INTERNATIONAL RESEARCH JOURNAL OF PHARMACY
www.irjponline.com
ISSN 2230 – 8407

Research Article

PERFORMANCE ANALYSIS OF FIREFLY SEARCH FUZZY C-MEANS (FSFCM) FOR DETECTING LUNG CANCER NODULES

Amudha K *, Balu S [2], Sakthivel K [2]
Department of Electronics and Communication Engineering, Kongunadu College of Engineering and Technology, Trichy, Tamil Nadu, India
[2]Department of Computer Science and Engineering, K. S. Rangasamy College of Technology, Namakkal, Tamil Nadu, India
*Corresponding Author Email: amudhak02@gmail.com

Article Received on: 12/08/17 Approved for publication: 16/09/17

DOI: 10.7897/2230-8407.089163

ABSTRACT

Computer-Aided Diagnosis system is used for diagnosing the lung cancer nodule from the chest Computer Tomography (CT) images. The System can automatically detect and diagnosis the lung cancer nodules with efficient accuracy and it also minimize the time taken by the radiologist for interpretation. The computer-aided diagnosis system is developed by combined techniques such as image denoising, segmentation, feature extraction and classification. CT lung image is denoised by the median filter. The system automatically segments region of interest from denoised lung computer tomography image using the proposed methodologies namely FSFCM. Then, the optimal initial cluster centres are created for the FSFCM to improve the segmentation results. This work presents a modified firefly search fuzzy C-means algorithm to produce efficient segmentation results.

Key Words: Segmentation, Fuzzy C-Means, Cluster, Computer Tomography images.

INTRODUCTION

Image Processing procedures are widely used in many medical areas for image improvement and better visualization**.** Lung cancer is the most common cancer in both men and women. Lung cancer is also known as lung carcinoma. Lung cancer is generally categorized as Small cell lung cancer (SCLC) and Non-small cell lung cancer (NSCLC). Lung cancer may be found as a mass or tumor on a chest computerized tomography of a patient[1].

The advancements in the medical field have led to introduce many imaging techniques for the diagnosis of lung cancer. Some imaging techniques are Computed Tomography, Radiography, Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET). CT and Radiography is used for ionization of radiation procedure whereas MRI and PET do not use the same. For the accurate diagnosis of the cancer, radiation procedures are best suitable. Radiologists are trained in such a way to perform radiation procedures[2]. According to American College of Radiology (ACR) recommendations, a CT scanner should meet or exceed some specifications to achieve acceptable clinical lung CT scans. It means scanning time does not take more than 2 seconds per slice or image and it has Minimum of 2 mm or less slice thickness. Then the inter scan delay does not exceed to 4 seconds and limiting spatial resolution must be measured to confirm that it meets the specifications of the unit manufacturer. Then the table pitch also not exceed to 2:1 for most CT scanners. CT scan can be used for detecting both acute and chronic changes in the lung parenchyma which means the internals of the lungs. This work presents a segmentation algorithm to detect tumor from the computer tomography images[3].

MATERIALS AND METHODS

Computer-Aided Diagnosis System

A generic CAD system consists of image denoising, segmentation, feature extraction and classification is shown in Figure 1. The image pre-processing step improves the quality of the input image through the application of noise removal, enhancing and standardization procedures. Pre-processing is essential because the effectiveness of proceeding steps like definition of ROI and feature extraction are dependent on the quality of the input images. Noise reduction is important to improve the quality of an image for better visualization[4].

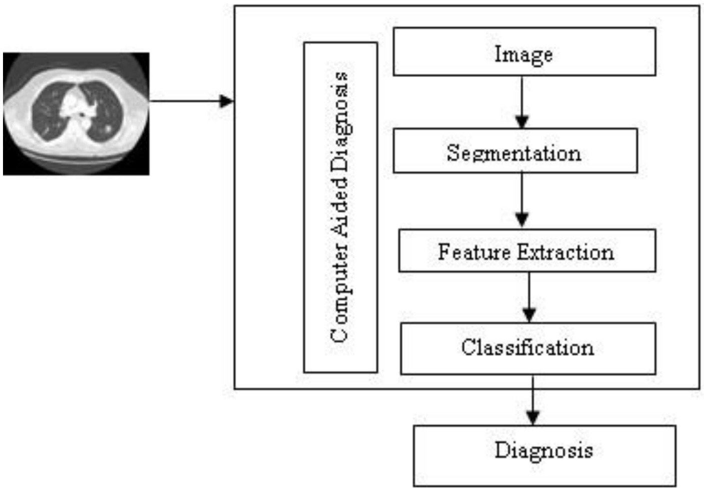


Figure 1: Computer-Aided Diagnosis System

The lung region is extracted by the image pre-processing techniques. Then a novel segmentation algorithm is used to obtain the region of interest (ROI). Once the region of interest is identified the features are extracted from the ROI and finally classification is illustrated by the case of cancer detection where the region is classified as either cancerous or non-cancerous. For that, the proposed system presented one segmentation algorithm to segment tumour effectively from the CT image.

METHODOLOGY

The proposed segmentation algorithm formed by Firefly Search algorithm and Fuzzy C-Means algorithm is shown in Figure 2.

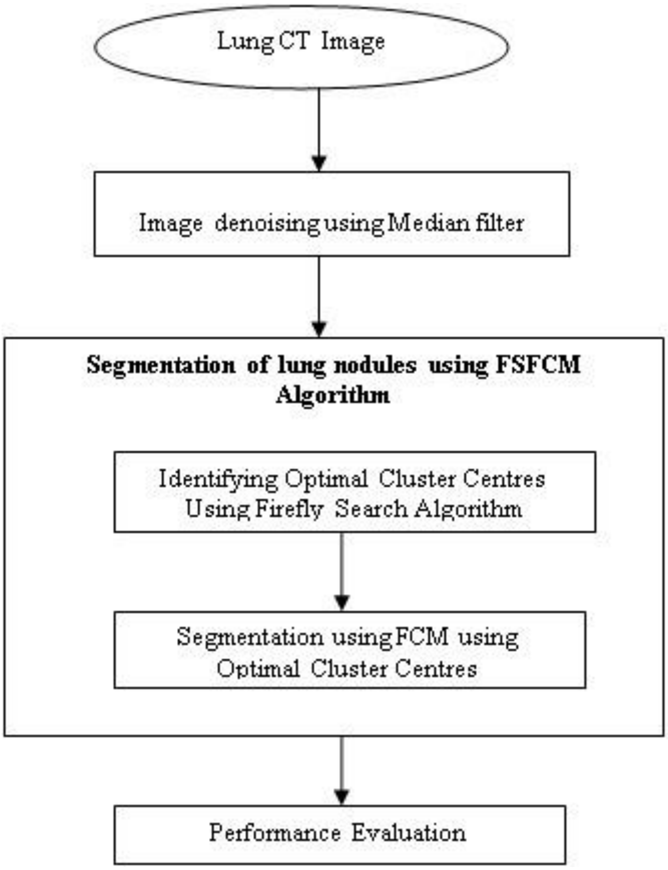


Figure 2: Flow of proposed segmentation algorithm

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. The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features important for further processing[5]. The noise present in the input CT image is reduced by employing a median filter given in Equation,

$$I'(x,y) = median_{3x3}\,[I(x,y)]$$

Where, I(x,y) represents original image and I'(x,y) represents the filtered image. The input CT image is shown in Figure 3(a) and the filtered image using a 3×3 median filter is shown Figure 3(b).

(a) Original CT image

(b) Denoised Image

Figure 3: Performance of Median Filter

The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing. Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise. Image denoising technique is used to remove noise artifacts[6].

SEGMENTATION

Image segmentation is the process of partitioning an image into regions by grouping together neighbourhood pixels based on the some predefined similarity criterion. The similarity criterion can be determined using specific properties or features of pixels representing objects in the image. In other words, segmentation is a pixel classification technique that allows the formation of regions of similarities in the image[7].

Firefly Search- Fuzzy C-Means (FSFCM)

A clustering approach Firefly Search with Fuzzy C-Means (FSFCM) is proposed that consists of two phases. In the first phase, the Firefly Algorithm examines the search space of the given data set to determine the near-optimal cluster centers and then those cluster centers give initial mean value for FCM algorithm[8]. In the second phase, the finest cluster centers that have been identified by the Firefly algorithm which are employed as the preliminary cluster centers for the FCM algorithm. Thus, the proposed clustering method consists of two phases as follows.

- 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 Firefly Algorithm.
- 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 given in equation,

$$A = S_1\{a_{1,}a_2,\ldots\ldots..a_d\}, S_2\{a_{1,}a_2,\ldots\ldots..a_d\}, S_3\{a_{1,}a_2,\ldots\ldots..a_d\}$$

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

Parameters Setting

The parameter setting of the firefly algorithm, (number of fireflies n = 110, max iteration =1000, β = 1 and γ =1) is carefully selected based on preliminary experiments, then the examination step of initialization phase will start and the solutions in every cluster centres will be initialized randomly.

Objective Function

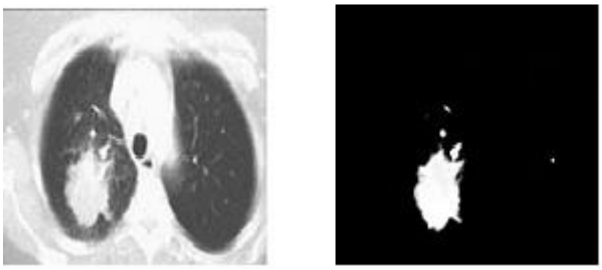
The fitness function is used to indicate a candidate solution either good or bad. The way of selecting the fitness function is a very significant matter in designing the proposed clustering algorithm, since the optimization solution and the performance of the algorithm count mainly on this fitness function[9].

In the proposed clustering algorithm, 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 we are comparing between two fireflies 'a' and 'b', if 'b' is brighter than firefly 'a', then firefly 'a' will move toward firefly 'b'. The proposed clustering algorithm was designed to enhance the performance of the traditional Fuzzy C-Means (FCM) in order to obtain more accurate segmentation process[10].

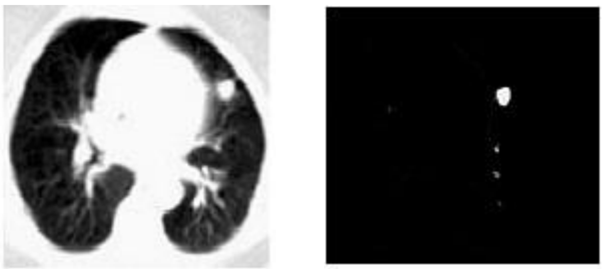
The hybridization step between Firefly Algorithm (FA) and Fuzzy C-Means is introduced to enhance the quality of the FA clustering results. The FCM have the ability to modify the cluster centres values till reaching the minimum variance, therefore obtaining more specific clusters. The pseudo code for the proposed system (FSFCM) is given below:

(Phase I)
begin
fireflies number (n)
Maximum number of generations (Max-Generation, iterations)
β, α and γ
Generate the initial population (n initial solutions) of the fireflies
$X_i = (1,2,3.............,n)$
The intensity of the light I_i *at* x_i *will be determined using the objective function value F(X)*
Determine the absorption coefficient γ
while (m <MaxGeneration)
for i = 1: n
for j= 1: n
if $(I_i>I_j)$
Move firefly 'i' towards 'j' in d-dimension
end if
Get attractiveness, which differs with distance 'r' via exp [-γr]
end for j
end for i
(Phase II)
Initialize FCM cluster center by new solution of FA
for t= to maxiterations do
if (abs (U(k+1)-U(k))<ε) then
break;
else
$U^{(k+1)}=U^{(k)}$
end if
end for
end

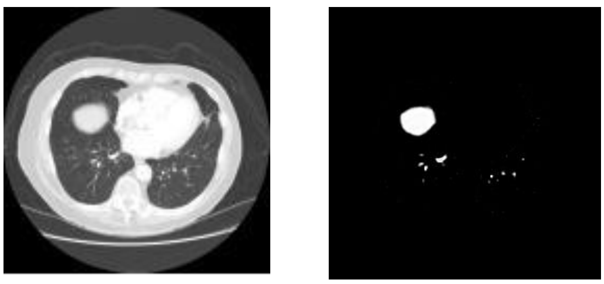
Segmentation is one of the key steps to empower a lung CAD algorithm to be successful. Without the segmentation, it is inefficient to extract the features over the entire CT lung image, which will cause too extensive computations and usually result in a poor classification[11]. The goal of segmentation is to find all suspicious regions that should contain as many cancers (benign or malignant) as possible, whereas the false positives will be excluded with a trained classifier using the additional features extracted from the suspicious segments. The final tumour segmentation (ROI) is shown in Figure 4.



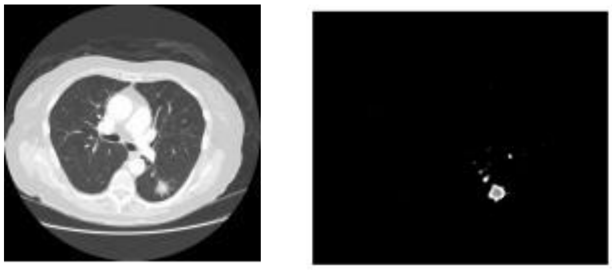
(a)Original CT Image 1 and its tumour segmentation



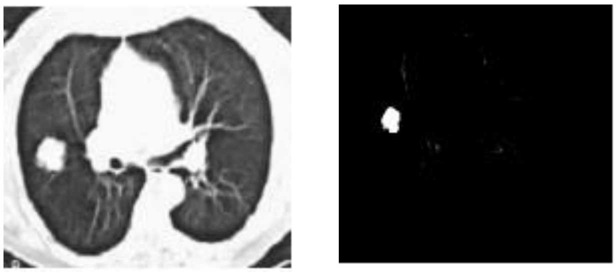
(b) Original CT Image 2 and its Tumour Segmentation



(c) Original CT Image 3 and its Tumour Segmentation



(d) Original CT Image 4 and its Tumour Segmentation



(e) Original CT Image 5 and its Tumour Segmentation

Figure 4: Performance of the proposed methodology (FSFCM)

PERFORMANCE EVALUATION

The important properties like Solidity, Perimeter and Area can be used to find the segmentation accuracy[12]. Solidity is a scalar value which specifies the proportion of the pixels in the convex hull which are also in the region, it is computed using the equation given by,

$$Solidity = \frac{A_s}{H}$$

Where, 'A_s' is the shape region and 'H' is the convex hull area of the shape. Area is the actual number of pixels in the region of interest. Convex Area is a Scalar that specifies the number of pixels in the Convex Image[13]. Convex Image is a Binary image (logical) that specifies the convex hull, with all pixels within the hull filled in is computed using equation given by,

$$Area = \sum_{i}^{n} P_i$$

Where 'P_i' is the pixels in the region of interest

Perimeter is the distance around the boundary of the region. Region props compute the perimeter by calculating the distance between each adjoining pair of pixels around the border of the region.

These three values are obtained for segmented images using FSFCM and Ground truth images. The ground truth image has been formed in discussion with an oncologist so that the lung is correctly marked. The percentage of segmentation is relative with respect to the ground truth image. The segmentation accuracy is calculated using the below equation to evaluate the quality of the proposed methodology.

$$A_s = \frac{S_{prop}}{O_{prop}} X100$$

Where S_{prop} is defined by,
= $(Solidity+Area+Perimeter)_{Ground\ Truth}$
O_{prop} is defined by,
= $(Solidity + Area +Perimeter)_{Ground\ Truth}$

This proposed methodology FSFCM is able to effectively segment lung region and tumour region from the lung CT images. The results are obtained is shown in Table 1 for five CT lung images and the performance output is shown in Figure 5 also Figure 6 shows the segmented lung nodules along with its corresponding ground truth images.

Table 1: Segmentation Accuracy for CT lung images

Image		Solidity	Area	Perimeter	Total	Accuracy in %
Image 1	Segmented	0.76535	43412	1541.4398	44954.2052	97.81115
	Unsegmented	0.76535	44216	1743.4398	45960.2052	
Image 2	Segmented	0.679	36759	1832.9068	38592.5858	95.09173
	Unsegmented	0.679	38553	2030.9068	40584.5858	
Image 3	Segmented	0.71355	30596	1307.0235	31903.7371	95.79303
	Unsegmented	0.8376	31897	1407.0235	33304.8611	
Image 4	Segmented	0.8376	42361	1151.918	43513.7556	95.44042
	Unsegmented	0.679	43759	1832.9068	45592.5858	
Image 5	Segmented	0.8915	37671	1198.918	38870.8095	95.7584
	Unsegmented	0.679	38759	1832.91	40592.5858	
Average Accuracy in %						**96.03408**

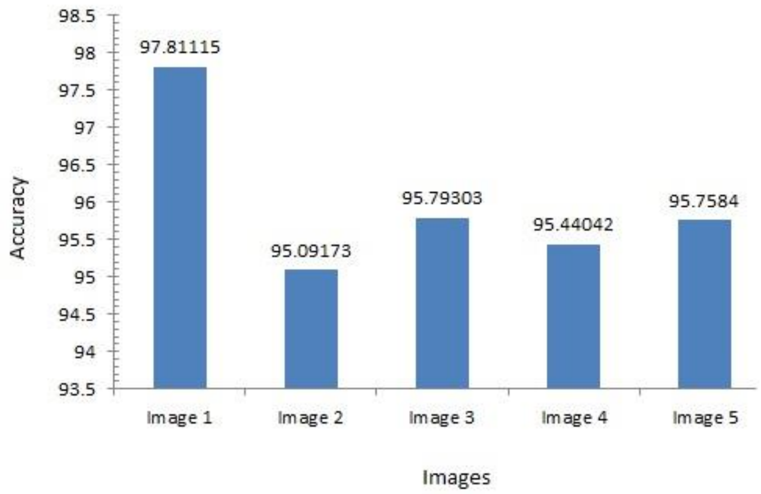
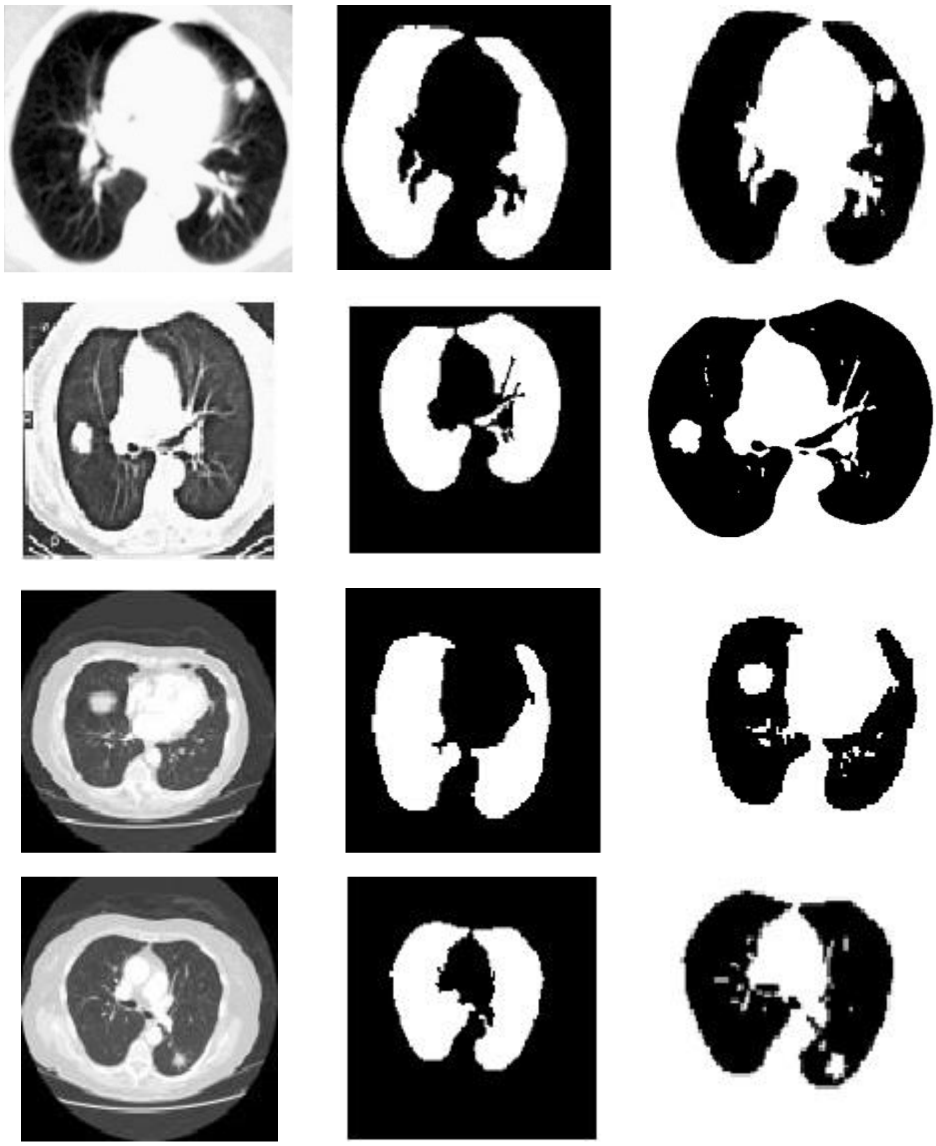


Figure 5: Output performance of CT lung images using FSFCM



(a) Original Image **(b) Ground Truth** **(c) Segmented Lung**

Figure 5: Segmentation Performance for various images along with its ground truth using the FSFCM

The properties like Solidity, Perimeter and Area is been used for measuring the segmentation accuracy. From the Table 1, it can be observed that the segmentation accuracy of the proposed methodology is 96.03% outperforms the state-of-art algorithm.

CONCLUSION

In this work, an automatic CAD system has been developed for the early diagnosis of lung cancer. The segmentation methodologies FSFCM and FSFMKM 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 96.03% has been achieved by using the FSFCM. The proposed scheme reduces the computational complexity of CAD system without compromising the performance of the system.

The difficulty in the early detection of lung cancer nodules 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 chapter 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. Shaik P S, Kavitha C. A review on computer aided detection and diagnosis of lung cancer nodules. International Journal of Computer Technology 2012; 3:393-400.
2. Hu S, Hoffman E A, Reinhardt J M. Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. IEEE Transactions on Medical Imaging 2001; 20:490-498.
3. Sluimer I, Prokop M, Van Ginneken B. Toward automated segmentation of the pathological lung in CT. IEEE Transactions on Medical Imaging 2005; 24:1025-1038.
4. Chonglun L. A new automatic seeded region growing algorithm. Proceedings of 6th International Conference on Image and Signal Processing (CISP) 2013; 543-549.
5. Gomathi M, Thangaraj P. A computer aided diagnosis system for lung cancer detection using machine learning technique. European Journal of Scientific Research 2011; 5l: 260-275.
6. Jiangdian S, Caiyun Y, Li F, Kun W, Feng Y, Shiyuan L, Jie T. Lung lesion extraction using a toboggan based growing automatic segmentation approach. IEEE Transactions on Medical Imaging 2016; 35:337-353.
7. Venkatasalam K, Rajendran P. Effective RBIR Fuzzy C-Means segmentation HAAR wavelet with user interactive multi threshold robust features vector. Asian Journal of Information Technology 2016; 15:223-231.
8. Gong M, Liang Y, Shi J, Ma W, Ma J. Fuzzy C-means clustering with local information and kernel metric for image segmentation. IEEE Transactions on Image Processing 2013; 22:573-584.
9. Saeid N, Amin A N, Saeid H, Abdolreza S, Farhad S. Optimization of KFCM clustering of hyper spectral data by particle swarm optimization algorithm. International Journal of Humanities 2013; 20:101-120.
10. Alomoush WS, Abdullah S S, Hussain R. Segmentation of MRI brain images using FCM improved by firefly algorithm. Journal of Applied Sciences 2013; 14:61- 66.
11. Faizal Khan Z, Kannan A. Intelligent approach for segmenting Computer tomography lung images using fuzzy logic with bit plane. Journal of Electrical Engineering Technology 2014; 9:742-750
12. Ganesh S, Harsha B. Lung segmentation and tumor identification from computer tomography scan images using support vector machine. International Journal of Scientific Research 2014; 3:1-6.
13. Mai M, Ayat K, Amr S. Support vector machine based computer aided diagnosis system for large lung nodules classification. Journal of Medical Imaging Human Informatics 2013; 3:214-220.

Cite this article as:

Amudha K *et al*. Performance analysis of firefly search fuzzy c-means (FSFCM) for detecting lung cancer nodules. Int. Res. J. Pharm. 2017;8(9):88-94 http://dx.doi.org/10.7897/2230-8407.089163

Source of support: Nil, Conflict of interest: None Declared