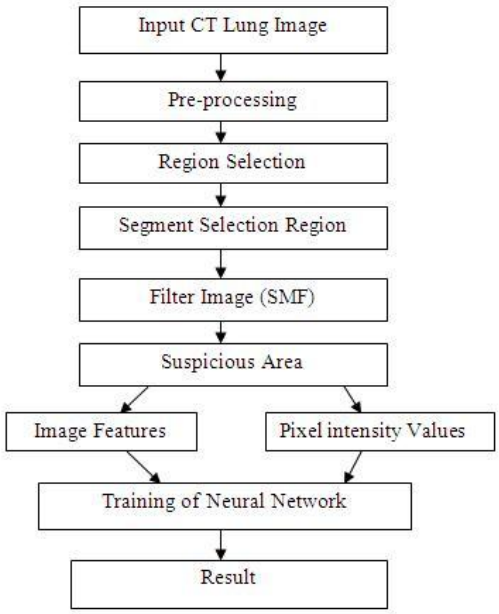


Figure 3. Pre-Processing *SMF*

A.Lung Region Segmentation

Lung masks were prepared using active shape models that are available with the database. These

images can be used to segment the CT lung region while the user can identify the scope when selecting suspicious points.

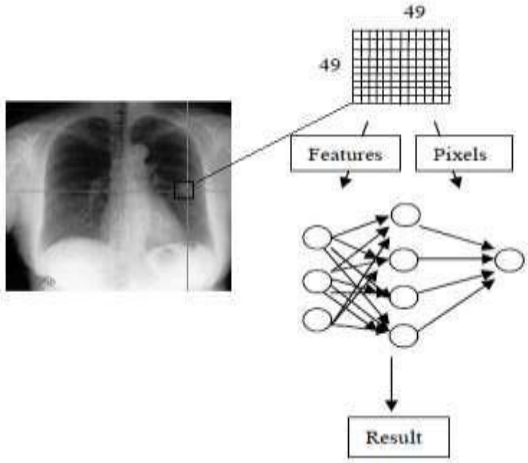


Figure 4. Layer Based Region Segmentation

The selection of suspicious region involves a 49x49 square mask which include nodules with 13 mm diameter with a resolution of 96 pixels per inch. In the image database, the size of a nodule varies between 8.9 mm and 29.1 mm with an average size of 17.4 mm. Because the size of lung nodule detected at an initial stage, is considered to be between 5mm and 20 mm.

Table 1: The calculated SMF of the filtered images using all types of filters

Filter Type	Output SMF
Average Filter	7.3133
Weighted Average Filter	4.2478
Gaussian Filter	6.4443
Selective Median Filter	8.5816
Wavelet Filter	8.8732
Wiener Filter	13.3969

B.Lung nodule segmentation

Thresholding technique is applied on the separated lung fields‟ images. The valley point value between the two peaks of the histogram is used as threshold value for segmentation of nodule (Fig. 2). In most of the cases it is seen that with this threshold value, the other lung field area apart from the nodule having similar gray (pixel) values is also segmented. Further the region labeling (in case of small-cell type lung cancer images (Fig. 3)) and region growing (for the non-small-cell type of lung cancer images) algorithms are applied for separating the nodules from the background. Morphological operations like dilation and erosion are used to remove the artifacts from the image while segmentation. The input patterns to the classification stage are the resulting features vectors from the features extraction and selection stage. The classifiers were trained and their performance was evaluated using the dataset of CT lung images.

Fig 5: (A) Lung Fields after Multiplication, (B) Image after Thresholding, (C) Separated Nodule

Region-growing is one of the conceptually simplest approaches to image segmentation. In this algorithm, neighbouring pixels of similar amplitude are grouped together to form segmented region. In the first stage of the process, pairs of quantized pixels are combined together in groups called atomic regions if they are of the same amplitude and are four connected. In the second stage, merging of the weak and common boundaries between the regions is carried out.

FigURE 6. Original Image, (B) Separated Lung Fields, (C) Separated Cancerous Portion

Region-labeling replaces each pixel by a negative number representing the label of the region to which the pixel belongs. The algorithm uses a list to keep track of pixels that are yet to be labeled. Insertion and removing of the pixel are the two operations carried out with reference to the list. Pixel is inserted at the end of the list and, pixel is removed from the front of the list.

C.Features extraction

Geometrical features like area, diameter, perimeter, and irregularity index have been estimated from the separated lung nodules. The number of pixels having the values „1"in the image array gives the area of the segmented tumor image. The algorithm is developed which estimates the area using bit quads, 2-by-2 pixel patterns. The number of boundary pixels in the tumor image is estimated as the perimeter of the tumor image. According to the morphology of tumor, the shape of the tumor is circular in nature.

To find out the irregularity in the circular shape, the circularity index is measured by using the equation like, $I = 4\pi A / p^2$: where, P is the perimeter of the tumor and A is area of the tumor in pixels. Geometrical features for the figure 5 (c) are included in table I. Texture or the contrast features are important features used in the classification of the lung cancers.

Contrast features are again classified under two categories, first order statistic and second order statistic. An advantage of using the wavelet transform coefficients is that the processing time/cost needed in the feature extraction stage is the least due to the compacted representation of the wavelet transform.

Table 2 Geometrical Features

S.No	Features	Value
1	Area	2815
2	Perimeter	226.85
3	Diameter	59.686
4	Irregularity index	0.69

D.ANN based classification

Further these estimated features are applied as input pattern to an expert system, which is designed to test the effectiveness of the input features so as to discriminate the lung cancer images. Artificial Neural Network (ANN) theory and practice suggest that, in a diagnostic application, the network should be trained with a balanced mixture of inputs from each diagnostic class. With this approach, a set of image samples consisting of 50% of small-cell-type, 50% of non-small-cell-type, and 50% of tuberculosis category are used for training and testing the network. In the presented work, a three layered, feed-forward artificial neural network is used.

It has the first layer as the input layer consisting of eight nodes, one hidden layer with 8 hidden nodes (h1 to h8), and three output nodes (o1, o2 and o3) in the output layer. The input layer uses 8 inputs as follows --- Avg. gray level (AGL), Std. deviation (STD), Smoothness (SMT), Third moment (THM), Uniformity (UNF), Entropy (ENT), Contrast (CNT), and Energy (ENG). The network is trained by using feed forward Back-Propagation algorithm.

For each training pattern presented to the input layer of the network, error at the nodes in the output layer of the network is estimated. Back-propagation algorithm refers to the propagation of error of the nodes from the output layer to the nodes in the hidden layers. These errors are used to update the weights of the network. The amount of weights to be added or subtracted to the previous weight is governed by delta rule.

IV. Conclusion and Future Work

Finally our proposed method analyse the various lungs image by reducing the noise using automatic segmentation algorithm in different region. The algorithm is based on the characteristic that the CT values of lung tissue have significant difference with the other tissue within the human body. We also proposed to enhance the CT Lung images

using Selective Median Filter (SMF) for improving the quality of the image by reducing the noise. We implement new neural network Multi-Level Classifier segmentation method to reduce the suspicious region from the enhanced CT scan image. We also designed a New Neural Network based Multi-Level Classifier for classifying the CT scan images using the extracted textural features. In future we analyse individual classifiers were then entered to a Multi-Level Classifier data fusion system for further improvement in the classification process.

REFERENCES

[1] Omar S. Al-Kadi IEEE Member and D. Watson: Susceptibility of texture measures to noise: An application to lung tumor CT images. BioInformatics and BioEngineering, 2008. 8th IEEE International Conference on BIBE 2008

[2] H. Shao, J.-W. Zhang, W.C. Cui, H. Zhao Automatic feature weight assignment based on genetic algorithm for image retrieval, in: IEEE International Conference on Robotics, Intelligent Systems and Signal Processing. 2003.

[3] B. Bhanu and Y. Lin: Object detection in multi-modal images using genetic programming, Appl. Soft Computing. 2004

[4] Nihad Mesanovic, Mislav Grgic, Haris Huseinagic, Matija Males, Emir Skejic , Muamer Smajlovic: Automatic CT Image Segmentation of the Lungs with Region Growing Algorithm 2011.

[5] Texture Analysis of Aggressive and Nonaggressive Lung Tumor CE CT Images Author: Omar S. Al-Kadi*, Member, IEEE, and D. Watson
IEEE Trans ON Biomedical Engineering, VOL. 55, NO. 7, July 2008

[6] Gonzales R.C. and Woods R.E., Digital Image Processing, 2nd Edition, New Jersey, Prentice Hall, 2004

[7] Tinku A., Ajoy K., "Image Processing Principles and applications," John Wiley & Sons, Inc., Hoboken, New Jersey, 2005.

[8] Bernd J.," Digital Image Processing," 6th revised and extended edition, Springer-Verlag Berlin Heidelberg, Printed in The Netherlands, 2005.

[9] Shi J., Malik J., "Normalizes Cuts and Image Segmentation," IEEE Trans., Pattern Anal. and Machine Intell., pp.888-905, 2000.

[10] Mancas M., Gosselin B., Macq B., "Segmentation Using a Region Growing Thresholding," Circuit Theory and Signal processing Laboratory, 2004.

[11]Y.Uchiyama, S.Katsuragawa, H.Abe, J.Shiraishi, F.Li, Q.Li, C.-T.Zhang, K.Suzuki and K.Doi. Quantitive computerised analysis of diffuse lung disease in high-resolution computed tomography, Med. Phys., pp. 2440-2453, 30 (9) September 2003