Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 12, December-2019

Multiple Step of Image Processing and ANN for Diagnosing Lung Diseases: Review Study

Prof. Dr. Hanan A. R. Akkar
Department of Electrical Engineer/UOT
University of Technology
Iraq/Baghdad

Asst. Lecturer Suhad Qasim G. Hadad
Departmant of Computer /UOT
University of Technology
Iraq/Baghdad

***Abstract*: - One of the critical physical ailments in recent years is lung cancer. Basically, it is caused by the initiation of rampant malignant abnormal cells grown in either both lungs or in one of them. The fatality rate that happens due to lung cancer is highest and is considered the leading cause of mortality out of all types of cancers. It is crucial to identify the ailment at the initial stage of the numerous types of cancer as it critical, to provide a higher survival rate for cancer sufferers. Hence, it is imperative for innovative techniques to be constructed. There were numerous soft computation methods utilized recently to identify the cancerous cells derived from medical image to provide an uncomplicated design, with minimal cost, no time consuming, and lastly offer positive outcomes to patients, specifically for patients in remote locations, with ordinary doctors (not specialist) to execute amazing activities to diagnose ailments. The current research introduces and compare between various review techniques and methods to categorize and diagnose pulmonary cancer ailments.**

Keywords- CT images, Cancer of the lung, Image processing, back propagation, Segmentation, Artificial Neural Networks (ANN).

I. INTRODUCTION

Numerous tools such as Sputum Cytology, or Computed Scan Tomography (CT), MRI scans or chest X-ray may be utilized to determine pulmonary cancer. Nevertheless, the mentioned techniques take a lot of time and occasionally are distant from the patients and are seen to be expensive. Moreover, the mentioned systems identify cancers in their developmental phase where the patient has little chance to survive. Prevalently, researchers are seriously attempting to identify pulmonic different kinds of cancer in their initial stages. The processing of images and Artificial Neural Network techniques provide the benefit for the enhancement of the research in ailments diagnostics [13]. In past few years, numerous images processing together with soft computing instruments were utilized worldwide for the enhancement of medical diagnostics

ANN processes are utilized to provide an extra precision for particular functions in comparison with those attained by the human brain. The current research is organized with five sections. In section two, common flowchart of pulmonary cancer recognition system indicated in figure (1), which explains the categorization stages for medical images. Section three comprises investigations with significant study for lung cancer identification and diagnostic schemes, with the description of algorithms and final result. While section four includes an outline, summary and observation for the categorization of lung cancer and its precision according to the methods that are explained in the previous sections .lastly, section five introduces a conclusion of this work.

II. COMMON DETECTION SYSTEMS

It has been noticed that the common methods of pulmonary cancer ailments and diagnostic process utilize various types of medical images shown in figure (1). Typically, the scheme comprises of five steps, which are itemized as the following [13]:

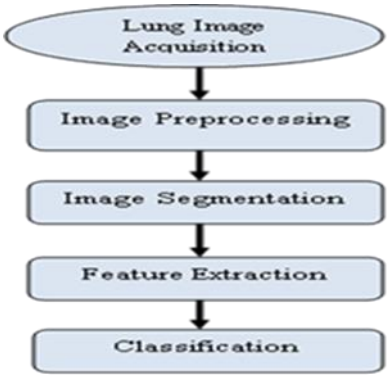

Figure 1: Common lung Cancer Detection Phases [13]

A. Image Capture
B. Image Pre-processing
C. Image Segmentation
D. Feature Extraction
E. Classification and investigation

A. Image Capture:

The capturing of image like computerized tomography (CT) or X-rays image is the initial phase for classification. CT images are the optimal selection that provides a minimal distinction. A CT Scan can be define as an X-ray that use computer assistance to gathered many X-ray to build cross sectional and three dimensional pictures for human body parts[11,13] . Moreover, the CT scan is a low hazard technique that is utilized for the diagnoses operations as it possesses a good resolution, uncomplicated, with minimal aberrations, in addition to the ease in modification and calculation of the mean for the scanned image [5, 11].

B. Image Pre-processing:

Objective of Images Pre-processing are the improvement of the illumination of pixels, where a single image may be isolated into significant characteristics that are inclusive of the internal features of these pixels. Numerous functions may be executed through the utilization of this process like denoising

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 12, December-2019

various kinds of noise from the image, enhancing the variances amongst the neighboring areas or characteristics, making the image less complicated by utilizing smoothing or removal features at specific scales, and keeping characteristics with particular needed scales. Numerous methods such as Median, Gabor, Wiener filter and morphological functions can be utilized in the manipulation of the image during the pre-processing phase [11].

C. Image Segmentation:

The core phase in image analysis is the segmentation of images. The segmentation process will allocate the image zones into varying objects or parts, like vascular system, organs, bones, and varying tissue kinds. The purpose of segmentation is to make it easier or modify an image symbol into a more meaningful and uncomplicated image for assessment to be utilized in diagnostic operations or in other phases. Suitably, segmentation may be explained by the division of the entire image area Z into a collective of minor segmentations, {Z1, Z2,… Zs}. A few techniques were being utilized in segmentation like watershed algorithm, Clustering techniques, thresholding, Edge detection, and connected component labelling algorithm [7].

D. Feature Extraction:

Feature Extraction is a major section of preprocessing, it utilizes algorithms and procedures to discover and isolate certain favored plan or area of a determined image. The core elemental facts are region, border and irregularity. All variables are determined with scales. In this stage it is eminent to classify and segregate areas of interest containing important characteristics attained from the lung image that establishes the nodule profile particularly for the probability of deduction of the presence of lung cancer could be made. Binarization and masking methods are utilized according to the lung analysis and lung data features [4, 6].

E. Classification and Investigation:

The concluding stages are resulting features of extraction, and diagnostics the diseases. It has been illustrated that numerous algorithms are utilized at this stage to arrange data into classified form by developing special algorithms which are recognized as classifiers. A duo of classification kinds can be utilized which are: supervised and unsupervised recognition and organization techniques [14]. Back Propagation (BP), Hopfield ANN, Adaptive Resonance Theory (ART), Convolution (CNN), Support Vector Machine (SVM), Artificial Neural Networks (ANN), are all supervised recognition and organization techniques. Unsupervised recognition and classifiers techniques include k-means clustering, Fuzzy Neural Network (FNN), Winner-take-all algorithm, in addition to Hebbian algorithm [2, 15].

III. LITERATURE REVIEW

The aim of researchers in this particular field is to develop methods that can realize initial phase of cancer. They attempt to improve the precision of the identification method by utilizing numerous processes like image processing, and recognition methods. Varying precision may be managed with utilizing different methods and processes techniques which will produce numerous outcomes. Various algorithms with an elaboration of the methods are shown in the following paragraphs:

Ganesan , Venkatesh , Rama , Palani, constructed a design for diagnosing cancer with X-rays lung images which were derived from one hundred lung cancer patients collected from various hospitals health organizations [1]. The data were trained and evaluated with multi-layer ANN back-propagation network it involved concealed layers that consist of four neurons X1, X2, X3 and X4, which signify the presence of lung cancer. The weights are computed and adjusted to minimize incorrect cases, and then the favorable set has been associated on the weightage of the first beginning layer. The implementation of back-propagation with suitable activation functions produced greater than 87% of accuracy, in addition the execution consumes a mere few seconds of duration [1].

Fatima Taher,Rachied S. Sammouda , suggested two segmentation algorithms for segmente sputum colour pictures. First Hopfield Neural Network (HNN), second technique was the C-mean Fuzzy (FCM) algorithm. Both algorithms were utilized in numerous varying areas like determination and identification, huge portions of these methods were constructed upon distance settings. Almost one thousand sputum color images were utilized to train and evaluate the previous methods. The utilization of the sputum segmenting images had illustrated a proper and ideal result in the primary phase of lung cancer , however the explorations done manually was time consuming, with errors, and it also required a highly qualified device to produce accurate outcome [2]. The outcome attained showed that segmentation using Hopfield Neural Network has greater precision and reliability in all situations, 97% of the image could be segmented successfully in nuclei cytoplasm area and background, the algorithm took short period to be achieved, and it didn't need more than 120 iterations to finish the segmentation. Fuzzy C-mean algorithm takes 50 iterations to finish the segmentation. Finally the totalities of the results are essential information for a computer supported analysis [2].

Ada, R. Kaur, objective of this study was to find characteristics of individual images by using a technique which was relied upon the features extraction and principal component analysis investigation, for the detecting of pulmonary in CT scans image [3]. Attainments of the CT images are derived from the Lung Image Database Organization (LIDO), Which is an important component of medical imaging titled as DICOM (Digital Imaging and Communication in Medicine) [5]. Preprocessing stages were initiated with image improvement method Histogram Equalization for the development and denoising the image. Feature extraction using binarization, and gray level co-occurrence approaches, these methods were implemented to achieve the general features of improved segmented image, features such as contrast, energy and maximum probability were extracted. Finally, they were gathered and made a comparison for the white pixels against certain thresholds to examine the regular, and an abnormal image of the lungs. When an image possesses lesser white pixels compared to threshold, this will indicate that it is a normal case. When the total of white pixels surpasses the limitations, the image is indicated as abnormal. The restriction quality of 255 was

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 12, December-2019

utilized as threshold. Principle component analysis (PCA) method was employed to obtain the last categorization .Some of the feature extracted were (for example): contrast = 0.6044, energy= 0.377, while homogeneity = 0.909[3].

Hanan Akkar, Samera Shams, presented a method for automatic lung cancer identification utilizing MATLAB a collaborative language which is a matrix that runs according to scientific and engineering mathematical computation and concept [1, 4]. In this work X-ray of lung image was used, the method required two stages for the purpose of testing the lung image that might be normal or constitute irregular features such as tumorous growth, or TB (Tuberculosis). First stage was managed for improving and denoising X-ray image, while second stage involved in the construction of ANN for the categorization of image. The training utilized back propagation algorithm combined with particle swarm optimization (PSO). For back propagation algorithm the design require an input layer, a hidden layer and output layer. Threshold value was executed with the entire neuron within the hidden layer and input layer, later a comparison with the initial layer was made. In condition when the outcomes do not match, the weight of previous relatives will be adjusted. Lastly, the neural particle swarm was applied to integrate back propagation in order to obtain optimal outcome [4].Figure (2) shows the X-rays image with pre-processing image.

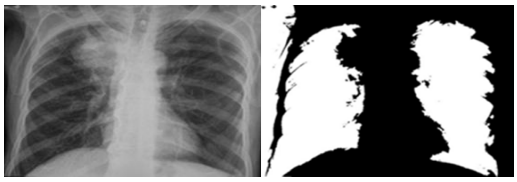

Figure 2: Two Image Shows Pre-processing stages for X-ray lung image [4]

Priyanka B., Asoke N., recommended isolation of pulmonary cancer into four stages (Stage I until IV) in CT image. The stages were set by tumorous growth, magnitude of the lymph node and location in the pulmonary imaging. Stages from I to IV were organized in sequence of strength will explained in the following [5]. Stage I: lung restricted cancer. Stages II and stage III: cancer was limited in chest. Stage IV: cancer of the chest has been grown malignant and spread to other parts of the body. Pre-processing which is the initial procedure was split into three phases: Smoothing, enhancement and segmentation Smoothing is executed through utilization of two processes, Median and Weiner filters. Median filter is signified as a non-linear technique utilized for the denoising purposes in image processing. Weiner Filter is considered as the optimal linear filters choice to skip extra noise and clouding from corrupted images [5]. Enhancement phase was directed through the operation of three methods .Gabor filter, auto-enhancement method and finally Fast Fourier transform process. Gabor filter has shown better results than other methods. Thresholding is an important and meaningful technique which is utilized in image segmentation.Additionally marker watershed technique was used for segmentation, which gave better result than thresholding. For the purpose of calculateing the possibility of the presence of cancer disease and feature extraction, two methods were introduced. First one was binarization, while the second was masking. A support vector machine which is a supervised learning model was utilized to achieve analysis on the data [5]. Pictorial representations for stages (I and IV) are shown below:

Stage I: the pictorial representation is shown in the figure (3) below :

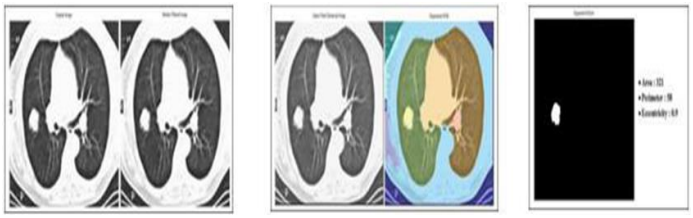

Figure 3: stage I: a) Original and median filtered Image b) Enhanced and segmented Image c) Binary Image [5]

Stage IV: the pictorial representation is shown in figure (4) below:

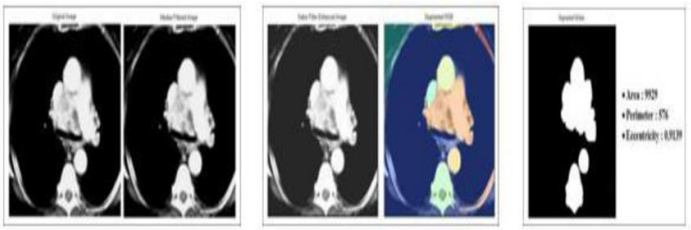

Figure4: stage IV: a) Original and median filtered Image b) Enhanced and segmented Image c) Binary Image [5]

Taolin, Hui, Shan, Xiuying**,** identification of Lung cancer was explored by the utilization of three dimension convolution network systems. Dataset of lung CT image which contained 1,397 critical patients' cases was used. Three dimension convolution networks were used with (11) layers which generated 12,544 neurons. Segmented lung volume was done by using Hounsfield Unit value, where the lung region has lower value with (-500), and the air value is (-1000) .Choosing threshold as (-400) will be satisfied in this work. The new volumes size of CT scan image will feed to the deep learning network and should include all the important features of the CT scan. Basically, convolution network AlexNet was used, the whole structure consisted of 11 layers with 5 convolution, 3 pooling layers and three full connections layers. The final structure takes 84 hours to train the network. The best result was gotten with dimension of input volume (128x128x20), the train instances were 1,300, and the test was 40 instances with 500 epochs [6].

S. Perumal, T. Velmurugan**,** provided a computer aided automated detection model for detection of lung cancer ailments. The input CT image was enhanced by applying adaptive histogram equalization which improves the contrast of the whole input image. Image segmentation was done by using watershed transform algorithm with marker control in order to divide the input image into fundamental regions that were useful to the next process [7]. Five levels of wavelet transform were used to get the feature extraction. Finally, classification was done by using three methods, which are artificial neural network, artificial bee colony, and the enhanced artificial bee colony. The result shows that EABC gives the highest accuracy with 92.4 for true rate, and 7.6 for

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 12, December-2019

false error rate [7]. Figure (5) shows original image and enhancement image with five levels wavelet transform.

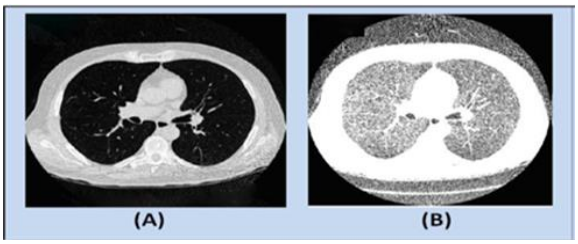

Figure5: (A) Original Image (B) Using 5 Levels Wavelet transform for Enhancement [7]

Ahmad, Fereshteh , Eric , LaRose, and Peggy, contribution of this work was the potential of joining several data of medical images combined from several sources to give more accurate diagnostic system . Data were collected from freely available site of data science, and also from private clinic of medical images. Two levels were suggested for classification, the first one was scale space, while the second was three dimension convolution networks. In the first level the images were resized to three different scales 50 × 50 × 20, 100 × 100 × 20, and 150 × 150 × 20, in order to save memory and time. Augmented data were generated by using two Gaussian noise between the layers to make the network more suitable for numerous noise and distortion .Meanwhile, in the second level three dimensions convolution networks were introduced with four layers ,three dimensions pooling , and fully connected network layer . Activation function was added, which was a nonlinear layer [Rectified liner unit ReLu], to overcome related problem. Finally softmax was used to calculate the target of each class. Various sequences of layers were used for implementation, such as 32x64 and 128 with filter size 3x3x3 [8].

Marios. Stergios, Lukas, Andreas, and Stavroula, recommended a technique of Convolutional Neural Network (CNN), constructed for categorization of interstitial lung diseases patterns. The recommended method comprises five convolutional layers with 2x2 kernels, with the utilization of Leaky ReLU for initiations, followed by pooling layers that were fitting with the size of the last feature maps and three compressed layers [9]. The final compressed layer possesses 7 outputs. Seven categories were taken into accounts which are: health wellbeing, micro nodes, association node, and others categories [9]. For the purpose of training and calculation dataset of 14696 images were utilized, resultants from 120 CT scans from various clinics [9].figure (6) shows the architecture of the proposed CNN for lung image classification.

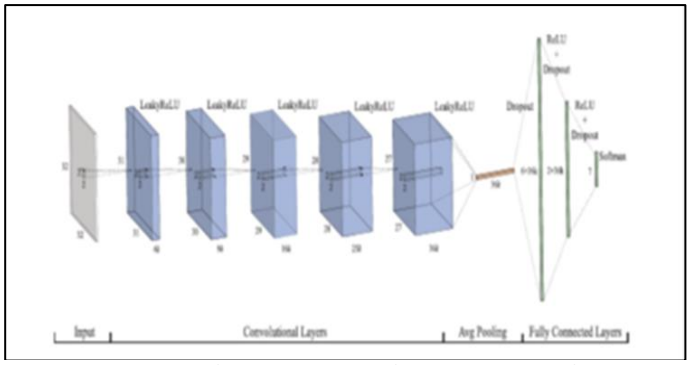

Figure 6: Architecture of the proposed CNN for lung image classification [9]

Sathiesh K., P.Santhoshini, introduced a design to detect lung cancer diseases, where CT scan image was collected from different hospital, threshold was used for segmentation. Feature extraction was done with two methods, gray level matrix structures and double tree compound wavelet. Generally, features that were extracted include dissimilarity, connection, energy and uniformity. Fuzzy technique was offered for clustering in which each part can belong to more than one group. Back propagation is finally introduced to classify lung cancer diseases for common and irregular region .The final accuracy was 90%, with sensitivity of 100%, and specificity of 79% [10].

IV. SUMMARY AND OBSERVATION

Image processing procedures with intelligent techniques were seen to be valuable for the initial estimation, and diagnosis of lung cancer.Table (1) indicates overview of the rate for the utilization of image processing and classification techniques through the implementation of different ANN techniques, with their operation degree for the detection of lung cancer. Figure (7) shows the performance of some technique used for classifying operations.

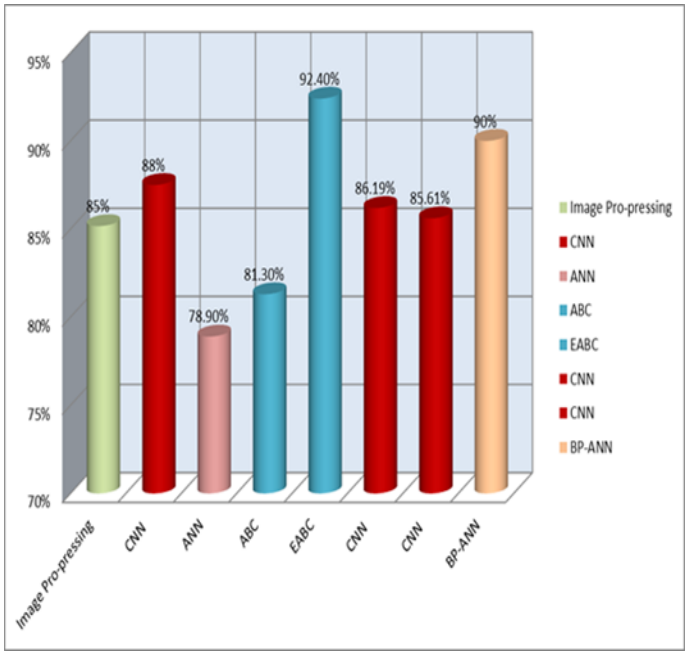

Figure7: performance and final accuracy of some techniques

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 12, December-2019

Table-1 Summary for classification of lung image for cancer diseases

Author Name	Year	Type of Image	Classification Technique	Accuracy
Ganesan, Venkatesh, , Rama , Palani,	2010	X-ray Image	ANN-Back-propagation Network	87%
Fatima Taher, ,Rachied Said Sammouda	2011	X-ray Image	Hopfield Neural Network ,and C-mean Fuzzy algorithms for segmentation images	Best result with Hopfield Neural Network
Ada, Rajneet Kaur	2013	CT Scan Image	feature extraction and Principal Component Analysis (PCA)	No. of White Pixels = 134011
Hanan. Akkar , Samera Shams	2014	X-ray Image	Back propagation Technique with particle swarm optimization with multilayer	87.359 %
Priyanka B. , Asoke N.	2017	CT Scan Image	Gabor Filter as image enhancement give best result Watershed filter for Segmentation give best result	80. 735 85.165
Taolin , Hui , Shan , Xiuying	2017	CT Scan image	11 Layers of Convolution Neural Network .best result was with input volume of (128x128x20), train is (1,300), test (40),epoch (500) .	87.5%
S. Perumal , T. Velmurugan	2018	CT Scan Image	ANN (Artfical Neural Network ABC(Artifical Bee Colony) EABC(Enhanced Artifical Bee Colony)	78.9% 81.3% 92.4%
Ahmad , Fereshteh , Eric , and Peggy	2018	CT Scan Image	Two kinds of approaches were used: scale space and three dimension Convolution Network with augmented data by using ReCall	86.19%
Marios.Stergio., Lukas, Andreas, and Stavroula,	2016	CT Scan Image	Convolution Neural Network, five layer used with seven output.	85.61%
Sathiesh. K, P. Santhoshini	2017	CT Scan Image	Artificial Neural Network – Back Propagation Neural used for classification .Fuzzy Neural used for clustering	90%

The performance of the algorithms in figure (7) and table (1) shows that:

1- Using X-rays as an input image, classification with ANN, Bp-ANN and image processing achieved precisions between 87% to 87.3%.

2- Using CT scan image as an input, Classification with CNN was achieved precisions between 85% to 87% depending on the size of convolution layers.

3- Using CT scan image as an input, classification with Artificial Bee Colony and Enhanced Artificial Bee Colony achieved highest rate 92%.

4- Using CT scan image as an input, classification with BP-ANN achieved 92%.

5- Using CT scan image or X-rays image as an input with various techniques of image processing achieved between 80% to 81%.

6- Using the Sputum color image. Segmentation with Hopfield Neural Network (HNN) technique and C-mean Fuzzy (FCM) algorithm. Both give good results for segmenting both nuclei and cytoplasm region.

V. CONCLUSION

Summary of varying processes for the detection and categorization of lung cancer was clarified. Explanation of different phases of the global diagram of processing system for lung cancer ailments is carried out through the utilization of medical images. Moreover, it has been discovered that the utilization of Computed Tomography (CT) images are the optimal choices when compared to Chest Radiograph (X-ray) or Magnetic Resonance Imaging (MRI scan) in addition to Sputum Cytology. The CT image provides visual representation with low dissimilarity, in addition to a high resolution, greater uncomplicatedness and minimal noise. In the case of segmentation, Marker Controlled Watershed algorithm produces better outcome in comparison with Gabor filtering. The utilization of ANN-Back Propagation, CNN, and EABC in categorization has generated positive outcome.

REFERENCES:

[1] N. Ganesan , K. Venkatesh , M. Rama , A. Palani," Application of Neural Networks in Diagnosing Cancer Disease Using Demographic Data", International Journal of Computer Applications , Volume 1-No.26 , 2010

[2] Fatima Taher,Rachied S. Sammouda , " Lung Cancer Detection by Using Artificial Neural Network and Fuzzy Clustering Methods", IEEE GCC Conference and Exhibition ,American Journal of Biomedical Engineering., February 19-22, 2011

[3] Ada, Rajneet Kaur," Feature Extraction and Principal Component Analysis for Lung Cancer Detection in CT scan Images". International Journal of Advanced Research in Computer Science and Software Engineering , volume 3 ,March 2013

[4] Hanan. Akkar, Samera Husain," Investigation of using particle Swarm Optimization and Neural Network for x-Ray Lung Disease ". Republic of Iraq Ministry of Higher Edu., Informatics Institute for Postgraduate Studies,2014

[5] Priyanka B., Asoke N., "Detection of Different Stages of Lungs Cancer in CT-Scan Images using Image Processing Techniques". International Journal of Innovative Research in Comp. and Comm. Eng., Vol. 5, May 2017

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 8 Issue 12, December-2019

[6] Taolin Jin, Hui Cui1, Shan Zeng, Xiuying Wang." Learning deep spatial lung features by 3D convolutional neural network for early cancer detection". The University of Sydney, Australia. College of Math. and Computer Science, China ©2017 IEEE

[7] S. Perumal T. Velmurugan, "Lung cancer detection and classification on CT scan images using enhanced artificial bee colony optimization". International Journal of Engineering & Technology, May 2018

[8] Ahmad P. Tafti, Fereshteh S. Bashiri, Eric LaRose, and Peggy Peissig." Diagnostic Classification of Lung CT Images using Deep 3D Multi-Scale Convolutional Neural Network". IEEE International Conference on Healthcare Informatics, ©2018 IEEE

[9] Marios Anthimopoulos,Lukas Ebner ,Andreas Christe,and Stavroula Mougiakakou,"Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network" .IEEE Transactions on Medical Imaging , vol. 35, no. 5, May 2016

[10] Sathiesh K., P.Santhoshini, "Lung Pattern Classification for Interstitial Lung Disease using ANN-BPN and Fuzzy Clustering", GRD Journals- Global Research and Development Journal for Engineering , Volume 2 , April 2017

[11] Kamil Dimililer , Buse Ugur, Yoney K. Ever ." Tumor Detection On CT Lung Images Using Image Enhancement". Online Journal of Science and Technology - January 2017

[12] Shraddha G. Kulkarni, Sahebrao B. Bagal." Techniques for Lung Cancer Nodule Detection: A Survey". (IRJET) Volume: 02 Issue: 09 , Dec-2015

[13] N.Bhaskar , Boddu Chaitanya ,"A Survey on Early Detection and Prediction of Lung Cancer ". (IJTIMES) , June-2018

[14] Bh. Malik, J. P. Singh, Veer B. P., Prashant Naresh." Lung Cancer Detection at Initial Stage by Using Image Processing and Classification Techniques". International Research Journal of Engineering and Technology (IRJET) ,Nov -2016

[15] Shriwas, Dikondawar." Lung Cancer Detection and Prediction by Using Neural Network". International Journal of Electronics &Comm., Publisher for Research, January 2015