Journal of Communications Technology, Electronics and Computer Science, Issue 3, 2015
ISSN 2457-905X

AN IMAGE BASED DIAGNOSTIC SYSTEM FOR LUNG DISEASE CLASSIFICATION

Dr. C.Bhuvaneswari
Assistant Professor and Head
Thiruvalluvar University College of Arts and Science,
Thiruvennainallur,villupuram District
bhuvana.csdept@gmail.com

***Abstract*— Model-based detection and classification of nodules are two major steps in CAD systems design and evaluation. A common health problem, lung diseases are the most prevailing medical conditions throughout the world. In this paper, Lung diseases are automatically classified as Emphysema, Bronchitis, Pleural effusion and normal lung.The lung CT images are taken as input, preprocessing is applied, feature extraction is done by various methods such as Gabor filter extracts the texture features, walsh hadamard transform extracts the pixel co-efficient values, and a fusion method is proposed in this work which extracts the median absolute deviation values. Feature selection including statistical correlation based methods and Genetic Algorithm for searching in feature vector space are investigated. Four types of the classifiers are used where the Multi-Layer Perceptron Neural Network (MLP-NN) classifier with proposed fusion feature extraction method, genetic algorithm feature selection method gives promising result of 91% accuracy than J48, K- Nearest Neighbour and Naïve bayes classifiers.**

Keywords—classification; classifiers; detection; feature extraction; feature selection; fusion; Genetic Algorithm; lung nodules; K- Nearest Neighbour; textures.

I. INTRODUCTION

Medical image processing requires a comprehensive environment for data access, analysis, processing, visualization, and algorithm development. Lung disease detection and classification involves a series of steps. A lot of research efforts have been directed towards the field of medical image analysis with the aim to assist in diagnosis and clinical studies [1] The work of Beigelman-Aubry et al. [2] presented evaluation of nodule detection and its response time when performed by radiologists with and without use of a computerized system. Nodule pixels are often brighter than the surrounding areas but in some cases, the difference in grey levels is not at all significant.

Lung diseases are primary cause for most of the disabilities which also leads to death in the world. The chest CT shows the first important modality of the assessment of the diseases. There are many types of the disease that causes the lung infection such as inflammatory lung diseases, chronic obstructive pulmonary disease(COPD),Emphysema, Chronic Bronchitis, pleural effusion, Interstitial lung diseases, bronchitis and lung carcinoma.

In this study, the dataset of lung diseases such as the emphysema, pleural effusion, bronchitis and normal lung are taken. After resizing of the image, preprocessing by the median filter and morphologically smoothening are carried out. Feature extraction techniques such as the Gabor filter, Walsh hadamard transform, and proposed fusion method are performed. The feature selection methods such as correlation based feature selection and genetic algorithm which is used to optimally initialize the cluster centers are applied. Then various classifiers such as Multi-Layer Perceptron Neural Network (MLP-NN), J48, Naïve bakes, KNN classifiers are used to classify the images, the performance measures of the precision, recall and the classification accuracy are evaluated and the results are tabulated.

II. RELATED WORKS

Medical image analysis is a complex task in which a human expert makes extensive use of the knowledge of anatomy and imaging techniques. Specially, the automatic segmentation of chest radiographs is challenging problem from a computer vision point of view. Because there are large anatomical variations from person to person and the most important problem is that radiographs are projection images and thus contain superimposed structures. Computed tomography (CT) for the body has been available since 1975.

By reducing the dimensionality of the input set correlated information is eliminated at the cost of a loss of accuracy [3]. Feature selection deals with selecting a subset of features, among the full features, that shows the best performance in classification accuracy [4]. Delorme et al. [5] have used a pixel-based approach to classify lung tissue in six classes using local texture measures. Medical Data Mining is a promising area of computational intelligence [6] applied to an automatically analyze patients records aiming at the discovery of new knowledge useful for medical decision-making. In feature selection problems, each feature subset is represented by a binary string [7]. One is assumed to have knowledge of the spectral properties of the original signal and the noise, and one seeks the linear time-invariant filter whose output would come as close to the original signal as possible.[8] A template matching algorithm [9] and genetic algorithms [10] have been used to detect lung nodules. Liang et.al., proposes a computer-aided diagnostic system for emphysema that segments the lungs into multiple square regions and classifies the segmented regions into 5 classes of severity[11]. Chiou et al [12] proposed application of neural network based hybrid system for lung nodule detection, which based on artificial neural network architectures were developed for improving diagnostic accuracy and speed for lung cancerous pulmonary radiology. The configuration of the HLND system included the following processing phases; data acquisition and pre-processing, in order to reduce and to enhance the figure-background contrast, quick selection of nodule suspects based upon the most prominent feature of nodules, the disc shape and completed features pace determination and neural classification of nodules. A supervised back propagation trained neural network was developed for recognition of the derived feature curve, a normalized marginal distribution curve. Walsh-Hadamard transform needs less time for feature extraction of EEG signals. EEG signals basis studies shows that these signals have binary nature [13]. In [14] first images were enhanced through Gabor filter. It has given better results than other enhancement techniques. They only worked on the colored image enhancement and not extract the nucleus region and even not the cell region. In Features Extraction stage to obtain the general features of the enhanced and segmented image by binarization. From the above related works the combination of various lung diseases are taken and their pathological conditions are discussed with the radiologist and the proposed system is designed.

III. MATERIALS AND METHODS

The evaluation of the proposed evolutionary approach have been performed by using a set of 200 lung images which comprise of a combination of emphysema, bronchitis, pleural effusion, and normal lung respectively. The Computed Tomography (CT) images are obtained from the Sri Manakula Vinayagar Medical College and Hospital Madagadipet. These images are captured using Philips MX16EVO which uses exclusive technologies such as the EVO Eye algorithm and MAR technology for excellent image quality with reduced noise which delivers performance and productivity that are at the top of its class. The anonymizing of Dicom Images is done and the discussions regarding the lung diseases are made with the radiologist for guidance.The proposed work is divided in to four phases. In the first phase input images are taken, resized, converted to grey scale image and preprocessing is done where the median filter and the morphological smoothening are applied. In the second phase feature extraction method is carried out, the third phase of work is feature selection, where the choice of methods including the Genetic Algorithm(GA) to select the optimal solution and correlation based feature selection are done, finally in the fourth phase the selected optimal features are given as input to various classifiers to classify the images in to respective diseases.

IV. PROPOSED METHODOLOGY

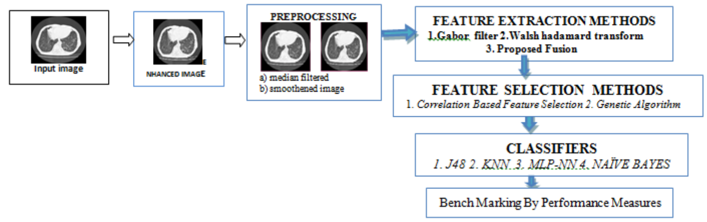

Fig. 1. Block diagram of the proposed work

Fig. 1 illustrates the block diagram for the proposed work with the methodologies used.

A. *Image enhancement and preprocessing*

Image pre-processing can significantly increase the reliability of an optical inspection. The CT lung disease image is taken as an input image. The image is sized to 128 by 128 of 16 x16 window and gray scale image is extracted. Median filter and morphological smoothening filter technique are applied to remove the noise from the images.

B. Input CT image

The lung CT images have low noise when compared to scan image and MRI image. The main advantage of the computed tomography image is it has better clarity, low noise and distortion. A combination of lung disease images are taken for preprocessing. The size of the input images is 256×256 pixels. The enhanced images are obtained by converting in to the gray scale image.

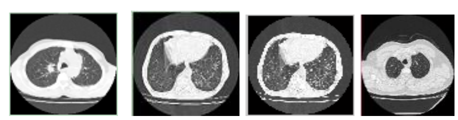

Fig.2. Input image of various lung diseases

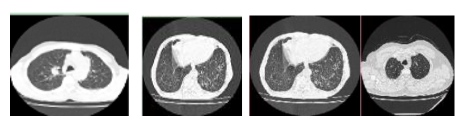

Fig. 3. Enhanced image

Fig. 2 shows the input CT images of lung diseases. The Fig 3 shows enhanced images used for effective preprocessing steps that will be used for the median and morphological smoothening filter.

C. Preprocessing techniques

Preprocessing phase of the images is necessary to improve the quality of the images and make the feature extraction phase more reliable.

D. *Median filter*

An image containing salt-and-pepper noise will have dark pixels in bright regions and bright pixels in dark regions that can be removed by the median filter. The filtered image will be used for the further processing techniques The formula for the median filter is given as

Y(t)=median((x(t-T/2),x(t-T1+1),…,x(t),…,x(t+T/2)) (1)

Where y is the median filter at the moment t,
t is the size of the window of the median filter ,
x - image,
T – pixel location ,
T1- the pixel at first row and first column (r1,c1).

The major advantage of using the median filter are resilient to statistical outliers, incurs less blurring and simple to implement

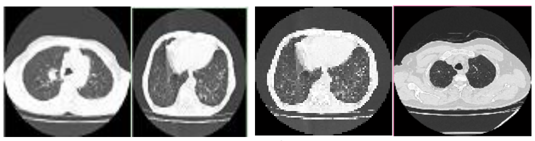

Fig. 4. Median filtered image

The figure 4 represents the median filtered images of lung CT.

E. Morphological smoothened image

Image smoothening technique uses the linear Gaussian noise reduction technique, the *dilation* where the flat structuring element defined as the maximum value, *erosion* as the minimum value *opening* which suppresses the bright details smaller and the *closing* which suppresses the dark details are computed. The formula for erosion and dilation are given as follows,

$$U(X \ominus B) = \left\{(x,z) \,\middle|\, U\left(B^{1}_{x,z}\right) \subset U(X)\right\}$$
$$U(X \oplus B) = \left\{(x,z) \,\middle|\, U\left(B^{2}_{x,z}\right) \subset U^{c}(X)\right\} \quad (2)$$

The goal of the Morphological smoothening filter is to filter out noise that has corrupted a signal. It is based on a statistical approach.

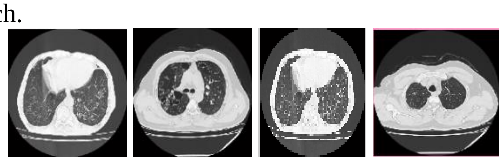

Fig. 5. Morphologically smoothened image

Fig. 5 represents the morphological smoothened images which helps to extract the features.

V. FEATURE EXTRACTION

It is important to recognize and extract interesting features for an exacting task in order to decrease the complexity of processing. The extraction of the features from an image can be done using a variety of image processing techniques. Texture analysis is a quantitative method that can be used to quantify and detect structural abnormalities in different tissues.

The following methods such as Gabor filter which is used and works in the spatial domain will return the texture values. The Walsh hadamard transform will work on the frequency domain and returns the pixel co-efficient values. The proposed fusion technique is the combination of the Gabor and Walsh hadamard transform using median absolute deviation technique. The images are spilt in to 15 windows where the **max, min, average and median** features are calculated for each window using different feature extraction methods which results in total of 60 features of each image.

A. 2D Gabor filter

Gabor filter is a linear filter used for edge detection. Frequency and orientation representations of Gabor filters are found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. A set of Gabor filters with different frequencies and orientations are used for extracting useful features from an image.

A 2-D Gabor filter over the image domain(x, y) is given by

$$G(x,y) = \exp\left(-\frac{(x-x_0)^2}{2\sigma_x^2} - \frac{(y-y_0)^2}{2\sigma_y^2}\right)$$
$$\mathrm{X}\exp\left(-2\pi i\left(u_0(x-x_0)+v_0(y-y_0)\right)\right) \quad (3)$$

where

(x_0, y_0) is location in the image,

(u_0, v_0) specifies modulation which has

frequency $\omega_0 = \sqrt{u_0^2 + v_0^2}$ and

orientation $\theta_0 = \arctan\left(v_0 / u_0 \right)$

σ_x and σ_y are standard deviation of Gaussian envelope

The Gabor filter will get the orientation as input and calculate the feature vector, the orientation taken are horizontal, vertical and the angles represented by pi/values. The features are extracted and used for the feature selection process.

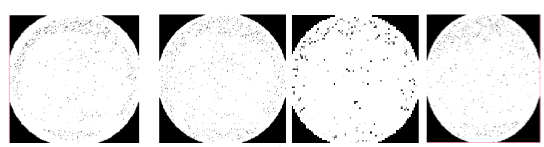

Fig. 6. Gabor filtered image

Some properties of Gabor filters are it is a tunable bandpass filter, similar to a windowed Fourier transform, satisfies the lower-most bound of the time-spectrum resolution. It's a multi-scale, multi-resolution filter that has selectivity for orientation, spectral bandwidth and spatial extent. The output of the gabor filter are the texture values.

B. Walsh-Hadamard Transform (WHT)

The WHT is a suboptimal, non-sinusoidal, orthogonal transformation that decomposes a signal into a set of orthogonal, rectangular waveforms called Walsh functions.
It is involved with periodic finite series and provides a spectrum whose period contains the same sample number as the temporal sequence. It uses fast and efficient methods to compute algorithms in a few operations and could be extended to multidimensional signals. The WHT is given by

$$WX_K = \frac{1}{N}(H_M X_k) \quad (4)$$

and the corresponding inverse transform as,

$$X_K = (H_M W X_k) \quad (5)$$

Where the square and symmetric Hadamard transform matrix H_m of order m is recursively defined as,

$$H_m = \begin{pmatrix} H_{\frac{m}{2}} & H_{\frac{m}{2}} \\ H_{\frac{m}{2}} & -H_{\frac{m}{2}} \end{pmatrix} \quad (6)$$

for m>1 and $m=2^k$ with $H_1=[1]$.Since H_m contains only the +1 or -1 entry, the transformation requires only real additions and subtractions. The CT image is given as input and the pixel coefficient values are extracted as output.

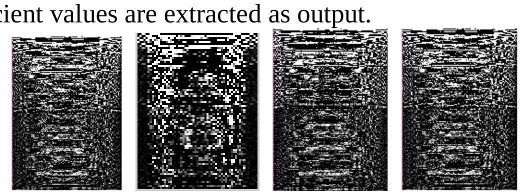

Fig. 7. walsh hadamard transform.

The feature vectors are stored in the excel file in .csv extension.

C. Proposed fusion techniques

Fusion is the combination of the different method to yield best results. Multi-sample refers to fusion of multiple samples extracted from the same source like using multiple images from the lung CT scan and Multi-algorithm means fusion of multiple processing methods for individual samples which combines multiple feature extraction methods. Features from Gabor and WHT are fused using Median Absolute Deviation (MAD) methods. The mean and median absolute deviations are measures of dispersion, or spread around the median. The absolute deviation from the median is the difference of the value .In this study, the features extracted from and Gabor filter are fused using template level fusion(multiple templates combine to form a single template). WHT extracts the features from the frequency domain and the Gabor captures the salient visual features corresponding to spatial localization, orientation selectivity, and spatial frequency.
Features from Gabor and WHT are fused using Median Absolute Deviation (MAD) between the two features.

$$let\ W_i = \{w_{i,1}, w_{i,2}, ..., w_{i,n}\} \text{ be the WHT coefficient}$$

$$let\ G_i = \{g_{i,1}, g_{i,2}, ..., g_{i,n}\} \text{ be the gabor coefficients}$$

The fused feature vector $X_i = \{x_{i,1}, x_{i,2}, ..., x_{i,n}\}$ is obtained by normalizing the feature vector to obtain $W_i^{'}$ and $G_i^{'}$ using MAD and taking the average of the same. The mean absolute deviation is calculated by taking the mean of the absolute deviations from the median: $(\sum (|X_n\text{-}MED|))/N$. (7)

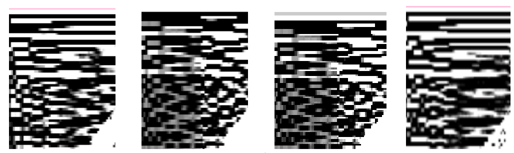

Fig. 8. Proposed fusion technique.

Four features such as **max,min ,average and median** are extracted from each window and **60 features** are extracted for each image (15 windows) using the fusion technique. The main advantages of proposed feature fusion is, it derives the most discriminatory information from the original feature set and also eliminates redundant information from the feature set. The redundant values in the Gabor and the Walsh hadamard methods are eliminated and the features are extracted.

VI. FEATURE SELECTION

Feature selection is the process of choosing subset of features relevant to particular application and improves classification by searching for the best feature subset, from the fixed set of original features according to a given feature evaluation criterion. The multivariate filter model search of the feature selection method i.e. correlation based feature selection method and genetic algorithm approach are taken here.

A. Correlation based feature selection (CFS)

One of the filter based multivariate model search which is models feature dependencies, Independent of the classifier, better computational complexity than wrapper methods is correlation based feature selection. CFS searches feature subsets according to the degree of redundancy among the features. The evaluator aims to find the subsets of features that are individually highly correlated with the class but have low inter-correlation. Correlation coefficients are used to estimate correlation between subset of attributes and class, as well as inter-correlations between the features.

$$r_{xc} = \frac{K\overline{r_{zi}}}{\sqrt{K + K(K-1)\overline{r_{ii}}}} \quad (8)$$

Where r_{xc} is the correlation between the summed feature subsets and the class variable

k is the number of subset features

r_{zi} is the average of the correlations between the subset features an the class variable

r_{ii} is the average inter-correlation between subset features.

The output of the feature selected (ie) 40 features are selected from CFS method.

B. *Genetic algorithm*

The task of the feature extraction and selection methods is to obtain the most *relevant information* from the original data and represent the information in a *lower dimensionality* space. Genetic algorithm is an evolutionary approach which can be used for the feature selection technique that selects the efficient features alone .The general block of the genetic approach is given by the illustrated,

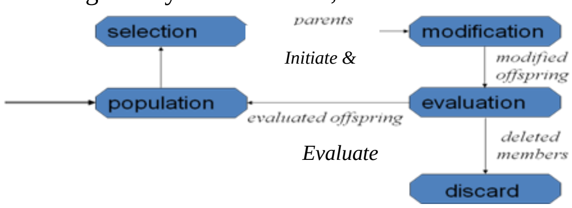

Fig. 9. Genetic approach.

The working principal used in the GA begin by randomly creating its initial population of 30.Solutions are combined via a crossover operator to produce offspring, two point crossover with crossover rate of 0.7 thus expanding the current population of solutions. The individuals in the population are then evaluated via a fitness function Root mean square error(RMSE) and the less fit individuals are eliminated to return the population to its original size. The process of crossover, evaluation, and selection is repeated for a predetermined number of generations or until a satisfactory solution has been found or condition fails. A mutation operator is generally applied to each generation in order to increase variation. In the feature selection formulation of the genetic algorithm, individuals are composed of bit strings: a 1 in bit position indicates that feature should be selected; 0 indicates this feature should not be selected. The mutation type is bit flip (0 will become 1 or vice versa) and mutation rate is 0.01and the stopping condition is RMSE threshold of 0.001 or 500 operations with generation of 500, fitness condition for RMSE and the crossover operator is two point. 30 features are selected from genetic algorithm feature selection which is fed in to classifiers.

VII. CLASSIFIERS

The “classifier" referred as the mathematical function, which is implemented by a classification algorithm, will maps input data to a category of supervised learning. Four types of the classifiers are used for the proposed work. The dataset comprise of 200 images of four different diseases are used by 10 cross fold validation.

A. Naive bayes classifier

Naive Bayes classifier is used for classifying the images and the performance measures are calculated and the comparative study of the successful feature selection algorithm is chosen. The extracted features are classified to the most likely class. Learning in Naïve Bayes is simplified by assuming that the features are independent for a given class.

$$P(C_i|V) = \frac{P(V|C_i)P(C_i)}{P(V)} \tag{9}$$

Where a) V is the feature value of dimensions of lung images.
b) P (Ci|V) is the probability of disease like emphysema, pleural effusion, bronchitis and normal for the given features like max, min, average and median for 15 windows.
c) P(x) is the probability of feature values.

The simplified version of the naïve bayes is given by the expression

$$P(V|C_i) = \prod_{k=1}^{n} P(x_k|C_i) \tag{10}$$

The attributes of classes such as for class 1 to 4 0.25, 0.33, 0.13, 0.29 are computed respectively and the mean, standard deviation, weighted sum, precision of the classifier are calculated up to 125 attributes.

B. J48

The J48 Decision tree classifier developed by Ross Quinlan which works, in order to classify a new item, it first needs to create a decision tree based on the attribute values of the available training data. So, whenever it encounters a set of items (training set) it identifies the attribute that discriminates the various instances most clearly. This feature describes about the data instances so that we can classify them the best is said to have the highest information gain.

The class label of 4 are taken as the 4 leaves and size 7 with the sigmoid nodes of the values for the node up to 3 with node 34 and sigmoid node of 4 to node 34 which ranges to all the input values are calculated .The threshold value of the node is computed as 3.3838 by the classifier.

C. Multi-layer perceptron neural network

A neural network consists of units (neurons), arranged in layers, which convert an input vector into some output. Each unit takes an input, applies a (nonlinear) function to it and then passes the output on to the next layer. Generally the networks are defined to be feed-forward: a unit feeds its output to all the units on the next layer, but there is no feedback to the previous layer. Weights are applied to the signals passing from one unit to another, and it is these weights which are tuned in the training phase to adapt a neural network to the particular problem at hand. This is the learning phase. Neural networks have found application in a wide variety of problems The back propagation learning algorithm is widely used for multi-layer feed forward network. The back propagation algorithm is using a weight adjustment based on the sigmoid function, like the delta rule.

An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. In multilayer perceptron each neuron uses a *nonlinear* activation function which was developed to model the frequency of action potentials, or firing, of biological neurons in the brain. The activation is simply a weighted sum function. The outgoing signals can be adjusted according to loss and prior probability value. In the output layer, there is one neuron to represent the classification result. The sum for each hidden node is sent to the output layer and the highest values wins.

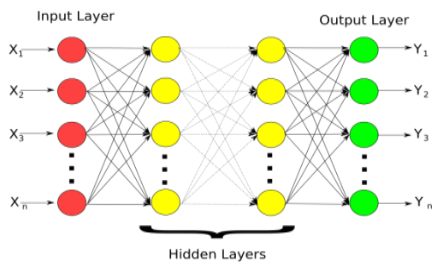

Fig. 10.Architect of the Multilayer perceptron.

This method has influenced to categories as the four types of the images as normal, pleural effusion, bronchitis, emphysema. The testing and the training are done in the 10 fold methods.

C. k-NN

The ***k*-nearest neighbor's algorithm** is a non-parametric method for classification and regression predicts objects' "values" or class memberships based on the *k* closest training examples in the feature space. *K-NN* is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. The *k*-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its *k* nearest neighbors. If *k* = 1, then the object is simply assigned to the class of that single nearest neighbor.
The 60 attributes and class index of 125 are taken. The values of the column are taken as 1 to 16, 17 to 32, 33 to 48 and 49 to 59 order and the values are calculated .the log score value of bayes, Bayesian Dirichlet equivalence for uniform joint distribution, Minimum description length ,entropy are calculated and values of the class label are computed.

VIII. EXPERIMENT AND RESULTS

The proposed method is implemented in matlab 7.10.The dataset includes 200 images (55 47,49,52 for four class labels respectively) where 125 are used for training and 75 images for the testing .In order to process the image contrast enhancement was performed to get the clear image. The preprocessed images are median filtered for noise removal , morphologically smoothened and given as a input to the feature extraction process which involves Gabor filter,walsh hadamard transform and proposed fusion method and the statistical features of min,max,median,average of 15 windows are extracted from the preprocessed images with 60 features.

The feature selection is done by the correlation based feature selection and genetic algorithm method where the relevant features are extracted and the images are classified using the classifiers. The evaluation of the testing instances using the various methods are evaluated and given by 4 classifiers.

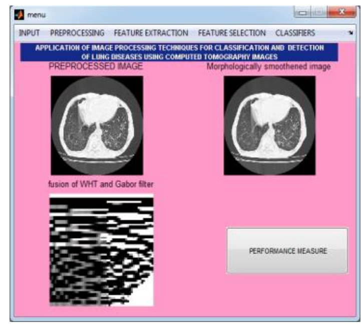

Fig.11. Output screen

IX. PERFORMANCE MEASURES

The performance measures such as *Precision* is the proportion of the examples which truly have class *x* among all those which were classified as class *x*. **Precision (P) = tp/(tp+fp) and Recall (R) = tp/(tp+fn).**

The results of the precision and the recall for the Gabor filter, Walsh hadamard transform , proposed Fusion techniques are presented in the table 1.

TABLE 1: Precision and recall of Gabor, wht and fusion

Technique	Gabor Precision	Gabor Recall	WHT Precision	WHT Recall	FUSION Precision	FUSION Recall
CFS_NB	0.76	0.76	0.72	0.7	0.77	0.79
CFS_J48	0.72	0.72	0.72	0.72	0.72	0.72
CFS_KNN	0.68	0.68	0.64	0.62	0.64	0.62
CFS_NN	0.8	0.81	0.79	0.78	0.85	0.85
GA_NB	0.77	0.78	0.73	0.73	0.73	0.73
GA_J48	0.75	0.78	0.72	0.71	0.83	0.83
GA_KNN	0.72	0.71	0.75	0.75	0.88	0.86
GA_NN	0.84	0.86	0.75	0.78	0.92	0.91

The figure 12 represents the precision and the recall values for the classification.

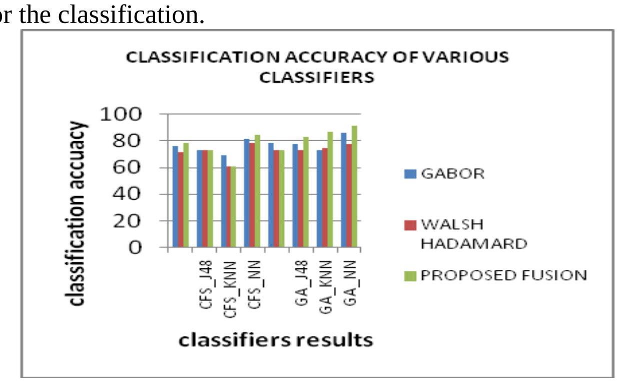

Fig. 12. Classification Accuracy of various classifiers.

The *True Positive (TP)* rate is the proportion which were classified as class *x*, among all which truly have class *x*, which is equivalent to *Recall*.

The *False Positive (FP)* rate is the proportion which were classified as class *x*, but belong to a different class, among which are not of class *x*.

The true positive ,true negative, false positive, false negative for the given set of the images are tabulated as the following .The four classifiers and their respective values for the classification are listed in the table 2.

TABLE 2: classification of lung diseases by various classifiers

classifiers	TP	TN	FP	FN
Naive Bayes	80	32	7	6
J48	84	33	4	4
MLP-NN	88	34	2	1
KNN	79	32	7	7

In this paper we have used four classifiers and compared their performance. So far in the literature the lung diseases are dealt with one or two diseases only but in our work the 4 lung diseases images are classified. In the previous studies the lung cancer images are classified in major of the literature. Basu et al [15] proposed developing a classifier model for lung tumors in CT-scan images, this method investigation, a large group of2D and 3D image features which were hypothesized to be useful were evaluated for effectiveness in classifying the tumors with 68% accuracy, showing new features.
Either Gabor filter or statistical features are investigated in the previous studies and also the genetic algorithm alone is used for feature selection. So our work differs from previous methods in proposed feature extraction, selection and also we have used four classifiers and found a more suitable classifier for classification of lung diseases.

X. CONCLUSION

Lung diseases are the most common disease in the world that can be of various types .The input image is the lung diseases CT dataset where the 200 images of different diseases are taken and the preprocessing of the images are done where the features such as mean, median, max and min are extracted in 15 sub windows totaling 60 features for each image are found and the results of the different methods are calculated.The features extracted are fed in to feature selection method and the classifier are used to classify the images and the result shows that the proposed feature fusion method with genetic algorithm feature selection using MLP_NN classifier gives 91% of classification accuracy. This work has been monitored by the radiologist. Thus this work has given a successful automated diagnosing method which helps to diagnose the disease from the CT scans and also mutually assist the radiologists.

REFERENCES

1. C. Beigelman-Aubry, P. Raffy, W. Yang, R.A.Castellino,P.A. Grenier, “Computer-aided detection of solid lung nodules”, American Journal of Roentgenology ,Vol :189,pg:948–955, ISSN:1546–3141.
2. Duncan J.S.,Ayache N., “ Medical Image Analysis Progress Over two decade and challenges ahead”, IEEE Transaction on Pattern Analysis and Machine Intelligence , Vol 22, No 1 ,PP. 85 – 106,2000.
3. Addison, J F D, MacIntyre J (editors) “Intelligent techniques: A review”, Springer Verlag (UK) Publishing Company, 1st Edition, Chapter 9, September 2003.
4. Jain and D. Zongker, “Feature Selection: Evaluation,Application, and Small Sample Performance”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 2, pp. 153-158, 1997.
5. Delorme S, Keller-Reichenbecher MA, Zuna I, Schlegel W, van Kaik G.,”Usual interstitial pneumonia quantitative assessment of high-resolution computed tomography findings by computer-assisted texture based image analysis”, Investigation Radiology pg:566–74, 1997.
6. Zakaria Suliman Zubi and Rema Asheibani Saad, “Using Some Data Mining Techniques for Early Diagnosis of Lung Cancer”, Proceedings of the 10th WSEAS international conference on Artificial intelligence Knowledge Engineering and Data Bases, pp. 32-37, 2011.
7. M.J Martin-Bautista and M-A Vila, “A Survey of Genetic Feature Selection in Mining Issues”, Proceedings of the Congress on Evolutionary Computation, 1999, Vol. 2, pp. 1314-1321
8. Weiner Filtering, “ http://www.owlnet.rice.edu”.
9. M.N. Gurcan, B. Sahiner, N. Petrick, H.P. Chan, E.A. Kazerooni, P.N. Cascade and L.Hadjiiski, “Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computer-aided diagnosis system”, Medical Physics, vol. 29, N 11, pp.2552, 2002.
10. Y. Lee, T. Hara, H. Fujita, S. Itoh and T. Ishigaki,” Nodule detection on chest helical CT scans by using a genetic algorithm”, Proceeding in Intelligent Information Systems, pp. 67 - 70,1997.
11. Liang, T. K., Tanaka, T., Nakamura, H., & Ishizaka, A,” A neural network based computer-aided diagnosis of emphysema using CT lung images”, In *SICE, 2007 Annual Conference* (pp. 703-709), IEEE, 2007.
11. Chiou, Y. S., Lure, Y. F., Freedman, M. T., & Fritz, S. ,”Application of neural network based hybrid system for lung nodule detection.”, In *Computer-Based Medical Systems, Proceedings of Sixth Annual IEEE Symposium on* (pp. 211-216). IEEE, *1993*..
12. E.J. Speckmann and C.E. Elger, “Introduction to the Neurophysiological Basis of the EEG and DC Potentials”,

Chapter 2 in E.Neidermeyer and F. Lopes da Silva, Electroencephalography, 5th edition, Lippincott Williams & Wilkins, New York: 2005.

13. Mokhled S. AL-TARAWNEH, “Lung Cancer Detection Using Image Processing Techniques”, Leonardo Electronic Journal of Practices and Technologies, ISSN 1583-1078, Issue 20, January-June 2012 .
14. Basu, S., Hall, L. O., Goldgof, D. B., Gu, Y., Kumar, V., Choi, J.,& Gatenby, R. A. “Developing a classifier model for lung tumors in CT-scan images”, In *Systems, Man, and Cybernetics (SMC), IEEE International Conference on* (pp. 1306-1312). IEEE, *2011*.
15. Roy, K., Chaudhuri, C., Kundu, M., Nasipuri, M., & Basu, D. K..,” Comparison of the Multilayer perceptron and the nearest neighbor classifier for handwritten digit recognition”,. Journal of Information Science and Engineering,vol:21(6),pp:1245-1257, 2005.
16. Hassan, M. R., Hossain, M. M., Bailey, J., & Ramamohanarao, K.,”Improving k-nearest neighbour classification with distance functions based on receiver operating characteristics.” In Machine Learning and Knowledge Discovery in Databases (pp. 489-504), Springer Berlin Heidelberg, 2008.
17. Mona Sharma,Markos Markou,Sameer Singh,”Evaluation of texture methods for image analysis”, Pattern Recognition Letters,pp.1-8
18. M. Sonka, V. Hlavac and R. Boyle,” Image processing, analysis and machine vision”, PWS publishing, SanFrancisco, 1999.
19. Jain and D. Zongker, “Feature Selection: Evaluation,Application, and Small Sample Performance”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 2, pp. 153-158, 1997.
20. Armato SG, Giger ML, MacMohan H.” Automated detection of lung nodules in CT scans: preliminary results”,. Med Phys; 28:1552–61, 2001.
21. Lee Y, Hara T, FujitaH, Itoh S, Ishigaki T.,”Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique”. IEEE Transaction Medical Imaging 20:595–604, 2001.
22. Sheba.O, A Sukesh Kumar, “Automated Diagnosis of Macular Degeneration through Image Processing”, IE(I) Journal-CP,PP. 3-6, Vol 89, November 2008 .
23. K. Cederlund, U. Tylen, L. Jorfeldt, and P. Aspelin, “Classification of emphysema in candidates for lung volume reduction surgery,” Chest, vol. 122, pp. 590–596, 2002.
24. H. O. Coxson, R. M. Rogers, K. P. Whittall, Y. D’Yachkova, P. D. Par´e, F. C. Sciurba,and J. C. Hogg, “A quantification of the lung surface area in emphysema using computed tomography,” American Journal of Respiratory and Critical Care Medicine, vol. 159, no. 3,pp. 851–1073, 1999.
25. K. J. Park, C. J. Bergin, and J. L. Clausen, “Quantitation of emphysema with threedimensional CT densitometry: Comparison with two-dimensional analysis, visual emphysema scores, and pulmonary function test results,” Radiology, vol. 211, no. 2, pp. 541–547,1999.
26. D.Wormanns, S. Diederich, M. G. Lentschig,F.Winter, andW. Heindel, “Spiral CT of pulmonary nodules: interobserver variation in assessment of lesion size,” European Radiology,vol. 10, pp. 710–713, 2000.
27. J. P. Ko, H. Rusinek, E. L. Jacobs, J. S. Babb, M. Betke, G. McGuinness, and D. P. Naidich,“Small pulmonary nodules: volume measurement at chest CT — phantom study,” Radiology,vol. 228, no. 3, pp. 864–870, 2003.
28. E. Bomsztyk and H. Rusinek, “Volumetric analysis of lung nodules using a hybrid algorithm,” in Proceedings of the SPIE, vol. 5370, 2004, pp. 22–29.
29. N. Xu, N. Ahuja, and R. Bansal, “Automated lung nodule segmentation using dynamic programming and EM based classification,” in Proceedings of the SPIE, vol. 4684, 2002, pp. 666–676.
30. R. Uppaluri, E. Hoffman, M. Sonka, G. Hunninghake, and G. McLennan, “Interstitial lung disease. A quantitative study using the adaptive multiple feature method,” American Journal of Respiratory and Critical Care Medicine, vol. 159, no. 2, pp. 519–1073, 1999.