IPASJ International Journal of Electronics & Communication (IIJEC)
Web Site: http://www.ipasj.org/IIJEC/IIJEC.htm
A Publisher for Research Motivation..... Email: editoriijec@ipasj.org
Volume 3, Issue 1, January 2015 ISSN 2321-5984

LUNG CANCER DETECTION AND PREDICTION BY USING NEURAL NETWORK

[1.]Raviprakash S. Shriwas, [2.] Akshay D. Dikondawar

[1.]M.E. in Digital communication Assistant professor in EXTC Department
J.D.I.E.T. Yavatmal

[2.]Final Year(Electronics and Telecommunication) Jawaharlal Darda Institute of Engineering and Technology, Yavatamal, Maharashtra

ABSTRACT

Nowadays cancer has become huge threat in human life .there are many types of cancer, Lung cancer is one of the common types causing very high mortality rate. The best way of protection from lung cancer is its early detection and prediction .The detection of lung cancer in early stage is a challenging problem, due to the structure of the cancer cells, where utmost of the cells are overlapped with each other. It is a computational procedure that sort images into groups according to their similarities. In this Histogram Equalization is used for preprocessing of the images and feature extraction process and neural network classifier to check the condition of a patient in its early stage whether it is normal or abnormal. The performance is based on the correct and incorrect classification of the classifier.

1. INTRODUCTION

Cancer is the most serious health problem worldwide. Lung cancer is one of the most dangerous cancers in the world, with the least survival rate after the diagnosis, with increase in the number of deaths every year gradually. Lung cancer is cause due to uncontrolled growth of abnormal cells in one or both of the lungs .The possible treatments are surgery, chemotherapy, and radiotherapy. The earlier the detection gives the higher the chances of successful treatment. It is considered to be the main cause of cancer death worldwide, and it is hard to detect in its early stages because its symptoms appear only in the advanced stages causing the mortality rate to be the highest among all other types of cancer. More people die because of Lung cancer than other types of cancer such as breast, skin, colon, and prostate cancers. There is significant proof indicating that the early detection of lung cancer will decrease mortality rate [2]. The latest estimates according to the recent survey provided by world health organization indicates that around 7.6 million deaths in whole world each year because of this type of cancer. There are many techniques to diagnose lung cancer, like Chest Radiography (x-ray), Computed Tomography (CT), Magnetic Resonance Imaging (MRI scan) .but, most of these techniques are costly and time consuming. And most of these techniques are detecting the lung cancer in Its advanced stages. Hence, there is a great need of a new technology to diagnose the lung cancer in its early stages. Image processing techniques provide a good class tool for cultivating the manual analysis. Artificial neural networks offer a completely different approach to problem solving and they are sometimes called the sixth generation of computing. The objective of this research is to apply neural networks and their associated analysis techniques to Health care, specifically to the management of lung cancer patients. A number of medical researchers utilized the analysis of sputum cells for early detection of lung cancer, most recent research relay on quantitative information, such as the shape, size and the ratio of the affected cells [4].

2. LITERATURE SURVEY

Carcinogenesis (the creation of cancer), is the process due to which normal cells are transformed into cancer cells. (i.e. uncontrolled and dangerous cell growth). Cancer is the general name for over 100 medical conditions involving uncontrolled and dangerous cell growth. There are a number of cancers to which only males are susceptible or to which only females are susceptible. Smoking is the one of the common risk for lung cancer, due to the harmful carcinogens found in tobacco smoke. Lung cancer is mostly related to smoking or use of tobacco. As Compared to nonsmokers, men who smoke are about 23 times more likely to develop lung cancer and women who smoke are about 13 times more likely to develop lung cancer[2]. About 90% of lung cancer deaths in men and almost 80% in women causes due to

smoking. Neural networks are important tools for cancer detection and monitoring. An initial diagnosis called early diagnosis is made based on the demographic and clinical data of the patient. About more than 30% cancer deaths are preventable. Curing cancer has been a major aim of medical researchers for decades. Artificial neural networks now are used in many fields. They have become well established as practical, multipurpose, strong computational methodologies with solid theoretic support and with strong potential to be effective in any discipline especially in medicine. Over the last two decades, a great amount of research work has been conducted for automated cancer diagnosis. Chiouet .(2003) intended an artificial neural network based hybrid lung cancer detection system named HLND, which can improve the accuracy of diagnosis. It includes the processing phases, such as pre-processing to enhance the figure background contrast of image, selection of nodule suspects quickly based upon the most noticeable feature of nodules and complete feature space determination and neural classification of nodules. Gutte, Henrik (2007) developed a completely automated method based on image processing techniques and artificial neural networks for the interpretation of combined Fluorodeoxyglucose (FDG) positron emission tomography (PET) and computed tomography (CT) images for the diagnosis and staging of lung cancer. Kenji Suzuki(2003) investigated a pattern-recognition technique based on an artificial neural network (ANN). Penedo et al (1998) developed a system that employed an artificial neural network to detect suspicious regions in a low-resolution image and employed another artificial neural network to deal with the curvature peaks of the suspicious regions, which was used in the detection of lung nodules found on digitized chest radiographs. Bartfay (2006) proposed a neural network model by Utilizing data on patients from two National Cancer Institutes of Canada clinical trials, he compared analytical accuracy of neural network models and logistic regression models on risk of death of limited-stage small-cell lung cancer patients [2].

3. LUNG CANCER

Lung cancer is a disease of multiplying and growing of abnormal cells into a tumor. Cancer cells can be carried away from the lungs in blood or lymph fluid that surrounds by lung tissue [3]. Metastasis refers to cancer spreading beyond its site of origin to other parts of the body.

Types of Lung Cancer

Cancer that starts in the lung is called primary lung cancer. There are different types and these are divided into two main groups such as

- Small cell lung cancer
- Non small cell lung cancer

4. METHODOLOGY

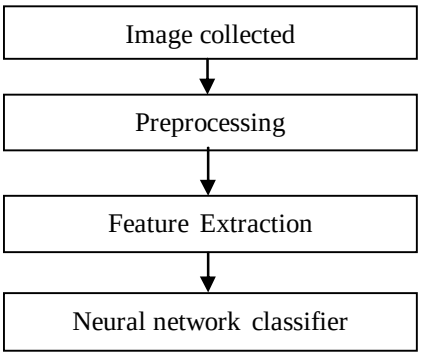


Fig.4.1. Methodology of work

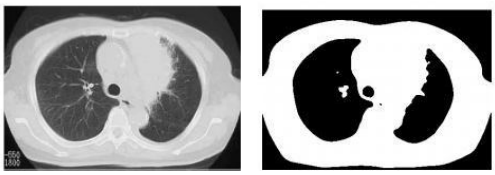
In the LCDS (Lung cancer detection system) system the contrast and color of the images are enhanced. After that the nucleuses in the images are segmented by thresholding. All these are simple digital image processing techniques. After that, LCDS develops morphologic and colorimetric techniques to extract features from the images of the nucleuses. On this basis, neural network classifier is employed to analyze those features to judge whether cancer cells exist in the specimens or not. Similarly, if there are cancer cells, the cancer cell type is identified.

4.1 Image Collected

The CT scan images of lung cancer consisting of small cell type of lung cancer images and non-small cell type of lung cancer images .Having each of the size 512 × 512 pixels in size. The digitized images are stored in DIACOM format with a resolution of 8 bits per plane.

4.2 Preprocessing Of Image

In the image Pre-processing stage it contains smoothing, enhancement, and segmentation. The aim of image enhancement is to improve the interpretability or observation of information in images for human viewers, or to provide better input for other automated image processing techniques[5].Most of the pre-processing is done with the help of MATLAB software. Generally, the quality of image is affected by different artifacts due to non-uniform intensity, variations, motions, shift, and noise [6]. Thus, the pre-processing of image aims at selectively removing the redundancy present in scanned images without affecting the details which that play a key role in the diagnostic process. Hence, Histogram-Equalization becomes play the important step in preprocessing of image. Therefore each image is preprocessed to improve its quality .Image segmentation is an essential process for most image analysis successive tasks.



Input Image Output Image

Fig. 4.2. Shows the Histogram Equalization on CT scan

4.3 Image Enhancement

The image enhancement is defined as to improve the quality of image, so that the resultant image is better than the original image[7] , the process of improving the quality of a digitally stored image by manipulating the image with MATLAB™ software. It is quite easy, for example, making an image lighter or darker, or to increase or decline contrast of the image.

Image enhancement techniques can be divided into two categories such as:

1. Spatial domain techniques, which operate directly on pixels.
2. Frequency domain techniques, which operate on the Fourier transform of an image.

4.4 Image Segmentation Technique

In this step the lung fields are separated from the rest of the things. Segmentation divides an image into its fundamental regions or objects. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is usually used to locate objects and boundaries (lines, curves, etc.) in images. More specifically, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share definite visual characteristics. The result of image segmentation is a set of segments that cooperatively cover the entire image, or a set of outlines removed from the image (edge detection). Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture [8].

4.5 Feature Extraction

Image features Extraction stage is an important stage that uses algorithms and techniques to detect and separate various desired portions or shapes (features) of a given image. With this necessary feature required for analysis is extracted to predict the probability of lung cancer presence, the following two methods are used such as Binarization and GLCM, both methods are based on strong facts that related to lung anatomy and information of lung CT imaging.

4.5.1 GLCM (Grey Level Co-Occurrence Method)

The GLCM is a tabulation of how often different combinations of pixel brightness values (grey levels) occur in an image. Firstly we create gray-level co-occurrence matrix from image by using graycomatrix function in MATLAB software. From this we can calculate texture measures from the GLCM. The features extracted using this method are (contrast, energy, entropy, maximum probability, correlation, cluster shade, cluster prominence, dissimilarity, autocorrelation, sum variance, sum entropy, difference entropy, information measures).

4.5.2 Binarization Approach

Binarization approach has been applied for detection of cancer. In this we extract the number of white pixels and check them against some threshold to check the normal and abnormal lungs. we started to count the black pixels for normal and abnormal images to get the average which will be denoted later as threshold then each image black pixels will be compared to this threshold, whether it is greater, then it is normal, else the opposite If the number of the white pixels of a new image is less that the threshold, then it indicates that the image is normal. otherwise, if the number of the white pixels is greater than the threshold value, which shows that the image in abnormal. Combining Binarization and GLCM approaches together will lead us to take a decision whether the case is normal or abnormal.

IPASJ International Journal of Electronics & Communication (IIJEC)
Web Site: http://www.ipasj.org/IIJEC/IIJEC.htm
A Publisher for Research Motivation.....
Email: editoriijec@ipasj.org
Volume 3, Issue 1, January 2015
ISSN 2321-5984

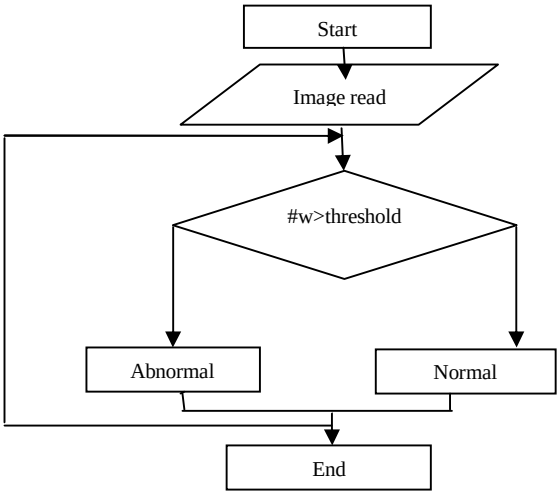


Fig.4.5.1. Binarization check method

5.NEURAL NETWORK CLASSIFIER

Supervised feed-forward propagation neural network ensemble used as a classifier tool. Neural network differs in various ways from traditional classifiers such as Bayesian and k – nearest neighbor classifiers. One of the main differences between them is linearity of data. Traditional classifiers like Bayesian and k – nearest neighbor requires linear data to work correctly. But neural network also works for non-linear data because it is simulated on the observation of biological neurons and network of neurons. Wide range of input data for training makes neural network to effort with higher accuracy level, or in other words a small set of data or large set of similar data makes system to be biased .Thus neural network classifier requires a large set of data for training and also long time to train to reach the stable state. But once the network is trained it works as fast as biological neural networks by propagating signals as fast as electrical signals . The architecture of the neural network contains of three layers such as input layer, hidden layer and output layer. The nodes in the input layer are linked with a number of nodes in the hidden layer. Every input node linked to each node in the hidden layer. The nodes in the hidden layer may linked to nodes in another hidden layer, or may be to an output layer. The output layer consists of one or more response variables .

The Stages Performed in Neural Network Classifier:-

- Create feed-forward propagation network.
- Train neural network with the training samples and the group defined for it.
- Simulate the neural network to check the particular selected input sample has cancer or not.
- From the outcomes of network and the samples trained in network, classification rate is calculated.

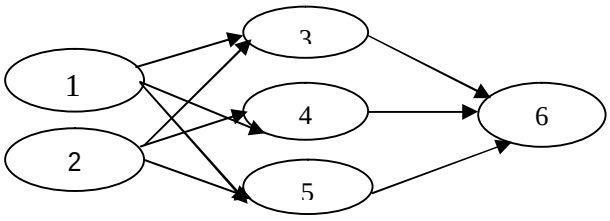


Fig. 5.1. feed-forward propagation network

ADVANTAGES

1] It requires less input of knowledge about the problem than other approaches.
2] It is skilled of implementing more complex Separating of feature space.
3] It is amenable to high-performance parallel processing implementation of image.
4] It works for both linear and non-linear data.
5] It works with high accuracy.

6.DISADVANTAGES

1] The extensive amount of training is required.
2] Slower operation when implemented as a simulation on a conservative computer.
3] The unavailability of a detailed understanding of the decision making process that is being used.

7. CONCLUSION

Neural Network Algorithm is implemented using open source and its performance is compared to other classification algorithms. It shows the best results with providing highest TP Rate and lowest FP Rate and in case of correctly classification, it gives the 96.04% result as compare to other classifiers. Neural network model is a diagnostic system that performs at an accuracy level is constructed. In this process, the performance of neural network structure was investigated for lung cancer diagnosis problem. People can be checked for lung cancer disease quickly and painlessly and thus detecting the disease at an early stage. This indicates that neural network can be effectively used for lung cancer diagnosis to help oncologists. The prediction could help doctor to plan for a better medication and provide the patient with early diagnosis.

8.FUTURE SCOPE

The million order dataset can be selected and image classification can be done on larger dataset. With improved size of dataset various issues such as uploading data, managing feature set, increased execution time of classification algorithms etc. could be considered. More image features can be extracted for better classification. Various combinations of previous features can be used to correctly classify medical data.

References

[1] Rajneetkaur, Ada, ”Early Detection and Prediction of Lung Cancer Survival using Neural Network Classifier”, International Journal of Application in Engineering and Management .vol-2,june2013.
[2] Dr.N.Ganeshan,Dr.k.Venkatesh,Dr..Rama,A.Malathi Palani, ”Application of Neural Networks in Diagnosing Cancer Disease Using Demographic Data”, International Journal of Computer Applications,vol- 1,2010 .
[3] Anita Chaudhary,SonitSukhraj Singh, “Lung Cancer Detection Using Digital Image Processing”, International Journal of Research in Engineering & Applied Sciences,vol-2
[4] Almas Pathan, Bairu.K.saptalkar, “Detection and Classification of Lung Cancer Using Artificial Neural Network”, International Journal on Advanced Computer Engineering and Communication Technology Vol-1 Issue:1.\
[5] D. Kim, C. Chung and K. Barnard, "Relevance Feedback using Adaptive Clustering for Image Similarity Retrieval," Journal of Systems and Software, vol. 78, pp. 9-23, Oct. 2005.
[6] Kakeda, S. et al., “Improved Detection of Lung Nodules on Chest Radiographs Using a Commercial Computer-Aided Diagnosis System”, American Journal of Roentgenology, 182, February 2004, pp. 505-510.
[7] Dr.S.A.PATIL,M. B. Kuchanur, ” Lung Cancer Classification Using Image Processing,” International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 3, September 2012.
[8] B. Zhao, G. Gamsu, M. S. 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).