International Journal of Combined Research & Development (IJCRD)
eISSN:2321-225X;pISSN:2321-2241 Volume: 5; Issue: 4; April -2016

Lung Cancer Detection on Chest Radiographs Using Image Processing

Kavita [1], Dr.Channappa Bhyri[2], Dr.Kalpana Vanjerkhede[3]
[1]PG Student, Department of Electronics and Instrumentation Engineering,
[2,3]Professor, Department of Electronics and Instrumentation Engineering,
Poojya Doddappa Appa College of Engineering, Kalaburagi, Karnataka, India,
[1]kgkstar.in@gmail.com, [2]channubhyri@yahoo.com, [3]hodit@pdaengg.

Abstract— **Among all types of cancer, lung cancer is considered as one of the most common cause of death throughout the world. The earlier detection of lung cancer is a challenging problem due to structure of cells, where most cells are overlapped each other. In our project we are developing CADe(Computer Aided Detection) scheme. The purpose of developing CADe scheme is to detect the lung nodule by VDE (virtual dual energy) technique, where ribs and clavicles are suppressed with SVM(support vector machine)classifier. The results of the work show that the VDE technique is going to improve the quality of the image. Thus we are aiming to get the more accurate results by using CADe scheme.**

Keywords— **CADe(computer-aided detection), VDE (virtual dual energy), Lung nodule.**

I. INTRODUCTION

The goal of our project is to present a critical review of major CADe(Computer-Aided Detection) systems for lung cancer in order to identify challenges for future research. Lung cancer occurs when a malignant (cancerous) tumour grows inside the lungs, in structures such as the bronchi (small tubes that connect the windpipe to the inner surfaces of the lungs where gas transfer takes place). Like many other types of cancer, lung cancer is capable of spreading (metastasizing) to other parts of the body. For detection of lung cancer at an early stage, chest radiographs (CXRs) are used far more commonly for chest diseases because they are the most cost-effective, routinely available, and dose-effective diagnostic tool. Because CXRs are so widely used, improvements in the detection of lung nodules in CXRs could have a significant impact on early detection of lung cancer. Initially imaging was done with help of X-ray in which X-rays having different energy levels that exposure to a patient's body but which may effects the other soft tissue. For this dual energy technique was used in that two X-rays were merged together and form a single energy level. This dual energy technique is now used in software also that is called VDE (Virtual Dual Energy). In this project we are detecting nodules candidates based on morphologic filtering technique. A support vector classifier is employed for classification of the nodule candidates. In this study, MATLAB tool will be used throughout every phase of the operation.

Lung cancer is a disease of multiplying and growing of abnormal cells into a part commonly called as tumour. Cancer cells can be carried away from the lungs in blood or lymph fluid that surrounds by lung tissue. Metastasis refers to cancer spreading beyond its site of origin to other parts of the body. Type of Lung Cancer that starts in the lung is called primary lung cancer.

Smoking is the most common type of cause due to which lung cancer occurs; the survey give the statistics that more than 80% of cases are result of this. The reason for this is as that cigarette smoke contains many chemicals which affect our body's way of filtering the air in the lungs. Smoke not only irritates lungs but also produce the mucus in large amount which can paralyses the cilia an hair like tiny structure whose purpose is to clean the lungs from dust and dirt. So this leads to accumulation of the toxic dust or dirt due to smoke to blockage and congestion the lungs as paralysis of lung functioning. This kind of symptoms is generally to this extra mucus means chronic bronchitis also called with the name of 'smoker's cough'. Asthma is also one of the major disease causes due to this cigarette smoke like as discussed above about the congestion air passage and filtering.

II. RELATED WORK

All researchers have aim to develop such a system which predict and detect the cancer in its early stages. Also tried to improve the accuracy of the Early Prediction and Detection system by pre-processing, segmentation feature extraction and classification techniques of extracted database. The major contributions of the research are summarized below.

Jun WEI, Yoshihiro Hagihara, et al.[1] have presented the performance of a computer-aided diagnosis system depends on the feature set used in it. This paper shows the results of image feature selection experiments. They evaluated 210 features to look for the optimum feature set. For the purpose, a forward stepwise selection approach was employed. The area under the receiver operating characteristic (ROC) curve was adopted to evaluate the performance of each feature set. Analysis of the optimally selected feature set is given and the experiments using 247 chest x-ray images. Ada, Rajneet Kaur et al. [2] they have used Histogram Equalization and is used for pre-processing of the images and feature extraction process and neural network classifier to check the state of a patient in its early stage whether it is normal or abnormal. After that they predict the survival rate of a patient by extracted features. Experimental analysis is made with dataset to evaluate the performance of the different classifiers. The performance is based on the correct and incorrect classification of the classifier. All experiments are conducted in WEKA data mining tool. K. Suzuki,et al. [3] they developed an image-processing technique for suppressing the contrast of ribs and clavicles in chest radiographs by means of a multi resolution massive training artificial neural network (MTANN). An MTANN is a highly nonlinear filter that can be trained by use of input chest radiographs and the corresponding "teaching" images. They employed "bone" images obtained by use of a dual-energy subtraction technique as the teaching images. They have used a validation test database consisting of 118 chest radiographs with pulmonary nodules and an independent test database consisting of 136 digitized screen-film chest radiographs with 136 solitary pulmonary nodules collected from 14 medical institutions in this study. Dasu Vaman Ravi Prasad [4] have developed image quality and accuracy is the core factors of this research, image quality assessment as well as improvement are depending on the enhancement stage where low pre-processing techniques is used based on Gabor filter within Gaussian rules. Relying on general features, a normality comparison is made. In this research, the main detected features for accurate images comparison are pixels percentage and mask labeling.

III.SYSTEM ARCHITECTURE

CADe systems for detecting lung nodules are usually consists of two methods one is direct segmentation of lung nodule, this method can be preferred when ribs and clavicles are less. If ribs and clavicles are more then we will go through another method called creation of VDE, in that we can suppress the ribs and clavicles permanently but it takes more time and we can also be preferred direct segmentation but it does not suppress the ribs and clavicles permanently so there is chance of getting wrong output.

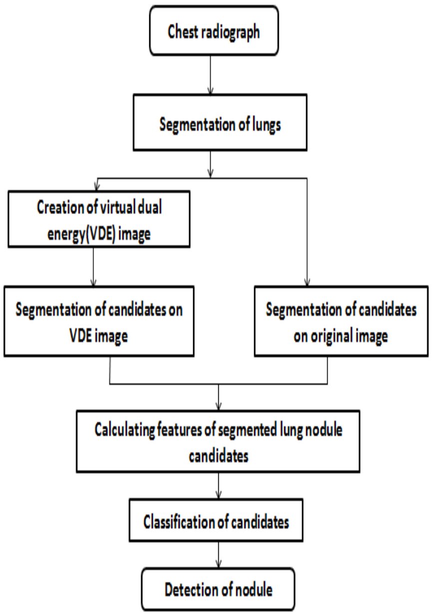


Fig.1 Block diagram for our CAD scheme

Firstly, we acquired the chest radiograph. Further we had made a segmentation of that image, it is necessary when ribs and clavicles were less, if those are more then we will go for creation of VDE. Segmentation and VDE were done by using morphological operation method and this creation of VDE can also improve the quality of data. Using histogram and threshold value we can done segmentation process. After segmentation we can detect the nodule candidates by using gradient edge detection method. Further that segmented candidate will be given for feature extraction by using GLCM (gray level co-occurrence) method then directly we will do the classification of candidates by using ANN (Artificial neural network). Finally we will get an output

IV.SYSTEM MODULE

This system composed of five different modules. For users of CADe systems, it is important to have a basic understanding of these modules in order to understand its operation.

- Segmentation of lungs and creation of VDE (virtual dual energy) by using morphological operation method.
- Detection of nodule candidate on VDE image by using gradient edge detection method.
- Segmentation of lung nodule candidates on Original & VDE image by using Otsu's method.
- Feature extraction from VDE & original image by using bounding box algorithm and GLCM (gray level co-occurrence) method.

5.1 Segmentation of lungs

In this module, firstly we take an input image that is chest radiograph and that image will be converted into an RGB to gray the respective histogram will displayed. To clear the boundaries in the binary image and subtracts from original binary image. Again clear/erase the boundaries pixels from the original gray scale image. Apply thresholding to isolate the lung tissue from the gray scale image to binary image. Extract the two largest blobs from the gray scale image and eliminate the unwanted blob to get two largest regions.

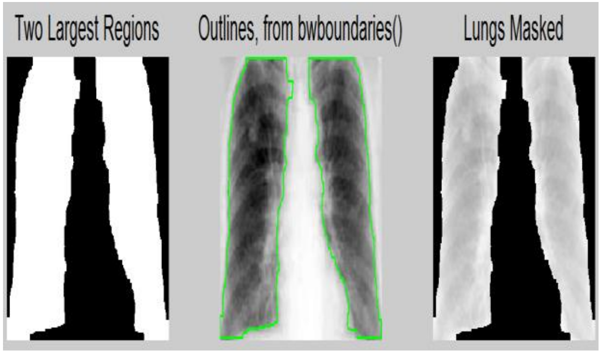
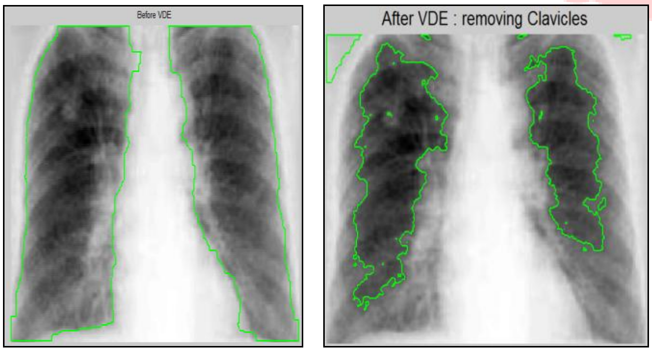


Fig.2 Segmented image

5.2 Creation of VDE

In this module, the performance of VDE will be analysed. After segmentation, mark the clavicle and non clavicle points on that segmented image to get the pixel values at the marked points and prepare feature vector and label vector. SVM(support vector machine) is a highly nonlinear filter that can be trained by use of input CXRs. Apply label matrix on segmented lung image to get an VDE image.



(a) (b)

Fig. 3: (a) VDE image. (b)After removing ribs and clavicles

5.3 Detection of nodule candidate on VDE image

In this module, Sobel method helps to discover the edges in an image; it does so by seeing the image gradient. Image gradient is the change in the intensity of the image. The intensity of the image will be of maximum value where there is a separation of two dissimilar regions thus an edge must exist there. The gradient will be greater where the intensity value is very large. The Sobel operator uses this greatest value to detect edges in an image The Sobel method finds edges using the Sobel approximation to the derivative. It returns edges at those points where the gradient of I is maximized.

5.4 Segmentation of lung nodule candidates on Original & VDE image

Binary image is obtained by Otsu method. To avoid the effect of borders on image processing, we have to remove borders. An opening morphological operation is used to clear borders of image to isolate lung area from borders of the image. After that, connected components of this binary image are extracted, and then two largest components with most area are selected as initial estimation of lungs then the convex hull is used as a mask to apply Otsu method on enhanced image. At the end, a morphological dilation is applied to reduce False Negative (FN) area with introduced structuring element is applied.

Some nodules had similar characteristics to those of bones in terms of the shape, the size, the contrast, and the orientation. The features of these nodules may be suppressed in VDE image. Detected nodules in the nodule candidate detection step may be misclassified as non-nodules based on the features in the original image or VDE image alone.

5.5 Feature extraction from VDE & original image

Finally, shape, gray-level, texture, and specific FP features were extracted from nodule candidates. A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial Dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM and then extracting statistical measures from this matrix. The number of gray levels in the image determines the size of the GLCM. The gray-level co-occurrence matrix can reveal certain properties about the spatial distribution of the gray levels in the texture image.

5.6 Classification of Candidates by using neural network

Finally, shape, gray-level, texture, and specific FP features were extracted from nodule candidates. A nonlinear ANN was employed for classification of the nodule candidates. We selected this classifier because its generalization ability is relatively high with a small number of training samples. The ANN classifier was trained/tested with a leave-one-out cross-validation test.

V. RESULTS

In this section, we present some experimental results to demonstrate the performance of the CADe scheme, which incorporated the VDE technique. We first defined how to train the SVM to create the VDE images for the CADe scheme. Next, the sensitivity for nodules candidate detection on VDE images with different rib contrast was presented and compared to that on the original image. In our CADe schemes, some features extracted and selected for each CADe scheme were given in Table I. From the table, we can see that the selected features for image. Comparing sample 1 and sample 2, the area and perimeter values are less in sample 1 so it concluded that cancerous cells are more in sample 1.

Table 1: Geometrical features of sample 1

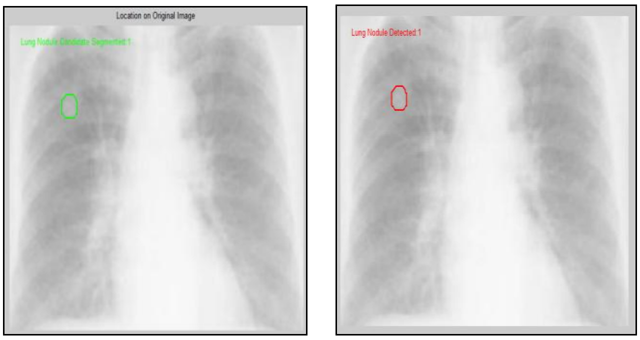
	PARAMETERS	VALUES
SAMPLE 1	Area	528
	Perimeter	81.94

Table 2: Geometrical features of sample 2

	PARAMETERS	VALUES
SAMPLE 2	Area	170
	Perimeter	44.97

Table 3: Experimental results

SL.NO	NAME	RIGHT	LEFT	RESULT
1	X ray	no	yes	Detected
2	X ray1	no	yes	Detected
3	X ray 2	no	yes	Detected
4	X ray 3	no	yes	Not Detected



(a) (b)

Fig.4 (a) Lung nodule candidate segmented. (b) Lung nodule detected

VI. CONCLUSIONS

An image improvement technique is developing for earlier disease detection and treatment stages; the time factor was taken in account to discover the abnormality issues in target images. Image quality, accuracy and timing are the core factors of this research. The proposed technique is efficient for segmentation principles to be a region of interest foundation for feature extraction obtaining. The proposed technique gives very promising results comparing with other used techniques. Relying on general features, a normality comparison is made. The main detected features are area and perimeter.

REFERENCES

1.H. Yoshida, “Local contralateral subtraction based on bilateral symmetry of lung for reduction of false positives in computerized detection of pulmonary nodules,” IEEE Trans. Biomed. Eng., vol. 51, no. 5, pp. 778–789, May 2004.
2. Ada[1], Rajneet Kaur[2] “Early Detection and Prediction of Lung Cancer Survival using Neural Network Classifier”, (IJAIEM)Volume 2, Issue 6, June 2013.
3.Suzuki,J.Shiraishi,H.Abe,H.MacMahon,andK.Doi,“False-positive reduction in computer-aided diagnostic scheme for detecting nodules in chestradiographs by means of massive training artificial neural network,” Acad. Radiol., vol. 12, pp. 191–201, Feb. 2005.
4. Dasu Vaman Ravi Prasad,“Lung cancer detection using image processing techniques”, International journal of latest trends in engineering and technology.(2013)
5. B. van Ginneken, B. M. ter Haar Romeny, and M. A. Viergever, “Computer-aided diagnosis in chest radiography: A survey,” IEEE Trans. Med. Imag., vol. 20, no. 12, pp. 1228–1241, Dec. 2001.
6.J.Wei, Y. Hagihara, A. Shimizu, and H. Kobatake, “Optimal image feature set for detecting lung nodules on chest X-ray images,” in Proc. Comput.Assisted Radiol. Surg., 2002, pp. 706–711. 11
[7]. Suzuki K, Abe H, Li F, Doi K. Proc. SPIE Med. Imag. San Diego, CA: 2004. Suppression of the contrast of ribs in chest radiographs by means of massive training artificial neural network; pp. 1109–1119.
[8]. Chen S, Suzuki K, MacMahon H. Development and evaluation of a computer-aided diagnostic scheme for lung nodule detection in chest radiographs by means of two-stage nodule enhancement with support vector classification. Med. Phys. 2011;38:1844–1858.
[9]. Shiraishi J, Katsuragawa S, Ikezoe J, Matsumoto T, Kobayashi T, Komatsu K, Matsui M, Fujita H, Kodera Y, Doi K. Development of a digital image database for chest

radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. AJR Amer. J. Roentgenol. 2000 Jan;174:71–74.
[10]. Shiraishi J, Li Q, Suzuki K, Engelmann R, Doi K. Computer-aided diagnostic scheme for the detection of lung nodules on chest radiographs: Localized search method based on anatomical classification. Med. Phys. 2006 Jul;33:2642–2653.
[11]. Hardie RC, Rogers SK, Wilson T, Rogers A. Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs. Med. Image Anal. 2008 Jun;12:240–258.
[12]. Cootes TF, Hill A, Taylor CJ, Haslam J. Use of active shape models for locating structures in medical images. Image Vis. Comput. 1994;12:355–365.
[13]. Serra J. Image Analysis and Mathematical Morphology. Vol. 1. Academic; New York: 1982.
[14]. Vincent L, Soille P. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Trans. Pattern Anal. Mach. Intell. 1991 Jun;13(6):583–598.
[15]. Egan JP, Greenberg GZ, Schulman AI. Operating characteristics, signal detectability, and the method of free response. J. Acoust. Soc. Amer. 1961;33:993–1007.