International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 02 Issue: 09 | Dec-2015 www.irjet.net p-ISSN: 2395-0072

Techniques for Lung Cancer Nodule Detection: A Survey

Shraddha G. Kulkarni[1], Sahebrao B. Bagal[2]

[1] *PG student, Department of electronic and telecommunication, Late G. N. Sapkal College of engineering, Anjaneri, Nasik, Maharashtra, India*

[2] *Head: Department of electronic and telecommunication, Late G. N. Sapkal College of engineering, Anjaneri, Nasik, Maharashtra, India*

---***---

Abstract - ***Lung cancer is proving to be a shattering threat to human-beings which is more common in people who used to smoke. Out of 100 different types of cancers observed in human body this is the third largest found cancer with less survival rate. Early detection of lung cancer can increase the chance of survival among people. Various image processing and soft computing techniques can be used to determine cancer cells from medical images. Most commonly CT-images are used for processing because of their high resolution, better clarity, low noise and distortions. This paper focuses on different techniques that have been proposed to provide detection of lung cancer nodules. This survey attempt to summarize techniques such as pre-processing, segmentation, feature extraction, classification used in the lung cancer detection systems.***

Key Words: *Image processing, Soft computing, CT, Segmentation, Feature extraction.*

1. INTRODUCTION

Lung cancer is the type of cancer that begins in the lungs. Lung cancer is considered to be the main cause of cancer death worldwide, and it is difficult to detect in its early stages because 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 any other types of cancer such as breast, colon, and prostate cancers. There is significant evidence indicating that the early detection of lung cancer will decrease mortality rate [1].

Based on the statistics by the American Cancer Society, it is believed that there are 2,20,000 new cases, 1,60,000 deaths per year and the 5-year survival rate for all stages is 15% only[2]. The various factors that influence the 5-year survival rate are stage of cancer, type of cancer, other factors like symptoms, general health etc. However the symptoms of lung cancer do not appear until cancer spreads to other areas, thus leading to 24% chances of lung cancer detection in early stages [3]. So there is need of an accurate early detection of lung cancer system to increase the survival rate [4].

In medical Imaging different types of images are being used, but for the detection of lung diagnosis Computed Tomography (CT) images are being preferred because of better clarity, low noise and less distortion. One more main feature of CT scan images is that it is very easy to calculate the mean and variance of CT scan images [5].

The further sections in this paper are as follows, Section 2 gives generalized structure of lung cancer detection system using medical images that are explained using figure 1, Section 3 includes various surveys on variety of lung cancer detection systems and finally in Section 4 conclusions was made.

2. GENERALIZED LUNG CANCER DETECTION SYSTEM

The generalized structure of any lung cancer detection system using medical images is as shown in figure 1.The system mainly consist of five sections those are listed as below,

1. Image Acquisition
2. Image Pre-processing
3. Image Segmentation
4. Feature Extraction
5. Classification

2.1 Image Acquisition

Image acquisition is the first process. CT images are well preferred as offers visualization of low contrast or small volume nodules by diminishing the slice thickness. Lung CT images can be acquired from publicly available databases namely Early Lung Cancer Action Program (ELCAP) [6], Lung Image Database Consortium (LIDC) [7] or Medical Image Database [8].

2. 2 Image Pre-processing

The aim of this process is an improvement of the image data that suppresses unwanted distortions or enhances some features important for further processing. With CT images, they already contain lesser amount of noise. But for getting more accurate result we will pre-process the image with help of filters [9]. The following two process are used for image pre-processing,

1. Image Smoothing
2. Image Enhancement

Image smoothing basically suppresses the noise or other small fluctuations in the image while enhancement technique is used to improve the interpretability or perception of information in images for human viewers, or to provide better input for other automated image processing techniques [10]. Several technique such as Median filter, Gabor filter, histogram equalization can be used for image pre-processing.

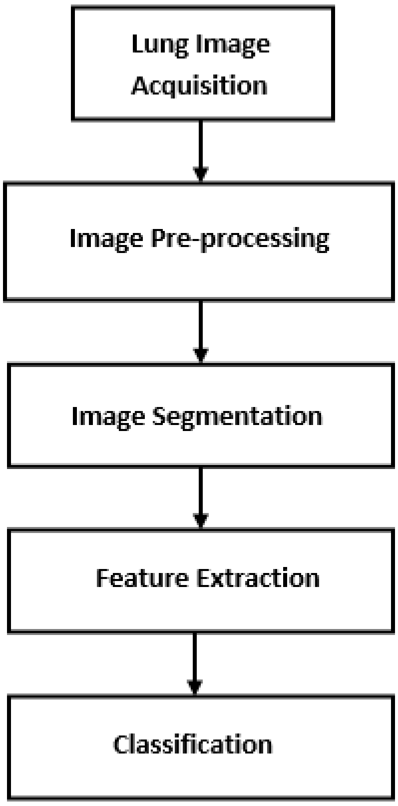

Figure-1 Generalized lung cancer detection system using medical images

2.3 Image Segmentation

Image segmentation is a crucial process for most image analysis consequent tasks. Especially, most of the existing techniques for image description and recognition are highly depend on the segmentation results. 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 [11]. Methods such as thresholding, watershed algorithm, region growing algorithm can be used for segmentation purpose.

2.4 Feature Extraction

This stage is an important stage that uses algorithms and techniques to detect and isolate various desired portions or shapes of a given image. When the input data to an algorithm 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. As the lung cancer tumors are generally spherical in shape, basic characters of feature extraction are area, perimeter and eccentricity [12].

2.5 Classification

After the structure is analyzed, each and every region identified is evaluated individually (scoring) for the probability of a True Positive (TP). Several methods exists for the classification process. Some of them are, rule based methods, minimum distance classifier, cascade classifier, Bayesian classifier, Multilayer perception, Radial Basis Function network (RBF), Support Vector Machine (SVM), Artificial Neural Networks, Fuzzy logic etc.

3. REVIEW ON LUNG CANCER DETECTION SYSTEMS

Fundamentally any lung cancer detection system using medical image processing follow same flow as mentioned in section 2. But use of variety of algorithms and techniques lead to difference in accuracy and results. Such techniques and some of their details are given below,

3.1 Image Smoothing using Median Filtering

A. Kulkarni et al. [13] proposed a system on lung cancer detection using CT images in DCOM format. Image smoothing was done by Median filter. It reduces blurring of edges. The advantage of using median filter in the system is that it is not affected by individual noise spike, eliminates impulsive noise quite well and it does not blur edges much and can be applied iteratively.
Gabor filter is used for enhancement purpose as it gives better result compared to Fast Fourier Transform and auto enhancement. Image presentation based on Gabor function constitutes an excellent local and multi-scale decomposition in terms of logons that are simultaneously localization in space and frequency domain. Marker

controlled Watershed algorithm is used for segmentation purpose. Area, Eccentricity and parameter were extracted features on the basis of which classification was done. For classification purpose supervised learning model, Support vector machine was implemented. Finally depending on classification results the stage of lung cancer is determined. The advantage of this method is that the median filter gives more accurate results compared to Gaussian, mean and wiener filter. The system achieves commendable performance by applying M-C watershed for segmentation and SVM for classification but the disadvantage is that it can only process DICOM format images.

3.2 CLAHE and Fuzzy Clustering Method

K. Punithavathy et al. [14] suggested a methodology for automatic lung cancer detection in PET/CT images. For pre-processing, contrast level adaptive histogram equalization (CLAHE) technique is used along with Wiener filtering. Morphological closing and opening operations are performed for accurate extraction of lung ROI. Feature classification is done using fuzzy clustering method. FCM is simple, unsupervised and soft clustering method and it retains more information of the image compared to hard clustering method.

The advantage of their method is that the Morphological operations enable accurate lung ROI extraction and reduce the search space and the texture analysis yielded number of significant texture features. These features served as input to the FCM classifier helps in accurate detection of the lung cancer. Results of the proposed methodology are promising with an overall accuracy of 92.67%.

3.3 Level Set-Active Contour Model

A. Amutha et al. [15] offered a level set-active contour model with minimizer function for lung tumor diagnosis and segmentation. Active contour model is able to find the accurate boundary of the tumor, whose energy depends on its spatial positioning and shape changes. Along with active contour modeling, level set equations are incorporated for the accession of segment uniformity criterion defined over the given classification.

The method possess active contour modeling along with level set algorithm to amend the performance. The lung image is denoised by Kernel Based Non-Local Neighborhood denoising method with different denoising functions such as exponential function kernel, cosine function kernel, flat kernel, Gaussian, Turkey-bi-weight and wave kernel function, and then processed with the best kernel function. Characteristics of images such as contrast, energy, entropy, variance and homogeneity are considered that paves a way for appropriate classification results.

The classification of lung image is made by the trained neural network based on Bayes Classification known as Multivariate Multinomial Distributed Bayes Classification which categorizes the image under normal and abnormal stages. The main advantage of the system is that it formulates Level set equations in such a way that the dependency of active contour on gradient is minimized. Another advantage is that the implemented segmentation algorithm holds the properties of both level set and active contour methods. Also the combination of two segmentation methods tremendously reduces the computation time and internal energy, and also amends segmentation energy. The location of mass boundaries are well detected and preserved by the system, independent of gradient. Furthermore the methodology can also segment the lung field with pathology of variant forms more precisely.

3.4 Neuro - Fuzzy Classifier

A. Tariq et al. [16] proposed computer-aided diagnosis (CAD) system for automated detection of pulmonary lung nodules in computed tomography (CT) images. The system implemented median filter for denoising and gradient mean and variance based method for extraction of background. The threshold segmentation is done on the basis of optimal thresholding. Sobel gradient operator in the horizontal and vertical directions is implemented for edge detection. Features extraction is done for features such as area, energy, eccentricity, entropy, mean and standard deviation.

For classification purpose, the feature vector is fed to a hybrid classifier based on neural network and fuzzy known as 'nuero fuzzy classifier'. It contains two sub-networks i.e. fuzzy self-organizing network and Multi-Layer Perceptron (MLP) in a cascaded way. In NFC the feature vector is given as input to fuzzy layer to generate pre-classification vector which is given to MLP for classification of test sample. The results of system is accurate and effective which also facilitates the detection of small nodules along with the developed one which lead to early diagnosis of lung cancer.

3.5 Hybrid Multi-layered GMDH type Neural Network

In the study done by T. Kondo et al. [17], the regions of the lung cancer were recognized and extracted automatically by using the revised GMDH-type neural network. Multi-detector row CT (MDCT) images of the lung are used in the study. In the recognition procedure, the revised GMDH-type neural network is organized to recognize the lung regions and then the regions of the lung cancer are extracted.

In this method the revised GMDH-type neural network algorithm using the knowledge base for the medical image diagnosis was proposed and it was applied to the medical image diagnosis of the lung cancer and the results of the revised GMDH-type neural network were compared with those of the conventional sigmoid function neural network trained using the back propagation algorithm. The revised GMDH-type neural network architecture fitting the characteristics of the medical images is organized using the knowledge base for the medical image diagnosis. Furthermore, the neural network architecture is selected from three types of neural network architectures such as the sigmoid function neural network, the radial basis function (RBF) neural network and the polynomial neural network using the knowledge base system.

In GMDH-type neural network structural parameters such as the number of layers, the number of neurons in hidden layers and useful input variables are automatically selected to minimize prediction error criterion defined as PSS. In the case of the conventional neural network, we obtain many different output images for various structural parameters of the neural network and many iterative calculations of the back propagation are needed for various structural parameters in order to find more accurate neural network architecture. It was revealed that the revised GMDH-type neural network algorithm was accurate and a useful method for the medical image diagnosis of the lung cancer.

3.6 CLA (Cellular Learning Automata)

N. Hadavi et al. [9] presented a technique for automatic detection of lung cancer by using cellular learning automata. Image enhancement was performed using Gabor filter. Thresholding technique was used for image segmentation because of its advantages such as fast processing and easy influence. Features are extracted as nodule size, shape, contrast and the region for analysis. The new technique used Cellular automata is a mathematical Model. It is composed of lattice of cells where each cell has a set of stats and local rules governing them. There are three types of cellular automata; one, two and three dimensional cellular automata and the two dimensional CA was used in this study.

Cellular learning Automata (CLA) model is obtained from developing the cellular automata with appending a learning automaton to each cell. CLA model is designed for systems where their components according to experiences of themselves and other components experiences are trained, and CLA have the capability to improve their behavior. In this model cellular automata randomly select one action from their limited possible action vector, when the automata doing the action, environment responds a signal to automata, and with this signal automata upgrade the vector of action. The system used obtained pattern as an environment for learning the CLA. Initially each cell (Learning Automaton) chooses an action from its possible action vector, after that based on selected actions a pattern sends response signal to lattice. Learning automata (LAs) update their action vector. In this algorithm a CLA can be adopted by unknown environment and in next step performs more favorable action. The cellular learning automata model include numbers of automata with interaction, also it works as a multi-agent system and because it has the ability to learn from environment is a self- organizing system. Also if cellular learning automata are well trained, the model is capable of reduced rate of error and enhance the system's reliability. Hence these are the most prominent advantages of this model.

3.7 Binarization and Masking Approaches

B. Patil et al. [18] proposed an approaches to predict the probability of lung cancer presence. First approach is Binarization and the second is masking. The images used in analysis were in standard JPEG format hence for converting them in Grey level 'Otsu's' method was used. Marker-controlled watershed segmentation was used and it gave accuracy of 85.27%. For prediction of lung cancer Binarization approach which is depends on the fact that the number of black pixels is much greater than white pixels in normal lung images is used. So the counting starts the black pixels for normal and abnormal images to get an average that can be used later as a threshold, if the number of the black pixels of a new image is greater than the threshold, then it indicates that the image is normal, otherwise, if the number of the black pixels is less than the threshold, it indicates that the image is abnormal. Another method for prediction is masking which depends on the fact that the masses are appeared as white linked areas inside ROI (lungs), as they increase the percent of cancer presence increase. Out of these two methods Binarization gives better results.

3.8 Marker-Controlled Watershed Algorithm

S. Kanitkar et al. [19] introduced a novel approach for detection of lung cancer using image processing. The Gaussian filter is used to smooth the input image in the preprocessing stage so that it removes high frequency components from the image. As well as, in the pre-processing stage, Gabor filter is used for enhancement and thresholding and Marker-Controlled watershed transform is used for the segmentation purpose. The features such as average intensity, perimeter, area and eccentricity are extracted from the detected tumor. Watershed segmentation is used to extract the region minimum value from an image. It determines the corresponding to the dividing line with the least value. Dividing line in the image gives the rapid change of boundary. This transform finds catchment basins and watershed edge lines in the image. It treats the image as a plane, where light pixels are high and dark pixels are low. The important drawback associated to the watershed transform is the over segmentation that usually results. Hence to overcome the drawbacks of this watershed segmentation i.e. over segmentation, the marker based watershed segmentation technique is used. It can segment boundaries from an image. Morphological operations are performed on the watershed segmented image to get final segmented image. The proposed marker controlled watershed segmentation technique separates the touching objects in the image. It provides best identification of the main edge of the image and also avoids over segmentation. Hence it is efficient for segmentation.

CONCLUSION

In this paper we survey different techniques for lung cancer detection system. The generalized structure of lung cancer detection system using medical images is also described. Out of the various techniques discussed here, we found that Marker Controlled Watershed algorithm is superior over Thresholding in case of segmentation. Also Median filtering, Level Set-Active Contour Model, CLA and Binarization enhances the performance of lung cancer detection system.

REFERENCES

[1] S. S. Parveen, Dr. C. Kavitha, "A Review on computer aided detection and diagnosis of lung cancer nodule", International Journal of Computers & Technology, Volume 3 No. 3, Nov-Dec, 2012.
[2] American Cancer Society (ACS), "Report on Lung Cancer," 2014
[3] A. Jamal, et.al, "Cancer statistics, 2007"CA Cancer J Cln, vol 57, no.1 pp. 43-66, 2007.
[4] K. Kancherla, Srinivas Mukkamala, "Early Lung Cancer Detection using Nucleus Segmentation based Features", IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2013
[5] J. Kaur, Nidhi Garg, Daljeet Kaur, "A survey of lung cancer detection techniques on CT scan images", International Journal of Scientific & Engineering Research, Volume 5, Issue 6, June-2014
[6] Early Lung Cancer Action Program(ELCAP) : http://via.cornell.Edullungdb.html
[7] Lung Image Database Consortium(LIDC): https://imaging.nci.nih.gov/ncia/login.jsf
[8] Medical Image Database: MedPix : http://rad.usuhs.edu/medpix/index.html
[9] N. Hadavi Md. Jan Nordin, Ali Shojaeipour, "Lung Cancer Diagnosis Using CT-Scan Images Based on Cellular Learning Automata" , IEEE International Conference on Computer and Information Sciences (ICCOINS), 3-5 June 2014
[10] A. Kulkarni, A. Panditrao, "Classification of Lung Cancer Stages on CT Scan Images Using Image Processing",IEEE International Conference on Advanced Communication Control and Computing Technologies, 2014.
[11] M. S. Al-Tarawneh, "Lung Cancer Detection Using Image Processing Techniques", Leonardo Electronic Journal of Practices and Technologies, Issue 20, January-June 2012, p. 147-158
[12] V. A. Gajdhane, L.M. Deshpande, " Detection of Lung Cancer Stages on CT scan Images by Using Various Image Processing Techniques", Volume 16, Issue 5, Ver. III (Sep – Oct. 2014), PP 28-35
[13] A. Kulkarni ,A. Panditrao, "Classification of lung cancer stages on CT scan images using image processing", IEEE,International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), 2014
[14] K. Punithavathy, M. Ramya, S. Poobal "Analysis of Statistical Texture Features for Automatic Lung Cancer Detection in PET/CT Images", International Conference on Robotics, Automation, Control and Embedded Systems – RACE 2015, 18-20,February 2015.
[15] A. Amutha, R. Wahidabanu, "Lung Tumor Detection and Diagnosis in CT scan Images", International conference on Communication and Signal Processing, April 3-5, 2013.
[16] A. Tariq, M. Usman Akram, M. Younus Javed, "Lung Nodule Detection in CT Images using Neuro Fuzzy Classifier", Fourth International Workshop on Computational Intelligence in Medical Imaging (CIMI),2013.
[17] Tadashi Kondo, Junji Ueno, Shoichiro Takao, "Medical Image Diagnosis of Lung Cancer by Hybrid Multi-layered GMDH-type Neural Network Using Knowledge Base", International Conference on Complex Medical Engineering (ICME), July I - 4, 2012.
[18] B. Patil, Sanjeev N. Jain, "Cancer Cells Detection Using Digital Image Processing Methods", International Journal of Latest Trends in Engineering and Technology (IJLTET), Vol. 3 Issue 4 March 2014.
[19] S. Kanitkar, N. Thombare , S. Lokhande "Detection of Lung Cancer Using Marker-Controlled Watershed Transform", International Conference on Pervasive Computing (ICPC), 2015.