International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-7, Issue-5S4, February 2019

Segmentation of Lungs from Chest X-rays using Firefly Optimized Spatial FCM(FASFCM)

Ebenezer Jangam, Rahul Kumar, Rajesh Dwivedi, Vishnu Kumar

ABSTRACT--- Segmentation of lungs from chest x ray is a non-trivial task required as a preprocessing step for detection of different diseases like cardiomegaly, tuberculosis, pneumonia. High accuracy in segmentation of lung results in high accuracy of detection of diseases from lungs. For the past four decades multiple techniques were proposed for automatic segmentation of lungs. In this paper, we propose a hybrid segmentation technique based on Bat optimized fuzzy c-means clustering algorithm. The output of the fuzzy c-means is given to level set to finalize the segmentation of the lungs. The performance of the proposed technique is evaluated using two public chest x ray datasets: JRST and Montgomery County. JRST contains 247 chest X-rays and MC dataset contains 138 chest X-rays. The Jaccard coefficient for the proposed segmentation technique is 95.1 which is on par with the state of art segmentation techniques.

Keywords: Use about five key words or phrases in alphabetical order, Separated by Semicolon.

1. INTRODUCTION

In 1983, Fuzzy c-means Clustering [1] was developed by J.C. Bezdek to cluster the real datasets which do not have definite boundaries between the clusters. Fuzzy c-means clusters the data points such that a data point can belong to more than one cluster with a determined degree. This determined degree is called the membership value of data points and membership value remains in range (0, 1). Even though fuzzy c-means is one of the most commonly used techniques, there are some problems. The main problems associated with fuzzy c-means are (i)randomly generated cluster centers are used for initialization and (ii) probability of getting trapped in local minima is high. In spite of the popularity of fuzzy c-means, the drawback of the algorithm is that it can be trapped at the local optimum rather than global optimum[4]. Firefly optimization was proposed by X S Yang In this paper, we propose Randomized Fuzzy c-means clustering with Maximin Distance Classifier[3] based on Bat optimization. Firefly Algorithm[5] has a property of achieving a high rate of converging at global optimum. Bat optimization is swarm based optimization technique. There are many other swarm based optimization techniques like PSO [6], ACO [7], GA [8], Simulate annealing[9] etc.

X-ray was discovered in 1895[14] which became most common means to diagnose multiple diseases like Tuberculosis, Pneumonia, Cardiomegaly and other lung related diseases. Computers are available In 1960s researchers published articles about radiology report analysis using computer[7]. In 1970s, focus was upon the detection of abnormalities in chest x-ray using a computer.

The traditional chest analysis is the most prevalent radiological procedure, making up a minimum of a third of all exams in the radiology division. Moreover, Pulmonary diseases like pneumonia, tuberculosis, emphysema and lung cancer can be screened based on the chest radiograph[28]. But, computerized interpretation of a chest radiograph is extremely challenging due to presence of superimposed anatomical structures. The complexity of computerized analysis of chest x-ray along with their prevalence in radiology department is the main reason for the researchers to concentrate on the development of computer algorithms to assist radiologists in reading chest images.

Segmentation of lungs is the essential step in the Computer Aided Detection(CADe) and Computer Aided Diagnosis(CADx) of chest x-rays. It is the basic step performed in the automatic tuberculosis screening[27]. It is the part of automatic pneumonia screening. It is used as the first step to detect cardiomegaly(enlargement of heart). It is the preliminary step in lung nodule detection from chest x-rays using computer algorithms. To find out the abnormalities in the lungs from chest x-rays, the primary task is to delineate the lungs from the chest x-rays[18], [11], [20], [17], [16].

Researchers proposed a wide range of techniques for automatic lung segmentation, rotation[19] and foreign object detection[33].

Some of the techniques used for lung segmentation are thresholding, region growing, neural networks, active contours, pixel classification, structured edge detection, adversial networks, graph cuts, game theory. Although there are a variety of existing techniques for lung segmentation, researchers are investigating for the novel techniques which can result in high accuracy. Higher the accuracy in segmentation of the lungs, higher is the accuracy in classification and detection of diseases like cardiomegaly, pneumonia and other lung related diseases. In recent times, hybrid techniques are being investigated to segment the lung regions with high accuracy.

Revised Manuscript Received on February 11 , 2019.

Ebenezer Jangam, Department of CSE, Vignan Foundation for Science Technology and Research, Andhra Pradesh, India. (E-mail: ebenezer.jangam@gmail.com)

Rahul Kumar, Department of CSE, Vignan Foundation for Science Technology and Research, Andhra Pradesh, India.

Rajesh Dwivedi, Department of CSE, Vignan Foundation for Science Technology and Research, Andhra Pradesh, India.

Vishnu Kumar, Department of CSE, Vignan Foundation for Science Technology and Research, Andhra Pradesh, India.

We propose a hybrid lung segmentation technique which is the combination of fuzzy c-means and level set algorithm. The performance of the proposed technique is compared with the existing techniques. The accurate detection of lung boundary results in high accuracy in estimating heart boundaries and thereby increases the accuracy in the detection of cardiomegaly (enlargement of heart).

The content of the paper is organized in the following manner. Section 2 contains details about the proposed hybrid lung segmentation technique. Section 3 has the information about the public datasets and performance metrics used for evaluation of proposed lung segmentation technique. The performance of proposed technique is compared with existing techniques in Section 4. Segmented lung boundaries are used for cardiomegaly detection in Section 5. Section 6 concludes the paper by giving possible future directions.

2. SEGMENTATION OF LUNGS USING FIREFLY OPTIMIZED SPATIAL FCM AND LEVEL SET ALGORITHM

The proposed segmentation algorithm contains three steps. In the first step, firefly optimization is used for clustering the pixels

2.1. Firefly Optimization

Firefly algorithm was based on the behavior of fireflies and their movement. Three central concepts in Firefly algorithm are •Any firefly can be attracted to any other firefly irrespective of sex

- A firefly will be attracted towards a brighter firefly and the less brighter one moves towards more brighter one
- The objective function decides the brightness of the firefly Attractiveness of a firefly varies with distance according to the following equation.

$$\alpha_0 = \alpha \exp(-\gamma)$$

2.2 Spatial Fuzzy c-means clustering

Fuzzy c-means clustering is one of the popular algorithms used for segmentation of medical images[12], [5], [1]. It clusters pixels of given image into different classes such that each pixel may belong to more than one class with an associated probability. In fuzzy clustering, centroid and scope of each cluster are estimated by minimizing the cost function. The origin of FCM is K-means algorithm where N objects are assigned to N classes(K<=N) with minimum value of cost function.

In the context of medical images, N is total number of pixels in the image . These N pixels should be assigned to one of the clusters. Centre of the cluster should be finalized. According to FCM clustering, the pixels nearer to the centroid will receive high values and the pixel far away from the centroid will receive low values.

2.3. Level set algorithm

When level set methods are used for image segmentation, an image is segmented based on boundaries which change dynamically. When fuzzy clustering is used for segmentation, it employs pixel based classification[23]. Level set algorithm embeds active contours into a time dependent PDE function f(t, x, y). The zero level set C(t) is used to approximate the active contours.

$$(t, x, y) \text{ is } \begin{cases} < 0 \text{ if } (x, y) \text{ is inside } C(t) \\ = 0 \text{ if} (x, y) \text{ is at } C(t) \\ > 0 \text{ if} (x, y) \text{ is outside } C(t) \end{cases}$$

2.4. Hybrid algorithm for lung segmentation

Hybrid algorithm for lung segmentation uses the combination of spatial fuzzy clustering and level set algorithm to achieve better accuracy. After preprocessing, the fuzzy clustering algorithm is used to classify the pixels which belong to region of interest. The outcome of fuzzy clustering algorithm is utilized by the level set segmentation algorithm for proper initialization. The outcome of fuzzy clustering is also useful to estimate the controlling parameters of level set segmentation. In other words, Level set segmentation algorithm finds out the contours of interest in a medical image from the outcome of fuzzy clustering algorithm.

Let R_k be the region of interest which is obtained from the fuzzy cmeans clustering algorithm. Then the level set function can be initialized using the following equation.

$$\Phi_0 \,(x,y)=-4\varepsilon(0.5-B_k) \qquad (4)$$

where is a constant regulating Dirac function. Bk is a binary image obtained from the equation Bk =Rk $\geq b_0$, where b_0 is an adjustable threshold.

The parameters controlling the level set segmentation are selected to get optimal results by using iterative method. The initial values are chosen according to the method outlined in [12].

3. PERFORMANCE EVALUATION

3.1. Data Sets

A. Public datasets of chest x-ray for segmentation and disease detection

The following are the public datasets available for segmentation of chest x rays.

- JSRT/SCR dataset [29]
- MC dataset[10]

The following dataset contains chest x-rays labelled with multiple diseases.

- Chest x-ray 14 dataset [29]

1) SCR Dataset: JSRT dataset is a Chest X ray image database of 247 chest radiographs[26].JSRT is the public dataset for lung nodule detection. SCR dataset was made public in order to promote comparison of techniques proposed for segmentation of anatomical structures[29]. SCR dataset is the most common dataset used to evaluate segmentation of anatomic structures (lungs, heart, clavicles) in a CXR as shown in Table II.
2) Montgomery County dataset: Montgomery County(MC) dataset contains 138 PA CXRs in this dataset which are collected under TB control program. 80 CXRs are considered to be normal and 58 are

Published By: Blue Eyes Intelligence Engineering & Sciences Publication

abnormal with manifestations of TB[10]. Manual segmentation on images of MC dataset was performed and binary lung masks were generated. Montgomery dataset was primarily made available for tuberculosis screening, but it is useful for segmentation of lung fields.

3) Chest X-ray8 Dataset: Chest X-ray 8 dataset is a massive dataset released publicly in 2017 [31] for detection of multiple diseases. Initially, it was a dataset of Chest X-ray images with eight different disease labels and no finding label. Along with images and labels, there was other information about patients' visit, gender, etc. We selected 150 chest x-rays with label cardiomegaly and 150 chest x-rays which are normal with no labels. A customized dataset with 300 chest x-rays is chosen for evaluation of technique for the detection of cardiomegaly.

3.2 Performance Metrics

The Jaccard similarity coefficient Ω and Dice coefficient are the commonly used metrics to measure the performance of a segmentation technique.

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

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

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

$$\boldsymbol{Dsc} = 2\boldsymbol{Tp}/(2\boldsymbol{Tp} + \boldsymbol{Fp} + \boldsymbol{Fn})$$

where TP (true positives) is the count of pixels which are classified correctly, FP (false positives)is the count of pixels which are identified as part of the object but they belong to background in reality, and FN (false negatives) are the pixels which are classified as background but the fact is that they belong to the object.

3.3 Comparision of the performance of the proposed technique with the existing lung segmentation techniques

Level set algorithm used the spatial fuzzy c-means clustering algorithm to improve the segmentation accuracy. The proposed technique was evaluated using JSRT SCR dataset and MC dataset.

A. Comparision of performance of the proposed lung field segmentation technique on JSRT SCR dataset

Using proposed hybrid technique, lung region is extracted from all the images in JSRT SCR dataset. Average values are computed. The results are compared with the other lung segmentation techniques in Table 1. Our proposed segmentation technique has recorded an overlap of 95.6 ± 1.5 and DSC of 97.6±1.2 which is better than human observer.

Highest accuracy is 96.3±1.2 when lower order adaptive region growing technique[4] is used. Human observer accuracy is e calculated as 94.6±1.8 and more than half of the segmentation techniques genarated an accuracy more than human observer.

RESULTS & DISCUSSIONS

Method	**Ω**	**DSC**	**time in sec**
Lower Order Region Growing[4]	96.3 ±1.2	98.3±0.7	3-20
SEDUCM [32]	95.18±1.8	97.4±1.2	<0.14
Proposed Method	95.6±1.5	97.6±1.2	25
SIFT-Flow [3]	95.4±1.5	96.7±0.8	20~25
MISCP [18]	95.1±1.8	-	13~28
Hybrid voting [29]	94.9±2.0	-	>34
Local SSC [23]	94.6±1.9	97.2±1.0	35.2
Human observer [29]	94.6±1.8	-	-
Inverted Net [15]	94.6	97.2	7.1
PC Post-processed [29]	94.5±2.2	-	30
ASM tuned [29]	92.7±3.2	-	1
ASM _SIFT [29]	92.0±3.1	-	75

Table 1. Comparison of performance of proposed method with existing techniques using JSRT dataset

B. Comparision of performance of proposed lung field segmentat ion technique on Montgomery County Dataset

Using proposed hybrid technique, lung region is extracted from all the images in MC dataset. Average values are computed. The results are compared with the other lung segmentation techniques in Table 2. Proposed hybrid technique has recorded a 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 2.

Only a few segmentation techniques are evaluated using Montgomery County [4], [3], [6]. Lower order region growing approach has reported high accuracy of 96.6±1.8 as shown in Table 2. SCAN Technique has recorded an accuracy of 91.4 ± 0.61 with MC data set against 94.7 ± 0.4 using JSRT SCR dataset.

Method	**Ω**	**DSC**	**time in sec**
Lower Order Region Growing[4]	96.6 ±1.8	97.8±0.5	/
SIFT-Flow [3]	94.1±3.4	96.0±1.8	<0.14
Proposed Method	93.5±2.1	95.8±1.5	21
SCAN[6]	95.4±1.5	96.7±0.8	20~25

Table 2. Comparison of performance of proposed method with existing techniques using MC dataset

CONCLUSION

Using Firefly based Spatial Fuzzy c-means clustering for intialization of parameters in Level set method yeilded an average overlap is 95.6 for JRST dataset, which is on par with the state of art segmentation techniques. For MC dataset the overlap is 95.8, which is on par with the state of art segmentation techniques used for lung segmentation from chest x rays.

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

Other optimization techniques can be applied to compare and improve the results further as a future work. Investigation is needed to compare the performance of different lung segmentation techniques with datasets of different sizes.

REFERENCES

1. Cai, W., Chen, S., Zhang, D.: Fast and robust fuzzy cmeans clustering algorithms incorporating local information for image segmentation. Pattern recognition 40(3), 825–838 (2007)
2. Candemir, S., Jaeger, S., Palaniappan, K., Antani, S., Thoma, G.: Graph-cut based automatic lung boundary detection in chest radiographs. In: IEEE Healthcare Technology Conference: Translational engineering in health & medicine. pp. 31–34 (2012)
3. Candemir, S., Jaeger, S., Palaniappan, K., Musco, J.P., Singh, R.K., Xue, Z., Karargyris, A., Antani, S., Thoma, G., McDonald, C.J.: Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. IEEE transactions on medical imaging 33(2), 577– 590 (2014)
4. Chondro, P., Yao, C.Y., Ruan, S.J., Chien, L.C.: Low order adaptive region growing for lung segmentation on plain chest radiographs. Neurocomputing 275, 1002– 1011 (2018)
5. Chuang, K.S., Tzeng, H.L., Chen, S., Wu, J., Chen, T.J.: Fuzzy c-means clustering with spatial information for image segmentation. computerized medical imaging and graphics 30(1), 9–15 (2006)
6. Dai, W., Doyle, J., Liang, X., Zhang, H., Dong, N., Li, Y., Xing, E.P.: Scan: Structure correcting adversarial network for organ segmentation in chest x-rays. arXiv preprint arXiv:1703.08770 (2017)
7. Giger, M.L., Chan, H.P., Boone, J.: Anniversary paper: History and status of cad and quantitative image analysis: the role of medical physics and aapm. Medical physics 35(12), 5799–5820 (2008)
8. Ibragimov, B., Likar, B., Pernus, F., Vrtovec, T.: Accurate landmark-based segmentation by incorporating landmark misdetections. In: Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on. pp. 1072–1075. IEEE (2016)
9. Ibragimov, B., Likar, B., Pernus, F., et al.: A game theoretic framework for landmark-based image segmentation. IEEE Transactions on Medical Imaging 31(9), 1761–1776 (2012)
10. Jaeger, S., Candemir, S., Antani, S., Wang, Y.X.J., Lu,´ P.X., Thoma, G.: Two public chest x-ray datasets for computer aided screening of pulmonary diseases. Quantitative imaging in medicine and surgery 4(6), 475 (2014)
11. Karargyris, A., Siegelman, J., Tzortzis, D., Jaeger, S., Candemir, S., Xue, Z., Santosh, K., Vajda, S., Antani, S., Folio, L., et al.: Combination of texture and shape features to detect pulmonary abnormalities in digital chest x-rays. International journal of computer assisted radiology and surgery 11(1), 99– 106 (2016)
12. Li, B.N., Chui, C.K., Chang, S., Ong, S.H.: Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation. Computers in biology and medicine 41(1), 1–10 (2011)
13. Li, X., Luo, S., Hu, Q., Li, J., Wang, D., Chiong, F.: Automatic lung field segmentation in x-ray radiographs using statistical shape and appearance models. Journal of Medical Imaging and Health Informatics 6(2), 338–348 (2016)
14. Mould, R.F.: A century of X-rays and radioactivity in medicine: with emphasis on photographic records of the early years. CRC Press (1993)
15. Novikov, A.A., Lenis, D., Major, et al.: Fully convolutional architectures for multi-class segmentation in chest radiographs. IEEE Transactions on Medical Imaging (2018)
16. Rao, N.G., Kumar, V.V., Krishna, V.V.: Texture based image indexing and retrieval. IJCSNS International Journal of Computer Science and Network Security 9(5), 206– 210 (2009)
17. Rao, N.G., Sravani, T., Kumar, V.V.: Ocrm: optimal cost region matching similarity measure for region based image retrieval. Int J Multimedia Ubiquitous Eng 9(4), 327 (2014)
18. D. Seghers, D. Loeckx, F. Maes, D. Vandermeulen and P. Suetens, "Minimal Shape and Intensity Cost Path Segmentation," in IEEE Transactions on Medical Imaging, vol. 26, no. 8, pp. 1115-1129, Aug. 2007
19. Santosh, K., Vajda, S., Antani, S., Thoma, G.R.: Edge map analysis in chest x-rays for automatic pulmonary abnormality screening. International journal of computer assisted radiology and surgery 11(9), 1637–1646 (2016)
20. Seghers, D., Loeckx, D., Maes, F., Vandermeulen, D., Suetens, P.: Minimal shape and intensity cost path segmentation. IEEE Transactions on Medical Imaging 26(8), 1115–1129 (2007)
21. Seghers, D., Loeckx, D., Maes, F., Vandermeulen, D., Suetens, P.: Minimal shape and intensity cost path segmentation. IEEE Transactions on Medical Imaging 26(8), 1115–1129 (2007)
22. Sethian, J.A., et al.: Level set methods and fast marching methods. Journal of Computing and Information Technology 11(1), 1–2 (2003)
23. Shao, Y., Gao, Y., Guo, Y., Shi, Y., Yang, X., Shen, D.: Hierarchical lung field segmentation with joint shape and appearance sparse learning. IEEE transactions on medical imaging 33(9), 1761–1780 (2014)
24. Shi, Y., Qi, F., Xue, Z., Chen, L., Ito, K., Matsuo, H., Shen, D.: Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics. IEEE Transactions on Medical Imaging 27(4), 481–494 (2008)
25. Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T., Komatsu, K.i., 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. American Journal of Roentgenology 174(1), 71–74 (2000)
26. Vajda, S., Karargyris, A., Jaeger, S., Santosh, K., Candemir, S., Xue, Z., Antani, S., Thoma, G.: Feature selection for automatic tuberculosis screening in frontal chest radiographs, Journal of medical systems 42(8), 146 (2018)
27. Van Ginneken, B., Stegmann, M.B., Loog, M.: Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. Medical image analysis 10(1), 19–40 (2006)
28. Wang, C.: Segmentation of multiple structures in chest radiographs using multi-task fully convolutional networks. In: Scandinavian Conference on Image Analysis. pp. 282–289. Springer (2017)
29. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M.: Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on. pp. 3462–3471. IEEE (2017)
30. Yang, W., Liu, Y., Lin, L., Yun, Z., Lu, Z., Feng, Q., Chen, W.: Lung field segmentation in chest radiographs from boundary maps by a structured edge detector. IEEE journal of biomedical and health informatics 22(3), 842– 851 (2018)