Texture-Based Identification of Interstitial Pneumonia Patterns in Lung Multidetector CT

Mr. Anup R. Aswar **Ms. Kunda P. Nagarikar** **Prof. P. D. Pawar**

***Abstract*—Identification and characterization of diffuse parenchyma lung disease (DPLD) is very challenging in computer aided scheme for CT lung image analysis. In this study, an automated scheme for volumetric quantification of interstitial pneumonia (IP) patterns, a subset of DPLD, is presented, utilizing a Multidetector CT (MDCT) dataset. Initially, lung-field segmentation is achieved by automated gray-level thresholding followed by a texture-based border refinement step. Then identification and characterization of IP patterns is formulated as a two-class pattern classification of LP into normal, and infected by means of *k*-nearest neighbor voxel classification, using co-occurrence matrix which is used for feature extraction and feature is selected using high variance.**

***Index Terms*—Image segmentation, image texture analysis, respiratory system.**

I. INTRODUCTION

The medical term used to describe the actual functioning parts of a human or animal lung is Lung parenchyma. It contains the alveolar walls as well as the blood vessels. Pneumonia is an inflammation of the lungs. It is usually caused by infection with viruses or bacteria and less commonly other microorganisms and certain drugs. The symptoms include a cough, chest pain, fever, and difficulty breathing. . It can have more than 30 different causes. Understanding the cause of pneumonia is important because pneumonia treatment depends on its cause.[7]

Computed tomography (CT) has become the wide choice for lung imaging because high-resolution CT (HRCT) scan protocols allow visualization of limited portion of lung parenchyma (LP) 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. 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, characterisations of LP into normal and abnormal tissue types.

This paper presents a computer-aided scheme for the identification and characterization of interstitial pneumonia(IP) in MDCT. The method is differentiated from previously reported schemes by employing an LF segmentation stage adapted to IP patterns affecting lung borders, by means of gray-level thresholding combined with texture refinement step. The features are extract by co-occurance matrix and feature selection is incorporate by variance method. Subsequently normal and infected images are identified and characterized by employing *k*-NN classification.[1]

II. 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.

M.H Fazel Zarandi 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.

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.

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][4]

III. SYSTEM ARCHITECTURE

In this strategy, 3D Histogram thresholding is used for lung filed segmentation and Border Refinement performed by texture analysis. After that the features are extract using Co-

occurrence matrices and final step consist of K-NN classifier which is used to classify the features into normal and infected one. The figure1 given below gives the idea about Sequence of Events required for Disease Identification,[2]

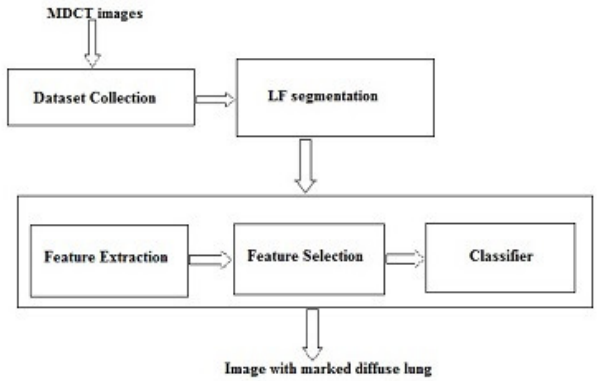

Figure1: System architecture

IV. METHODS

A. DATASET

This step includes gathering of MDCT scans of various patients for analysis. A dataset consisting of 10 MDCT scans in which five samples are of normal patients and five patients diagnosed with IP secondary to connective tissue diseases. 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 pneuminia identification and characterization. Remaining MDCT scan of a normal patient (one out of five) was used to control of the proposed system performance on LP identification and characterization.

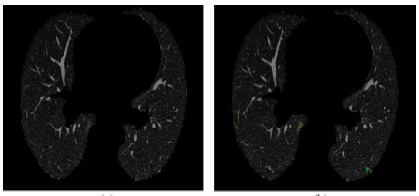

Figure2: Normal lung Image

B. SEGMENTATION

1. Lung field segmentation

The LF segmentation algorithm is used as a pre-processing step in CAD schemes of lung disease. The method consist of a two-stage LF segmentation technique adapted to IP pattern affecting the lung border. The first stage of the algorithm adopts a histogram thresholding LF segmentation algorithm. 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.

In this study, segmentation is carried out using histogram thresholding. There are different types of thresholding like local, global and optimal thresholding. The local thresholding is convenient to use.

The steps for local thresholding are as follows:

1) Select the threshold value T and segment it.
2) After segmentation we get output:
   G1 consisting of all pixels with gray level values >T.
   G2 consisting of pixels with values <=T.
3) The average of each set is computed.
4) Compute a new threshold value T=(1/2)(mean1 +mean2)
5) Keep repeating steps 2 to 4 until new threshold value match before one.[1][2][3]

C. IP PATTERN IDENTIFICATION AND CHARACTERIZATION

IP pattern identification and characterization is accomplished by subjecting LP volume to two-class voxel classification based on texture analysis. Specifically, a k-NN classifier is employed to assign a label of normal and infected to each LP voxel.[1]

A. 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 of gray levels in an image. There are two types of GLCM- 2D and 3D co-occurrence. 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 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 co-occurrence matrix features were calculated within the LP volume which are 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. 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]

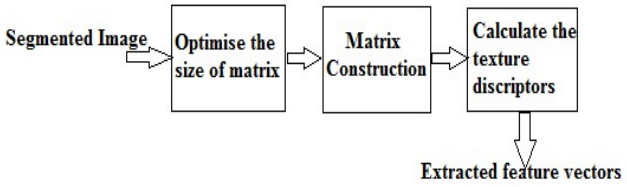

Figure3: Sequence of events in Co-occurrence Matrix Feature Extraction.

B. 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 two classes (interstitial pneumonia, normal 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 feature is constructed.

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

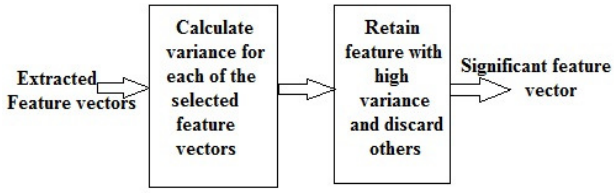

Figure4: Sequence of events in Step-wise Discriminant Analysis.

C. 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 toassign to each LP voxel a label of normal and infected using as inputs to 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.[8]

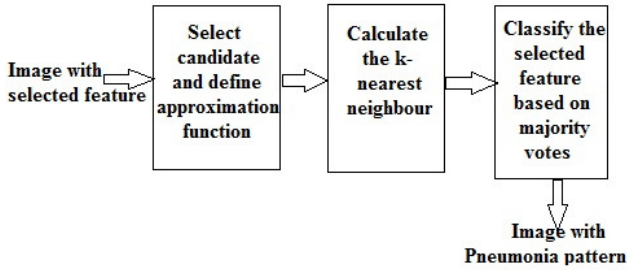

Figure5: Sequence of events in K-NN classification

Euclidean distance is calculated using following formula

$$d(p, q) = \sqrt{\sum(p_i - q_i)^2}$$

Where p_i=feature vectors of training image.
q_i = feature vectors of test image.

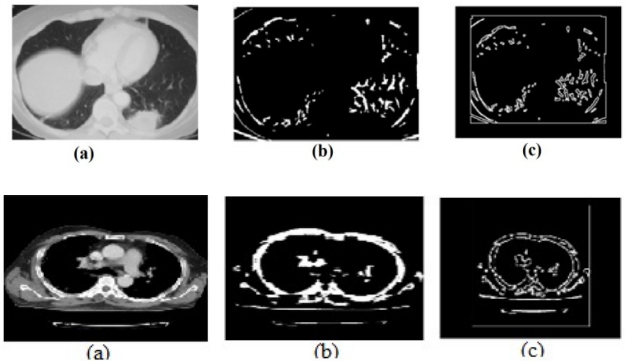

Figure 6: (a)Original image, (b) Segmented image, (c) Output of texture refinement step

IV. RESULT

For various different input MDCT scans, the system was able to correctly classify the input images and report whether the organ is infected with Interstitial Pneumonia or not. The 2fig 1 shows, the original image of lung and the segmented image after applying the Histogram Thresholding algorithm. After that the texture refinement step is used for texture analysis. The various feature are extracted and one of the feature is selected with high variance And the image is classified as normal or infected using K-NN classifier.

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.

ACKNOWLEDGMENT

We are extremely grateful to Dr. S.M. Gulhane Head of the Department of Electronics and Telecommunication Engineering, for extending the facilities of the Department towards our project and for his unstinting support.

We express our deep gratitude to our guide, Prof P.D. Pawar for guiding us through every phase of the project. We appreciate her thoroughness, tolerance and ability to share her knowledge with us. We thank her for being easily approachable and quite thoughtful. Apart from adding her own input, she has encouraged us to think on our own and give form to our thoughts. We owe her for harnessing our potential and bringing out the best in us. Without her immense support

through every step of the way, we could never have it to this extent.

We would like to thank Dr Patil mam for providing us the input datasets and helping us with the manual analysis of the CT images. We thank our parents, family, and friends for bearing with us throughout the course of our project and for the opportunity they provided us in undergoing this course in such a prestigious institution.

REFERENCES

[1] Panayiotis D. Korfiatis, Anna N. Karahaliou, Alexandra D. Kazantzi, Cristin Kalogeropoulou, 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.Avinash Kautham , 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, "fuzzy Image processing for diagnosing inflammation in pulmonary biopsies", Transaction e: industrial engineering, vol. 16, no. 2, sharif university of Technology, december 2009, pp 3-13.

[5] Ryszard S. Choras, "Image Feature Extraction Techniques and Their Applications 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)

[8] Oliver Sutton "Introduction to k Nearest Neighbour Classification and Condensed Nearest Neighbour Data Reduction."