OPEN ACCESS

International
Journal
Of Modern Engineering Research (IJMER)

Neural Network Based Lung Cancer Detection And Prediction

Swapnil K.Kamble[1], Ashish A. Patil[2], Manisha A. Deshmukh[3], Rohini Onkare[4]

Electronics & Telecommunication Department,
Padmabhushan Vasantraodada Patil Institute Of Technology, Budhgaon, Sangli, India.

ABSTRACT:- *Lung cancer is the most common type of cancer among various cancers with the highest molarity rate. In this paper, the survey has been proposed on Medical Image processing schemes for cancer detection. Lung cancer killed many people in recent years. Early diagnosis of lung cancer can help doctors to treat patients and keep them alive. Early diagnosis facilitates identification of treatment phases and increases success rates in treatment. The most common way to detect lung cancer is using computed tomography (CT) images. Uses lung CT images as input to the Computer Aided Diagnosis (CAD) and based on an algorithm helps doctors to perform an image analysis. CAD system provides segmentation of nodules on the lobes with neural networks and ensures the lung cancer tumor and classified it into the stages.*

Keywords:- *Image processing; Computer aided diagnosis;Neural network; Classification in stages.*

I. INTRODUCTION

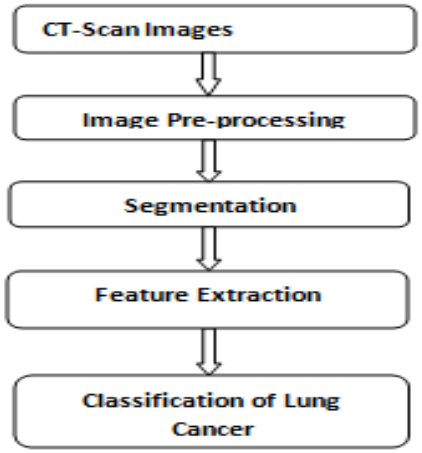
Cancer is a group of diseases characterized by the uncontrolled growth and spread of abnormal cells. If the spread is not controlled, it can result in death. Cancer is caused by both external factors (For examples tobacco, infectious organisms, chemicals, and radiation) and internal factors (inherited mutations, hormones, immune conditions, and mutations that occur from metabolism). These causal factors may act together or in sequence to initiate or promote the development of cancer. Cancer starts when cells in a part of the body start to grow out of control. Cancer cell growth is different from normal cell growth. Instead of dying, cancer cells continue to grow and form new abnormal cells. Cancer cells can also grow into other tissues, something that normal cells can't do. Growing out of control and invading other tissues are what makes a cancer cell. [1] Cells become cancer cells because of changes to their DNA. DNA is in every cell and it directs all its actions. In a normal cell, when DNA is damaged the cell either repairs the damage or dies. In cancer cells, the damaged DNA is not repaired, but the cell doesn't die like it should. Instead, the cell goes on making new cells that the body doesn't need. These new cells all have the same damaged DNA as the first cell does. [2] Cancer cells often travel to other parts of the body where they can grow and form new tumors that crowd out normal tissue. This happens when the cancer cells get into the body's blood cell or lymph vessels. The process of cancer spreading is called metastasis [3].The World Health Organization Estimatesin the India during 2012 nearly 7 lakhs peoples die due to the cancers in every years. While more than 10 lakhs people newly diagnose. In 2012 year 4.77 lakhs men's & 5.37 lakhs women's were diagnosed with the cancer In India.

II. MATERIAL AND METHODS

Computer Aided Diagnosis (CAD System)

The systems that are created by the integration of computer & medical science are called CAD system. For diagnosis lung diseases, first time in 1990 CAD system was used. The success of a particular CAD system can be measured in terms of accuracy of diagnosis, speed, and automation level. A schematic diagram of a typical CAD system for lung cancer is shown in Figure. The segmentation of lung tissues on chest images is a pre-processing step in developing the CAD system in order to reduce the search space for lung nodules. Next, detection and segmentation of lung nodules from the available search space are mandatory steps. Lastly, the classification of the detected nodules into stages. Classification of the detected nodules is a major component in CAD systems for detection and diagnosis of lung nodules in CT. In CAD systems for detection (often abbreviated as CADe), a classification component categorizes the nodule candidates identified in the previous step into nodules or no nodules (i.e., normal anatomic structures), whereas a CAD system for diagnosis (often abbreviated as CADx) classifies detected nodules (either by a computer or a radiologist) into stages.Chest X-ray image has been used for detecting lung cancer for a long time. The early detection and diagnosis of pulmonary

nodules in chest X-ray image is performed by radiologists. However, the radiologist's sensitivity for the detection of small pulmonary lesions is limited. In the lung cancer screening, radiologists may miss some cancers during interpretation of CT images. To reduce radiologists' workload, computer-aided detection (CAD) systems are used. Computer-aided diagnosis (CAD) has been proven to be a very effective approach as assistant to radiologists for improving diagnostic accuracy, therefore may be used in the first stage of examination in the near future. The use of a computer-aided detection (CAD) system can provide an effective solution by assisting radiologists in increasing the scanning efficiency and potentially improving nodule detection. The designed CAD system provides segmentation of nodules on the lobes with neural networks model of Self-Organizing Maps (SOM) and ensures classification between benign and malignant nodules with the help of ANN (Artificial Neural Network). The objective of this study is identifying all nodules from the chest CT lung images and classifying these nodules into cancerous (Malignant) and non-cancerous (Benign) nodules, to reduce the false positive rate using Image processing techniques and Neural Network techniques. Improved method of segmentation based on soft computing will be used for lung cancer detection.



A. **Image Pre-Processing**

By image pre-processing we will improve the quality of data through the application of methods for diagnosis. Filters such as Median filter or Laplacian filter or Gaussian filter will be observed & best suited will be used. Image standardization, i.e. gray scale transformation can be used if necessary.

1 .Image Standardization

It is nothing but improvement of an image appearance by increasing dominance of some features or by decreasing ambiguity between different regions of the image. Image enhancement processes consist of a collection of techniques that seek to improve the visual appearance of an image or to convert the image to a form better suited for analysis by a human or machine. Histogram processing is the act of altering an image by modifying its histogram. Common uses of histogram processing include normalization by which one makes the histogram of an image as flat as possible. This is also known as contrast enhancement. Intensity transformation functions based on information extracted from image such as enhancement, compression, segmentation and description.

The Histogram of digital image with the intensity levels in the range **[0, L-1]** is a discrete function.
h(rk) = nk
Where
• ***rk is*** the intensity value.
• ***nk*** *is* the number of pixels in the image with intensity ***rk***.
• ***h(rk)*** is the histogram of the digital image with Gray Level.

2. Segmentation ofthe Region of Interest (ROI)

We will divide the image into regions of similar attributes by using improved segmentation methods. Most of the segmentation methods are adhoc.
In proposed work different soft computing segmentation methods may be used such as SOM or K means clustering, etc. for better improvement.

3. Feature Extraction

Feature extraction is the important step in lung cancer detection. In the feature extraction & selection step, the features of ROI were extracted to differentiate benign/malignant tumors in lung CT images. The differentiation of tumors can be done by the help of statistical & shape of the tumors. Malignant nodules are more complex & irregular in shape as compared to the benign nodules, whereas benign are has rounder shape & well defined borders. The malignant nodules show relatively higher variance values. An image feature Extraction stage is 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. Digital Image Processing

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. Nowadays, image processing is among rapidly growing technologies. It forms core research area within engineering and computer science disciplines too.

Image processing basically includes the following three steps

i. Importing the image via image acquisition tools.
ii. Analysing and manipulating the image.
iii. Output in which result can be altered image or report that is based on image analysis.

There are two types of methods used for image processing namely, analog and digital image processing. Analog image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. Digital image processing techniques help in manipulation of the digital images by using computers. The three general phases that all types of data have to undergo while using digital technique are pre-processing, enhancement, and display, information extraction.

B. **Artificial Neural Network**

The first neural network, "Perceptron", was created in 1956 by Frank Rosenblatt. ANN is collections of mathematical models that emulate the real neural structure of the brain. ANN consist input layer, hidden layer and output layer. The feed-forward neural networks are the simplest type of artificial neural networks devised.

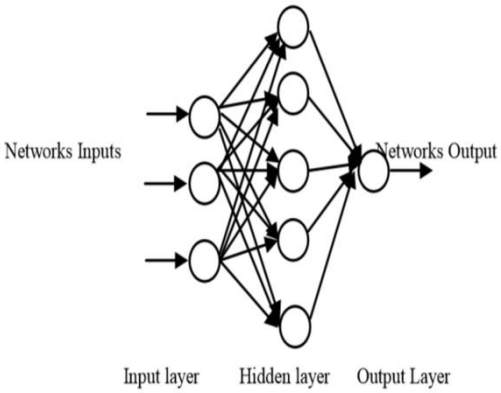


Fig 1. Neural network image

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well.Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyse. This expert can then be used to

provide projections given new situations of interest and answer "what if" questions. Other advantages include:

i. Adaptive learning: An ability to learn how to do tasks based on the data given for training or initial experience.
ii. Self-Organisation: An ANN can create its own organisation or representation of the information it receives during learning time.
iii. Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.
iv. Fault Tolerance via Redundant Information Coding: Partial destruction of a network leads to the corresponding degradation of performance. However, some network capabilities may be retained even with major network damage.

Multi-layer feed-forward perceptron model was used in ANN. Back-propagation algorithm was also utilized to train the network. Levenberg-Marquardt method was used as learning method. In addition, performance of network was calculated according to mean square error (MSC) rule.

III. RESULT

The MATLAB software is used in this paper for the performance evaluation of CAD system. In this study a lung CT scan images are used. In the first step we are obtaining the binary images of gray scale images. Then we are using the median filter to remove the noise and the unwanted contents from the image. Then that image is dilated to obtain region of interest (ROI). Dilation is one of the two basic operators in the area of mathematical morphology. The other being erosion. It is typically applied to binary images, but there are versions that work on gray scale images. The basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels (*i.e.* white pixels, typically). Thus areas of foreground pixels grow in size while holes within those regions become smaller. Here we have used thresholding technique to obtain ROI in BW. Image thresholding is a simple technique, yet effective, way of partitioning an image into a foreground and background. This image analysis technique is a type of image segmentation that isolates objects by converting gray scale image into binary image. Image thresholding is most effective in image with high levels of contrast. There are two algorithms used in thresholding is histogram and multi-level thresholding.

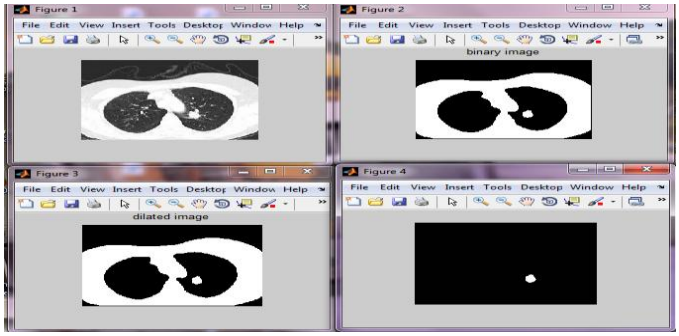


Fig 2.a) Original image b) binary image c) dilated d) extracted cancer nodule.

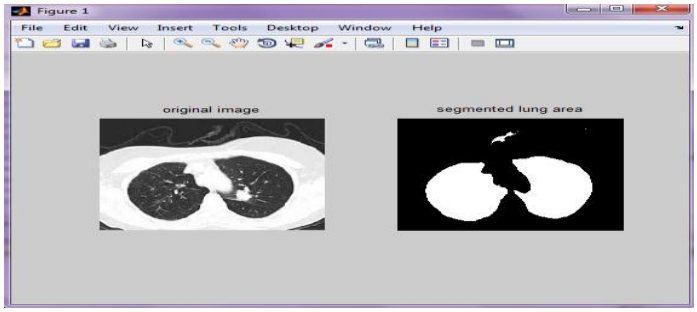


Fig 3. a) Original image b) segmented lung area.

By using the thresholding technique we find the BW area of lung cancer nodule and the BW area of lung. And finally BW area is calculated to obtain proportional area which is used as an input to the ANN. Here we are using the three layers ANN one is input layer, second is hidden layer and third is output layer. We are using two hidden layers five nodes in each. ANN was first trained by the training data and then tested with the data. ANN was compares the input value with the training data and classified into the stages at the output.

IV. CONCLUSION

The CAD system is an integrated structure since it includes pre-processing, segmentation and feature extraction. The CAD system allows successful detection of lung cancer in early stages. To obtain more accurate results we divided our work into three stages: Image Enhancement stage, Image Segmentation stage and Features Extraction stage. Lung cancer is the most dangerous disease in the world in the process of detection the lung cancer this disease plays a very important role to avoid serious stages and to reduce its growth in the world. Neural network model is a diagnostic system that performs at an accuracy level is Constructed. The prediction could help doctor to plan for a better medication and provide the patient with early diagnosis.

REFERENCES

[1]. Artificial Neural Network-Based Classification System for Lung Nodules on CT Scans Images. (IEEE 2014)
[2]. Feature Extraction and Classification of Lung Cancer Nodule using Image Processing techniques (IJERT-2013).
[3]. Lung cancer detection from CT images using image processing technique(Available online at www.ijkrcsms.com
[4]. Artificial Neural Network-Based Classification System for Lung Nodules on CT Scans Images. (IEEE 2014)
[5]. Feature Extraction and Classification of Lung Cancer Nodule using Image Processing techniques (IJERT-2013).
[6]. Lung cancer detection from CT images using image processing (IEEE 2012)
[7]. Disha Sharma, Gagandeep Jindal, "Identifying Lung Cancer Using Image Processing Techniques", International Conference on Computational Techniques and Artificial Intelligence (ICCTAI'2011), vol: 17, pp: 872-880,2011
[8]. Non-Small Cell Lung Cancer, Available at: http://www.katemacintyre. foundation.org/pdf/non-small-cell.pdf, Adapted from National Cancer Institute (NCI) and Patients Livingwith Cancer (PLWC), 2007, (accessed July 2011).
[9]. Nunes É.D.O., Pérez M.G., Medical Image Segmentation by Multilevel ThresholdingBased on Histogram Difference, presented at 17th International Conference on Systems, Signals and Image Processing, 2010.