International Journal of Advanced and Innovative Research (2278-7844) / Volume 7 Issue 9

A NOVEL LUNG NODULE DETECTION TECHNIQUE FOR LUNG CANCER IN CT IMAGES

[1]M.Mary Adline Priya, [2]Dr.S.Joseph Jawhar

[1]Research Scholar, [2]Professor

[1]Faculty of Information and Communication Engineering,

Arunachala College of Engineering for Women, Kanya Kumari, India

michaelpriya21@gmail.com

***Abstract*— Lung cancer represents one of the main causes of death among all the possible diseases, being the leading cause of death among all the different cancers. From all the imaging modality, computed tomography is one of the most sensitive and adequate imaging technique to identify malignant nodules due to its image quality, resolution and level of detail. Accurate lung nodule segmentation from computed tomography (CT) images is of great importance for image-driven lung cancer analysis.**

***Keywords*— Feature extraction, computed tomography, lung cancer, clustering.**

I. INTRODUCTION

The accurate investigation and diagnosis is a challenging task in medical field in managing various diseases. Over the years, Computer-Aided Diagnosis (CAD) systems have emerged in order to help the specialists in the lung cancer detection and diagnosis, facilitating their work. In general, the CAD systems try to identify malignant nodules inside the lung parenchyma using either chest radiography (Chen & Suzuki, 2013; Lee, Goo, Park, Lee, & Jin, 2012; Pereira, Fernandes, Mendonça, & Campilho, 2007) or CT imaging (Van Ginneken, 2008). There is a large variability in CAD systems for lung cancer diagnosis that adopt different strategies, that can be organized in five consecutive main tasks: the lung parenchyma extraction, for segmenting the lung region; the lung nodule candidates detection, obtaining a group of structures that are suspicious of being a nodule; the refinement and final lung nodule detection, for reducing the number of false positives from the suspicious nodule candidates set; the segmentation of the nodule region, facilitating further local analysis; and the final benign/malignant differentiation.

II. RELATED WORKS

Accurate segmentation of lung region is important since, the nodules present on it may be on the boundary of the lung parenchyma. So such lung nodules may lost and this reduces the detection accuracy, if the entire lung is not segmented accurately.

Lin DT et al proposed a novel threshold based segmentation approach for segmenting lung region present in the CT lung images. In their work, during preprocessing they have used a 5x5 median filter for removing the noise present in it. The Foreground region is then separated by omitting the rim of the image along with the background regions. In [11] segmentation is done by thresholding each image by an optimal threshold derived by comparing the curvature of the lung boundary along with the ribs. A combination of background removal operator together with iterative gray level thresholding is used by Atonally et al. for segmenting the lung region. In their work, due to the presence of noise the background was not well eliminated well. Pu et al. proposes an adaptive border marching algorithm to segment the lung region and reduces the under segmentation ratio. They used the gray-level thresholding to obtain the lung regions and then follow flood-filling methodology to remove any non-lung regions present after the thresholding. Ozekes et al segmented the lungs of the CT images using Cellular Neural Networks trained by genetic algorithm. In their work, the lung regions were specified using the 8 directional searches and +1 or -1 value were assigned to each vowel. In the work proposed by Cao Lei et al, a rough image of lung was acquired by combining optimal thresholding together with mathematical morphology. A self fit segmentation algorithm was then applied on the segmented result to obtain a final refined output.

III. PROPOSED METHOD

A. Image acquisition

The lung images used in this work are available in the public database LIDC-IRDI. Some principles, such as identity protection, patient consent and lung cancer diagnosis, are defined by the publisher. The database contains 1012 CT scan series and 290 radiography scan series. Each image contains information regarding the nodules, such as location, size, quantity and malignancy. The scans are diagnosed by four specialists and all diagnostics are indicated in the auxiliary files. Some scans contain conflicting information between the image header and the auxiliary files. Therefore, only 833 CT scans were chosen to be part of the candidate database. The images are divided into two databases: nodule and non-nodule candidates. Fig.1 depicts some examples of nodule candidates used in the proposed methodology. Fig. 2 demonstrates examples of non-nodule candidates.

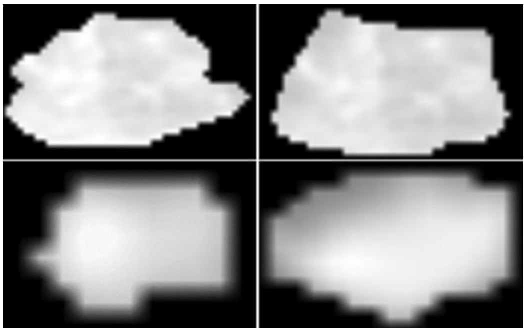

Fig. 1. 2D view of nodule candidates.

Fig. 2. 2D view of non-nodule candidates.

B. Edge detection

The Ant colony optimization algorithm is used for edge detection. The edge detected image is given as the input to black circular neighbourhood algorithm. The assumption is that all the edges are black pixels. Here the edges represent the edges of the lungs along with the nodules edges. As the lung edges are fine lines and nodules are identified in the form of clustered black pixels, black circular neighbourhood algorithm is applied.

C. Black circular neighbourhood algorithm

The black circular neighbourhood algorithm is used to identify the lung nodule pixels. The nodules identified from the CT image are clustered as benign and malignant based on shape, size and intensity of the nodule. The black circular neighbourhood algorithm is used to find the centre pixel of clustered black pixels. Initially, clustered black pixels are identified using the 4-connected and 8-connected properties of the pixels.

The proposed algorithm detects nodules whose diameter is up to 5 units. It detects both spherical and elliptical nodules. From the black pixel clusters, the centre pixel is identified to be the centre of lung nodules. The block diagram of the proposed work is given in Fig.3.

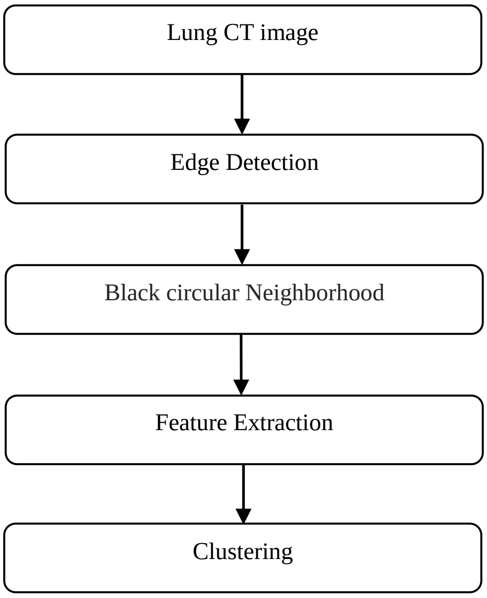

Fig.3. Block diagram of proposed work

D. Feature extraction

The extraction of features occurs in this phase. This is a necessary step for posterior classification. The proposed methodology uses three techniques for feature extraction. The first one is the ACs model, the second is the RD model, and the third combines features from the ACs and the RD. Directional texture (the rose diagram model) exploits the advantages of both the directional and the intensity information in the images (Rivera, Castillo, & Chae, 2015). Analysing that information in a local neighbourhood diminishes the influence of noise in the feature extraction because the gradient works with the edges, regardless of single pixel intensity variation. The use of artificial life concept in image classification resulted in promising results by Zhang and Chen (2004). The ACs method provided high level statistical features that represent consistencies in the textures of the images. Both techniques utilize different types of analyses inside the texture. If used together, they may reduce weaknesses and the classifier has the advantages of both methods.

E. Genetic algorithm based Clustering

The genetic algorithm based clustering is used to cluster the lung nodules. GA based clustering is chosen to reduce the distances between the cluster center and other objects of the cluster.

The feature matrix with shape, size and intensity features is used to cluster the lung nodules. Genetic algorithm based clustering is used for this purpose. The feature matrix is initially given to the k-means algorithm to cluster the nodules and the cluster centres are identified. The cluster centres are refined based on the genetic algorithm.

The searching capability of genetic algorithm is used in an effective way to find the appropriate cluster centers such that a similarity metric of the resulting clusters is optimized. The selection, crossover, and mutation processes of the genetic algorithm help to refine the cluster centers. It gives the lung nodules as benign and malignant based on the lung's size (small or large), shape (spherical or elliptical) and intensity. The selection of fitness function is very essential step in genetic algorithm. It helps to accurately place the cluster centers. Similarity calculated by cross-correlation coefficient normalized covariance is used as the fitness function.

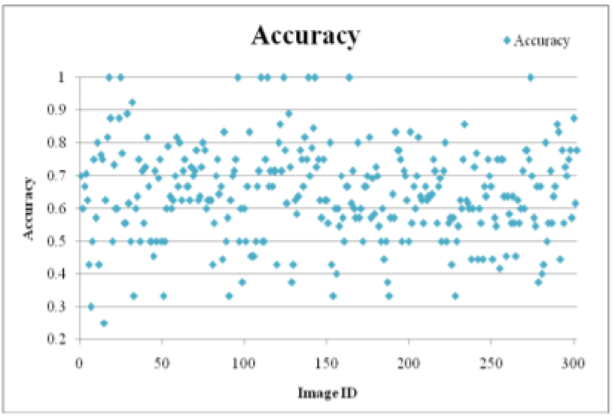

Fig. 4. Accuracy for genetic algorithm based clustering

IV. Conclusion

In this work, a black circular neighborhood algorithm was used to extract lung nodules from edge detected CT images. The extracted nodules are clustered by using Genetic algorithm (GA). GA based algorithm is proposed to reduce the number of false positive (FPs) and give more accuracy. The overall accuracy is improved. This result is higher than the hierarchical based clustering algorithm.

References

1. Shahid Eqbal Dr. M. A. Ansari, 2015, "Medical Image Feature Extraction for Computer Aided Diagnosis of Lung Cancer," International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 6.
2. R. Kishore, 2015, "An Effective And Efficient Feature Selection Method For Lung Cancer Detection," International Journal of Computer Science & Information Technology (IJCSIT) Vol 7, No 4.
3. Jaffar, M.A. Hussain, A. Nazir, M. Mirza, A.M, 2008, "GA and Morphology Based Automated Segmentation of Lungs from CT Scan Images," Computational Intelligence for Modelling Control & Automation, IEEE International Conference, pp. 265 – 270.
4. M.Gomathi Dr. P.Thangaraj,2010, "A New Approach to Lung Image Segmentation using Fuzzy Possibilistic C-Means Algorithm", International Journal of Computer Science and Information Security, Vol. 7, No. 3.
5. Kanazawa, K., M. Kubo and N. Niki, 1996, "Computer aided diagnosis system for lung cancer based on helical CT images," Proceedings of the 13th International Conference on Pattern Recognition, Newport Beach, CA, USA, pp: 381-385. ISBN: 0- 8186-7282-X.
6. Penedo, M.G, M.J. Carreira., A. Mosquera and D. Cabello, 1998, "Computer-aided diagnosis: a neuralnetworkbased approach to lung nodule detection," IEEE Trans. Med. Image., 17: 872-880. ISSN: 0278- 0062.
7. Yamamoto, T., Y. Ukai, M. Kubo, N. Niki and H. Satou, et al., 2000, "Computer aided diagnosis system with functions to assist comparative reading for lung cancer based on helical CT image," Image Process. Int. Conf. Proceed,1: 180-183. ISBN: 0- 7803-6297-7.
8. Yim, Y., H. Hong and Y.G. Shin, 2005, "Hybrid lung segmentation in chest CT images for computer aided diagnosis," Proceedings of 7th International Workshop on Enterprise networking and Computing in Healthcare Industry, June 23-25, Seoul Nat. University, South Korea, pp: 378-383. ISBN: 0-7803- 8940-9.
9. Jose L. Marroquin, Edgar Arce Santana, and Salvador Botello, 2003, "Hidden Markov Measure Field Models for Image Segmentation," IEEE transactions on pattern analysis and machine intelligence, VOL. 25, NO. 11.
10. S. Kumar Raja, ChiranjitMukhopadhyay, and T. V. Ramachandra, 2011, "Hybrid Bayesian Classifier for Improved Classification Accuracy," IEEE geoscience and remote sensing letters, vol. 8, no. 3.
11. M. Gomathi, P. Thangaraj, 2011, "A Computer Aided Diagnosis System for Lung Cancer Detection using Machine Learning Technique," European Journal of Scientific Research ISSN 1450-216X Vol.51 No.2 (2011), pp.260-275.
12. A.Bharathi, A.M.Natarajan, 2011, "Cancer Classification using Modified Extreme Learning Machine based on ANOVA Features," European Journal of Scientific Research ISSN 1450-216X Vol.58 No.2 (2011), pp.156-165.
13. SaleemIqbal, Khalid Iqbal, 2011, "Lungs Segmentation for Computer Aided Diagnosis," International Journal of Academic Research Vol. 3. No. 5.
14. Nisar Ahmed Memon, Anwar MajidMirza, and S.A.M. Gilani, 2006, "Segmentation of Lungs from CT Scan Images for Early Diagnosis of Lung Cancer", World Academy of Science, Engineering and Technology 20.
15. NingXua, NarendraAhujaaand Ravi, "Automated Lung Nodule Segmentation Using Dynamic Programming

and EM Based Classification”, UrbanaChampaign, IL 61801 b Siemens Corporate Research, Inc. Princeton, NJ 08540.

16. Murat Ceylan, Y¨Uksel ¨Ozbay, o. Nuriuc¸an, Erkanyildirim, 2010, “A novel method for lung segmentation on chest CT images: complex-valued artificial neural network with complex wavelet transform,” Turk J Elec Eng& Comp Sci, Vol.18, No.4.
17. Giorgio De Nunzio, Eleonora Tommasi, Antonella Agrusti, Rosella Cataldo, Ivan De Mitri, Marco Favetta, Silvio Maglio, Andrea Massafra, Maurizio Quarta, Massimo Torsello, Ilaria Zecca, Roberto Bellotti, Sabina Tangaro, Piero Calvini, Niccolò Camarlinghi, Fabio Falaschi, Piergiorgio Cerello, and Piernicola Oliva, “Automatic Lung Segmentation in CT Images with Accurate Handling of the Hilar Region”, Journal of digital imaging, Vol 24, No 1, pp 11-27, 2011
18. J. Quintanilla-Dominguez, B. Ojeda-Magaña, M. G. Cortina-Januchs, R. Ruelas, A. Vega- Corona, and D. Andina, "Image segmentation by fuzzy and possibilistic clustering algorithms for the identification of microcalcifications," Sharif University of Technology Scientia Iranica, vol. 18, pp. 580–589, Received 21 July 2010; revised 26 October 2010 accepted 8 February 2011.
19. Eero Salli, Hannu, J, Aronen, Sauli Savolainen, Antti Korvenoja & Ari Visa 2001, ‘Contextual Clustering for Analysis of Functional MRI Data’, IEEE transactions on Medical Imaging, vol. 20, no. 5, pp.403-m 414, 2001.
20. Linda G. Shapiro and George C. Stockman (2001): “Computer Vision”, pp 279-325, New Jersey, Prentice-Hall, ISBN 0-13- 030796-3.
21. B. S. Morse, Lecture 18: Segmentation (Region Based), 1998-2000.
22. S. G. Armato, G.McLennan, L. Bidautb,et.al“The lung image database consortium(LIDC) and image database resource initiative (IDRI): A completereference database of lung nodules on CT scans,” Med. Phys., vol.38,pp.915–931,2011.