# Segmentation of Lungs from Chest X-ray using Euler Number-based Thresholding, Morphological Operators and Greedy Snakes

**Ebenezer Jangam, A. Chandrasekhar Rao, Uppalapati Srilakshmi, D. Yakobu**

***ABSTRACT--- A Computer-Aided Diagnosis (CAD) system is required to precisely detect diseases from the given chest x-ray. Lung segmentation is the basic step performed in the detection of diseases from the chest x-ray. In this paper, we use euler number-based thresholding method for lung region segmentation from CXR images. Morphological operations and greedy snakes are used to improve the accuracy of segmentation. The proposed method is experimented on two datasets: JRST and India. JRST contains 247 chest X- rays and India set contains 100 chest X-rays. An overall accuracy of 96.25% was achieved. The proposed method is compared with state of art methods and it gives high accuracy and high performance.***

***Keywords: Boundary detection ;Chest radiography ; Chan-vese ;Lung field segmentation ;Snake segmentation***

## I. INTRODUCTION

Although Computed Tomograph (CT)and Magnetic Resonance Imaging (MRI) can be more efficient than X-ray, the latter is more method. Discriminant function was used to classify cardiac silhouettes. Sezaki and Ukena[9] in 1973 designed a practical instrument for automated mass screening of heart disease. CTR was generally available and doctors often rely on CXR for making quick decisions in emergency situations. Currently medical doctors perform preliminary diagnosis for heart diseases based on chest X-ray images (CXR). The manual process is not only time consuming but also error prone. It is difficult to analyse the chest X-ray images if they are huge in number, which is a common situation in populous countries. To overcome the difficulties, computerised analysis of CXR images can be adopted. CAD improves the diagnostic accuracy and can assist the medical doctors to come to the right conclusion.

In the first step, the lung region is separated from the X-ray image. Researchers used different techniques to separate the lung region. Method of segmentation largely depends on the image which is going to be segmented. Histogram based thresholding methods are most commonly used. But the limitation of this method is that accurate threshold is not guaranteed. Euler number-based thresholding was used for real time applications [2, 3]. For chest X-ray segmentation, Euler min max function was used[4].

An automated diagnosis for rheumatic heart disease was developed by HALL et al. [7] and KRUGER et al. [8]. They computed CTR and other cardiac parameters to locate cardiac boundaries using a gray-scale threshold computed by a scheme that detects the vertical boundary of the rib cage and the heart by analysing the horizontal profiles. .

Recent work related to chest x-rays is automated approach [10, 11, 17]. However, the existing approaches [12] have accuracy range varying from 73% to 86%, which needs to be improved in order to detect the cardiomegaly using computer. Here, in this paper, all the computations are performed by our self-designed image analysis tool: MedIT which is specifically designed to detect and predict the onset of Cardiomegaly.

In this paper, details about a CAD tool for detection of early symptoms of cardiomegaly are given. The paper is organised as follows. Section 2 describes the proposed method for lung segmentation. Section 3 presents the self-designed partially automated software Medit for analysing CXR and experimental results on JSRT dataset. Later sections present the results and conclusions.

## II. PROPOSED METHOD

The first step towards automated computation of Cardiothoracic Ratio is to create a binarized CXR image having lungs extracted out from the background. Many techniques can be employed to carry out lung segmentation. Euler number-based thresholding technique is used to carry out image segmentation. Once the lungs are isolated, then the image can be analysed for detecting cardiomegaly.

In order to enhance the quality of the segmented images, the CXR images need to undergo a pre-processing phase. The chest area needs to be cropped out from CXR images if the CXR contains unnecessary background. The contrast of the image is enhanced using histogram equalization. 2-D Gaussian operator is used not only to get smooth image but also to preserve the edge features. Using the Gaussian operator, noise is also removed.

**Revised Manuscript Received on February 11 , 2019.**

**Ebenezer Jangam** Department of Computer Science, Vignan's Foundation for Science Technology & Research, Guntur, Andhra Pradesh, India.

**A. Chandrasekhar Rao** Department of CSE, IIT(ISM) Dhanbad

**Uppalapati Srilakshmi** Department of Computer Science, Vignan's Foundation for Science Technology & Research, Guntur, Andhra Pradesh, India.

**D. Yakobu** Department of Computer Science, Vignan's Foundation for Science Technology & Research, Guntur, Andhra Pradesh, India.

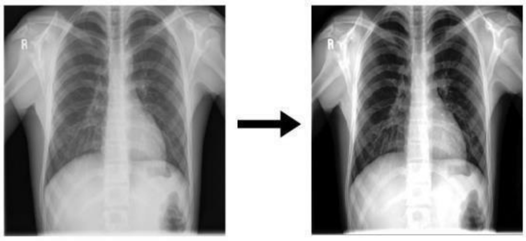

**Figure 1: Pre-processing the Lung image. Left image is the image before pre-processing and the right image is the image after pre-processing.**

*3.2 Lung Segmentation using Euler numbers*

Segmentation divides the image into a set of regions R, which consists of homogenous, non-overlapped, connected subregions as given in equation.

$$R = \{R_i: i = 1,2,3, \dots N\}$$

The union of all sub regions forms the original image which is shown in equation

$$I = R_1 \cup R_2 \cup R_3 \dots . \cup R_N$$

The regions $R_i$ should be connected for all i=0,1,2...N and each region $R_i$ should be homogenous. Different adjacent regions $R_i$ and $R_j$ should be disjoint i.e.

$$R_i \cap R_j = \emptyset$$

Lung field segmentation in Chest X-ray is a problem focused by research groups over the past decade. Different kinds of solutions were proposed. These solutions can be broadly categorized into categories like rule-based methods, pixel classification-based methods, deformable model-based methods and hybrid methods. Here we use Rule-based segmentation method involving thresholding or morphological operations. In-spite of existence of various approaches for segmentation, Thresholding is the most common method used. Thresholding converts an input image I to a binary image B.

To find T value histogram-based approaches were used. But the main disadvantage is that coherency of image is not guaranteed. Holes and extraneous pixels may appear in the segmented image. To preserve the coherency of image, we propose euler number based technique to find the threshold T. Although euler number based thresholding was applied for real time applications [18], it is not used for image thresholding.

The Euler number of an image captures the topological structure of that image [19]. The euler number does not vary with the operations like translations, rotations, scale changes, affinities, projections and even some non- linear transformations such as deformation of the shapes contained in the image.

Mathematically, the Euler number of a binary image can be calculated either by using global where C is the number of regions of the image (number of connected components of the object) and H is the number of holes in the image (isolated regions of the images background).

Euler number E of a binary image can be calculated using local computations as shown in equation.

$$(t) = {}^{1}/_{4}\,[q_1(t) - q_3(t) - 2q_d(t)]$$

where E(t) is the threshold value which is used to obtain the binary image from a grey level image, $q_1$ denotes the number of 2x2 matrices in the image with one 1 and remaining 0's. There are four different possible matrices which count as $q_1$. $q_3$ denotes the number of 2x2 matrices in the image with three 1's and remaining one 0. There are four such different possible matrices which could be counted as $q_3$. $q_d$ denotes the number of diagonal 2x2 matrices. There are two different possible matrices of $q_d$ type.

Equations 5 and 6 are expected to give same Euler number E for a given binary image. For chest X-ray, it is expected to separate two lung regions from the given image using segmentation technique. Therefore, the expected euler number is 2, since the expected number of connected components are 2 and the expected number of holes are 0. Using equation 5, Euler number E can be calculated in the following manner. Since equations 5 and 6 are expected to give the same Euler number E for a given binary image, Equation 6 can be made equal to 2.

$$E = c - h$$

It has been proved that the graph containing different threshold values on X-axis and corresponding euler numbers on Y-axis for a given image is decaying exponential. Hence, for a given euler number, a corresponding threshold value can be found. With this observation, second case is ruled out and Th is a singleton set and contains single value which is the required threshold value.

Algorithm Lung Segmentation Algorithm:

1. Once the grayscale image is obtained, it need to be converted into binary image. Thresholding converts an input image I to a binary image B according to equation 4.
2. The value of T is found out using Euler number-based thresholding. T is the threshold corresponding to euler number 2 for a given chest X-ray. The Euler number is calculated by using E = C - H, Where E denotes Euler number, C denotes number of connected components and H denotes number of holes. In the given chest X-ray, there are two connected components without holes. Hence Euler number is E = 2 - 0 = 2.
3. Remove the dark region at the four corners of the CXR by using Breadth First Search Algorithm.
4. Smoothen the lung boundaries by some erosion and dilution process using disk as the structuring element.
5. Erosion is an operation which is applied on a binary image B by a structuring element S (denoted B S). It generates a new binary image Be = B S. Erosion removes a layer of pixels from the inner and outer boundaries of the region.


*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*

6. Dilation is an operation which is applied on image B by a structuring element S. The effect of Dilation is opposite to erosion. Dilation adds a layer of pixels to both the inner and outer boundaries of regions.
7. Initialise the points of snake by selecting random points on the boundary obtained in step 6. Initialize the three controlling parameters of the snake (α,β and γ).
8. Snake energy for each control point is minimised individually. Greedy snake assumes that total energy is minimised when energy at each control point is minimised.
9. Analyse the image obtained for diagnosis of various diseases.

## III. EXPERIMENTAL RESULTS

### *3.1 Segmentation Tool*

Based on the proposed algorithms, we segmented lung regions in matlab and designed a software named MedIT. Snapshots of the software are shown below.

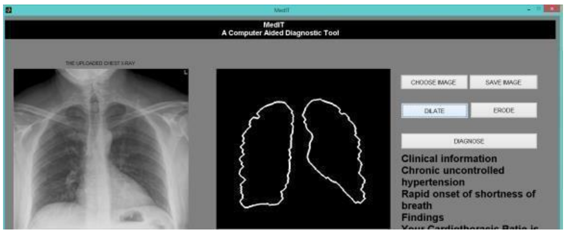


**Figure 6: Lung segmentation from the CXR image.**

Further fine tuning of the morphed image can be done using Dilate and Erode option buttons provided in the software. Dilation and Erosion are the most basic morphological operators. Pixels are added to the boundaries of objects in an image using Dilation whereas pixels are removed from object boundaries using Erosion.

### *3.2 Chest X-Ray Image Dataset*

The proposed algorithm is evaluated using the following CXR dataset.

JSRT DataSet: JSRT in cooperation with Japanese Radiological Society has developed a Chest X-ray image database of 247 chest radiographs with and without nodule. The images are collected from thirteen distinct institutions in Japan and 1 in the USA in 1988 and made it as a public dataset. Out of 247 images, 154 CXR images have lung nodules, while 93 are normal with no nodules.

The dataset is later modified as SCR dataset by including manual masks for segmentation of lungs, heart and clavicles. Researchers frequently used the dataset to evaluate the performance of the segmentation techniques.

### *3.3 Test for Accuracy*

If a system can differentiate the patient and healthy cases correctly, then the system is said to have high efficiency. In order to approximate the accuracy of a system, the proportion of true positive and true negative should be calculated in all evaluated cases.

Jaccard coefficient gives the overlap between Ground truth and segmented image. This can be calculated using:

$$\mathbf{\Omega} = Tp/(Fp + Tp + Fn)$$

Dice coefficient is another metric used to compare the performance of segmentation techniques.

$$Dsc = 2Tp/(2Tp + Fp + Fn)$$

where True positive (Tp) denotes the number of cases which are correctly identified as patient False positive (Fp) denotes the number of cases which are incorrectly identified as patient True negative(Tn) denotes the number of cases which are correctly identified as healthy False negative (Fn) denotes the number of cases which are incorrectly identified as healthy

| Method | Ω | DSC | time |
|---|---|---|---|
| SEDUCM [20] | 95.18±1.8 | 97.4±1.2 | <0.14 |
| Proposed Method | 95.2±1.6 | 97.1±1.0 | <11 |
| SIFT-Flow [21] | 95.4±1.5 | 96.7±0.8 | 20~25 |
| MISCP [22] | 95.1±1.8 | / | 13~28 |
| Hybrid voting [23] | 94.9±2.0 | / | >34 |
| Local SSC [24] | 94.6±1.9 | 97.2±1.0 | 35.2 |
| Human observer [23] | 94.6±1.8 | / | / |
| Inverted Net [25] | 94.6 | 97.2 | 7.1 |
| PC Post-processed [23] | 94.5±2.2 | / | 30 |
| ASM tuned [23] | 92.7±3.2 | / | 1 |
| ASM _SIFT [23] | 92.0±3.1 | / | 75 |

**Table 1: Comparison of performance of proposed method with the existing lung segmentation techniques**

### *3.4. Results*

Using proposed hybrid technique, lung region is extracted from all the images in JSRT dataset. Average values are computed. The results are compared with the other lung segmentation techniques in Table 1. Proposed hybrid technique has recorded an overlap of 93.5 ± 2.1 and DSC of 95.8 ± 1.5 which are comparable to the performance of state of art techniques listed in Table 1.

## IV. CONCLUSION

Using Euler number-based thresholding for initialization of parameters in snake segmentation method yielded an average overlap is 95.6 for JRST dataset, which is on par with the state of art segmentation techniques. The performance of the proposed technique is on par with the state of art techniques for lung segmentation. It has given better performance than the human observer. Hence it can be used in Computer Aided Diagnostic system for chest x-ray analysis.

The proposed hybrid lung segmentation method can be used to estimate heart boundaries. It can be used as initial step for lung disease detection.


*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*

## REFERENCES


1. N. Nakamori, K. Doi, H. MacMAHON, Y. Sasaki, and S. Montner, "Effect of heart-size parameters computed from digital chest radiographs on detection of cardiomegaly: Potential usefulness for computer-aided diagnosis.," Investigative radiology, vol. 26, no. 6, pp. 546–550, 1991.
2. M.-H. Chen and P.-F. Yan, "A fast algorithm to calculate the euler number for binary images," Pattern Recognition Letters, vol. 8, no. 5, pp. 295– 297, 1988.
3. S. B. Gray, "Local properties of binary images in two dimensions," IEEE Transactions on Computers, vol. 100, no. 5, pp. 551–561, 1971.
4. L. Wong and H. Ewe, "A study of lung cancer detection using chest x-ray images," in Proc. 3rd APT Telemedicine Workshop, Kuala Lumpur, vol. 3, pp. 210–214, 2005.
5. B. Van Ginneken, B. T. H. Romeny, and M. A. Viergever, "Computer-aided diagnosis in chest radiography: a survey," IEEE Transactions on medical imaging, vol. 20, no. 12, pp. 1228–1241, 2001.
6. H. Becker, W. Nettleton, P. Meyers, J. Sweeney, and C. Nice, "Digital computer determination of a medical diagnostic index directly from chest x-ray images," IEEE Transac tions on Biomedical Engineering, no. 3, pp. 67–72, 1964.
7. D. Hall, G. Lodwick, R. Kruger, S. Dwyer, and J. Townes, "Direct computer diagnosis of rheumatic heart disease 1," Radiology, vol. 101, no. 3, pp. 497–509, 1971.
8. R. P. Kruger, J. R. Townes, D. L. Hall, S. J. Dwyer, and G. S. Lodwick, "Automated radiographic diagnosis via feature extraction and classification of cardiac size and shape descriptors," IEEE Transactions on Biomedical Engineering, no. 3, pp. 174–186, 1972.
9. N. Sezaki and K. Ukena, "Automatic computation of the cardiothoracic ratio with application to mass screening," IEEE Transactions on Biomedical Engineering, no. 4, pp. 248–253, 1973.
10. A. H. Dallal, C. Agarwal, M. R. Arbabshirani, A. Patel, and G. Moore, "Automatic estimation of heart boundaries and cardiothoracic ratio from chest x-ray images," in SPIE Medical Imaging, pp. 101340K–101340K, International Society for Optics and Photonics, 2017.
11. L. Cong, L. Jiang, G. Chen, and Q. Li, "Fully automated calculation of cardiothoracic ratio in digital chest radiographs," in SPIE Medical Imaging, pp. 1013432–1013432, International Society for Optics and Photonics, 2017.
12. H. MacMahon, K. Doi, H.-P. Chan, M. L. Giger, S. Katsuragawa, and N. Nakamori, "Computer-aided diagnosis in chest radiology," Journal of thoracic imaging, vol. 5, no. 1, pp. 67– 76, 1990.
13. K. Nickol and A. Wade, "Radiographic heart size and cardiothoracic ratio in three ethnic groups: a basis for a simple screening test for cardiac enlargement in men," The British journal of radiology, vol. 55, no. 654, pp. 399–403, 1982.
14. Y. Mensah, K. Mensah, S. Asiamah, H. Gbadamosi, E. Idun, W. Brakohiapa, and A. Oddoye, "Establishing the cardiothoracic ratio using chest radiographs in an indigenous ghanaian population: a simple tool for cardiomegaly screening," Ghana medical journal, vol. 49, no. 3, pp. 159–164, 2015.
15. J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K.-i. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, "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,", American Journal of Roentgenology, vol. 174, no. 1, pp. 71–74, 2000.
16. Esmail, Hanif and Oni, Tolu and Thienemann, Friedrich and OmarDavies, Nashreen and Wilkinson, Robert J and Ntsekhe, Mpiko,"Cardio-thoracic ratio is stable, reproducible and has potential as a screening tool for HIV-1 related cardiac disorders in resource poor settings," Public Library of Science, vol. 11,no. 10,pp. 6349,2016.
17. Candemir, Sema and Jaeger, Stefan and Lin, Wilson and Xue, Zhiyun and Antani, Sameer and Thoma, George, " Automatic heart localization and radiographic index computation in chest x- rays," Proc. of SPIE Vol, vol. 9785, pp. 1-17, 2016.
18. Lakhani, Paras and Sundaram, Baskaran, "Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks," Radiological Society of North America,pp. 16- 26,2017.
19. He, Li-Feng and Chao, Yu-Yan and Suzuki, Kenji," An algorithm for connected-component labeling, hole labeling and Euler number computing", Journal of Computer Science and Technology Springer,vol. 28,no. 3,pp. 468-478, 2013.