Neural Network Based Approach for Detection of Abnormal Regions of Lung Cancer in X-Ray Image

Vinod Kumar
Research Scholar
Dept. of Computer Science & Applications
Kurukshetra University, Kurukshetra

Dr. Kanwal Garg
Assistant Professor
Dept. of Computer Science & Applications
Kurukshetra University, Kurukshetra

ABSTRACT

Cancer is the most familiar disease that affect both men and women. The survival rate of lung cancer is extremely poor. To increase the survival rate of cancer patient, it is essential to detect it very early stage which enables many treatment options with reduced risk. Now a days, the image processing mechanisms are used in a number of medical profession for improving detection of lung cancer. This paper presents a neural network based approach to detect lung cancer from raw chest X-ray images. The author use an image processing technique to remove noise using various filters and segment the lung to detect abnormal regions in the X-ray image and extracted regions that demonstrate area, perimeter and shape characteristics of lung nodules. These shape features are considered as the inputs to train a neural network and to verify whether a region is a malignant nodule or not. This research work concentrate on detecting nodules, early stages of cancer diseases, appearing in patient's lungs. Most of the nodules can be observed after carefully selection of parameters. The training dataset of X-ray images of lung cancer are processed in three stages to attain more quality and accuracy in the observational results.

Keywords - Artificial Neural Network, Image processing, Lung Nodule and MATLAB.

I. INTRODUCTION

1.1 Medical Background

Lung cancer was rare at the beginning of the 20th century, is now a global problem, which is most frequent cancer in the world. The survival rate can be improved by discovering the existence of cancer in early stage. Early stage can be performed in inhabitants screening. Chest projection radiography is the most common screening mode [1]. It has been exposed in the Early Lung Cancer Action Projects that the conventional chest X-Ray used for the finding of pulmonary nodules. The pulmonary nodule appears in lung as a spherically shaped mass, which is distorted by surrounding anatomical structures [12]. The complexity for detecting lung nodules in X-Ray radiographs are nodule sizes, density and contrast alteration [15]. The diameter of a nodule can be in between a few millimeters up to several centimeters. Some nodules are rather denser than the neighboring lung tissues. Hence, the visibility on a X-Ray radiograph is reduced. The nodule can be found anywhere in the lung field as a result of contrast variation to the background.

1.2 Lung cancer

Abnormal cells multiplying and developing into a tumor caused lungs cancer. Cancer cells can be taken away from the lungs in blood, or lymph fluid that surrounds lung tissue. Metastasis spread has occurred to any distant organs or lymph nodes away from the chest through the bloodstream. Common distant organs are opposite lung, brain, liver, bones and adrenal glands. Cancer that originates in the lung is called primary lung cancer [15]. There are Small Cell Lung Cancer (SCLC) and Non Small Cell Lung Cancer (NSCLC). SCLC type cancerous is diagnosed about 20 out of every 100 lung cancers. These cancer cells are small in size that caused by smoking and is very rare for someone who has never smoked. This type of lung cancer often spreads very early. NSCLC are collected together because they act in a similar way and respond to behavior in a different way to small cell lung cancer.

II. PRESENT WORK

The lung regions extraction comes up into two different categories; either rule-based or pixel classification based category [7]. Most of the proposed method belong to rule-based category [9,10], where a progression of phase, tests and rules are used in the extraction process. Techniques make use of thresholding, region growing, edge detection and ridge detection, morphological operations, fitting a geometrical models or functions are dynamic programming. On the other hand, there is one more approach employed in lung regions extraction process based on pixel classifications, where each pixel in the X-Ray is classified into lung or its background. Classifiers are several types of neural networks, trained with a variety of local features including intensity, location and texture measures [7]. Computer aided diagnosis systems can be divided into two groups [9] :density-based and model-based approaches. Taking into considering the fact that lung nodules have relatively higher densities than those of lung parenchyma, density-based detection methods employ techniques such as multiple thresholding, region-growing, locally adaptive thresholding in combination with region growing, opening and closing, using the histogram, the top 20% gray values considered as preliminary cancerous candidate regions, using the histogram, the normal tissues are removed, then elliptical-shaped regions, which is in general characterize abnormalities, are detected, and fuzzy clustering is used to identify nodule candidates in the lungs. For the model-based detection approaches, the relatively compact shape of a small lung nodule is taken into account while establishing the models to identify nodules in the lungs [11]. Techniques such as morphological filters and the anatomy based generic model have been proposed to identify sphere shaped small nodules in the lung. Nodule candidates are detected using template matching or a modified Hough transform in which edge pixels vote for circles that could cause these edges. After getting segmentation results, different features should be extracted to be used in the diagnosis phase where sets of rules are put together to distinguish between true and false cancerous candidates. Different features were extracted in the different research papers depending on the approaches used by the authors in the diagnosis phase. In some methods, uniformity, connectivity and position features were extracted [13]. The features such as size, circularity and mean brightness of region of interests (ROIs) were extracted [14]. Area, thickness, circularity, intensity, variance, localization and distance from the lung wall are extracted features [8]. The underlying idea of developing a pulmonary nodule diagnosis system is not to hand over the diagnosis to a machine, but rather that a machine algorithm take action as a support to the radiologist and points out spots of suspicious objects, so that the overall sensitivity is boosted. The proposed diagnosis system come across main objectives, which are improving the quality and accuracy of diagnosis, increasing therapy success by early detection of cancer, avoiding unnecessary biopsies and reducing radiologist interpretation time [16,17].

III. PULMONARY NODULE DIAGNOSIS SYSTEM

A pulmonary nodule appears in lung X-Ray as a spherically shaped mass which can be distorted by surrounding anatomical structure and there is no limitations on size or distribution in lung tissue [5]. The pulmonary nodule is classified into certain categories; nodule is connected to pleural surface, other connection to neighboring vessels by thin structure [16]. Pre-diagnosis approaches help to locate the risk of lung cancer disease in very early stage [2].

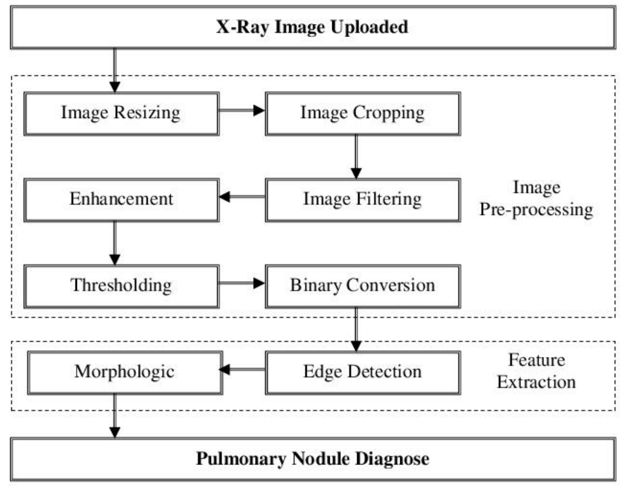


fig. 1 : Pulmonary Nodule Diagnosis System

Some of approaches used in pre-diagnostic are Artificial Neural Based Learning Process [3], Rule Based Learning Technique, Supervised Learning Methods, Fuzzy System, Expert System and Genetic Algorithm. In this paper, the author use ANN based learning method. In this paper, the original JPEG X-Ray image of 130 cases in various size and contrast are used for both healthy and cancerous patient which are collected from various reputed medical institute and hospitals [17] and it is further to classify the tumor as benign or malignant [16].

Pulmonary nodule diagnosis system uses resizing, cropping shown in fig. 2 (a) and applying median, gaussian filters to smooth the X-Ray images and the contrasts are enhanced. The lungs in the images are segmented by applying the Otsu’s thresholding. After binary conversion, pulmonary nodule detection system works-out morphologic technique to extract features include the perimeter, area and shape. On the behalf of this operation, cancerous identification of a lungs nodule is employed to analyze those features to evaluate whether cancer cells exist or not. Moreover, if there are cancers cells exist, then its type is identified. The entire process of lung nodule diagnosis system is shown in fig 1.

2.1 Image Pre-processing

It is necessary step to improve quality of X-Ray image by denoising, enhancement of structure and its contrast. Mean, Median, Laplacian and Gaussian filters are used for denoising after resizing and cropping of X-Rays. The process is adopted to enhancing the edge of image structure include unsharping and enhancing image contrast by histogram equalization [4]. Different type of tissues such as bone, muscle and fat have number of particulars in a X-Ray.

2.2 Denoising

In image processing, many methods are implemented for denoising like Mean, Median, Blurring, Gaussian and Laplacian filters. Image denoising algorithms may be the oldest practice. The input X-Ray is a normal RGB image that is not supported in MATLAB and convert it in grey scale. Then the grey scale image contain noises such as white noise, salt and pepper noise etc. White noise is the most common problems in image processing. The main idea of any filter is to calculate pixel weights depending on how similar their colors are [1]. The noise can be removed by using median filter from X-Ray.

2.3 Image Filtering

Median filtering is a nonlinear common enhancement digital filtering technique for removing noise without reducing the sharpness of the image [6]. It is usually applied in digital image processing for smoothing of the lung boundaries in thresholding. Author applied filtering by assigning 5X5 pixel. Linear filtering is used to remove certain types of noise. Gaussian or averaging filter are the appropriate for the removing of grain noise as each pixel get the average of pixels in its neighborhood and local variation caused by reducing of grain noise [10].

2.4 X-Ray Enhancement

Histogram manipulation is effectual practice for X-Ray image enhancement. Histogram of X-Ray as shown in fig. 2 (b) dividing the interval between the minimum and maximum pixel value into equally spaced bins. Count the number of pixels corresponding to each bin. The shapes of histogram is depending on the size of intervals. Histogram equalization as shown in fig. 2 (c) is that all the image intensities should be equally divided bins [16]. To calculate new assign value k for each brightness level j in the original X-Ray image

$$k = \sum_{i=0}^{j} \frac{N_i}{T} * I_{max}$$

Binary conversion is applied later than the process of enhancement as 8-bit X-Ray image altered into 2-bit gray scale image. If the pixel value in image is greater than threshold value, it shows ‘0’ (black) and if less than threshold then it show ‘1’ (white).

2.5 Thresholding

The cancerous nodules in the X-Ray image appear in low contrast and the non-nodule area show neither too bright nor too dark. The author use a multi-level thresholding to classify any point (x,y) in the image f(x,y) as belonging to object class if $T1<f(x,y)\leq T2$ to the other object if $f(x,y)>T2$ and to the background if $f(x,y)\leq T1$, where T1 and T2 are two threshold values and limits T1 = 120 and T2 = 170. upon observing all the X-Ray images in database, pixels within cancerous nodule are in range 125 to 158. the pixel value less than T1 and greater than T2 as background set to zero. The pixels whose values lies between T1 and T2 (foreground pixels) retain their pixel values. After this process, convert it into binary form by setting all the foreground pixel values equal to 255.

2.6 Lung Segmentation

Lung segmentation is very useful to extract the features from lung X-Ray. It may classify image pixel into anatomical regions such as bone muscles and blood vessels. In X-Ray image, segmentation have been segmented through pixel by pixel multiplication of lung mask. Lungs can be easily separated from other anatomic structure by binary thresholding m1(x,y) = Thr(f(x,y)) as shown in fig. 2 (d). After thresholding, the background is eliminated by suppressing all adjacent to image edges by flood filling. The matrix value put on view as 1 used for lungs and 0 for background.

2.7 Edge Detection

It is a set of connected pixels that lie on the boundary between two regions and detected by sobel methods because of its accuracy. Convolve the image g(r,c) to get smooth image and applying the threshold values to find the edge.

2.8 Lung Nodule Sighting

In this, several operation are applied such as thresholding and morphological operator like erosion and dilation. Erosion operator makes a region smaller while dilation operator enlarges a region [16]. Lung mask is smoothed by morphological closing with line element.

$$e(x_0 y_0) = \sqrt{1 for (x-x_0)^2 + (y-y_0)^2} \leq 4:5mm$$
$$0 \quad for \quad else.$$

The values in pixels are different. The smooth mask is s(x,y) = m(x,y) * e(x,y). Image f(x,y) is multiplied by smooth mask s(x,y) element by element g(x,y) = f(x,y) * s(x,y). From segmented lung X-Ray image, a nodule candidate is sighted, high densities appears as zero value (black) irregularities on the lung edge after thresholding. Now, subtraction of threshold image t(x,y) and smoothed mask s(x,y). Getting result j0(x,y)=s(x,y)-t(x,y).

$$b(x, y) = s(x, y) + (s(x, y) \oplus el(x, y))$$
$$\text{and } j(x,y)=g(x,y)*b(x,y)$$

Here, nodule candidate not located on lung boundary are eliminated from j0(x,y). Lung boundary is generated from the smoothed mask by morphological dilation and addition of X-Ray, where e is a line element. The mask j0=s(x,y) then multiplied element by element by boundary mask b(x,y).

2.9 Feature Extraction

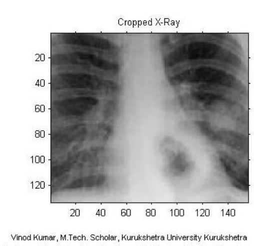
The feature act as the basis for classification process on binary X-Ray image. Here, three features are extracted as area, perimeter and shape. Area, perimeter and shape are a scalar value. Area allocate the actual number of over all summation of pixels (value 1) and perimeter provide the real number of inter-connected outline of the nodule pixel in the binary X-Ray image.

$$Shape = \frac{(4x\prod xArea)}{(Perimeter)^2}$$

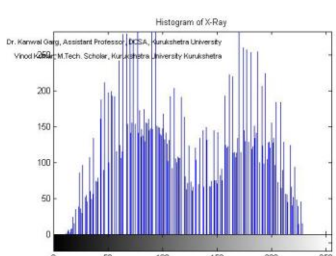
After getting the values from area and perimeter, shape can be calculated by the formula as shape = (4*pi*Area)/(Perimeter)^2.

IV. RESULTS AND DISCUSSION

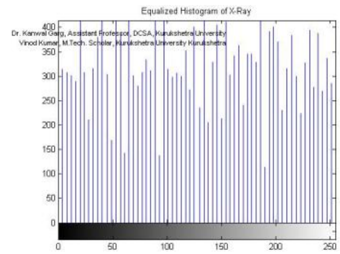
For experimentation of the proposed technique, the lung X-Rays are obtained from various prestigious medical college/hospital that provides the chance to do the recommended research. The experimentation data consists of 130 lung X-Ray images, are passed to the proposed pulmonary nodule detection system. The diagnosis rules are generated, forward through the learning process and then passed to the proposed system. At that time, the proposed system will go through its processing phases and finally it will detect the pulmonary nodule in the present uploaded X-Ray. On one hand, author have developed a pulmonary nodule diagnosis system for early detection of lung cancer using X-Ray images, which a high level of sensitivity has been attained. This prevents the system from hindering the radiologist's diagnosis. On the other hand, the proposed system is capable of detecting pulmonary nodules with area, perimeter and shape. This shows that the system is capable of detecting lung nodule in primarily stages as shown in fig. 2(a), 2(b), 2(c), 2(d), 2(e), 2(f), 2(g), 2(h), 2(i), 2(j), 2(k) and 2(l) . Hence, the patient's survival rate will improve because of early diagnosis. Artificial Neural Network results are shown in fig. 3(a), 3(b), 3(c), 3(d), 3(e) and 3(f). Classify between cancerous and non-cancerous candidate nodules. In the last, experiment results of 130 cancerous patients shown in fig. 4(a), 4(b) and 4(c) by graphical representation. Various record of epochs and their cumulative results shown when running NPR tool in MATLAB.



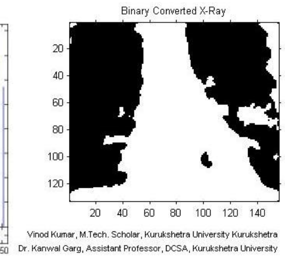
(a)



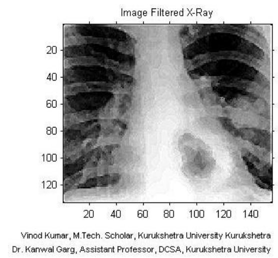
(b)



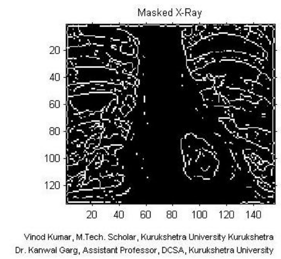
(c)



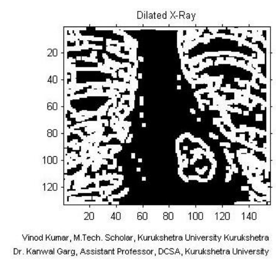
(d)



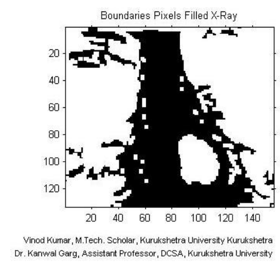
(e)



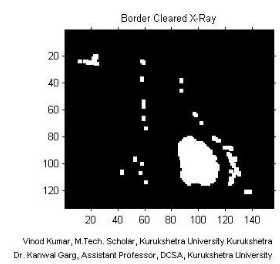
(f)



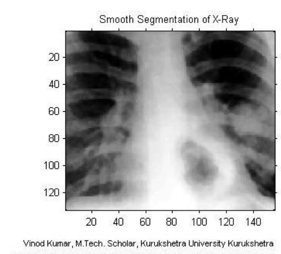
(g)



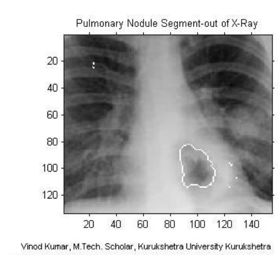
(h)



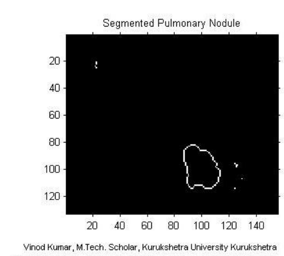
(i)



(j)

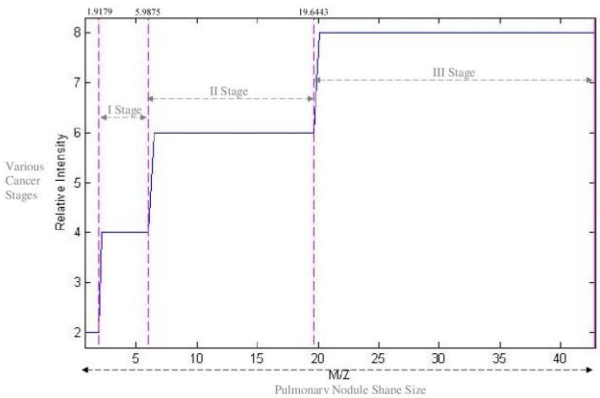


(k)

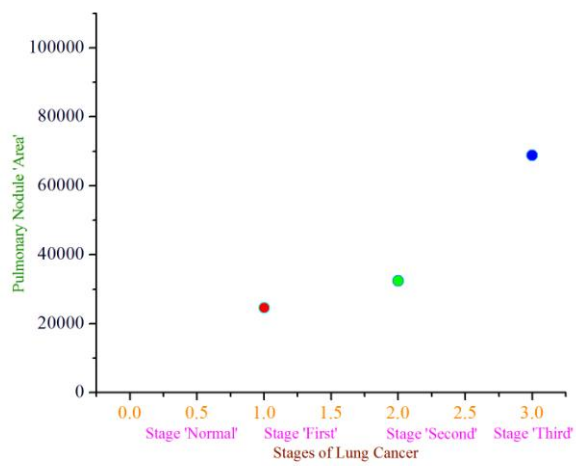


(l)

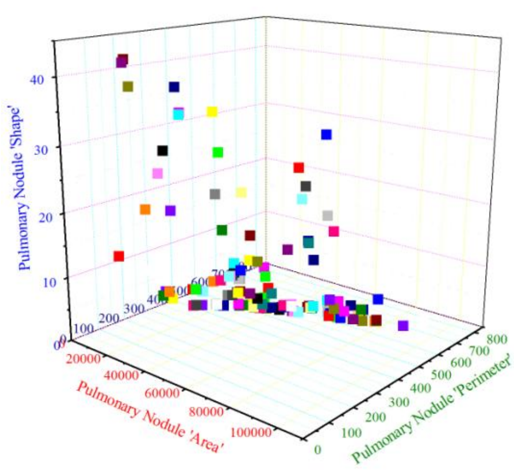
fig. 2 : Segmentation steps : (a) Resized & cropped X-Ray, (b) X-Ray Histogram, (c) Histogram Equalization, (d) Binary Conversion, (e) Flood Filled X-Ray, (f) Masking of X-ray, (g) Dilation of X-ray, (h) Boundaries Filled X-ray, (i) Border Cleared X-Ray, (j) Smooth Segmented X-ray, (k) Pulmonary Nodule Segment-Out, (l) Segmented Nodule



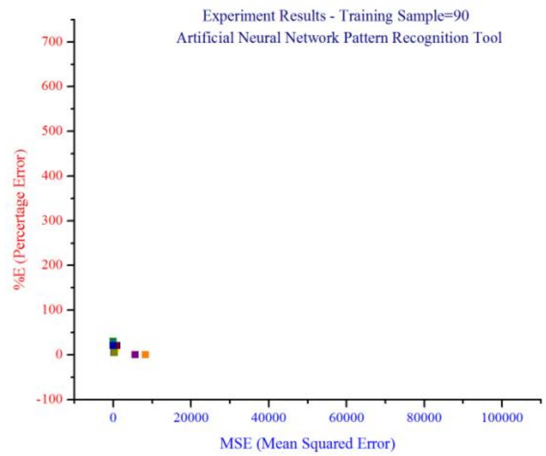
(a)



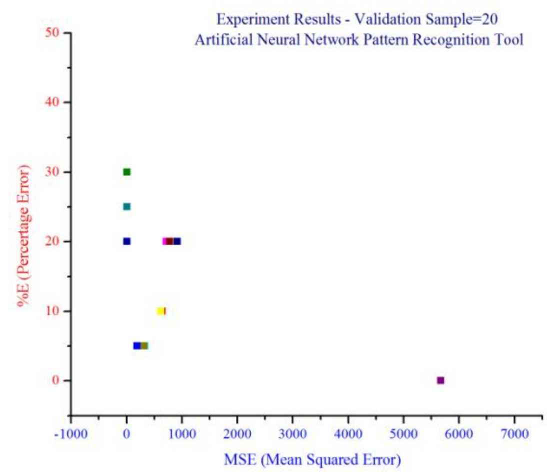
(b)



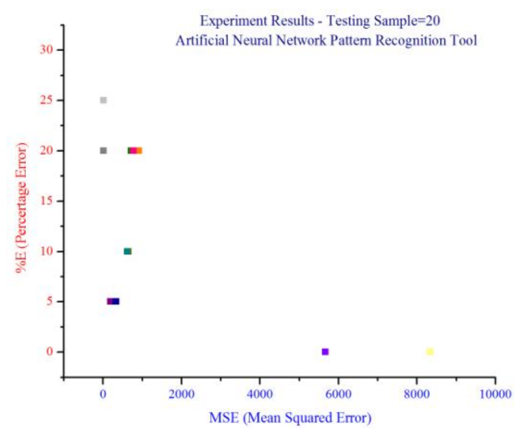
(c)



(d)



(e)



(f)

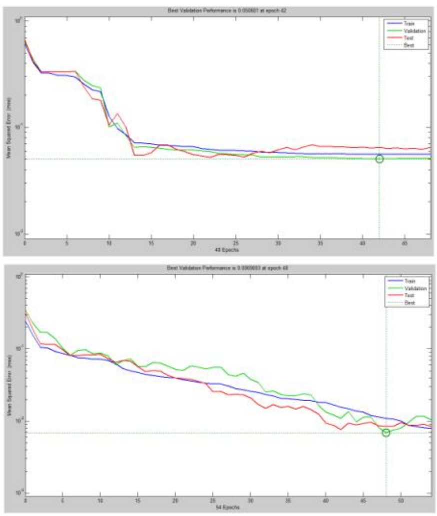
fig. 3 : Neural Network Pattern Recognition results : (a) Nodule Shape-Wise Stages, (b) Nodule Area-Wise Stages, (c) Nodule Area, Perimeter and Shape, (d) ANN Training Sample '90', (e) ANN Validation Sample '20', (f) ANN Testing Sample '20'

ISSN: 2278-0181
Vol. 1 Issue 5, July - 2012

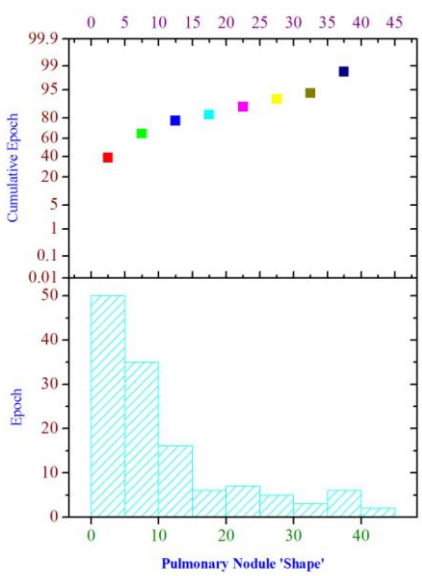
Exp.	Training Sample = 90		Validation Sample = 20		Testing Sample = 20	
	MSE	%E	MSE	%E	MSE	%E
I	5.60036e2	16.66666e0	5.06005e2	15.00000e0	6.49021e2	10.00000e0
II	2.03973e2	1.11111e0	2.58125e2	5.00000e0	3.24432e2	5.00000e0
III	4.33407e2	0	3.76674e2	5.00000e0	1.88855e2	5.00000e0
IV	4.51986e3	0	3.89442e2	5.00000e0	3.37171e2	5.00000e0
V	5.78186e2	15.55555e0	3.65883e2	10.00000e0	7.18531e2	20.00000e0
VI	6.41422e2	7.77777e0	1.02980e1	20.00000e0	6.22875e2	10.00000e0
VII	2.59429e2	2.22222e0	2.10619e2	5.00000e0	3.23680e2	5.00000e0
VIII	8.18304e2	17.77777e0	1.01453e1	20.00000e0	9.15543e2	20.00000e0
IX	2.19786e2	2.22222e0	3.18744e2	0	5.67028e3	0
X	3.76301e2	6.66666e0	3.35215e2	10.00000e0	7.79741e2	20.00000e0
XI	9.45833e2	27.77777e0	1.54557e1	45.00000e0	1.01051e1	30.00000e0
XII	7.17872e2	13.33333e0	1.15484e1	30.00000e0	1.12362e1	25.00000e0
XIII	2.36567e1	35.55555e0	1.33411e1	20.00000e0	1.33073e1	20.00000e0
XIV	**1.08425e2**	**0**	**6.90026e3**	**0**	**8.34194e3**	**0**

Table 5.1: Investigational Results of Artificial Neural Network Pattern Reorganization applying on 130 cancerous sample dividing for training 70%, validation and testing 20%.

(a)



(b)



(c)

fig. 4 : Experimental Results : (a) Mean Squared and Percentage Error with Training, Validation and Testing Sample in the 14 number of conduct experiments put into operation on 130 cancerous sample, (b) Training performance using Back Propagation Neural Network (c) ANNPR concerning on shape of nodule and getting results of 130 cases with epochs & its cumulative epochs subsequently whole processing.

V. CONCLUSION

In this paper, the author developed an automatic pulmonary nodule diagnosis system for early detection of lung cancer by studying lung X-Ray images using a number of steps. The approach starts by extracting the lung regions from lung X-Ray image using several image processing techniques in MATLAB including binary image, erosion, dilation, gaussian filter and median filter. Start with binary image including of thresholding technique that is used in the initial steps in the extraction process to convert X-Ray image into binary image, which is faster and user-independent. After the extraction step, the region growing segmentation algorithm is applied on extracted lung regions. Then the shape of nodule is calculated using shape formula with the help of area and perimeter of nodule. Finally, the extracted features helps to find the cancerous and non-cancerous candidate in X-Ray images. To differentiate the cancerous nodules from other suspected nodule area from X-Ray images, an artificial neural network using back propagation is developed. This consists of classifying the suspicious regions of pulmonary nodules.

REFERENCES

International Journal:

[1] B. Magesh et al., "Computer Aided Diagnosis System for the Identification and Classification of Lessions in Lungs", International Journal of Computer Trends and Technology, ISSN:2231-2803, IJCTT May-June 2011.

[2] K. Balachandran, Dr. R. Anitha, "Supervisory Expert System Approach for Pre-Diagnosis of Lung Cancer", International Journal of Advanced Engineering & Applications, Jan. 2010.

[3] Azian Azamimi Abdullah, Syamimi Mardiah Shaharum, "Lung Cancer Cell Classification Method using Artificial Neural Network", Information Engineering Letters, ISSN:21604114, Volume 2, Number 1, March 2012.

[4] M. Egmont-Petersen, D. de. Ridder and H. Handels, "Image Processing with Neural Networks - A Review", Pattern Recognition 35, 2279-2301, 2002.

[5] Qing-Zhu Wang, Ke Wang, Yang Guo and Xin-zhu Wang, "Automatic Detection of Pulmonary Nodules in Multi-slice CT Based on 3D Neural Network with Adaptive Initial Weight ", DOI 10.1109/ICICTA.2010.751, 978-0-7695-4077-1/10 $26.00 ©2010 IEEE.

[6] R. H. Chan, C. W. Ho and M. Nikolova, "Salt and Peeper Noise Removal by Median Type Noise Detector and Edge Preserving Regularization", Department of Mathematics-Hong Kong, pp. 1-7, 2003.

[7] B. V. Ginneken, B. M. Romeny and M. A. Viergever, "Computer Aided Diagnosis in Chest Radiography : A Survey", IEEE, Transactions on Medical Imaging, Vol 20, No. 12, 2001.

[8] S. A. Patil, Dr. V. R. Udupi, Dr. C. D. Kane, Dr. A. I. Wasif, Dr. J. V. Desai, A. N. Jadhav, "Geometrical and Texture Features Estimation of Lung Cancer and TB Image using Chest X-Ray Database", 978-4244-4764-0/09/$25.00 © 2009 IEEE.

[9] D. Lin and C. Yan, " Lung Nodules Identification Rules Extraction with Neural Fuzzy Network", IEEE, Neural Information Processing, Vol. 4, 2002.

[10] B. Zhao, G. Gamsu, M. . Ginsberg, L. Jiang and L. H. Schwartz, " Automatic Detection of Small Lung Nodules on CT Utilizing a Local Density Maximum Algorithm", Journal of Applied Clinical Medical Physics, Vol. 4, 2003.

[11] Cheran S. C., Gargano G., "Computer Aided Diagnosis for Lung CT using Artificial Life Models", Seventh International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2005.

International Conferences:

[12] Amal Farag, James Graham, Salwa Elshazly and Aly Farag, "Statistical Modeling of the Lung Nodules in Low Dose Computed Tomography Scans of the Chest", 17th International Conference on Image Processing, Proceeding of IEEE September 2010.

[13] A. El-Baz, A. A. Farag, Ph.D., R. Falk, M. D. and R. L. Rocco, M. D., "Detection, Visualization and Identification of Lung Abnormalities in Chest Spiral CT Scan : Phase I", Information Conference on Biomedical Engineering, Egypt, 2002.

[14] Yamamoto T., Ukai Y., Kubo M., Niki N., Satou H. And Eguchi K., "Computer Aided Diagnosis System with Functions to Assist Comparative Reading for Lung Cancer based on Helical CT Image", Image Processing, 2000 International Conference on Proceedings, Vol. 1, pp. 180-183, 2000.

Books:

[15] Heine Hansen, "Text book of lung cancer", 2nd ed., National University Hospital, Copenhagen, Denmark, ISBN-13: 978 0 415 385107 © 2008.

Thesis:

[16] Vinod Kumar, "Pattern Identification by Image Processing for Early Detection of Lung cancer", M. Tech. (Computer Science & Engineering) Dissertation, Department of Computer Science & Applications, Kurukshetra University, Kurukshetra, 2012.

Proceedings Papers:

[17] Vinod Kumar, Dr. Kanwal Garg and Rakesh Kumar, "Pulmonary Nodules Diagnosis from X-Ray Imaging using Image Processing", International Conference on Recent Trends of Computer Technology in Academia, ICRTCTA, J.R.N. Rajasthan Vidyapeeth University, Udaipur 21-23 April 2012. [0017]

AUTHOR'S PROFILE

Er. Vinod Kumar is a M. Tech. Student (Computer Science & Engineering), in the Dept. of Computer Science & Applications, Kurukshetra University, Kurukshetra. His research interest lies in image processing, data base and image mining.

Dr. Kanwal Garg is working as an Assistant Professor in the Dept. of Computer Science & Applications, Kurukshetra University, Kurukshetra with a teaching experience of 11 years. His research interests includes data base, data mining, web mining and warehousing. He has published 34 research papers in international and national journals, conferences and seminars. He has attended many workshops and faculty development programme.