LUNG NODULE DETECTION SYSTEM
Kalim Bhandare and Rupali Nikhare
Department of Computer Engineering, Pillai Institute of Technology, New Panvel, Navi Mumbai, India

ABSTRACT:
The Existing approach consist of CAD scheme with Virtual Dual Energy. CXRs where ribs and clavicles are suppressed with massive-training artificial neural networks (MTANNs) To reduce rib-induced FPs and detect nodules overlapping with ribs, we incorporated the VDE technology in our CADe scheme. The VDE technology suppressed rib and clavicle opacities in CXRs while maintaining soft-tissue opacity by use of the MTANN technique that had been trained with real dual-energy imaging. A nonlinear support vector classifier was employed for classification of the nodule candidates. The use of VDE technology, the sensitivity and specificity of the CADe scheme for detection of nodules and especially subtle nodules.

The proposed approach concentrates on detecting nodules, early stages of cancer diseases, appearing n patients lung. Most of the nodules can be observed after carefully selection of parameters. This scheme shows an efficient lung nodule segmentation through thresholding and watershed segmentation. The proposed system also uses the VDE technology suppressed rib and clavicle opacities in CXRs while maintaining soft-tissue opacity by use of the MTANN technique that had been trained with real dual-energy imaging. But instead of support vector machine the classifier used here is Artificial neural network i.e Gradient Descent and Back propagation algorithms. This scheme also gives the stages of the cancer based on the features extracted of the nodules. So depending on the stages the radiologist or the physician can identify and give proper treatment.

Keywords: Nodule, VDE, MTANN.

[1] INTRODUCTION
Lung cancer is a one of the major cause of death in the world. The main primary types are (SCLC) small-cell lung cancer and (NSCLC) non-small-cell lung cancer. Lung cancer is a stage which consists of unmanageable growth of tissues and cell. If left undiagnosed, this growth can spread beyond the lung by process of metastasis into nearby tissue or other parts of the body. The common diagnosis process includes surgery, chemotherapy and radiotherapy. Overall, 15% people treated with lung cancer can survive five year after the diagnosis.

Lung Cancer Types
Cancer that starts from lung is called primary lung cancer. There are different types and these are divided into two major groups

Small cell lung cancer
About 20 out of every 100 lung cancers patients are diagnosed of this type. Small cell lung cancer
LUNG NODULE DETECTION SYSTEM

is called so because these cancer cells are very small cells that are mostly filled with the nucleus. This type of cancer is usually caused by smoking. It is very rare for someone who has never smoked to develop it. Small cell lung cancer normally spreads quite early on and so doctors often suggest treatment like chemotherapy treatment rather than surgery. There are mainly three types of non small cell lung cancer. These are clubbed together because they behave in a similar manner and respond to diagnosis in a different manner to small cell lung cancer. The early detection of lung cancer can increase in all 5-year survival rates from 14 to 49% for cancer patients. Therefore, in image processing computer-aided diagnosis (CAD) systems are used to detect the presence of lung cancer cell in a CT-Scan images of patients. The typical CAD system for lung cancer diagnosis mainly include four steps: segmentation of the lung, detection of nodules, segmentation of nodules and diagnosis. The main aim of this study is to identify the various nodules detection techniques that are used by present existing CAD system for lung nodule detection at starting stage. As well, we also outline the advantages and disadvantages of the existing system. The main aim of CAD system is to suppress ribs and calvies to detect nodules by reducing false positive ratio caused by ribs. Our aim was to develop a CADe scheme with better sensitivity and specificity by means of "virtual dual-energy" (VDE) CXRs where ribs are suppressed with a massive-training artificial neural network (MTANN).

2. PROBLEM STATEMENT
The main aim for nodule detection in Chest Radiography is to detect nodules that overlap with ribs and calvies and to reduce the false positive ratio. The purpose is to develop the CAD scheme with increased sensitivity and specificity by use of Virtual Dual Energy and suppression is done by using MTANN (Massive Training Artificial Neural Network). Input given is the normal Chest Radiography and the Expected output is whether the nodule is detected. The nodule detection is done by using three different classifiers, and it also shows which classifier gives the best result i.e. the comparison is done between three classifiers after detection of the nodule based on the region of the nodule one can identify whether the nodule is cancerous nodule or non-cancerous the stages of the cancer is detected.

3. METHODOLOGY
The main aim of lung nodule detection system is to develop the CAD scheme with improved sensitivity and specificity by use of Virtual Dual Energy and suppression is done by using MTANN (Massive Training Artificial Neural Network). In this first single lung image is taken as input which is pre-processed by applying filter. LCDS system uses convolution filters with Gaussian pulse to smooth the cell images. The contrast and color of the images are enhanced. Then the nucleuses in the images are segmented by thresholding. All of those are simple digital image processing techniques. After that, LCDS utilizes morphologic and colorimetric techniques to extract features from the images of the nucleuses. The extracted morphologic features include the perimeter, area, roundness, and rectangleness of the nucleus. The extracted colorimetric features include the red component, green component, blue component, illumination, saturation, difference between red and blue components, and proportion of blue component of the nucleus.

Also the red component, green component, and blue component of the entire image are included as colorimetric features. On this basis, a lung cancer cell identification module is employed to analyze those features to judge whether cancer cells exist in the specimens or not. Moreover, if there are cancer cells, the cancer cell type is identified.
3.1 Input for the system

Single x-ray image is given as input to the system, to check whether given input image is affected by nodule or normal lung image. To check the region of the nodule. Based on the region stages are identified.

3.2 Preprocessing
3.2.1 Smoothing

It suppresses the noise or other small fluctuations in the image equivalent to the suppressions of high frequencies in the frequency domain. Smoothing also blurs all sharp edges that bear important information about the image. To remove the noise from the images, median filtering is used. Median filtering is a non-linear operation often used in image processing to reduce salt and pepper noise. In general, the median filter allows a great deal of high spatial frequency detail to pass while remaining very effective at removing noise on images where less than half of the pixels in a smoothing neighborhood have been affected.

![Figure 2. Smoothing (a) Original Image (b) Median Filtered Image](image)

3.2.2 Thresholding

The CT scan can be separated into two types of voxels, characterized by the density differences between the two anatomical structures. The high-density regions primarily consist of the body surrounding the lung cavity, whereas the low-density regions contain the lung cavity, the air surrounding the body, and other low-intensity regions. To extract the lung volume, we need to segment the low-density regions in the initial stage. For lung image segmentation, when using a histogram method, a fixed threshold value is needed in order to separate the low-density lung parenchyma from the surrounding lung anatomy, though the availability of different scanning protocols makes the selection of an appropriate threshold a challenging task. The appropriate threshold is selected using global thresholding algorithm.

3.2.3 Segmentation

Segmentation on Original Image: Image segmentation is the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify or change the presentation of an image into something that is more meaningful and easier to analyze. Segmentation divides the image into its constituent regions or objects. The result of image segmentation is a set of segments that collectively cover the entire image or a set of contours extracted from the image.

3.3 Creation of Virtual Dual Energy Image

MTANN is a highly non-linear filter that can be trained by use of input CXRS and corresponding teaching images. Bone image obtained by use of a dual energy radiography system were used as the teaching images. VDE images used to suppressed ribs and clavicles by use of MTANN technique.
3.4 Feature Extraction

This stage is an important stage that uses techniques to detect and isolate various desired portions or shapes of a given image. When the input data to a technique is too large to be processed and it is suspected to be notoriously redundant, then the input data will be transformed into a reduced representation set of features. The basic characters of feature are area, perimeter and eccentricity. These are measured in scalar.

3.5 Classification

Classification using Artificial Neural Network (ANN), Support vector machine (SVM), Kohonen self-organizing map. Once feature is extracted from the feature extraction module the output is given to the classifier as the input. Here the ANN’s Gradient Descent and back propagation algorithm is used for classification. It is designed to test the effectiveness of the input features so as to discriminate the lung nodule images.

3.5.1 Algorithm for ANN (Gradient descent)

Set of images sample consisting of nodules and non-nodules are used for training and testing.

1. Here three layered gradient descent artificial neural network is used.
2. It has first layer as input layer consisting of 62 features as input.
3. One hidden layer with 62 nodes (h1 to h62).
4. Two output nodes in the output layer as nodule detected or nodule not detected.
5. Back propagation is used for training the network.

Training pattern presented to the input layer of the network, error at the nodes in the output layer of the network is estimated.

6. Back propagation refers to the propagation of error of the nodes from the output layer to the nodes in the hidden layers.

3.5.2 Algorithm for SVM (Support Vector Machine)
Support vector machine is a machine learning tool which is based on the data classification idea. Classification is performed by constructing n-dimentional hyper plane which separates data in two categories. The data separation can be either linear or non-linear. for the classification of nodule the non-linear SVM with Gaussian kernel is used. The classifier is used because of its generalization ability is high with small number of training samples the SVM classifier is trained/tested with a leave-one-out cross validation test.

3.5.3 Algorithm for KSOM
KSOM is an artificial neural network paradigm KSOM is being developed without the need of the teaching image. This process is called as self organizing feature map, because it perform mapping of an external signal space into the system infernal representation without human intervention it consist of three layers first input layer, competitive layer and output similarity are mapped into closeness relationship on the competitive layer and we compete single winning unit. It has two phases training phase and recall phase.

3.2.7 Stages
Staging involves evaluation of a cancer’s size and its penetration into surrounding tissue as well as the presence or absence of metastases in the lymph nodes or other organs. Staging is important for determining how a particular cancer should be treated, since lung-cancer therapies are geared toward specific stages. If region of nodule is less than 7000 pixel then it is termed as stage I, that means it is non-cancerous but have to take all precautions otherwise it may led to early detection of cancer, else if region of nodule is greater than 7000 it is termed as stage II that means it is cancerous.

<table>
<thead>
<tr>
<th>Stages</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1</td>
<td>Area&lt;7000 Pixels</td>
</tr>
<tr>
<td>Stage 2</td>
<td>Area&gt;7000 Pixels</td>
</tr>
</tbody>
</table>

CONCLUSION
In our proposed approach the computerized scheme for detection of lung nodules by incorporating VDE image in which ribs and clavicles were suppressed by an MTANN technique. After the pre-processing stage, which consists of smoothing, enhancement, thresholding, then segmentation is applied on the original chest radiography. Creation of VDE is done on the threshold image in which MTANN is used. Than on this VDE created image segmentation is applied that is watershed algorithm. Once segmentation of original image and VDE created image is done that both image is combined for positioning the
nodule. After that feature extraction is done i.e. we get the geometrical shape of the nodule like area, perimeter, diameter, irregularity index.

We can implement this technique on some more images increasing the number of images used for the processes can improve the accuracy also MRI PET images can be considered this technique. comparison can be done for all this images so, one can justify which type of images gives better result.

REFERENCES


