FEATURE BASED IMAGE PATCH APPROXIMATION FOR LUNG TISSUE CLASSIFICATION

Dr.B.Vijayakumari[1],S.Abinaya[2], C.Chinnathai Jegathambal[2], R.Priyanka[2], A.Sivaravathi[2]
[1] Associate Professor, [2]UG student
Department of Electronics and Communication Engineering,
Kamaraj College of Engineering and Technology, Virudhunagar-626001.
E-mail ID: vijayakumariece@kamarajengg.edu.in,abinayakash004@gmail.com
anupappu6@gmail.com,priyankamenon.r@gmail.com,sivarevathi.a@gmail.com

ABSTRACT

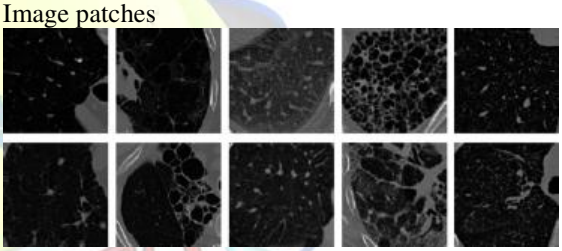
In this paper, the Interstitial Lung Disease represents a group of more than 150 disorders of the lung parenchyma has been considered. Most of these are harmful to the lung tissue and it eventually affect breathing .This work explains a new method for classification of lung tissue and identify the diseased pattern using feature based image patch approximation. From the computed tomography lung images, the image patches are taken which are used as an input. The Texture, Intensity, Gradient (T-I-G) descriptors are used for extracting the features. Based on the result of the feature vector and reference dictionary constructed, the image patches are classified. The patch approximation classifier is used for testing image patches and classifies the diseased tissue pattern. The outcome of this work will be helpful to the clinician for disease diagnosis.

Keywords: **Image patch, gradient, reference dictionary, texture.**

Introduction

A new classification methodfor five categories of lung tissues in Computed Tomography (CT) images, with feature-based image patch approximation is proposed. Two feature descriptors for higher feature descriptiveness, namely Gray Level Co-occurrence Matrix (GLCM) texture descriptor and Histogram of Oriented Gradients (HOG) gradient descriptor is deigned. Together with intensity features, each image patch is then labelled based on its feature approximation from reference image patches. And a patch approximative method is used for classification of image patches. The patch-wise labelling are then accumulated as probabilistic estimations for region-level classification. The Proposed method is evaluated on a publicly available ILD database and the real database collected showing encouraging performance improvements over the state-of-the-arts.

Image patches



(a) (b) (c) (d) (e)

Figure 1. Two example images of each tissue category. From left to right (a) normal, (b)Emphysema,(c) ground glass, (d) fibrosis, and (e) Micro nodules.

In image classification method, there are many classifiers are available for tissue classification such as SVM classifier,KNN classifier etc. SVM classifiers supports much fewer feature vectors than reference dictionary size and KNN classifier supports pair-wise distance computation. The result obtained are not sufficient upto expected level. Hence, it is difficult for image patch classification to identify the diseased pattern**.**

2. Proposed Methodology

To overcome these problems a new method is proposed, the image patch classification is done by using patch approximation classifier. There are two main stages in this proposed methodology. They are Feature Extraction using feature descriptors and image patch classification.

In this proposed methodology, the image is approximated as patches. Input images are segmented as left and right lung and for every image four patches are collected based on Region of Interest.

Patch approximation algorithm is used for image patch classification for input patch images. The input image patch is processed for feature extraction purpose before classification. The feature are extracted with three descriptors. Texture, intensity and gradient. The feature extraction is observed by patch approximation algorithm.

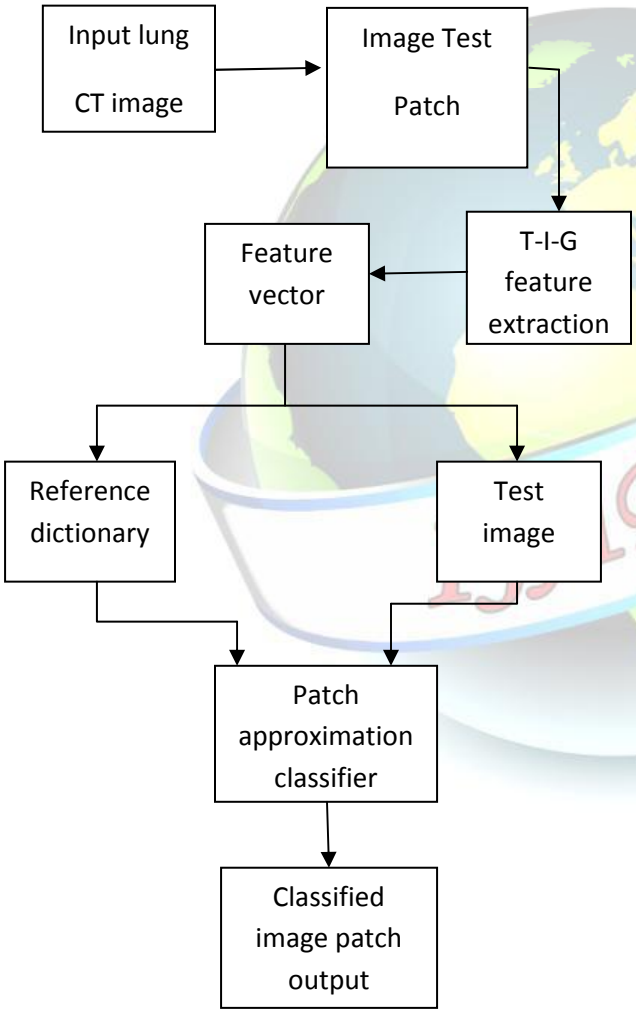


Figure 2. Block Diagram of Proposed Methodology

3.2.1 INPUT IMAGE

Input image is taken as gray level image. Because it has few level(0-255) compared to RGB image.

3.2.2 T-I-G FEATURE EXTRACTION

In feature extraction, first patch lung images are taken. GLCM and HOG is applied to patch lung images for texture feature and gradient feature extraction. Based on our visual analysis of lung patch images, it is observed that texture, intensity and gradient distribution of soft tissues within an image patchare quite informative and discriminative for different categories of lung tissues. Therefore, a patch-wise feature set, combining texture and gradient features are extracted for each image patch.

3.2.3 TEXTURE DESCRIPTION

A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. (The texture filter functions, described in Texture Analysis cannot provide information about shape, i.e., the spatial relationships of pixels in an image).

After getting GLCM matrix, the properties like contrast, correlation, energy and homogeneity are convolved. These statistics provide information about the texture of an image. The following table lists the statistics.

Contrast

It measures the local variations in the gray-level co-occurrence matrix.

$$\Sigma_{i,j}|i-j|^2 p(i,j) \longrightarrow (1)$$

Correlation

Measures the joint probability occurrence of the specified pixel pairs.

$$\sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j} \longrightarrow (2)$$

Where μ is the mean and σ is the standard deviation.

Energy

Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment.

$$\sum_{i,j} p(i,j)^2 \longrightarrow (3)$$

Homogeneity

Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

$$\sum_{i,j} \frac{p(i,j)}{1+|i-j|} \longrightarrow (4)$$

where p is the number of gray-level co-occurrence matricesin GLCM.

3.2.4 INTENSITY DESCRIPTION

Intensity information, is quite an additional information for lung CT images, e.g., normal tissues usually exhibit darker appearances. Therefore, the intensity difference is observed by calculating maximum, minimum andaverage intensity values for each image patches.

3.2.5 GRADIENT DESCRIPTION

Gradient distribution of an image is a different type of feature in complementary to the texture and intensity features. It is potentially very useful for discriminating pathological and normal lung tissues, since the former type often contains small segments that are less common in the normal lung.
Among the various types of gradient-based features, the HOG feature has been suggested as very effective for complex image patches.

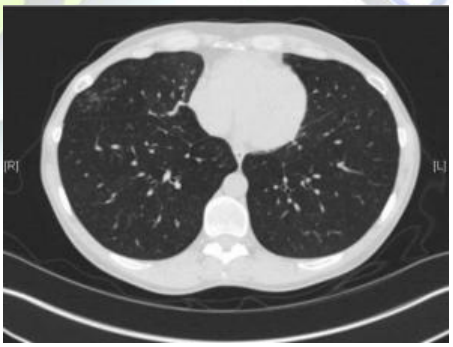
3.2.6 APPROXIMATE PATCH CLASSIFICATION

The next step is to classify each image patch into one of the five tissue categories. Considering that lung images normally exhibit quite different patterns even within the same tissue category. Therefore,a classification scheme that is especially effective in handling such issues. Thus a data-adaptive and Non-parametric approach, namely the Patch approximation method is used to classify an image patch based on the closeness of approximation by other image patches from each tissue category.

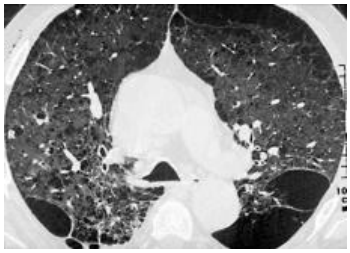
RESULT AND DISCUSSION

The proposed algorithm is evaluated with the database consists of Computed Tomography lung images. For this project work real time database is collected from Computed Tomography in Aarthi scan centre in Kovilpatti. The proposed algorithm is evaluated with the real database of twenty image patches which includes five diseased categories. The algorithm is implemented using MATLAB(2013a).The feature extracted image patches are compared with image patches in reference dictionary.Sample image patches of these five diseased pattern separately for left and right lung are shown in figure 3.

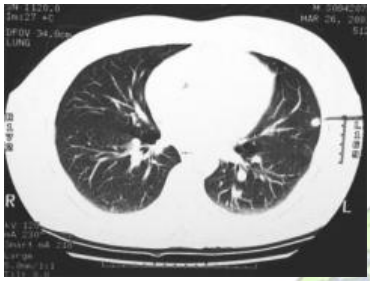
Fig 4.1: Database images



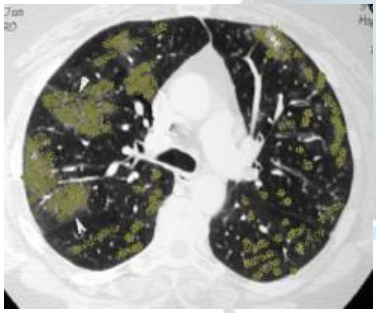
(a)Normal lung CT image



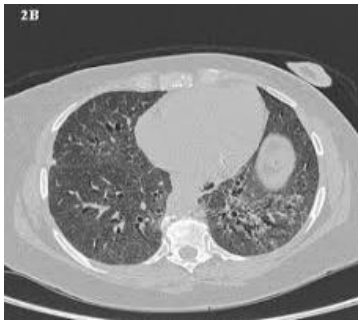
(b)Emphysema lung CT image



(c)Micronodules lung CT image



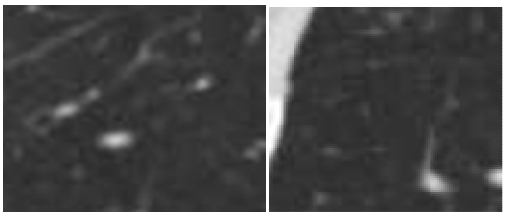
(d)Groundglass lung CT image



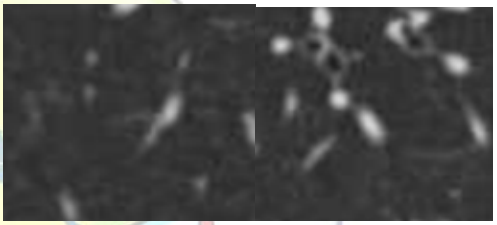
(e) Fibrosis lung CT image

The input lung tissues are then seperated as 20 image patches. The image patches are given below:

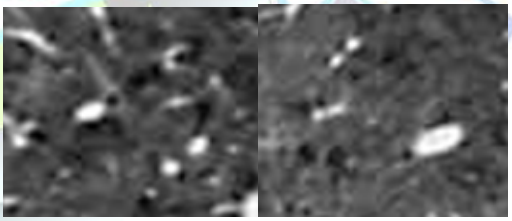
Fig 4.2. Image patches:



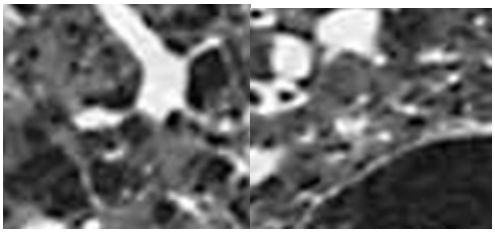
(a)Top left normal (b) Top right normal



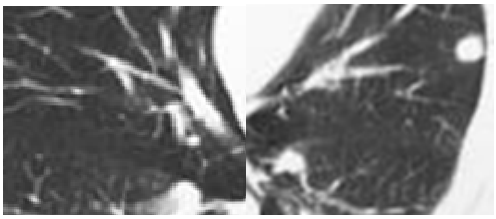
(c)Bottom left normal(d) Bottom right normal



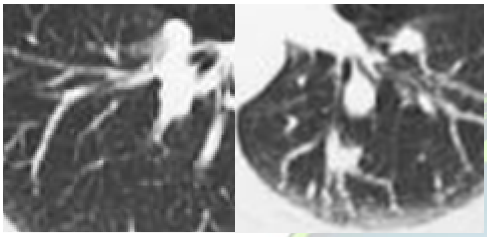
(e)Top left emphysema (f) Top right emphysema



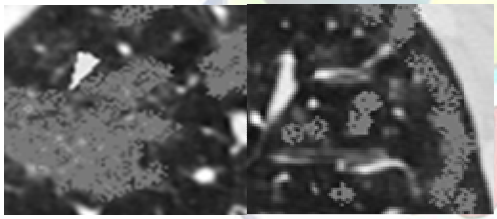
(g)Bottom left(h)Bottom rightemphysema emphysema



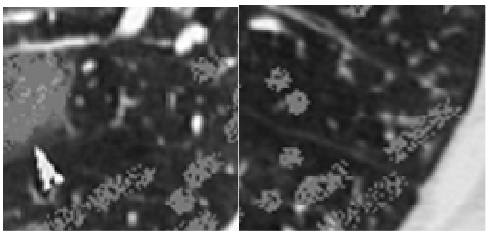
(i)Top left micro nodules(j) Top right micronodules



(k)Bottom left micro nodules (l)Bottom right micro nodules

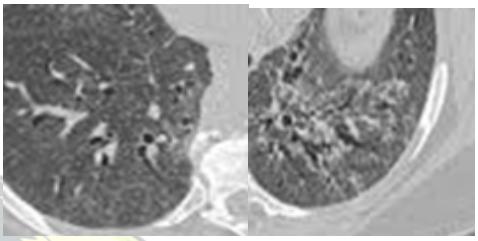


(m)Top left ground glass (n)Top right ground glass



(o)Bottom left groundglass (p)Bottom right groundglass

(q) Top left fibrosis (r) Top right fibrosis



(s) Bottom left fibrosis (t) Bottom right fibrosis

4.1 RESULTS OBTAINED FOR DISTANCE METRICS

Euclidean Distance Description

For each pixel in the input image taken, the distance transform assigns a number that is the distance between that pixel and the nearest nonzero pixel of input image.

The formula to calculate Euclidean distance is

D =bwdist(BW).

As an example, the (Euclidean) distance between points that taken from input test image patch (2, -1) and (-2, 2) is found to be

$$\begin{aligned} \text{dist}((2,-1),(-2,2)) &= \sqrt{(2-(-2))^2 + ((-1)-2)^2} \\ &= \sqrt{(2+2)^2 + (-1-2)^2} \\ &= \sqrt{(4)^2 + (-3)^2} \\ &= \sqrt{16+9} \\ &= \sqrt{25} \\ &= 5. \end{aligned}$$

The obtained Euclidean distance for given input test images are as follows.

FIGURE 4

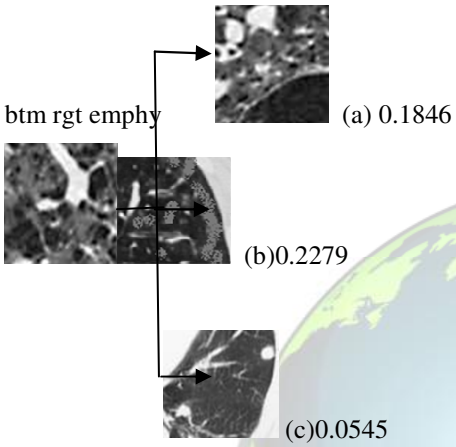


Fig 4.a.Sample patch of Bottom right Emphysema

From the figure 4.a, the Bottom right Emphysema is choosen as test patch is matched against other three reference image patch from reference dictionary are bottom left emphysema, top right ground glass and top right micro nodules.

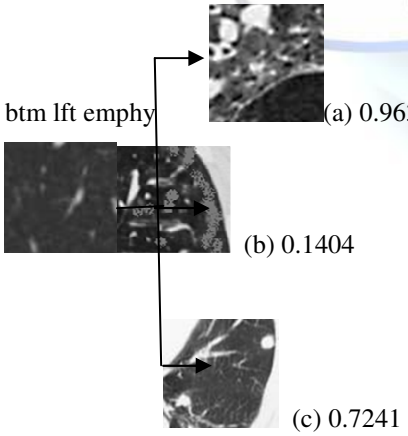


Fig 4.b.Sample patch of Bottom left Emphysema

From the figure 4.b, the Bottom left Emphysema is choosen as test patch is matched against other three reference image patch from reference dictionary are bottom left emphysema, top right ground glass and top right micro nodules

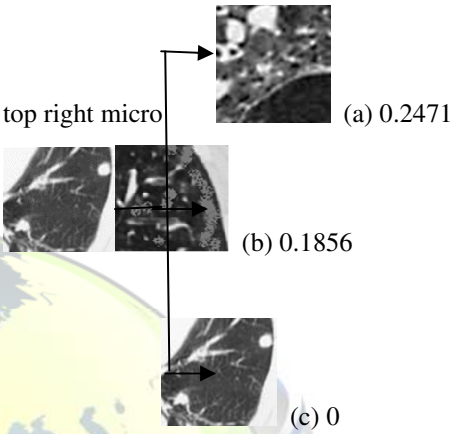


Fig 4.c.Sample patch of Top right Micro nodules

From the figure 4.c, the Top right Micro nodules is choosen as test patch is matched against other three reference image patch from reference dictionaryare bottom left emphysema, top right ground glass and top right micro nodules.

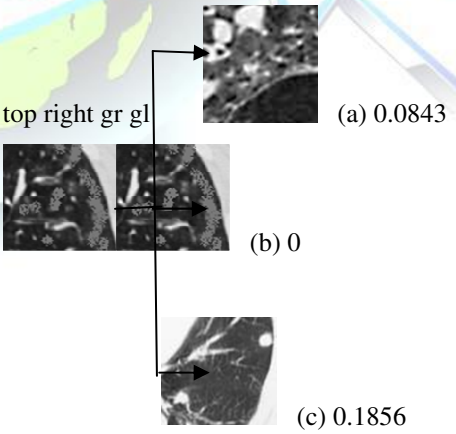


Fig 4.d.Sample patch of Top right Ground Glass

From the figure 4.d, the Top right Ground Glass is choosen as test patch is matched against other three

reference image patch from reference dictionary are bottom left emphysema, top right ground glass and top right micro nodules.

4.2 RESULTS OBTAINED FOR TEXTURE FEATURE EXTRACTION.

For texture featureextraction, GLCM is calculated that gives effective texture feature for an input test image patch. Graycomatrix is helpful in determining the GLCM. Graycomatrix creates the GLCM by calculating how often a pixel with gray-level value i occurs horizontally adjacent to a pixel with the value j. Each element (i,j) in glcm specifies the number of times that the pixel with value i occurred horizontally adjacent to a pixel with value j.

Table 4.1. GLCM texture description

Lung tissue	Contrast	Correlation	Energy	Homogeneity
Healthy	**0.0612**	**0.9953**	**0.6748**	**0.9896**
Emphysema	**0.0581**	**0.9939**	**0.6568**	**0.9849**
Ground glass	**0.0700**	**0.9849**	**0.7185**	**0.9779**
Fibrosis	**0.1644**	**0.9780**	**0.6633**	**0.9511**
Micro nodules	**0.0899**	**0.9921**	**0.6720**	**0.9731**

From the above table 4.1, 5 types of lung tissue rich texture information is observed that is very helpful in better classification performance**.** This GLCM texture feature description performance suggests that it can provide valuable assistance in the difficult task of texture analysis of lung tissue patterns in clinical routine with high reliability**.**

4.3 RESULTS OBTAINED FOR INTENSITY FEATURE EXTRACTION.

In order to make the classification as simple intensity descriptor is also used. It is tabulated in table 4.2.

Table 4.2 Intensity description

Lung tissue	Maximum intensity	Minimum Intensity	Average intensity
Healthy	**255**	**51**	**166.59**
Emphysema	**255**	**20**	**129.18**
Ground glass	**255**	**25**	**233.89**
Fibrosis	**255**	**0**	**226.73**
Micro nodules	**255**	**0**	**217.06**

It is inferred from table that both minimum and average intensity is providing some useful information. From the minimum intensity, it is difficult to differentiate fibrosis and micro nodules disease pattern. Hence, an additional information is observed from average intensity.

4.4 RESULTS OBTAINED FOR GRADIENT FEATURE EXTRACTION.

Gradient distribution of an image is a different type of feature in complementary to the texture and intensity features. It is potentially very useful for discriminating pathological and normal lung tissues, since the former type often contains small segments that are less common in the normal lung. Among the various types of gradient-based features, the HOG feature has been suggested as very effective.
The histogram of oriented gradients (HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. This method is similar to that of edge orientation histograms, scale-invariant feature transform descriptors, and shape contexts, but differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy.

The input test image patch is divided into small connected regions called cells, and for the pixels within each cell, a histogram of gradient directions is compiled. The descriptor is then the concatenation of these histograms. For improved accuracy, the local histograms can be contrast-normalized by calculating a measure of the intensity across a larger region of the image, called a block, and then using this value to

normalize all cells within the block. This normalization results in better invariance to changes in illumination and shadowing.

The first step of calculation in many feature detectors in image pre-processing is to ensure normalized color and gamma values. Instead, the first step of calculation is the computation of the gradient values. The most common method is to apply the 1-D centered, point discrete derivative mask in one or both of the horizontal and vertical directions.

Orientation binning:

The second step of calculation is creating the cell histograms. Each pixel within the cell casts a weighted vote for an orientation-based histogram channel based on the values found in the gradient computation. The cells themselves can either be rectangular or radial in shape, and the histogram channels are evenly spread over 0 to 180 degrees or 0 to 360 degrees, depending on whether the gradient is "unsigned" or "signed".

, **Table 4.3. HOG Feature Extraction**

TEST IMAGES	HOG Feature Extraction			
Bottom left normal	0.5017	0.6135	0.4829	-0.3724
Bottom right Emphysema	0.4165	-0.5986	0.5936	0.3404
Top right Ground glass	0.3938	0.6590	0.4597	-0.4463
Top right Micro nodules	0.4596	-0.5665	0.5978	0.3326

From the table 4.3, it is observed that HOG features for four test image patches whichrepresents objects by occurrences of gradient orientation in local portions. This feature extraction technique demonstrate effectiveness individually, better outcomes could be obtained.

4.4 RESULTS OBTAINED FOR PATCH APPROXIMATION CLASSIFIER.

In the patch approximation method Euclidean distance measure is used to find out the matching between the patches. Four test patches are matched against 20 other patches in the reference dictionary. It is tabulated in table 4.4.a to 4.4.e.

Table 4.4.a. Patch approximation for normal tissue

TEST IMAGES	REFERENCE DICTIONARY			
	Bottom left normal	**Bottom left emphysema**	**Top right emphysema**	**Top left normal**
Bottom Right emphysema	**1.1113**	**0.1846**	**0.1263**	**0.1881**
Bottom left normal	**0**	**0.9634**	**0.5584**	**0.1586**
Top right micro nodules	**0.7291**	**0.2471**	**0.0930**	**0.1709**
Top right ground glass	**0.1404**	**0.0843**	**0.3967**	**0.8285**

Table 4.4.bPatch approximation for Micro nodule

TEST IMAGES	REFERENCE DICTIONARY			
	Bottom right normal	**Top right normal**	**Top right micro nodules**	**Top left micro nodules**
Bottom Right emphysema	**0.8993**	**0.9225**	**0.0545**	**0.2291**
Bottom left normal	**0.1381**	**0.1111**	**0.7291**	**0.1424**
Top right micro nodules	**0.5039**	**0.6692**	**0**	**0.1856**
Top right ground glass	**0.0827**	**0.7958**	**0.1856**	**0.0212**

Table 4.4.c. Patch approximation for Ground Glass

TEST IMAGES	REFERENCE DICTIONARY			
	Bottom right micro nodules	Bottom left micro nodules	Top right ground glass	Top left ground glass
Bottom Right emphysema	0.2140	0.3003	0.2279	0.1062
Bottom left normal	0.8531	0.8119	0.1404	0.7338
Top right micro nodules	0.1274	0.1421	0.1856	0.0908
Top right ground glass	0.0712	0.0953	0	0.2064

Table 4.4.d. Patch approximation for fibrosis

TEST IMAGES	REFERENCE DICTIONARY			
	Bottom right ground glass	Bottom left ground glass	Top right fibrosis	Top left fibrosis
Bottom Right emphysema	0.3847	0.2245	0.2407	0.2445
Bottom left normal	0.0236	0.1405	0.6259	0.3330
Top right micro nodules	0.0407	0.2923	0.5732	0.1947
Top right ground glass	0.1182	0.0959	0.4018	0.0627

Table 4.4.e. Patch approximation for Emphysema

TEST IMAGES	REFERENCE DICTIONARY			
	Bottom right fibrosis	Bottom left fibrosis	Bottom right emphysema	Top left emphysema
Bottom Right emphysema	0.2960	0.0487	0	0.0860
Bottom left normal	0.0830	0.8938	1.1113	0.6969
Top right micro nodules	0.1960	0.0172	0.0545	0.1122
Top right ground glass	0.0600	0.0849	0.2279	0.2521

The result obtained from the table shows that the patch approximative algorithm gives better classification and high accuracy in results. This is because the Euclidean distance values for each test patches is low and gives good results , so it is suitable for identification of diseased pattern.

8. Conclusion

A classification method for lung CT images is presented in this paper. Five categories of lung tissues –normal, emphysema, ground glass, fibrosis and micro nodules—that are important for ILD disease diagnosis, are the main objects to be differentiated. Afeature-based image patch approximation method is designed. First, an image patch is represented as a feature vector, based on our proposed GLCM texture and HOG gradient descriptors. Then, the image patch is classified into one of the five tissue categories, using our proposed Patch Approximative Classifier based on reference image patches.

9. References

2.1 LiQ,CaiW,Feng DD, (april 2013)'Lung image patch classification with automatic feature learning', ConfProc IEEE Eng Med Biol Soc.

2.2 I.Sluimer,A.Schilhm,M.ProkopandB.vanginneken (april.2006)'Computer analysis of computed tomography scans of the lung ',IEEEtrans.MED.imag.vol.25,no.4,pp.385-405.

2.3 L.sorensen,S.bsankar,andM.Deveruijne ,(2012)'Quantitative analysis of pulmonary emphysema using local binary pattern.'IEEE TRANS.MED.IMAG.,vol.29,no.2,pp.559-569.

2.4C.S.mendoza,G.R.washko,J.C.ross,A.A.diaz,D. A.lynch,J.D.crapo (feb 2010)'Emphysema quantification In a multi scanner HRCT cohort using Local intensity distribution.Proc.INT.symb.bio med .image .,pp.474-477.

2.5A.Depeursinge,A.vargas,A.platon,A.Geissbuhier. (2012)Building a reference multimedia database for interstitial lung diseases. Comput.med.imag.graph.,vol.36,no.3,pp.227-238.

2.6A.Depeursinge,D.Dedille,A.Tlaton,P.A.PolettiA.Geissduhler, (july 2012)Near-affine-invariant texture learning for lung tissue analysis using isotropic wavelet frames.IEEE trans.inf.technol.biomed.,dol.16,no.4,pp.665-675.

2.7 P.D. Korfiatis, A.N. Karahaliou, A. D. Kazantzi, and L.I.Costaridou .(may 2012) 'Texture-based identification and characterization of pheumonia in lung multidetector CT.'Inf. Technol. Biomed., vol. 14, no. 3, pp. 675–680.

2.8 A.Depeursinge,A.F.Oncuvierto-Rodriguez D.V.Deville.,(2012)Lung texture classification using locally oriented riesz components.MICCAI Newyork springer 2011,vol.6893,LNCS pp.231-238.