Advances and Applications in Mathematical Sciences
Volume 20, Issue 11, September 2021, Pages 2599-2605

PERFORMANCE ANALYSIS OF AUTOMATED DIAGNOSIS OF LUNG PARENCHYMA USING A NOVEL APPROACH (CFSFCM)

S. SHAIK PARVEEN, P. HEMAVATHY, G. DEVIKA and R. SANDHA

Department of Computer Science
Mannar Thirumalai Naicker College
Madurai, Tamil Nadu 625004, India
E-mail: parveensh.81@gmail.com
hemamanigandan@gmail.com
gdevika2016@gmail.com
yazhini98@gmail.com

Abstract

In this article, a combination of firefly search fuzzy C means (CFSFCM) clustering approach is proposed for segmenting the parenchyma of the computed tomography (CT) lung images. In this approach, the lung parenchyma is considered as the region of interest (ROI). This method is carried out in four steps. First step is the pre-processing of CT lung images to remove the noise and artifacts present in it. The lung parenchyma is segmented using the proposed approach CFSFCM based on its ground truth. Third step features extraction from the obtained segmented image using haralick features. Finally, in the fourth step classification is done to identify the nodules using Support vector machine whether it is malignant or benign. This method is tested over 200 CT scans for finding the efficiency. From the experimental results, it can be observed that the proposed approach gives an overall accuracy of 97.51% for automatically detecting lung parenchyma.

I. Introduction

The latest global statistics from 2008 show that 12.7 million new cancer cases and 7.6 million cancer deaths occurred worldwide. This accounts for 13% of all deaths for that year, making cancer a common threat to all families. As technology becomes more efficient, a trend towards computer

2010 Mathematics Subject Classification: 68M20.
Keywords: Lung Parenchyma, Fuzzy C Means, Firefly Search, SVM.
Received October 13, 2020; Accepted November 8, 2020

aided diagnostic (CAD) tools for identification, prognosis prediction and reoccurrence likelihood is becoming a reality [1]. A wide range of different image modalities can be found in medical imaging. These modalities allow the radiologist to obtain a non-invasive, usually three-dimensional, view of the internal structures of the human body, such as heart, kidney, liver and spleen at a very good spatial resolution. Among the medical image modalities, Ultrasound (US), Computed Tomography (CT), Magnetic Resonance (MRI) and Positron Emission Tomography (PET) imaging have become of great interest in many research areas [6]. All these anatomical and functional imaging modalities are of extreme importance in several domains of medicine, for example, computer aided-diagnosis, pathology follow-up, treatment planning and therapy (surgery, radio-therapy, chemo-therapy etc.). Computer assistance plays an important role in all this kind of clinical application. Due to medical image analysis technology that grew rapidly in last decades, considerably facilities were brought for clinical examinations. Within medical image analysis techniques, segmentation and classification are some key problems that we will further refer to in this work.

II. Proposed System

Image Denoising

Image Denoising can significantly increase the reliability of an optical inspection. Several filter operations which intensify or reduce certain image details enable an easier or faster evaluation [2]. The aim of denoising is an improvement of the image data that suppresses unwanted distortions or enhances some image features for further processing. The input CT image is shown in Figure 1(a) and the filtered image using a 3×3 median filter is shown Figure 1(b).

Figure 1. (a) Original CT image (b) Denoised Image.

Combined Firefly Search Fuzzy C-Means (CFSFCM)

The proposed segmentation methodology uses an intelligent and dynamic approach called Combined Firefly Search Fuzzy *C*-Means (CFSFCM). In this approach, the firefly algorithm (FA) examines the search space of the given data set to determine the near-optimal cluster centers and then those cluster centers are given initial mean value for FCM algorithm. Then, the finest cluster centers that have been identified by the Firefly algorithm are employed as the preliminary cluster centers for the FCM algorithm.

The proposed segmentation methodology consists of two phases:

• In order to determine the optimal cluster centers, firefly inspects the search space of the given dataset and then the values of the cluster centers will be obtained using the FA.

• Starting the initialization of the Fuzzy *C*-Mean algorithm based on the evaluated results in the first phase and to overcome the drawbacks of Fuzzy *C*-Mean algorithm such as getting stuck in the local optimal and being susceptible to initialization sensitivity.

Identifying near-optimal cluster centres using firefly search

The cluster centres of the provided data set are encoded by each and every solution of the firefly search. The solution will be as in Equation (1):

$$A = (s_1\{a_1, a_2, \ldots, a_d\}, s_2\{a_1, a_2, \ldots, a_d\}, s_3\{a_1, a_2, \ldots, a_d\}) \tag{1}$$

where a_i is a numerical characteristic that explains a cluster centre and $a_i \in A$, where *A* is the collection of the feasible array of each and every pixel attribute. Consequently, each cluster centre s_i is defined by *d* numerical feature $a_1, a_2, \ldots, a_d$. As a result, every single solution has an actual size of $(c * d)$ where *c* represents the given number of clusters and indicates the feature number outlining the given data set.

In the proposed segmentation methodology, a firefly (brighter one) that has minimum fitness value, for each iteration will have the ability to affect and influence in the movements of the other fireflies. Therefore, when comparing between two fireflies *a* and *b*, if *b* is brighter than firefly *a*, then firefly *a* will move toward firefly *b*. The proposed approach is designed to enhance the performance of the traditional FCM in order to obtain more accurate segmentation process. The hybridization step between FA and FCM is introduced to enhance the quality of the FA clustering results. The FCM have the ability to modify the cluster centres values till reaching the minimum variance, therefore obtaining more specific clusters.

III. Results and Discussion

The efficiency of the proposed segmentation algorithm is evaluated by comparing the results of the proposed segmentation algorithm with the metric proposed by [7]. The important properties like Solidity, Perimeter and Area are used to find the segmentation accuracy of CFSFCM. The segmentation accuracy A_{seg} is calculated using Equation (2) to evaluate the quality of the proposed methodology is shown in Table 1.

$$A_{seg} = \frac{S_{prop}}{O_{prop}} \times 100 \tag{2}$$

where S_{prop} is defined in Equation (3) as

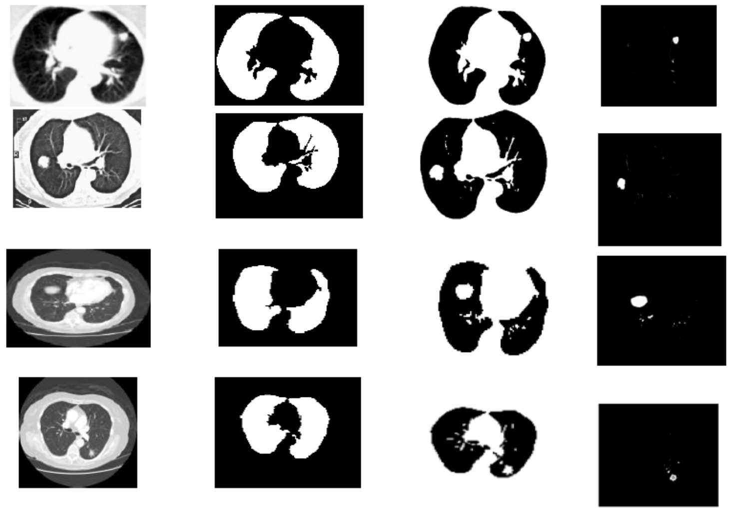
$$S_{prop} = (Solidity + Area + Perimeter)_{Segmented\ Image} \tag{3}$$

O_{prop} is defined in Equation (4) as

$$O_{prop} = (Solidity + Area + Perimeter)_{Ground\ Truth} \tag{4}$$

Table 1. Performance analysis of segmentation using CFSFCM.

Image		Solidity	Area	Perimeter	Total	Accuracy in %
Image 1	Segmented	0.76535	43412	1541.4398	44954.2052	97.81
	Unsegmented	0.76535	44216	1743.4398	45960.2052	
Image 2	Segmented	0.679	36759	1832.9068	38592.5858	95.09
	Unsegmented	0.679	38553	2030.9068	40584.5858	
Image 3	Segmented	0.71355	30596	1307.0235	31903.7371	95.79
	Unsegmented	0.8376	31897	1407.0235	33304.8611	
Image 4	Segmented	0.8376	42361	1151.918	43513.7556	95.44
	Unsegmented	0.679	43759	1832.9068	45592.5858	
Average Accuracy in %						**96.03**



(a) Original CT Image (b) Ground Truth (c) Segmented Lung (d) Tumor Segmentation

Figure 3. Results obtained for CT images along with ground truth using the Proposed Methodology (CFSFCM).

The most used kernel function for support vector machine is Radial Basis Function (RBF) because of their localized and finite responses across the entire range of real x-axis. The classification accuracy of Radial Basis Function (RBF) kernel was high [4]; also, the bias value and the error rate of Radial Basis Function (RBF) kernel were small when compared to other kernels. After the learning process is completed by providing several conditions, the proposed technique would be able to detect the presence of cancer in the lung parenchyma automatically. Table 2 and Figure 4 show the performance analysis for different kernels of SVM.

Table 2. Performance measures (Accuracy, Sensitivity, Specificity, PPV, NPV and F_Score in %) obtained using different kernels of SVM.

Measures	Linear	Polynomial	RBF
Sensitivity	92	94	98
Specificity	87	93	97
Accuracy	89.5	93.5	97.5
PPV	87.62	93.07	97.03
NPV	91.58	93.94	97.98
F_Score	89.75	93.53	97.51

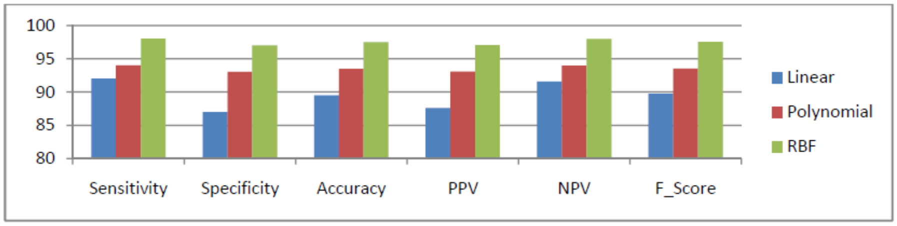


Figure 4. Performance measures (Accuracy, Sensitivity, Specificity, PPV, NPV and F_Score in %) obtained using different kernels of SVM.

IV. Conclusion

In this work, an automatic CAD system has been developed for the early diagnosis of lung cancer. The segmentation methodologies CFSFCM have been proposed for segmenting the lung region and tumor region in which the nodule exists can be identified to facilitate biopsy. An accuracy of 97.51% has been achieved by using the CFSFCM. The proposed scheme reduces the computational complexity of CAD system without compromising the performance of the system. The difficulty in the early detection of cancerous lung parenchyma was overcome in the proposed approach. This methodology provided a computer aided diagnosis system for early detection of lung cancer. This proposed approach increased the accuracy of the CAD (Computer Aided Detection) system. The methodology proposed in this article was capable of extracting the lung region and tumour region (ROI) from the lung CT images. This methodology was able to effectively extract tumour nodules from the lung CT images.

References

[1] S. S. Parveen and C. Kavitha, A review on computer aided detection and diagnosis of lung cancer nodules, International Journal of Computers and Technology 3 (2012), 393-400.

[2] S. Shaik Parveen and C. Kavitha, Detection of lung Cancer nodule using automatic region growing method, IEEE International Conference on Computing Communication and Networking, (2013).

[3] S. S. Parveen and C. Kavitha, Classification of lung cancer nodules using SVM kernels, International Journal of Computer Applications 95 (2014), 25.

[4] S. Shaik Parveen and C. Kavitha, Automatic Computer Aided Diagnosis System for

Detection of Lung Cancer Nodules Using Region Growing Method and Support Vector Machines (SVM), Research Journal of Applied Sciences 9 (2014), 299-307.

[5] S. Parveen and C. Kavitha, Segmentation of CT Lung nodules using FCM with firefly search algorithm, In: IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems (2015), 1-6.

[6] P. Hemavathy, G. Devika and S. S. Parveen, Efficiency Of Self Organizing Map In High Dimensional Data, International Journal of Research and Analytical Reviews (IJRAR) 6(2) (2019), 14-18.

[7] Z. Faizal Khan and Syed Usama Quadri, FEMD Algorithm for Effective Segmentation of CT Lung Images, International Journal of Computer Applications 111(8) (2015), 2-24.