Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 4058 - 4066
Received 16 February 2021; Accepted 08 March 2021.

Mammographic Image Segmentation using Automatic Evolutionary Algorithms

V. Sesha Srinivas[1]

Department of Information Technology RVR&JC CE, College of Engineering Guntur, India

vangipuramseshu@gmail.com

B. Satish Babu[2]

Department of Information Technology RVR&JC CE, College of Engineering Guntur, India

satishrvr@gmail.com

R S M Lakshmi Patibandla[3]

Department of IT, Vignan's Foundation for Science Technology and Research, Guntur, India

patibandla.lakshmi@gmail.com

ABSTRACT:

Breast cancer is noticed as the most common cancer and second main cause of cancerous deaths among women population in world wide. Many studies publicized that the early detection of breast cancer can cause longer survival rate. With the advent of new technologies in mammography, abnormality of masses can be effectively detected and diagnosed. Such Computer Aided Diagnosis and Detection(CADx, CADe) systems are automated and semi-automated systems help radiologists in early detection and diagnosis of breast cancer. The common phases of a CAD system are segmentation, feature extraction and classification. Efficient segmentation certainly can influence the subsequent phases of these systems. Clustering is a most common machine learning method in many segmentation applications including mammogram segmentation. Most of the segmentation methods are based on pixel intensity values and usually the segment with maximum intensity values is the region of interest. But in practice the complete segment may not contain abnormality. Abnormal mass is generally hard. This paper proposes a two phase evolutionary based segmentation combined with feature extraction for detection of regions of interests in mammograms using recent Automatic Clustering with simultaneous Feature Subset Selection for gray scale Image segmentation using Differential Evolution (ACFSDE).The proposed method has two phases; in the first phase suspected region is determined using automatic evolutionary intensity based segmentation. From the region hard mass area is determined using spatial information based segmentation. The second phase extracts textural features of each pixel by constructing GLCM which follows ACFSDE. Experiments are conducted on each image of MIAS database. The results demonstrated the accuracy and efficiency of the algorithm in identifying the masses of mammograms and the results are validated with ground truth.

Keyword: Evolutionary Algorithm, Mammograms, Segmentation

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 4058 - 4066
Received 16 February 2021; Accepted 08 March 2021.

1. Introduction

Mammography is becoming the most reliable method that help radiologist in early detection of abnormalities and treatment planning [1]. Several Computer aided detection methods are available in literature for the detection and classification of mammograms [2-7]. In all such methods image segmentation is an important issue. Segmentation plays an important role in wide variety of applications including remote sensing and medical imaging. Image segmentation is the process of grouping the pixels of image space into homogenous regions, with respect to specific characteristics. Real world applications may involve multiple characteristics. Mammogram segmentation can be posed as one of such problems, involves multiple characteristics.Clustering methods are one of the most commonly used techniques for image segmentation[8].

Image segmentation is an important characteristic in computer vision and machine learning. Segmentation plays an important role in wide variety of applications including remote sensing and medical imaging. Image segmentation is the process of grouping the pixels of image space into homogenous regions, with respect to specific characteristics. Segmentation is also used to find suspicious masses of the mammograms and helps radiologist for early detection and therapy planning.

Clustering methods are one of the most commonly used techniques for image segmentation [8]. K-means is the simple, efficient partition clustering and one of the top 10 clustering algorithms from past 50 years [9 - 11]. Sahiner et al used k-means for mass segmentation[12-13]. Li et al incorporated spatial information using adaptive thresholding[14-15]. The Fuzzy C-Means (FCM) is a soft clustering algorithm in which each element is associated to each cluster using a fuzzy membership [16]. Velthuizen, Chen and Lee used FCM with different objectives to find homogeneous regions with respect to grey-level values [17-18].

The increasing dimensionality with huge image size and number of images from different sources, is became a curse to data analysis and knowledge discovery and for clustering algorithms. The high dimensionality decreases the performance of the algorithm [19]. Such high dimensionality risk can be reduced using feature selection algorithms [20]. Feature subset selection is a process of selecting a good sub set of features by avoiding irrelevant features. Most of the feature selection algorithms are developed for supervised learning, rather than the unsupervised learning. It is believed that the unsupervised feature selection is more difficult due to the absence of class labels that can guide the search for the relevant information [21]. Recently, several algorithms have been proposed to address this issue for clustering and researchers are focusing on the concept of feature subset selection for clustering. Feature selection for clustering chooses a trivial subset of tangible features as of the data and formerlytracks the clustering algorithm merely on the certain features [22]. Srikrishna et al, 2013 has proposed an algorithm for Automatic Feature Subset Selection using Genetic Algorithm for clustering.

Evolutionary algorithms simulate the evolution across a sequence of generations/iterations of a population, a set of candidate solutions. A candidate solution is a vector, internally represented as a string of genes and is called chromosome or individual. Mutation and Crossover are two frequently used operators referred to as evolutionary approaches [23]. Differential Evolution (DE) is any of the furthermostprevailing stochastic real-parameter optimization algorithms in modernusage [24]. DE monitorsanalogous computational phases as in any customary evolutionary algorithm with specified operations such as crossover and mutation. Associated to

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 4058 - 4066
Received 16 February 2021; Accepted 08 March 2021.

former Evolutionary Algorithms DE is precisemodest to program. The topical studies on DE have exposed that DE offers anenhanced performance associated to former algorithms [25]. An Automatic Clustering using Differential Evolution (ACDE) algorithm [26] by Das, Ajith, in 2008, and applied on image segmentation [27]. Sanghamitra Bandyopadhyay et al.,remainedprogression of clusters exhausting point symmetry process. They have recycled a point symmetry centered cost function as unbiased function [28]. Medical image segmentation using genetic algorithm is demonstrated [29]. A DE based Automatic Fuzzy (Fuzzy ACDE) clustering is proposed by Das, Amith, 2009, by incorporating fuzzy concept in ACDE. K. KarteekaPavanet al., have proved the efficiency of ACDE in finding tissues in medical images [30]. A multi-objective particle swarm optimization (MOPSO) method remains proposed [31], for the problem of image segmentation. Suspected regions of the mammograms are detected by proposing a new micro-genetic algorithm with a texture proximity mask in 2015 [32].

2. Methodology

Image segmentation is a process of partitioning an image into homogeneous groups with respect to a specific characteristic i.e., intensity, textural information, shape etc. As most of the clustering methods for segmentation not consider spatial information and the complete region of interest may not contain hard mass. The proposed segmentation algorithm for mammograms contains 2 phases using Differential Evolution. Phase one is the preprocessing step that include, preprocessing, finding a segment containing higher intensity values. Phase2 follows texture based segmentation on the selected brightest segment from the first phase. Various textural features are extracted for each pixel by constructing GLCM. Appropriate features only considered for segmentation by applying feature extraction while clustering using Automatic Clustering with simultaneous Feature Subset Selection for gray scale Image segmentation using Differential Evolution (**ACFSDE)**. The steps of the proposed algorithm arepresented in the following figure Fig.1

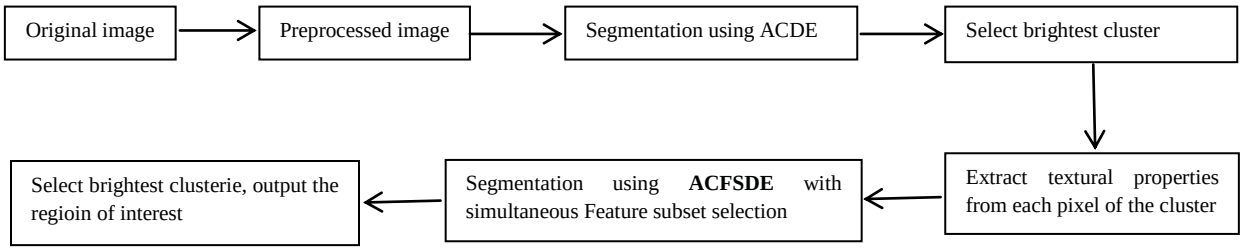


Fig1. Proposed Algorithm

The proposed two phase evolutionary segmentation for ROI of mammogram is as follows.

Step 1: After removing noise, artifacts, pectoral muscle etc., from each mammogram, apply automatic evolutionary intensity based segmentation using Automatic Clustering using Differential Evolution (ACDE).

Step 2: Find suspected area by selecting a segment with higher intensity value pixels.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 4058 - 4066
Received 16 February 2021; Accepted 08 March 2021.

Step 3: Construct GLCM for each pixel by selecting a neighborhood window of size 5x5 for each pixel.

Step 4: Generate a database, D, by extracting various textural features for each pixel.

Step 5: Reduce dimension of D by feature subset selection and find region of interest with simultaneous automatic evolutionary segmentation using ACFSDE.

Step6: Output the brightest segment.

ACFSDE is a recent variant to ACDE and is explained in the following section.

3. Automatic Clustering Simultaneous Feature Subset Selection of Gray Scale Images Using Differential Evolution (ACFSDE)

3.1 Feature Extraction

Being the clustering algorithm the EISDR is applied on images for segmentation of images. Here the work focuses on texture image segmentation. 16 textural features are extracted from each pixel of the image by constructing four different Gray Level Co-occurrence Matrices (GLCMs) of different orientations with 5x5 surrounding neighborhood window. A statistical scheme of exploratoryconstancy that reflects the spatial correlation of pixels is the gray-level co-occurrence matrix (GLCM), which is called as the gray-level spatial dependence matrix. The steps are as follows.

Step 1) For a given mxn input image construct Dataset of size [(mxn)x16]

1.1 for each pixel, consider a 5x5 neighborhood window and construct GLCM of specific orientation (space, direction, angle).

1.2 Extract four textural features, contrast, correlation, energy and homogeneity.

1.3 Repeat the two steps 1.1 and 1.2 for four times with different set of orientations of each GLCM.

Step 2) Apply EISDR on Dataset of size (mxn)x16

Step 3) Form the segmented image using the best chromosome.

The four textural features extracted from GLCM are Contrast, homogeneity, energy, and entropy described as follows.

Contrast: Approaches of the local variations in the gray-level co-occurrence matrix.

$$Contrast = \sum_{n=0}^{Ng-1} n^2 \sum_{|i-j|=n} P_d(i, j) \qquad (1)$$

Correlation: Methods of the joint probability existence of the stated pixel pairs.

$$Correlation = \frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Ng}(1-\mu_i)P_d(i, j)}{\sigma_i \sigma_j} \qquad (2)$$

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 4058 - 4066
Received 16 February 2021; Accepted 08 March 2021.

Energy: Delivers the sum of squared elements in the GLCM. Similarlyidentified as homogeneity or the pointednextinstant.

$$Energy = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng} P_d^2(i,j) \quad (3)$$

Homogeneity: Actions the intimacy of the dispersal of elements in the GLCM to the GLCM diagonal.

$$Homogeneity = \sum_{i=1}^{Ng}\sum_{j=1}^{Ng} \frac{Pd(i,j)}{1+|\ i-j\ |} \quad (4)$$

3.2 Chromosome Representation

For increased accuracy of clustering in high dimensional data the paper proposes a new Automatic Clustering Simultaneous Feature Subset Selection of gray scale Images using Differential Evolution (ACFSDE). The algorithm defines a new chromosome for optimal features and for optimal clusters.

A dataset (X) of size m x d one of the input where m is the number of pixels present in the input image and d is the features to be extracted (16), and another input to the algorithm is the maximum number of clusters (K_{max}). The chromosome is a vector of real numbers of dimension 16 + K_{max} + K_{max} × 16. The chromosome contains a set of sixteen thresholds to represent the active features, a set of K_{max} thresholds to represent the active centroids and a set of K_{max}centroids. The threshold entries are positive floating point numbers in [0,1], each of which controls the respective feature/ centroid is to be considered or not. The fi^{th} feature of the dataset is selected if T_{fi}>0.5 and j^{th} cluster center is selected for partitioning if $T_{c,j}$>0.5.

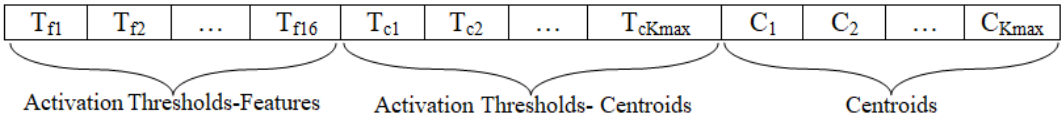


Figure 2: Chromosome Structure

3.3 Algorithm

The new ACFSDE is to invent optimal clusters spontaneouslythroughcertain subset of features. In this work Rand Index is used as the objective utility. The algorithm for the ACFSDE is as follows. Let X is a given data set with m elements each of with d dimensions.

Step 1: Initialize each chromosome to contain d + Max_k (randomly chosen) activation thresholds in [0, 1], and Max_k number of randomly selected centroids.

Step 2: To determine the Active Features in each chromosome which is having threshold value >0.5

Step 3:To determine the Active Cluster Centers which is having threshold value >0.5

Step 4: Conclude the centroids from each result setby active features only.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 4058 - 4066
Received 16 February 2021; Accepted 08 March 2021.

Step 5: for t = 1 to t_{max} do

a) Estimate its distance from all active cluster centers by using Euclidean distance metric for each data element X_i with chosen features.

b) Allocate X_i to closest cluster.

c)Estimate each resulteminence using Rand Index.

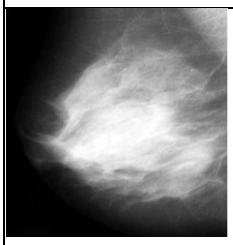
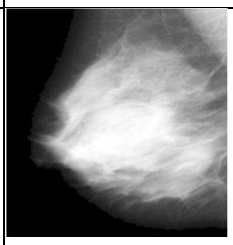
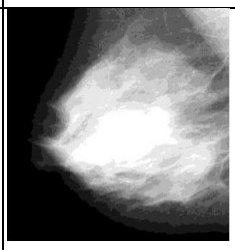
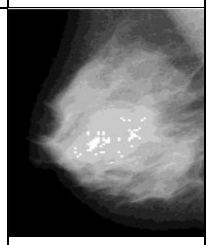
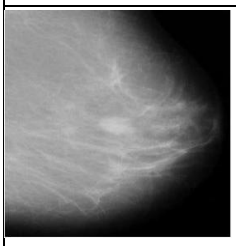
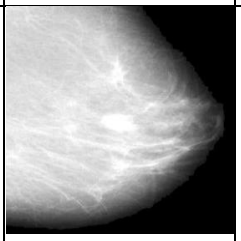
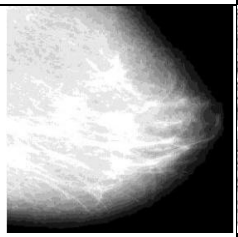
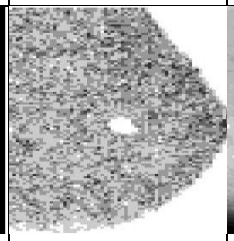
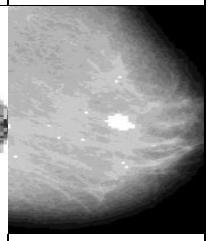
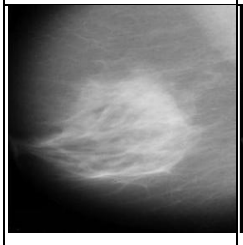
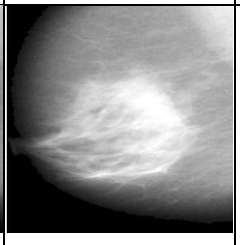
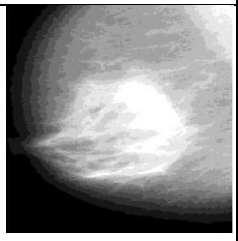
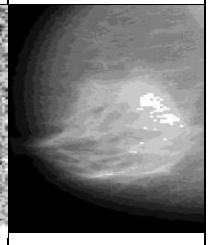
d) Update the populace by relating mutation and crossover operations as defined in DE algorithm which is described in the section 2.

Step 6) State the concludingresultattained by the globally best chromosome (one yielding the highest value of the fitness function) at time t = t_{max}.

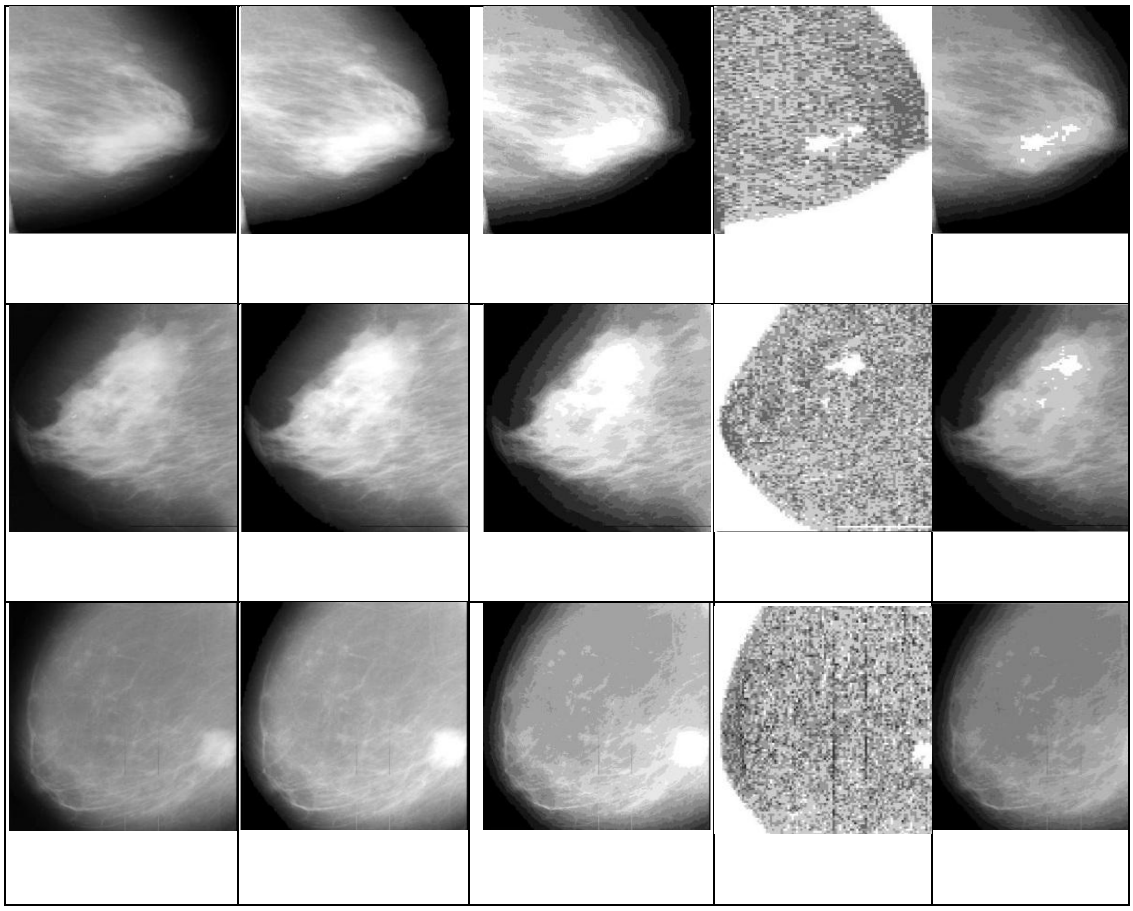
4. Experimental Results

The experiments are conducted on each image of MIAS database. The results in each phase of the proposed algorithm are reported in Table1.

Table 1: Results at different phases of the algorithm on MIAS dataset

Original	Preprocess	Intensity Based Clustering	Texture Based Clustering	ACFSDE Clustering
				
				
				

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 4058 - 4066
Received 16 February 2021; Accepted 08 March 2021.



According to the above table results our proposed algorithm projected the accurate results compared to existing ones.

5 Conclusions

A two phase automatic evolutionary segmentation is proposed using ACFSDE to find suspicious masses in the mammograms. The experimental results demonstrate the efficiency of the proposed algorithm in formative masses accurately from mammograms.

References

1. Jiang, J., Yao, B., &Wason, A.M. (2007). "A genetic algorithm design for micro calcification detection and classification in digital mammograms", Computerized Medical Imaging and Graphics, vol 31, 49–61.
2. Claudio Marrocco, Mario Molinara, CiroD'Elia,& Francesco Tortorella. (2010). "A computer-aided detection system for clustered microcalcifications", Artificial Intelligence in Medicine, vol 50, 23–32.
3. BrijeshVerma, Peter McLeod, & Alan Klevansky. (2009). "A novel soft cluster neural network for the classification of suspicious areas in digital mammograms", Pattern

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 4058 - 4066
Received 16 February 2021; Accepted 08 March 2021.

Recognition, vol42 , 1845 – 1852.

4. R S M Lakshmi Patibandla, Veeranjaneyulu,N.(2020), "A SimRank based Ensemble Method for Resolving Challenges of Partition Clustering Methods", Journal of Scientific & Industrial Research,Vol. 79, pp. 323–327.
5. Haralick , RM., Shanmugam, K., &Dinstein, I. (1973)."Textural features for image classification", IEEE Trans Syst Man Cybern ,vol 3,issue:6, 610–621.
6. Chang, HD., Shi , XJ., Min , R., Hu , LM., Cai , XP.,& Du , HN. (2006)." Approaches for automated detection and classification of masses in mammograms", Pattern Recognit ,vol 39, 646–668.
7. Ke, L., Mu, N.,&KangY, Y. (2010). "Mass computer-aided diagnosis method in mammogram based on texture features" Biomedical engineering and informatics (BMEI), 3rd international conference, IEEE explore, 146–149)
8. Jain, A.K., Murty, M.N., Flynn, P.J., 1999. Data clustering: a review. ACM: Computing Surveys 31 (3), 264–323.
9. MacQueen, J.B., 1967. Some methods of classification and analysis of multivariate observations. In: Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 281–297.
10. Wu, X., V. Kumar, J.R. Quinlan, J. Ghosh, D.J. Hand and D. Steinberg et al., 2008. Top10 algorithms in data mining. Knowl. Inform. Syst. J., 14: 1-37.
11. A.K. Jain., 2010. "Data Clustering: 50 Years Beyond K-Means" , Pattern Recognition letters, 31, 651-666
12. Sahiner, B., Chan, H.P., Wei, D., Petrick, N., Helvie, M.A., Adler, D.D., Goodsit, M.M., 1996. Image feature selection by a genetic algorithm: application to classification of mass and normal breast tissue. Med. Phys. 23, 1671–1684.
13. Sahiner, B., Chan, H.P., Petrick, N., Helvie, M.A., Goodsit, M.M., 1998. Computerized characterization of masses on mammograms: the rubber band straightening transform and texture analysis. Med. Phys. 25 (4), 516–526.
14. Li, L., Qian, W., Clarke, L.P., 1997. Digital mammography: computer assisted diagnosis method for mass detection with multi-orientation and multiresolution wavelet transform. Acad. Radiol. 4 (11), 724–731.
15. Li, L., Qian, W., Clarke, L.P., Clark, R.A., Thomas, J.A., 1999. Improving mass detection by adaptive and multiscale processing in digitized mammograms. In: Proceedings of SPIE, vol. 3661, pp. 490–498.
16. Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York.
17. Velthuizen, R.P., 2000. Computer diagnosis of mammographic masses. In: Workshop on Applied Imagery Pattern Recognition, pp. 166–172.
18. Chen, C.H., Lee, G.G., 1997. On digital mammogram segmentation and microcalcification detection using multiresolution wavelet analysis. Graph. Models

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 3, 2021, Pages. 4058 - 4066
Received 16 February 2021; Accepted 08 March 2021.

Image Process. 59 (5), 349–364.

19. Guyon, S. Gunn, A. Ben-Hur, and G. Dror, 2005, Result analysis of the NIPS 2003 feature selection challenge. In Advances in Neural Information Processing Systems (NIPS) 17, 545–552.
20. H. Liu, H. Motoda, and L. Yu, 2004. Selective Sampling Approach to Active Feature Selection. J. Artificial Intelligence, vol. 159, nos. 1/2,pp. 49-74.
21. M. H. C. Law,M. A. T. Figueiredo, and A. K. Jain, 2004. Simultaneous feature selection and clustering using mixture models. IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 9, 1154–1166.
22. Srikrishna, A., B. Eswara Reddy, and V. Sesha Srinivas, 2013. Automatic Feature Subset Selection using Genetic Algorithm for Clustering. Int J Recent Trends Eng Tech, vol. 9, 85-89.
23. Back, T., Fogel, D. B., &Michalewicz, Z, 1997. Handbook of evolutionary computation. New York: Oxford.
24. Swagatam Das, P. Nagaratnam Suganthan, 2011. Differential Evolution: A Survey of the State-of-the-Art. IEEE Transactions On Evolutionary Computation, vol. 15(1), 4-32.
25. Swagatam Das, Ajith Abraham and AmitKonar, 2009. Metaheuristic Clustering. Springer-Verlag Berlin Heidelberg. ISBN 978-3-540-92172-1, ISSN 1860949X.
26. Swagatam Das, Ajith Abraham, 2008. Automatic Clustering Using an Improved Differential Evolution Algorithm. IEEE Transactions On Systems, Man, And Cybernetics—Part A: Systems And Humans, vol.38(1), 218-237.
27. Swagatam Das, Amit Konar, 2009. Automatic image pixel clustering with an improved differential evolution. J. Applied Soft Computing, 9, 226–236.
28. Sanghamitra Bandyopadhyay and Sriparna Saha, 2008, A Point Symmetry-Based Clustering Technique for Automatic Evolution of Clusters. IEEE Transactions on Knowledge and Data Engineering, vol.20, no.11, 1441-1457.
29. Ujjwal Maulik, 2009. Medical Image Segmentation Using Genetic Algorithms. IEEE Transactions On Information Technology In Biomedicine, Vol. 13, No. 2, 166-173.
30. K. KarteekaPavan, V. SeshaSrinivas, A. SriKrishna, B. Eswara Reddy, 2012. An Automatic Tissue segmentation in Medical images using Differential Evolution. Journal of Applied Sciences,vol.12, issue.6, 587-592.
31. Melgani, F., Pasolli, E, 2013. Multi-objective PSO for hyperspectral image clustering. Computational Intelligence in Image Processing, Springer, Berlin, 265–280.
32. Wener Borges de Sampaio, Aristófanes Corrêa Silva, Anselmo Cardoso de Paiva, Marcelo Gattass, 2015. Detection of masses in mammograms with adaption to breast density using genetic algorithm, phylogenetic trees, LBP and SVM. Expert Systems With Applications, Elsevier, 8911–8928.