International Journal of Enhanced Research in Management & Computer Applications, ISSN: 2319-7471
Vol. 2 Issue 9, Nov.-Dec., 2013, pp: (40-47), Available online at: www.erpublications.com

Detection system for lung cancer based on neural network: X-Ray validation performance

Vinod Kumar[1], Anil Saini[2]
[1,2]Asst. Professor, Dept. of Computer Science & Applications
Kurukshetra University, Kurukshetra, Haryana, India

Abstract: In our daily life, cancer is well-known disease that causes of death in both men and women and understand about the survival rate of lung cancer which is extremely poor. To increase this survival rate of cancerous patient, it is primarily to detect at premature stage which enables many new options for the cancer treatment without risk. In this paper, the author represents Lung Cancer Detection System for finding of lung cancer by analyzing chest X-rays with the help of image processing mechanisms. This system assists to radiologists for their X-ray image interpretation 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 techniques to denoise, to enhance, for segmentation and edge detection in the X-ray image to extract the area, perimeter and shape of nodule. These extracted features are considered as the inputs of neural network to train and to verify whether the extracted nodule is a malignant or non-malignant. 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 are processed in three stages to attain more quality and accuracy in the processed examination.

Keywords: Artificial Neural Network, Medical Image Processing, Segmentation and MATLAB.

I. INTRODUCTION

Lung cancer was uncommon at the beginning of the 20th century, is now a worldwide dilemma, which is more recurrent cancer in the world. The Lung Cancer remains a leading cause of mortality [14]. 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 [2]. It has been exposed in the Early Lung Cancer Action Projects that the usual chest X-Ray used for the finding of pulmonary nodules. The pulmonary nodule occurs in lung as a spherically shaped mass which is faint by neighboring anatomical structures. The complexity for spotting lung nodules in X-Ray radiographs are nodule sizes, density and contrast modification. The diameter of a nodule can be in between a few millimeters up to some centimeters.

Cancer detection system has been imparted a new step in early diagnosis of lung cancer. The proposed system consists of five core level: (i) Lung region extraction (ii) Segmentation of extracted lung (iii) Nodule detection (iv) Feature extraction (v) Analysis using Neural Network. Some nodules are rather denser than the neighboring lung tissues. Hence, the visibility on an X-Ray radiograph is reduced. The nodule can be found anywhere in the lung field as a result of contrast variation to the background. A tumor caused a lung cancer by multiplying and developing of abnormal cells. Due to this fact, lung cancer survival rate is 15% in five years [4]. To the examination of cancer cells, a piece can be taken away from the lungs in blood, or lymph fluid that surrounds lung tissue. Through the bloodstream, metastasis can spread to any distant organs or lymph nodes away from the chest. Lung, brain, liver, bones and adrenal glands are common opposite distant organs.

Cancer that originates in the lung is called primary lung cancer [22]. There are Small Cell Lung Cancer and Non Small Cell Lung Cancer. Small Cell Lung Cancer is diagnosed about 20 out of every 100 lung cancers. These cancer cell are small in size often spreads very early and is caused by smoking and is very rare for someone who has never smoked. Non Small Cell Lung Cancer is collected together. It acts in a similar manner and react to behavior in a unusual way to small cell lung cancer.

2. RELATED WORK

Abdullah et al. [2012] [1] stated that the segmentation of the lung region due to the limitation regarding on the similarities of the intensity in the X-ray image. As for lung cancer nodule detection process, it does not seem to be the problem because of the absent of the similar intensity due to the lung segmentation done. It can be used in the lung cancer application, the system can also be used in the application such as the detection and classification of breast tumor in mammography images regarding on the higher variation of intensity present.

Zare et al. [2011] [23] declared that the approaches of content- based image retrieval (CBIR) using low level features such as shape and texture are investigated in order to create a framework that classify medical X-ray image automatically. Gray level Co- occurrence Matrix, Canny Edge Operator, Local Binary Pattern and pixel level information of the images act as image based feature representations. The performance of image classification offered by combining the promising features stated above is investigated. Experimental results using 116 different classes of 11,000 X-ray images showed 90.7% classification accuracy.

Gomathi et al. [2010] [11] expressed that Computer Aided diagnosing system which uses FPMC algorithm for segmentation to improve the accuracy. Rule based technique is applied to classify the cancer nodules after segmentation. For its better classification, the learning is performed with the help of Extreme Learning Machine.

Patil et al. [2009] [4] conveyed that image segmentation is important for medical image analysis. It helps to detect the absence or presence of disease in an image. The Gray Level Co-occurrence Matrix (GLCM) technique is used to estimate texture features. It is applied on two main types of lung cancer images, like Small-cell, Non Small Cell type and as well as on TB database.

Jia Tong et al. [2007] [22] acknowledged that several steps are followed to detect the cancer like segmentation of lung parenchyma, the detection of suspicious nodule candidates, the feature extraction and classification. Here the author used adaptive threshold segmentation, math morphologic, Gaussian filter, Hessian matrix algorithms.

3. METHODOLOGY

The proposed Lung Cancer Detection System can identify the appropriate cancerous effected regions by applying the following steps shown in Fig. 1. In lung X-Ray, pulmonary nodule appears as a spherically shaped mass [13]. It can be distorted by adjacent anatomical formation. There are no boundaries on size or spreading in lung tissue. The pulmonary nodule is categorized into a few groups; nodule is associated to pleural surface and connected to neighboring vessels by thin structure [16]. Pre-diagnosis approaches help to locate the risk of lung cancer disease in very early stage [9].

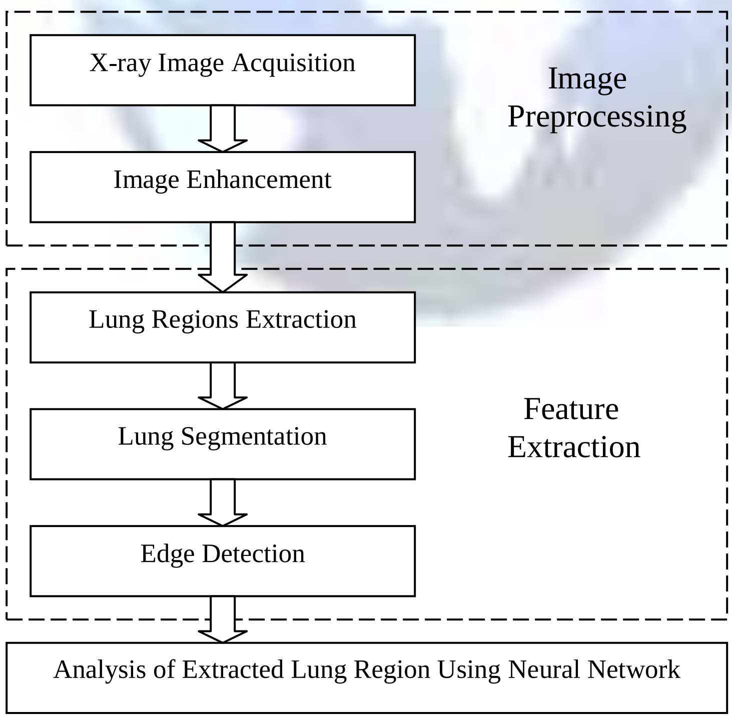

Fig. 1: Lung Cancer Detection System

In the pre-diagnostic of lung cancer, some of approaches used are Artificial Neural Based Learning Process, Rule Based Learning Technique, Supervised Learning Methods, Fuzzy System, Expert System and Genetic Algorithm.

3.1 IMAGE PROCESSING

In this paper, the author use ANN based learning method and used several original X-Ray images of various size and contrast for both cancerous and non-cancerous patients, which are collected from various reputed medical institute and hospitals [12] and it is further to classify the tumor as benign or malignant [16]. The diagnosis system uses resizing, cropping shown in Fig. 2(d) and applying median, gaussian filters Fig. 2(g) to smooth the X-Ray images and the contrasts are enhanced [15]. The images of lung are segmented by applying the Otsu's threshold. After binary conversion, lung cancer detection system works-out morphologic technique to extract features including the perimeter, area and shape of nodule [17]. On the behalf of this operation, cancerous classification of a lungs nodule is employed to study those features either cancer cells exist or not. In addition, if cancers cells stay alive, then its stage is identified. After completion of image acquisition, it follows that to enhance the co-related pixel scalar values in MATLAB program.

3.1.1 IMAGE ACQUISITION

To make better X-Ray image, the author put efforts to denoise, to enhance the structure and contrast. Once Median, Laplacian and Gaussian filters are used for denoise then the process is adapted to enhancing the edge of image structure contains unsharp and enhances the image contrast by histogram equalization [10]. There are different types of tissues like bone; muscle and fat have number of information in an X-Ray. Only the grey scale image contains noises such as white noise, salt and pepper noise etc. The most common problems in image processing are white noise. Here is the main suggestion of any filter is to calculate pixel weights. Where as the median filter is a nonlinear common enhancement digital filtering technique for removing noise without reducing the sharpness of the image [11]. The subsequent figures shown in Fig. 2(a), Fig. 2(b) and Fig. 2(c) : [20]

Fig. 2(a): Original X-ray

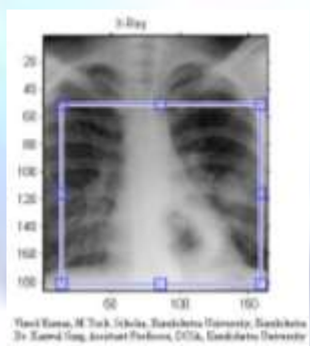

Fig. 2(b): To be cropped

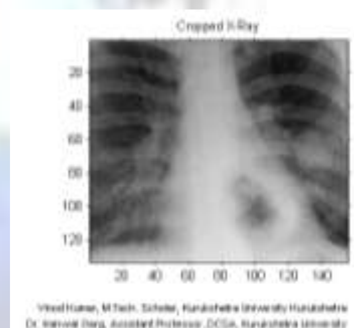

Fig. 2(c): Cropped X-ray

3.1.2 IMAGE ENHANCEMENT

It is commonly applied for smoothing of the lung boundaries shown in Fig. 2(j) in threshold in digital image processing. The author applied here for filtering by assigning 5x5 pixels. The linear filter is treating as to remove certain types of noise. Histogram operation is powerful practice for X-Ray image enhancement. The interval values between the minimum and maximum pixel divided into equally spaced bins in histogram. It count the number of pixels related to each bin. The shapes of histograms are depending on the size of intervals. The entire image intensities should be equally divided bins as shown in Fig. 2(e) histogram equalization.

The value k for each brightness level j in the original X-Ray image is determined by

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

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) [20].

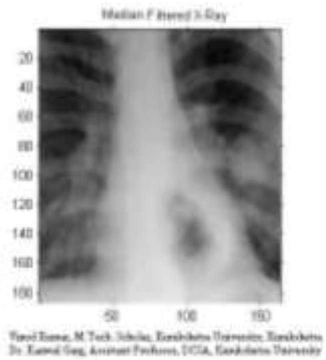

Fig. 2(d): Median Filtered X-ray

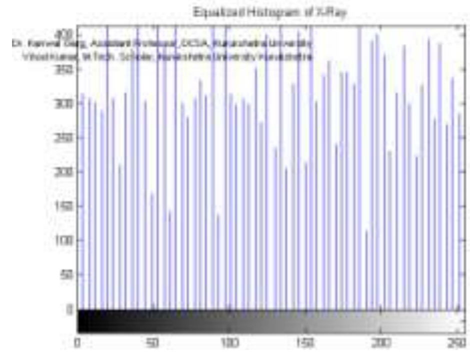

Fig. 2(e): Equalized Histogram of X-ray

3.2 FEATURES EXTRACTION

The author describes the feature extraction that 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 threshold 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.

3.2.1 LUNG REGIONS EXTRACTION

After observing all the X-Ray images in database, pixels within cancerous nodule are in range 125 to 158. The author stated that the pixel values are less than T1 and greater than T2 as background set to zero. The pixels whose values lie 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 [20].

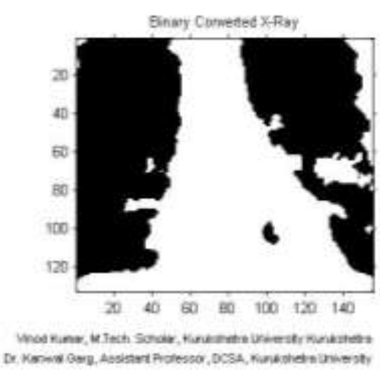

Fig. 2(f): Binary Conversion

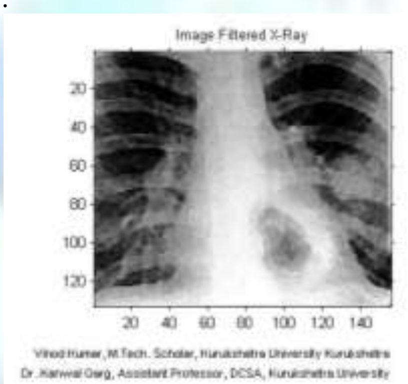

Fig. 2(g): Gaussian Filter

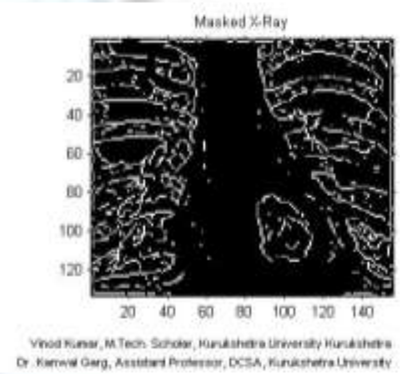

Fig. 2(h): Image Masking

The preliminary segmentation is very essential as if intact lung is not segmented correctly candidate nodules will be lost which may have important information about cancer into it. The different techniques are used for segmentation. One method is Threshold based segmentation but it has the problem with the selection of suitable threshold value. An additional technique is based on edge detection but the problem arises in the images related to edgeless or very noisy [18].
In this paper, the author used region based approach for segmentation. The region based segmentation is a technique for detecting the region directly. Chest X-ray images consist of different regions such as the background, lung parenchyma, heart, liver and other organ area [3]. The objective of lung region extraction is to branch out the lung region from other lung anatomy. Hence, in this paper, the segmentation is done by region growing algorithm and the growing of region start exactly from location of seed, appending those neighboring pixels that have predefined properties similar to seed.

3.2.2 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 has been segmented through pixel by pixel multiplication of lung mask show in Fig. 2(h). Lungs can be easily separated from other anatomic structure by binary

threshold m1(x,y) = Thr (f(x,y)) as shown in Fig. 2(f). After threshold, 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.
In this, several operation are applied such as threshold and morphological operator like erosion and dilation. Erosion operator makes a region smaller while dilation shown in Fig. 2(i) operator enlarges a region [016]. Lung mask is smoothed by morphological closing with line element.

$$e(x0\ y0) = \sqrt{1 \text{ for } (x - x0)2 + (y - y0)2 \leq 4} : 5 \text{ mm}$$
$$0 \text{ for else}$$

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 threshold [019]. 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).

$$b(x, y) = s(x, y) + (s\ (x, y) \circ ef(x, y))$$
$$\text{and } j(x, y) = g(x, y) * t(x, y)$$

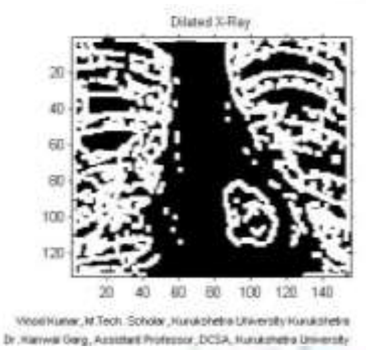

Fig. 2(i): Image Dilation

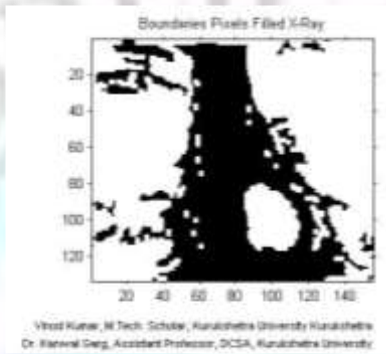

Fig. 2(j): Boundaries Filled

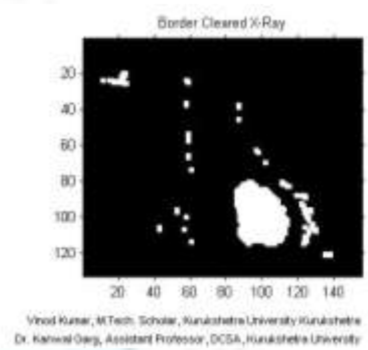

Fig. 2(k): Border Cleared

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).

3.2.3 EDGE DETECTION

The features are extracted on the basis of classification process on binary X-Ray image [1]. 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.

Here, three features are extracted as area, perimeter and shape. Area, perimeter and shape are a scalar value. Area allocates the actual number of over all summation of pixels (value 1) and perimeter provides the real number of inter-connected outline of the nodule pixel in the binary X-Ray image [20, 21].

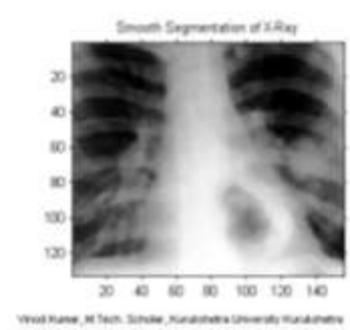

Fig. 2(l): Smooth Segment

Fig. 2(m): Nodule Segment

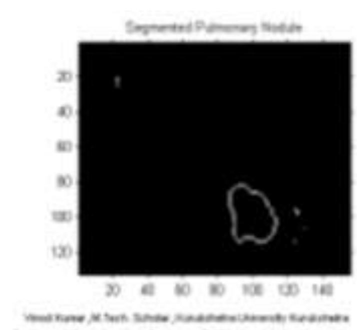

Fig. 2(n): Noisy Pixels

The edge detection method is implemented on the resulted X-ray. These features are used to develop diagnostic rules to detect cancer nodule. The features that are used in this study are area of candidate region mean intensity value of candidate region maximum draw able circle [5]. First of all, to eliminate isolated pixels, then to reject a nodule which does not co-related to candidate nodules. This is the cancerous region which is identified using the segmentation [7] as shown in the Fig. 2(m) and Fig. 2(n) shows some noisy pixel during the lung segmentation related to same pixel value identified in different site of the X-ray.

4. PERFORMANCE ANALYSIS USING NEURAL NETWORK

Artificial Neural Network is developed for diagnosis and classification of candidate nodules after applying training and testing process [6]. The neural network consists of three main layers, i.e. input layer, hidden layer and output layer. The author is employed the Back propagation algorithm currently. The suggestion of this algorithm is to reduce errors and produced by the difference between actual output and expected result accurately [7]. Firstly, the finest optimized neural network is obtained by varying various parameters of hidden nodes of network, i.e. training percentage for training ANN, number of epochs. As soon as the developed network becomes successful, it is then ready for classification process for the candidate nodule.

Neural network using MATLAB NNTOOL gives the results, which shown the ANN performance by graphical representation as shown in the following figures, Fig. 3(a), Fig. 3(b), Fig. 3(c), Fig. 3(d), Fig. 3(e), Fig. 3(f), Fig. 3(g), Fig. 3(h), Fig. 3(i), Fig. 3(j), Fig. 3(k), Fig. 3(l), Fig. 3(m) and Fig. 3(n). At this point, the author described figure's factors which exposed the performance of validation by artificial neural networks [20].

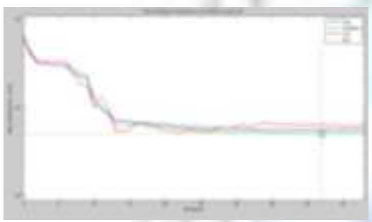
Fig. 3(a)

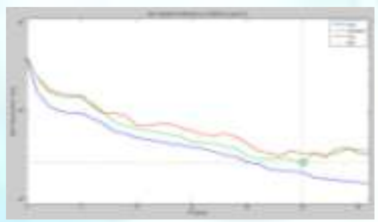
Fig. 3(b)

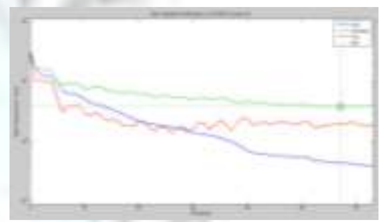
Fig. 3(c)

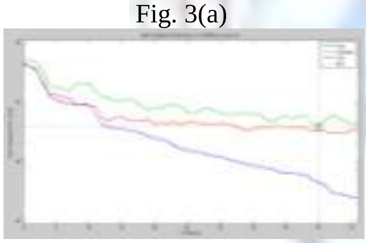
Fig. 3(d)

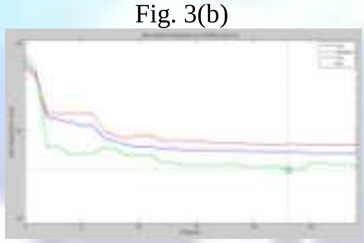
Fig. 3(e)

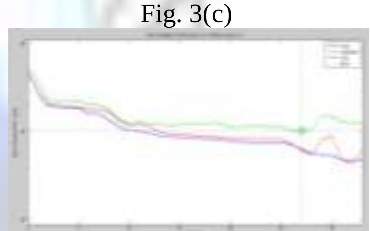
Fig. 3(f)

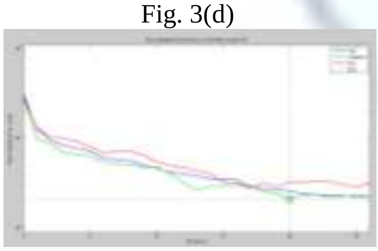
Fig. 3(g)

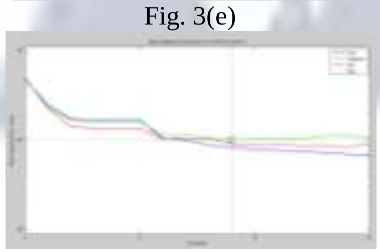
Fig. 3(h)

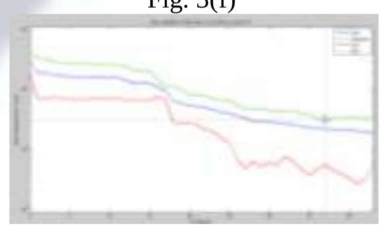
Fig. 3(i)

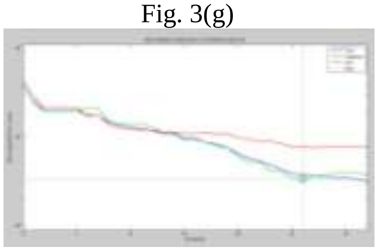
Fig. 3(j)

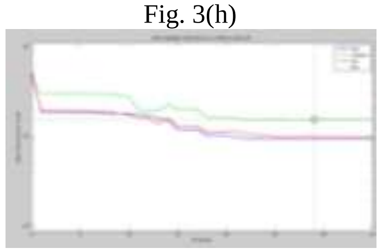
Fig. 3(k)

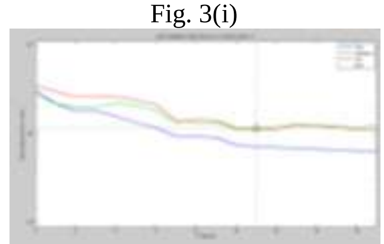
Fig. 3(l)

Fig. 3(m)

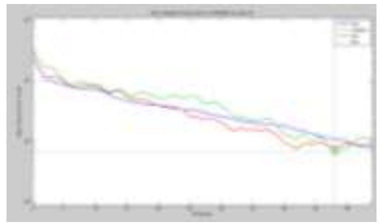

Fig. 3(n)

CONCLUSION

In this paper, the author developed an lung cancer detection 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 threshold 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 help 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. The principal advantages of Artificial Neural Network are their ability to discover the preferred information in data. The author suggests the best ANN technique with algorithm used for classification of lung cancer nodules in X-ray images. This system facilitates the radiologist and physician to recognize the suspicious nodules that increase the sensitivity of the diagnosis.

REFERENCES

[1]. 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.
[2]. 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.
[3]. 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.
[4]. 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.
[5]. 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.
[6]. 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.
[7]. D. Lin and C. Yan, “Lung Nodules Identification Rules Extraction with Neural Fuzzy Network”, IEEE, Neural Information Processing, Vol. 4, 2002.
[8]. Fatma Taher, Rachid Sammouda, “Identification of Lung Cancer Based On Shape and Color”- IEEE 2008.
[9]. K. Balachandran, Dr. R. Anitha, “Supervisory Expert System Approach for Pre-Diagnosis of Lung Cancer”, International Journal of Advanced Engineering & Applications, Jan. 2010.
[10]. M. Egmont-Petersen, D. de. Ridder and H. Handels, “Image Processing with Neural Networks - A Review”, Pattern Recognition 35, 2279-2301, 2002.
[11]. M. Gomathi , Dr. P. Thangaraj, “A Computer Aided Diagnosis System for Detection of Lung Cancer Nodules Using Extreme Learning Machine”, International Journal of Engineering Science And Technology Vol.2(10), 2010.
[12]. Penedo, M.G., Carreira, M.J., Mosquera, A. and Cabello, D., “Computer-Aided Diagnosis: A Neural-Network-Based Approach to Lung Nodule Detection”, IEEE Transactions on Medical Imaging, Pp: 872 – 880, 1998.
[13]. 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.
[14]. Ramnath Takiar, Deenu Nadayil, A Nandakumar, “Projection of number of cancer cases in India (2010-2020) by Cancer Groups”, Asian Pacific Journal of Cancer Prevention, Vol 11, 2010.
[15]. 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.
[16]. Vinod Kumar, Dr. Kanwal Garg, “Neural Network Based Approach for Detection of Abnormal Regions of Lung Cancer in X-Ray Image”, International Journal of Engineering Research & Technology, ISSN: 2278-0181, IJERT Vol. 1 Issue 5, July 2012.

[17]. 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.

[18]. 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.

[19]. Vinod Kumar, Dr. Kanwal Garg, 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.

[20]. Vinod Kumar (Research Scholar), Dr. Kanwal Garg (Guide) "Pattern Identification by Image Processing for Early Detection of Lung Cancer", Dissertation M.Tech. Computer Science & Engineering, Dept. of Computer Science & Applications, Kurukshetra University, Kurukshetra, June-2012.

[21]. Heine Hansen, "Text book of lung cancer", 2nd ed., National University Hospital, Copenhagen, Denmark, ISBN-13: 9780415385107, 2008.

[22]. Jia Tong, Zhao Da-Zhe, Wei Ying, Zhu Xin-Hua, Wang Xu, "Computer-Aided Lung Nodule Detection Based On CT Images", IEEE/ICME International Conference on Complex Medical Engineering, 2007.

[23]. Mohammad Reza Zare, Ahmed Mueen, Woo Chaw Seng, Mohammad Hamza Awedh, "Communication Systems and Networks Combined Feature Extraction on Medical X-ray Images" Third International Conference on Computational Intelligence, Bali, Indonesia, July 26-July 28, 2011.

AUTHOR'S PROFILE

Vinod Kumar is working as an Assistant Professor in the Dept. of Computer Science & Applications, Kurukshetra University, Kurukshetra with a teaching and technical experience of sixteen years. His research interests includes medical image mining, data mining, web mining and in the field of Networking. He has published four research papers in international journals and conferences. He has attended many workshops and faculty development program. He is a good counselor and motivator to the cancer patients last eighteen years as he was also victim of lung cancer and treated successfully by Living Foods, Naturopathy, Meditation and Yoga Science.

Anil Saini is working as an Assistant Professor in the Dept. of Computer Science & Applications, Kurukshetra University, Kurukshetra. His research area is Mobile Adhoc Networks, Programming Languages with Fuzzy Logic and Design and Analysis of Algorithm. He has published two research papers in international journals.