International Journal of Application or Innovation in Engineering & Management (IJAIEM)
Web Site: www.ijaiem.org Email: editor@ijaiem.org
Volume 3, Issue 12, December 2014 **ISSN 2319 - 4847**

Texture-Based Identification and Characterization of Pneumonia Patterns in Lung

Mr. Anup R. Aswar[1],Ms. Kunda P. Nagarikar[2], and Prof. Pragati D. Pawar[3]

[1]Mr Anup R. Aswar Electronics & Telecomm. Department
J.D.I.E.T. Yavataml

[2]Ms. Kunda P. Nagarikar Electronics & Telecomm. Department
J.D.I.E.T. Yavataml

[3]Prof. Pragati D. Pawar M.E. (Electronics)
J.D.I.E.T. Yavataml

ABSTRACT

Identification and characterization of diffuse parenchyma lung disease (DPLD) patterns is very difficult so an automated scheme for volumetric quantification of interstitial pneumonia (IP) patterns, a subset of DPLD, is presented, utilizing a multi detector CT (MDCT) dataset. Firstly, lung-field segmentation is achieved by 3-D automated gray-level Thresholding combined with an edge-highlighting wavelet pre- processing step, followed by a texture-based border refinement step. The vessel tree volume is identified and removed from lung field, resulting in lung parenchyma (LP) volume. Identification and characterization of IP patterns is formulated as a three-class pattern classification of LP into normal, ground glass, and reticular patterns, by means of k-nearest neighbour voxel classification, exploiting 3-D co-occurrence features. Performance of the proposed scheme in identifying and characterizing ground glass and reticular patterns was evaluated by means of volume overlap.

1.INTRODUCTION

The medical term used to describe the actual functioning parts of a human or animal lung is Lung parenchyma. It contain the alveolar walls as well as the blood vessels and the bronchi. If any part of the parenchyma get damaged or diseased, a person's life may be die. Pneumonia is a type of lung infection which makes the person very sick. Pneumonia may be caused by the bacteria, viruses or fungi. The symptoms of pneumonia are chest pain, cough, vomiting, difficulty in breathing and fever. Usually the sternness of pneumonia depends on the type of organism, cause of the inflammation, person's health and age. Computed tomography (CT) has become the wide choice for lung imaging. While high-resolution CT(HRCT) scan protocols allow visualization of limited portion of lung parenchyma (LP) (approximately 10%) and Multidetector CT (MDCT) allows acquisition of volumetric datasets with isotropic voxels, visualization, characterization, and quantification of the entire extent of lung anatomy, thus helps to characterization of diffuse parenchyma lung diseases (DPLDs) which are characterized by non-uniform distribution in the lung. The DPLDs is characterized by high inter and intra observer variability, It is complicated for image data reviewed due to lack of standardized criteria in assessing its complex and variable morphological appearance. Computer- aided diagnosis (CAD) schemes which automatically identify and characterize radiologic patterns of DPLDs in CT images have been proposed to improve management decisions. These are two stage in this systems that are as follows. 1)The segmentation of left and right LP based on gray-level methods. 2) Classification, identification, characterizations of LP into normal and abnormal tissue types. In 3-D DPLD texture analysis schemes, non-overlapping or overlapping sampling volumes of interest(VOIs) of LP are used. also 3-D texture analysis exploits first-order statistics , filter-based features ,co-occurrence matrices , run length matrices , 3-D local histograms, and fractal features , while classifiers, such as Bayessian , neural networks , support vector machines(SVMs), earth mover's distance, and *k*-nearest neighbours (*k*-NNs) are used.This paper presents a computer-aided scheme for the identification and characterization of interstitial pneumonia(IP) patterns in MDCT. The proposed method is differentiated from previously reported schemes by employing an LF segmentation. This stage is adapted to IP patterns (subset of DPLDs) affecting lung borders, by means of 3-D gray-level Thresholding combined with a lung-border voxel classification refinement step. Furthermore, the method incorporates a robust vessel-tree segmentation method utilizing an unsupervised Thresholding of responses produced by a 3-D multiscale enhancement filtering of vessels tubular structure. The identified vessel tree volume is removed from LF, to obtain the LP volume. Subsequently normal, ground glass, and reticular patterns are identified and characterized by employing *k*-NN voxel classification, exploiting 3-D co-occurrence analysis [1].

2.LITERATURE SURVEY

Korfiatis et al. have proposed a computer aided scheme for identification and characterization of Interstitial Pneumonia patterns in MDCT. First the segmentation of CT scan has been performed in two stages - lung field segmentation and vessel tree segmentation. The co-occurrence matrix for all regions are calculated from which eleven GLCM features are then extracted in five distances across four different directions, which provides the feature vector. This results in high dimensionality, which has been reduced through stepwise discriminant analysis (SDA). Finally, classification of image into three classes has been done using k-NN classifier.[1] M.H FazelZarandi et al. discussed the classical fuzzy approach to lung segmentation. Optimal smoothening of the input image has been done to reduce the noise followed by fuzzy c means clustering of the required lung tissues.[4] Arati S kurani et al. explained the feature extraction using Co-occurrence matrix 3 for 2D images and then moved on to how a similar technique can be used to construct co-occurrence matrix for volumetric data. It gives deep understanding of the Co-occurrence feature extraction. Francisco Moreno-Seco et al. proposed the KNN classification method. He discusses the advantages and disadvantages of KNN classification and proposed a modified classifier called K-NSN classifier which is more efficient than KNN classifier.[2] M.Gomathi et al. explained the fuzzy segmentation for lungs and gives four other varied techniques. The modified fuzzy segmentation technique is made use of in this system. The average weight used for redefining cluster centres gives better performance than the ordinary fuzzy segmentation[2].

3.SYSTEM ARCHITECTURE

The automated system under consideration is implemented using two different implementation strategies. These strategies are discussed below.

strategy 1.: In this strategy, 3D Histogram Thresholding is used for lung filed segmentation and Border Refinement performed using SVM classifier. After that the features are extract using 3D Co-occurrence matrices and irrelevant features are deleted using step-wise discriminant analysis. final step consist of K-NN classifier which is used to classify the features into healthy, interstitial pneumonia and other lung-diseased patterns.

Strategy2.: The strategy 2 is similar to the first strategy only there is a change in lung field segmentation. In this segmentation was done using Fuzzy segmentation technique . The co-occurrence matrices are used to extract features and the step-wise discriminant analysis was used to select the features. In the final step, K-NN classifier was used to classify the features into healthy, interstitial pneumonia and other lung-diseased patterns.

The most important drawback of standard FCM algorithm is it deals with image as separate points because the fuzzy function does not consider the spatial dependence. Now to decrease the noise effect during image segmentation, Fuzzy segmentation technique incorporates both the local spatial context and the non-local information into the standard FCM cluster algorithm using a novel dissimilarity index in place of the usual distance metric. Hence the 3D Histogram Thresholding and Border Refinement performed using SVM classifier is efficient than using fuzzy segmentation

The figure given below gives the idea about Sequence of Events required for Disease Identification[2][4].

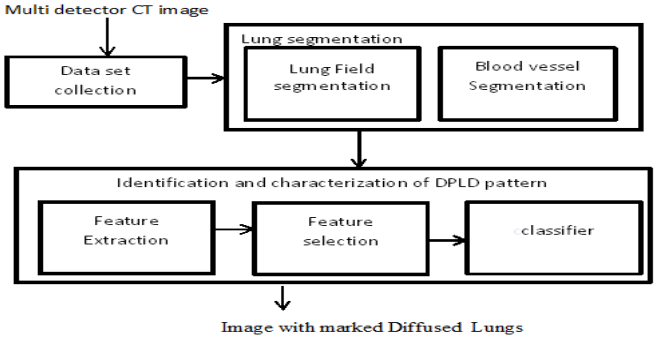


Figure 1.-System Architecture

4.MODULE OF THE STRATEGY

(1) Dataset Collection

This step includes collection of MDCT scans of various patients for analysis. The images collected are in the three different planes namely axial, spatial and coronal.

(2) Lung Segmentation

(a) Lung Field Segmentation

In this step the lung fields are removed from the rest of the image.

(b) Vessel Tree Segmentation

Once the lung fields have been extracted for analysis, the blood vessels are quantified to complete the segmentation of lungs.

(3) Feature Extraction

The necessary features required for analysis are extracted.

(4) Feature Selection

The relevant features are selected from the extracted features.

(5) Feature Classification

The selected features are categorised into reticular, ground class and normal patterns.

(6) Performance Evaluation

Compare the performances by using different algorithms by means of accuracy.

5.METHODS

5.1. DATASET

This step includes gathering of MDCT scans of various patients for analysis. The images gathered are in the three different planes namely axial, spatial and coronal. A pilot clinical case sample was obtained consisting of 14 MDCT scans in which four samples are of normal patients and ten patients diagnosed with IP secondary to connective tissue diseases, radiologically manifested with ground glass and reticular patterns. MDCT scans were obtained with a multi slice (16×)CT . Acquisition parameters of tube voltage, current, and slice thickness were 140 kVp, 300 mA, and 1.25 mm, respectively. The image matrix size was (512 × 512) pixels with average size of pixel was 0.89 mm. MDCT scans of five (out of ten) patients diagnosed with IP and MDCT scans of three normal patients were used to extract VOIs for training the k-NN classifier which is used for IP pattern identification and characterization. The training set of the classifier consisted of 350 cubic VOIs (21 × 21 × 21 pixels), which represent the patterns corresponding to reticular (150), ground glass opacities (100), and normal LP(100). MDCT scans of the remaining five patients were used for performance evaluation of the proposed method in identifying and characterization of ground glass and reticular patterns. Remaining MDCT scan of a normal patient (one out of four) was used to control of the proposed system performance on LP identification and characterization.

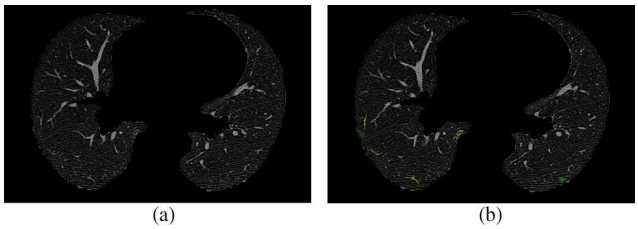


Figure 2-Normal lung Image

5.2. SEGMENTATION

5.2.1.1 Lung-Field Segmentation:

The LF segmentation algorithms is used as a pre-processing step in CAD schemes of lung disease. The method consist of a two-stage 3-D LF segmentation technique adapted to IP pattern affecting the lung border.The first stage of the algorithm adopts a 3-D histogram Thresholding LF segmentation algorithm combined with an edge-highlighting wavelet pre-processing step .However, the gray-level-based Thresholding algorithms are unable to correctly segmenting LF, in case of IPs affecting lung borders, since IPs are exhibit as tissue texture alterations. Hence a subsequent supervised texture classification refinement stage is used to deal with LF under-segmentation.

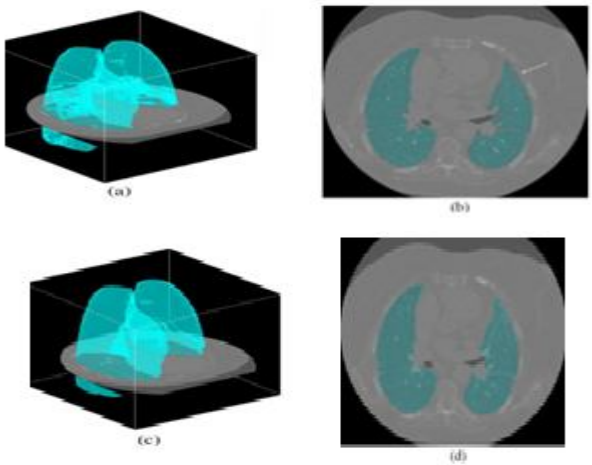


Figure3. LF segmentation example. (a) 3-D representation and (b) corresponding axial slice of segmented LF, provided by the first stage of the algorithm. (c) 3-D representation and (d) corresponding axial slice of segmented LF after application of the border refinement step. Segmented LF is depicted by cyan overlay.

Using an SVM classifier the iterative neighbourhoodlabelling of lung border voxels is performed. The SVM classifier assigns a label of LF or surrounding tissue (ST), using as inputs four first-order statistics (mean, standard deviation (SD), skewness, and kurtosis) extracted from a 7 × 7 × 7 pixel VOI centered at the voxel being labelled. Voxel labelling is initially applied on each border voxel of the initial LF volume, as provided by the 3-D gray-level-based algorithm, and subsequently, on its 18-connected neighbours. The initial LF volume is updated by adding voxels labelled as LF and removing voxels labelled as ST (muscle fat and bone). The process continues by checking every neighbouring voxel of an already labelled one, until the left and right LF volumes stay unaltered. The outermost voxels of corresponding unaltered LF volumes provide the final left and right LF borders. Coordinates of already labelled voxels are stored to avoid double-checking of neighbouring voxels during the LF volume updating[1][2][3]

5.2.1.2. Vessel-Tree Segmentation:

To reduce false-positive detections (items that are incorrectly labelled as positive) of vessels, accurate segmentation of vessel tree structure is required. For vessel-tree segmentation, a recently proposed algorithm in pulmonary CT angiography is accept, which exploits the tubular structure of the vessel tree. Particularly, a 3-D multiscale filter is applied on the segmented LF volume to enhance vessels and vessel bifurcation points and to reduce the noise present in the segmented LF volume, Along with the eigenvalues of the Hessian matrix at multiple scales (σ = 1, 2, . . . , 12 pixels). An expectation maximization algorithm is adopted to threshold the high response voxels at each scale, resulting in vessel segmentation. Finally, merging of vessel structure segments is performed to obtain final segmentation result. The segmented vessel tree volume is removed from the LF volume, resulting in identification of the LP volume[1][2][3]

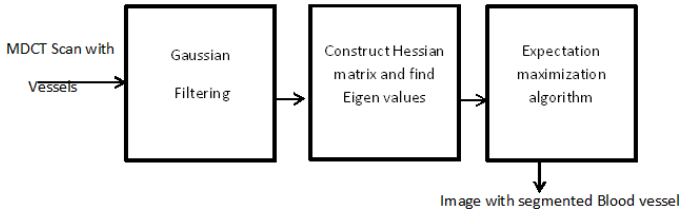
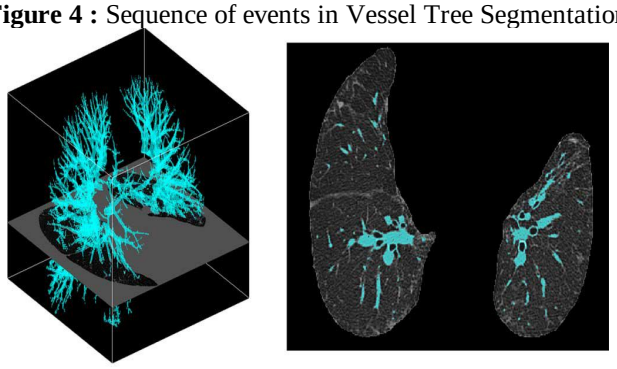


Figure 4 : Sequence of events in Vessel Tree Segmentation



(a) (b)

Figure.5 Example of vessel-tree segmentation. (a) 3-D representation and (b) corresponding axial slice of the segmented vessel tree depicted with cyan overlay.

5.3IP Pattern Identification and Characterization

IP pattern identification and characterization is accomplish /85by subjecting LP volume to three-class voxel classification based on 3-D texture analysis. Specifically, a k-NN classifier is employed to assign a label of normal, ground glass, or reticular to each LP voxel utilizing overlapping VOI sampling.[1]

5.3.1 3-D Gray-Level Co-occurrence Features:

The gray-level co-occurrence matrix (GLCM) is a well-established tool for texture analysis and characterizing the spatial distribution (second-order statistics) of gray levels in an image. Each and every element at location (i, j) of the co-occurrence matrix indicate the joint probability density of the occurrence of gray levels i and j in a particular direction θ and at a specified distance d from each other. The 3-D co-occurrence matrix stores the number of co-occurrences of pairs of gray levels i and j, which are separated by a distance d (in this study, d = 1, 2, . . . , 5 voxels) in 13 directions of a VOI . For each distance d 13 3-D co-occurrence matrix features were calculated from a sliding 21 × 21 × 21 pixel VOI within the LP volume (angular second moment, contrast, correlation, variance, inverse different moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, information measure of correlation 1, and information measure of correlation 2). The mean and range of each feature over the 13 co-occurrence matrices (corresponding to 13 directions) was calculated, comprising a total of 26 GLCM-based features for each distance d. In total, 130 features were calculated per VOI.[1][2][6][5]

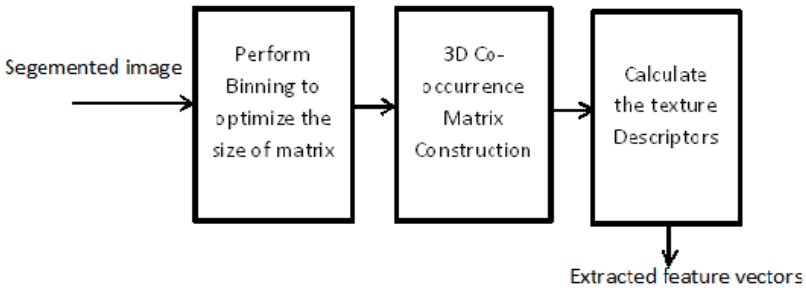


Figure 6 : Sequence of events in 3D Co-occurrence Matrix Feature Extraction

5.3.1.2. Feature Selection:

A statistical approach, the stepwise discriminant analysis (SDA) is adopted to reduce the dimensions of the feature vector (130). A covariance matrix for each of the features was constructed between all the three classes (interstitial pneumonia, healthy and other lung disease). The covariance matrix gives the difference between the corresponding feature values of the various classes. A threshold was set after trial and error, which selects a set of features that help to distinguish the various classes better. The idea was higher variance meant feature values farther apart, which made classification easier. Steps for Step-wise Discriminant Analysis are

(a) The Interclass Covariance matrix is for the features is constructed.

(b) From the covariance matrix the high features with high variance are Selected[1][2]

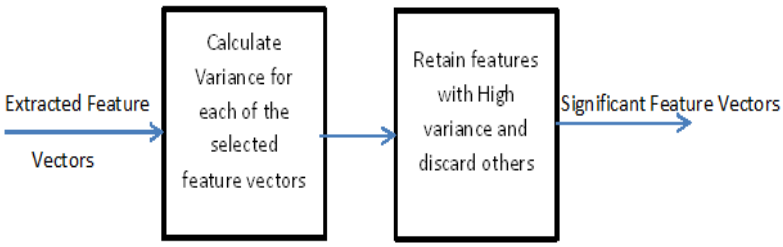


Figure7 : Sequence of events in Step-wise Discriminant Analysis

5.3.1.3 k-NN Classifier:

k-NN classification is one of the simplest supervised classification techniques in the field of statistical pattern recognition. In the present study, a k-NN classifier was used to assign to each LP voxel a label of normal, ground glass, or reticular, using as inputs the set of selected texture features. The k-NN classifies an unknown pattern according to the majority vote of its k-NNs. In this paper, the Euclidean distance was used as criterion. The number of neighbours (k) was selected based on the maximum correct classification rate, using a tenfold cross validation method. The classifier training dataset is partitioned into ten subsamples. Of the ten subsamples, one is retained as the testing sample, while the remaining nine subsamples are used as training data. The results from the ten folds are averaged to produce the generalized classification rate. A maximum rate of 0.932 ±□ 0.004 was achieved, for k = 10 NNs. Prior to classification, features were normalized to zero mean and unit variance. The parameters of the normalization were used to normalize the feature vector of an unknown pattern. [1][2]

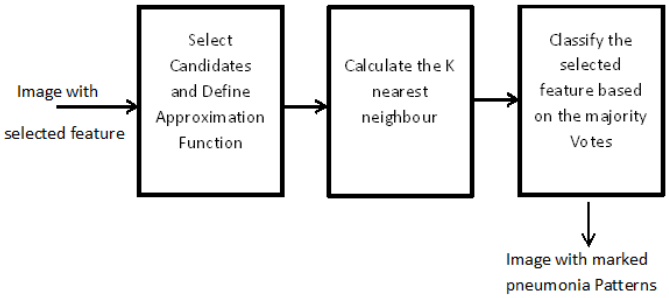


Figure 8 : Sequence of events in K-NN classification

5.4. Performance Evaluation Metrics

The performance of the proposed method in identifying and characterizing ground glass and reticular patterns was evaluated by means of volume overlap (VO), false-positive, and false negative fraction on five MDCT scans. For this purpose, a second radiologist with expertise in CT image interpretation, defined the voxel-exact ground truth for ground glass and reticular patterns by generating manual outlines within the segmented LP. For manual delineation, a tablet (Wacom Intuos3 Tokyo, Japan) was used with an active area of 305 × 305 mm with resolution of 0.2 lp/mm (lines per millimeter) and accuracy of □±0.25 mm. The metric of VO between "ground truth" (O), as defined by the second radiologist and computer (C)-derived borders, is given by as follows:

$$V_O = \frac{O \cap C}{O \cup C} \text{----------(1)}$$

To further quantify the fractions of the segmented results included or excluded with respect to the ground truth, the true positive fraction (TPF) and the false-positive fractions (FPFs) are also calculated [14]. They were defined as:

$$TP=\frac{O \cap C}{O} \text{---------- (2)}$$

And

$$FPF=\frac{O \cup C - O}{O} \text{---------- (3)}$$

Accurate segmentation is characterized by high TPF and low FPF. The value of VO is bound between zero (no overlap) and one (exact overlap)[1][2]

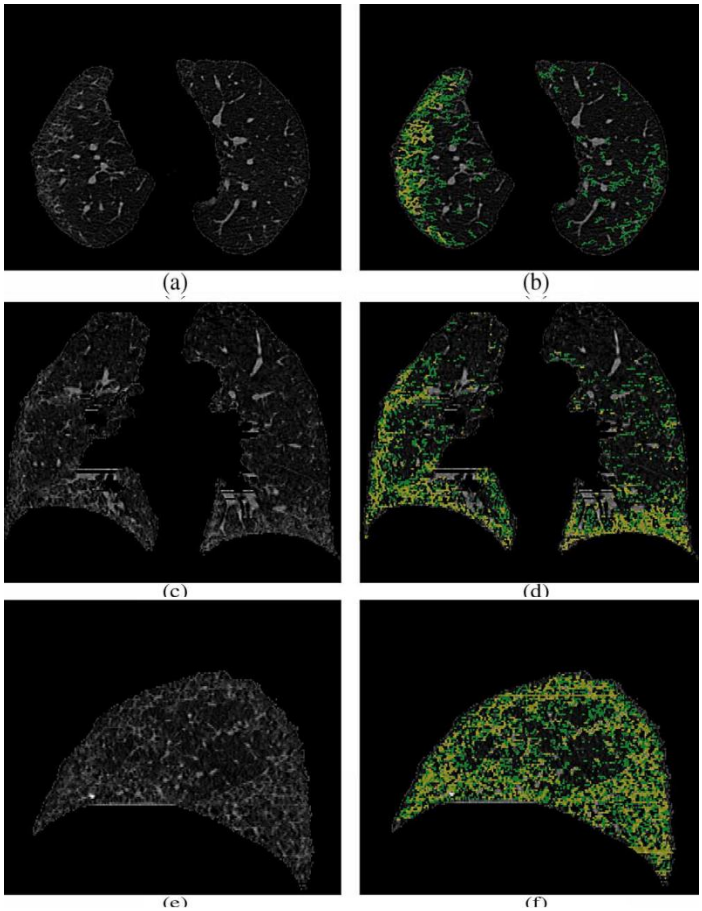


Figure 9. Application example of the proposed method on a case including both ground glass and reticular patterns in (b) axial, (d) coronal, and (f) sagittal representation. Yellow and green overlays correspond to reticular and ground glass patterns respectively. (a) Original axial, (c) coronal, and (e) sagittal slice data are also provided.

6.CONCLUSION

An automated system for identification and characterization of interstitial lung disease, as depicted on MDCT scans is presented. The system is based on an optimized data pre-processing step to isolate LP and on a three-class *k*-NN classification approach, utilizing 3-D co-occurrence features to classify LP voxels into three categories: normal, ground glass, and reticular. Preliminary results are promising, suggesting an accurate and reproducible system. Such systems are expected to assist radiologists in detection, characterization, and follow-up quantification of interstitial DPLDs.[1][2]

FUTURE SCOPE

This automated system can be extended to the diagnosis of other DPLDs. This class of disease patterns are very small and hence difficult to identify in the initial stages. This system can be enhanced to identify many more of the DPLDs. The third class which has been classified as other lung diseases consists of fibrosis and lung emphysema. Lung cancer, silicosis, tuberculosis, dermatomyositis, etc. can be included. In that case, training for the various diseases are to be included. Detailed Report generation from the statistics obtained from this system. The system can also have a report

generation module to give a formatted report to the patients. The various reports can be stored in the form of databases, which can be used for the classification also. So the learning changes from supervised to unsupervised learning. The 2D CT scan images can be viewed in a 3D axis by stacking of sequential 2D scans. This would prove useful in exact location of the disease patterns. More images can also be taken. Diagnosis of diseases for other organs using similar techniques and strategies can also be done.[2]

REFERANCES

[1] Panayiotis D. Korfiatis, Anna N. Karahaliou,Alexandra D. Kazantzi, CristinKalogeropoulou, and Lena I. Costaridou, Texture-Based Identification and Characterization of Interstitial Pneumonia Patterns in Lung Multidetector CT, IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, VOL. 14, NO. 3, MAY 2010

[2] S.R.AvinashKautham , Kavitha Ravi Analysis of Multi-detector CT scans for diagnosis of Interstitial Pneumonia Patterns

[3] Vaijinath V. Bhosle, Vrushsen P. Pawar**,** Texture Segmentation: Different Methods International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-5, November 2013

[4] M.H. fazelzarandi, m. Moeen, sh. Norouzzadeh and sh. Teimourian, "fuzzyImage processing for diagnosing inflammation in pulmonary biopsies",Transaction e: industrial engineering, vol. 16, no. 2, sharif university ofTechnology, december 2009, pp 3-13.

[5] Ryszard S. Choras, "Image Feature Extraction Techniques and TheirApplications for CBIR and Biometrics Systems", International Journal of Biology and Biomedical Engineering, Issue 1, Vol. 1, 2007, pp 5-7.

[6] S.Audithan, Dr.R.M.Chandrasekaran, "Document Image Segmentation based on gray Level Co-occurrence matrices and Feed Forward Neural Networks", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 8, No.5, August 2010, pp 3,4.

[7] Johns Hopkins Bloomberg School of Public Health, Pneumonia Progress Report 2012 International Vaccine Access Center (IVAC)

AUTHOR

Mr.Anup R. Aswarpersuing degree in Electronics & Telecommunication Engineering from J.D.I.E.T. Yavatmal

.

Ms. Kunda P. Nagarikarpersuing degree in Electronics & Telecommunication Engineering from J.D.I.E.T. Yavatmal

Prof. Pragati D. Pawar received the B.E. degrees in Electronics & telecommunicationEngineering from K.I.T.S, Ramtek and M.E. in Electronics from S.G.G.S. Nanded