ROBUST FEATURE ESTIMATION AND RANKING BASED IMAGE SEGMENTATION FOR LUNG CANCER DETECTION USING GABOR FILTERS

K.Sankar[1], *M. Prabakaran*

ABSTRACT

The process of identifying lung cancer from low quality X-Ray images has been discussed widely in medical sector. The problem of false positive results in identifying cancer introduce requirement of more sophisticated method. We propose a robust feature estimation and ranking method for the segmentation of lung images using which LCD can be performed. In the proposed approach the input image is filtered with gabor filter then , the lung region is identified and extracted using template matching techniques. The extracted lung region is sub sampled into number of small images called integral images. For each sub sampled integral image, we identify set of interest points where there is more valued pixel present in the integral image. The feature descriptor for each interest point is computed to represent the features around a point and integral image. The white matter deapthness is estimated using the gray values of the feature descriptor and ranked according to them between various integral images. Finally a top ranked region or integral image is selected and marked as identified location of LC. The proposed method reduces the time complexity and false positive results and produces efficient results.

[1]Research Scholar, Department of Computer Science, Karpagam University, Coimbatore, Tamil Nadu, India

[2]Assistant Professor, Department of Computer Science, Government Arts College, Ariyalur, Tamil Nadu, India

Keywords *- Image Segmentation, Template Matching, LCD, Gabor Filters.*

I. INTRODUCTION

The modern world suffers with various diseases which could be treated or non treated. Among them , the lung cancer which is the most deadly disease , which cannot be treated at the final stage. The lung cancer can be identified at the early stage, using small diagnosis of X-Rays. The X-Rays are low quality images, using which the medical practitioner can identify the presence of cancer at some tissues of lungs. Generally X-rays are gray scale images and will not look clear for visual perception. So that the practitioner cannot identify anything accurately and need some advanced medical support and tools. Also, there are automated systems and tools to support the decision for the medical practitioners. Those medical systems use various techniques to perform lung cancer detection. Still the medical system produces false positive or false negative results , so that there is a necessary for more sophisticated systems.

The segmentation is the process of separating a set of similar pixels to form a group and represent them in different color values than other pixels. The segmentation techniques uses various metrics to separate the pixels from others, for example color values, region properties and etc.. The segmentation techniques have been used for variety of medical problems and produces good results for most cases. Whatever the scans used to diagnose, the medical practitioner could not identify or locate a region exactly where the cancer present or he may ignore the presence of cancer. This is where segmentation plays and performs grouping of pixels and deviate them from other normal pixels to identify lung cancer.

The input image contains noise, blur so that the image has to be noise removed. The Gabor filter is the most efficient linear filter which could be used to remove noise from the image. The noise removed image can be further segmented to achieve the task of segmentation. We use the Gabor filter at different levels to get the quality image from the input image.

The input X-ray image contains noise and blurs, so that it has to be removed to achieve the required results. We compute mass estimation here, where the mass estimation is the process of computing the gray level density value or mass. The region of cancer affected pixels must have more gray levels and the mass of gray co-efficient will be higher in that region. We use this property of pixels to identify the location of lung cancer.

Template matching is a image processing technique where the shape features are trained for future usage. When there is an input image is submitted then the shapes present in the input image will be matched with the other templates trained earlier. Finally a template similar to the input shape will be identified as result.

II. Related Works

There are number of methods has been proposed earlier for lung cancer detection and we discuss few of them around the problem statement.

Methylation analysis in spontaneous sputum for lung cancer diagnosis [1], proposes A novel risk analysis is introduced, using the distinction between diagnostic and risk markers. Two independent sets were randomly composed from a prospectively collected sputum bank (Set 1: n = 98 lung cancer patients, n = 90 controls; Set 2: n = 60 lung cancer patients, n = 445 controls). Sputum cytology was performed for all samples. The following DNA hypermethylation markers were tested in both sets: RASSF1A, APC and cytoglobin (CYGB). Two statistical analyses were conducted: multivariate logistic regression and a risk classification model based on post-test probabilities.

Molecular profiling of small cell lung cancer in a Japanese cohort [2], conducted the present Shizuoka Lung Cancer Mutation Study to analyze genomic aberrations in patients with thoracic malignancies. We collected samples of SCLC from a biobank system and analyzed their molecular profiles. We assessed 23 mutations in nine genes (EGFR, KRAS, BRAF, PIK3CA, NRAS, MEK1, AKT1, PTEN, and HER2) using pyrosequencing plus capillary electrophoresis. We also amplified EGFR, MET, PIK3CA, FGFR1, and FGFR2 using quantitative real-time polymerase chain reaction (PCR) and the fusion genes ALK, ROS1, and RET using reverse transcription PCR.

Automated lung cancer detection by the analysis of glandular cells in sputum cytology images using scale space features [3], develop an automated lung cancer detection system which segments the cell nuclei and classifies the glandular cells from the given sputum cytology image using a novel scale space catastrophe histogram (SSCH) feature. Catastrophe points occur when pairwise annihilation of extrema and saddle happens in scale space. Unusual nuclear texture shows the presence of malignancy in cells, and SSCH-based texture feature extraction from nuclear region is done. From the input high-resolution image, the cellular regions are localized using maximization of determinant of Hessian, nuclei regions are segmented using K-means clustering, and SSCH features are extracted and classified using support vector machine and color thresholding.

Lung Nodule Detection in CT Images Using Thresholding and Morphological Operations [2], the lung CT image is subjected to various processing steps and features are extracted for a set of images. The processing steps include thresholding, morphological operations and feature extraction. By using these steps the nodules are detected and segmented and some features are extracted. The extracted features are tabulated for future classification.

Cell extraction from sputum images for early lung cancer detection [3], address this problem using two different methods, namely, a Rule-based method, and Bayesian classification. We describe the two methods and we compare their performances in terms of their behaviors with respect to color representation and color quantization.

Extraction and Segmentation of Sputum Cells for Lung Cancer Early Diagnosis [11], present here a framework for the extraction and segmentation of sputum cells in sputum images using, respectively, a threshold classifier, a Bayesian classification and mean shift segmentation. Our methods are validated and compared with other competitive techniques via a series of experimentation conducted with a data set of 100 images. The extraction and segmentation results will be used as a base for a CAD system for early detection of lung cancer which will improve the chances of survival for the patient. All the above discussed methods has the problem of false positive results, so that we propose a new robust feature estimation and ranking method for segmentation to identify lung cancer.

III. Proposed Method

The proposed method has three steps namely; Preprocessing, Gray-Co Mass Estimation, Segmentation, LC Detection. We discuss each of the stages deeply in the next chapters.

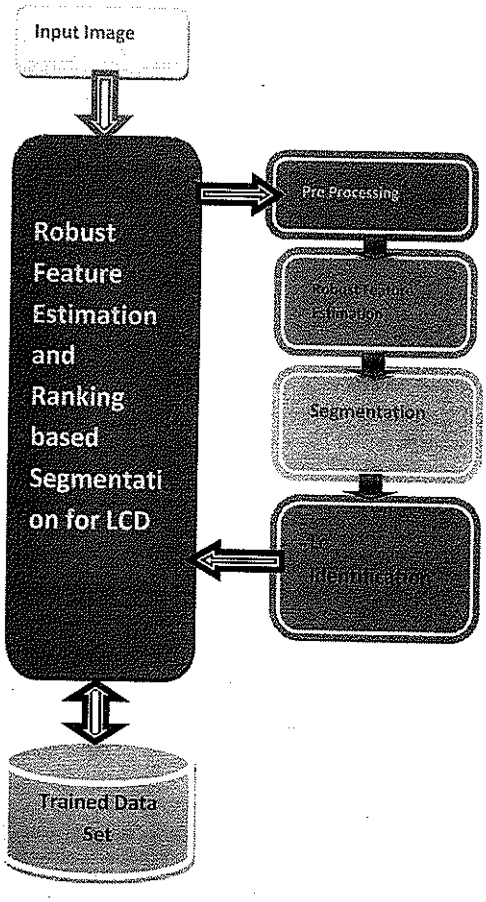


Figure1: Proposed System Architecture

3.1 Preprocessing:

The preprocessing is performed to increase the image quality, which helps the method to produce efficient results. The input image is applied with Gabor filter, which removes the noise present in the input image and the process is iterated for number of times. The Gabor filter is used to enhance the image quality which is a linear filter and used for edge detection. We apply the Gabor Filter with different frequency and orientation to extract the features of the image. This stage removes defects caused by the image acquisition process, for example, noise and lack of contrast. The preprocessed image is used for further processing to identify the lung cancer.

Input: image Img.

Output: Enhanced Image Img.

Step1: Initialize sinusoid factor λ, Orientation O, and phase offset p, standard deviation sd, gamma.

step2: compute sigmaX = gamma.

Step2: compute sigmaY = λ /gamma.

Step4: compute x axis rotaion Xr.

Step5: compute y axis rotation Yr.

Step6: **Ga** = exp(-.5*(Xr.^2/sigmaX^2+Yr.^2/sigmaY^2)).*cos(2*pi/sd*Xr+p);

step7. Stop.

3.2 Robust Feature Estimation:

The preprocessed image is applied with template matching technique to identify and extract the lung region from the input image. The extracted lung region is sub sampled into number of M×N size of integral images. The integral image are constructed in growing box filter and for each integral image constructed we compute the interest point and then we compute the feature descriptor which specifies the white pixel mass value around the interest point. This region growing approach will be performed up to few numbers of sizes. With the constructed feature descriptors, we estimate the gradient mass which specifies the presence of lung cancer in and around the interest point.

Algorithm:

Input : Preprocessed image Img, Template set Ts.

Output: Set of Interest points Ips and Gradient mass set Gms.

step1: For each Template T_i from Ts

Match T_i with regions of Image Img

if $\forall(Region(Img)\exists Ti)$ then

Extract the region from Img as Rlung = Region(Ti) $\exists Img$

end.

end.

Step2: for each size m From M

construct box filter of size m.

crop integral image of size m from Rlung.

IE = $\int_{m}^{m} Rlung(N \times (m \times m))$

for each pixel Pi from IE

compute gradient estimation Ge = $\int_{1}^{N} \frac{\sum NPi}{m}$

N- Total number of pixels present in IE.

m – size of box filter.

if Ge>previous then

keep the index of Pi.

end.

end

Select top valued gradient and pixel.

Select the pixel as interest point IP and compute region mass value Rmv.

$$\text{Rmv} = \int_{i=1}^{k} \frac{\sum Pi(IE)}{m} \times k$$

k- number of white pixels comes into the box filter of size m.

end.

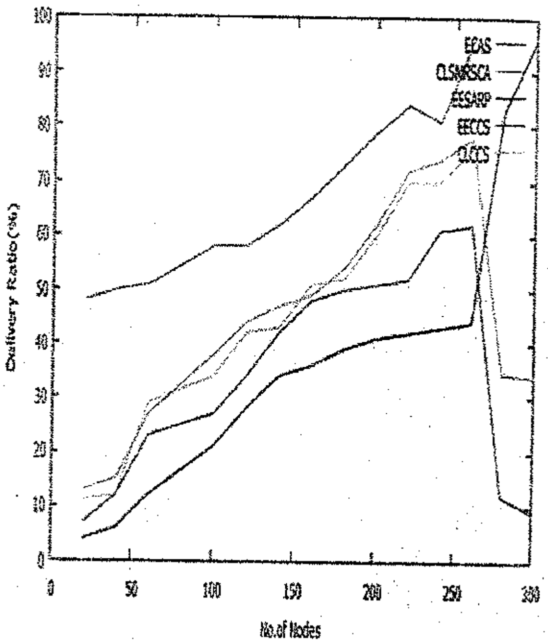
Step3: stop.

3.3 Segmentation:

The segmentation process is performed to differentiate and group similar pixels from other pixels. In our approach the segmentation is performed based on the gray scale values of the pixels. The cancer affected cells has more gradient values than other pixels , so that we represent them in white color and leaves other pixels with the same value what they have. The segmentation process generates number of regions and groups the similar pixels with more mass ratio to form a group.

3.4 Lung Cancer Detection:

The cancer from the image is detected and marked based on computed region mass value. We have computed region mass value for different regions of the image and we select the most region mass value from computed set of mass values. The interest point and window size and the pixels comes into this region will be marked with



different color to represent the cancer detected portion of the image.

Algorithm:

Input: Region Mass value Set RMS, Segmented image Img, Interest point set Ips.

Output: LCD marked image Oimg.

step1: select most maximum Rms value from RMS

Step2: identify the size of window of IE.

Step3: Identify the pixels comes to Ips.

Step4: mark those pixels with different color.

Step5: stop.

IV. RESULTS AND DISCUSSION

The proposed method has been implemented in Matlab and we have evaluated the proposed approach with various data sets. The proposed method has produced efficient results. The proposed Robust Feature estimation and ranking method for lung cancer detection has been evaluated with various data sets. The proposed method has been implemented on Matlab and tested with different 30 percent of images and for training we have used 70 percent of images of data set. The proposed method has produced efficient results and produced less time complexity also.

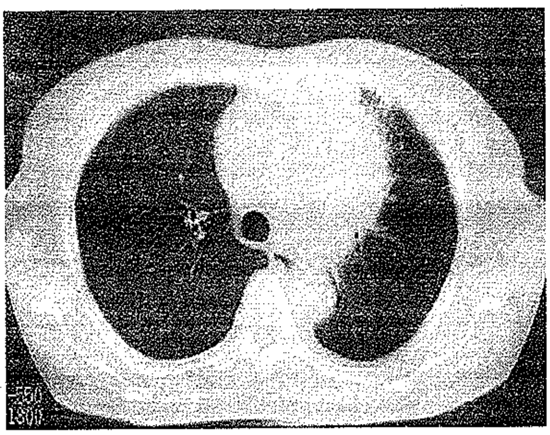


Figure 2: shows the gray scaled input image

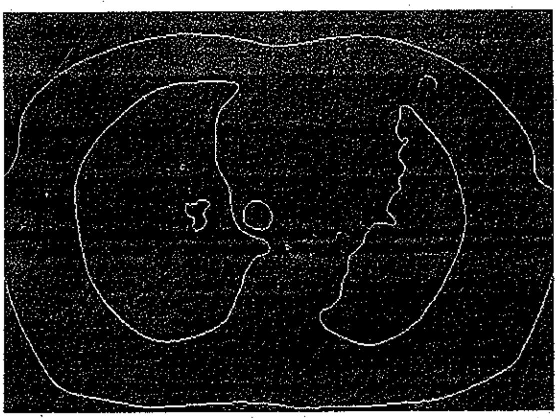


Figure 3: shows the edge detected image for template extraction.

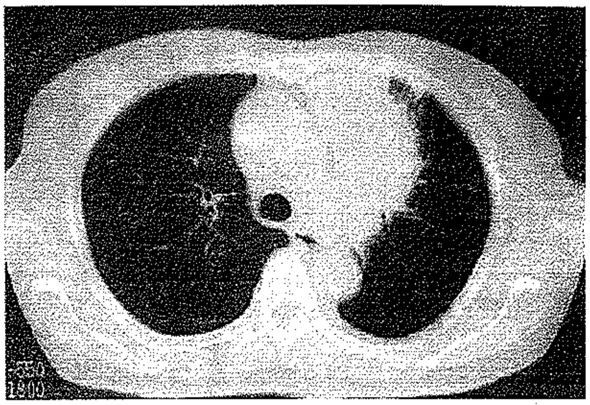
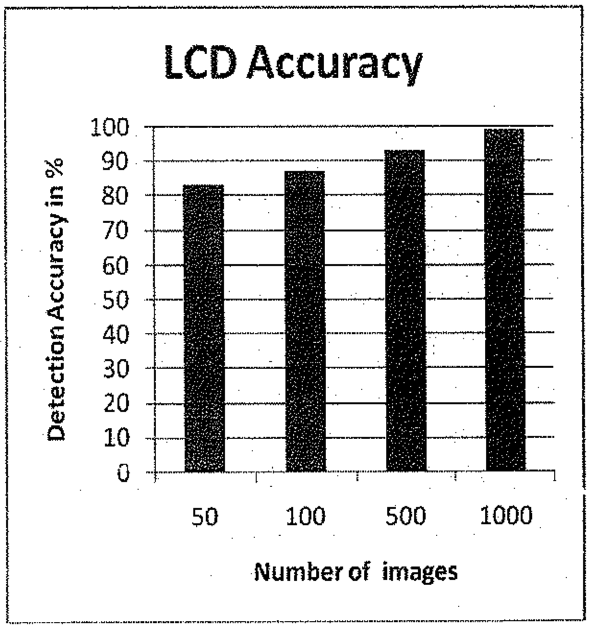


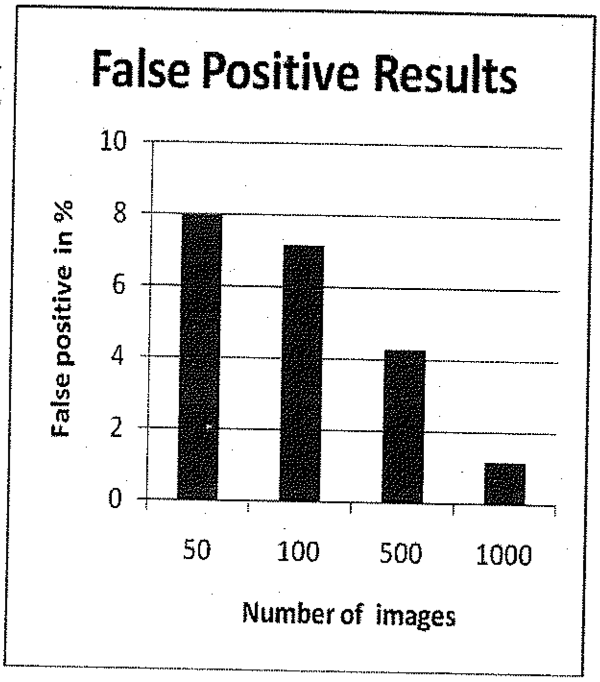
Figure 4: shows the cancer detected image.

Table1: shows the LCD accuracy



The table 1 shows the lung cancer detection accuracy with number of samples and it shows that the proposed method has good impact on LCD with growing size.

Table2: shows false positive ratio



The table2 shows the false positive results produced by the proposed method according to number of training images, and it shows clearly that the proposed method has produced efficient results with more training images.

V. Conclusion

We proposed a Robust Feature Estimation and ranking method for image segmentation to identify lung cancer using Gabor Filters, where each image is converted to gray scale and applied with Gabor filter to enhance the quality of image. The enhanced image is template matched to extract the region of lung from the input image. The extracted region image is converted into many number of sub sampling image called integral image. From generated integral image an interest point is identified which has more gradient mass around them and selected as interest point to compute the region mass value. The computed region mass values are ranked according to the value and the selected region is marked with different color to specify the location of cancer. Finally a region with cancer affected is selected and marked using cancer detection process which is performed according to the mass deviation estimation. The proposed method has produced efficient results and reduced the overall time complexity and false ratio.

References

1. Jasmijn Hubersa, Methylation analysis in spontaneous sputum for lung cancer diagnosis, Elsevier, Volume 84, Issue 2, May 2014, Pages 127–133
2. Kazushige Wakuda, Molecular profiling of small cell lung cancer in a Japanese cohort, Elsevier, Volume 84, Issue 2, May 2014, Pages 139–144.
3. S. Sajith Kecheril, Automated lung cancer detection by the analysis of glandular cells in sputum cytology images using scale space features, Springer, Signal, Image and Video Processing June 2013
4. Nikita Pandey ,A Novel Approach of Cancerous Cells Detection from Lungs CT Scan Images, International Journal of Advanced Research in Computer Science and Software Engineering, vol.2, issue 8, 2012.
5. V.Sudha, Lung Nodule Detection in CT Images Using Thresholding and Morphological Operations , International Journal of Emerging Science and Engineering (IJESE), Volume-1, Issue-2, December 2012.

6. Jiantao Pu, David S. Paik, Xin Meng, Justus E. Roos, and Geoffrey D. Rubin (2011) shape Break-and-Repair Strategy and Its Application to Automated Medical Image Segmentation', IEEE transactions on visualization and computer graphics, vol. 17, no. 1.

7. Panayiotis D. Korfiatis, Texture-Based Identification and Characterization of Interstitial Pneumonia Patterns in Lung Multidetector CT' , IEEE transactions on information technology in biomedicine, vol.14, no. 3, 2010.

8. Donner.C, Cell extraction from sputum images for early lung cancer detection, IEEE electro technical conference, 2012.

9. Antony Judice, A noval assignment of various bioimaging methods for lung tumor detection and treatment using 4-D and 2-D images, International journal of biomedical sciences, vol.9,issue.2,pp.54-60,2013.

10. Javed U, Detection of lung tumor in CE CT images by using weighted Support Vector Machines, International Bhurban conference on applied sciences and technology, pp.113-116, 2013.

11. Fatma Taher, Extraction and Segmentation of Sputum Cells for Lung Cancer Early Diagnosis, International journal of algorithms, 2013.

12. Disha Sharma, Gagandeep Jindal, "Computer Aided Diagnosis System for Detection of Lung Cancer in CT Scan Images", International Journal of Computer and Electrical Engineering, Vol. 3, No. 5, October 2011.

13. Samir Kumar Bandyopadhyay, "Edge Detection from CT Images of Lung", International Journal Of Engineering Science & Advanced Technology Volume - 2, Issue - 1, pg: 34 – 37, Jan-Feb 2012

14. Dr. H. B. Kekre, Ms. Saylee M. Gharge, "Image Segmentation using Extended Edge Operator for Mammographic Images", International Journal on Computer Science and Engineering, Vol. 02, No. 04, 2010.

AUTHOR'S BIOGRAPHY

K. Sankar, is working as Asst.Professor in Department of Computer Science , Shanmuga Industries Arts and Science College, Thiruvannamalai. He has completed both M.Sc., and B.Sc., in Nehru Memorial College, Puthanampatti, Trichirapalli. His area of research includes Digital Image Processing.

Dr. M Prabhakaran, is working as Assistant Professor, Department of Computer Science Government Arts College, Ariyalur, Tamil Nadu. He completed his M.Sc(CS) from Bharathidasan University, ME(CS) & MPhil(CS) from Vinayaka Mission's University and Phd(CS) from Vinayaka Mission's Unversity.He has more than 16 years of experience in his academic. He is supervising a lot of PhD Theses.He extends his guidence for the research activity by gving support for his fellow candidates. His research is focused on Image Processing, Data Mining, and Network Security.