IJCTA, 9(40), 2016, pp. 67-75

Segmentation of Mammogram Image using Multilevel Threshold and Gravitational Search Algorithm

T. Velmurugan* and M.P. Sukassini**

Abstract : Segmentation plays an imperative role in image processing. It splits the image into separate objects. Each object reveals information in the form of intensity, color or texture. Segmentation techniques differ from image to image depending upon the problem analysis of the image. Biomedical images are gray level images in which the intensity of pixels varies for each object present in it. Threshold is applied on biomedical images to extract the anomalous regions where the pixel value is high. Two level threshold partitions an image into two classes, whereas multilevel threshold partitions an image into multiple classes depending upon the threshold level. Automated selection of threshold is still a difficult task. Preprocessing or post processing of an image improves the quality of image by removing the noise present in it. Gravitational Search Algorithm (GSA) is a heuristic optimization method to find the best optimum solution for a problem. In this work,a new approach to segment abnormal regions of mammogram image is carried out using gravitational search algorithm method with multilevel threshold. The preprocessed image is used for segmentation. The preprocessing is done using morphology and median filter. The preprocessed image is applied as input to gravitational search algorithm along with threshold level. The pixels in the image are clustered based on the best value obtained from gravitational search algorithm. The execution time of GSA segmentation method is evaluated.

Keywords: Multilevel threshold, Gravitational Search Algorithm, Morphology, Median Filter, Mammogram Image Analysis.

1. INTRODUCTION

The growth of science and technology has changed the natural life of human. The productivity of foods is increased by using many chemicals which is harmful to human. The new hybrid seeds are introduced which is cultivated along with chemicals. The pesticides are used for pest control during farming. Human takes in all this harmful foods. The deficiency of proteins and minerals are replaced by artificial supplementary. This lead to different types of diseases in human by reducing the immunity in human body. The rate of growth of disease has increased as the new technology emerges. Cancer is the common disease found in recent years. Uncontrolled growth of cells in any parts of the body can become cancer and spreads to other parts also. Many scientific methods have been developed to identify the presence of cancer in human at the early stage and by regular medical checkup. Breast cancer is common in women which are analyzed by mammogram image. The breast is made up of glandular, fatty and connective tissue. Figure 1 shows the anatomy of the breast.

The breast has lobules which is glands that produce milk. The tubes that carry milk from the lobules to the nipple are duct. Fatty and connective tissues that surrounds and protects the ducts and lobules gives

* Associate Professor, PG and Research Department of Computer Science, D. G. Vaishnav College, Chennai, India *E-Mail: velmurugan_dgvc@yahoo.co.in,*

** Research Scholar, Bharathiyar University, Coimbatore, India. *E-Mail: sukassini.dgvc@gmail.com*

the shape to the breast. The brown area which is circular around the nipple is areola contains sweat glands that secrete moisture during breast feeding. Area at the centre of the areola is nipple. Cancer in breast is developed when cells in the breast grow out of control. These unusual growths of cells form a tumor. The malignant tumor can grow fast and spreads to other area of the body. Breast cancer is commonly found disease in women above 40 to 60.

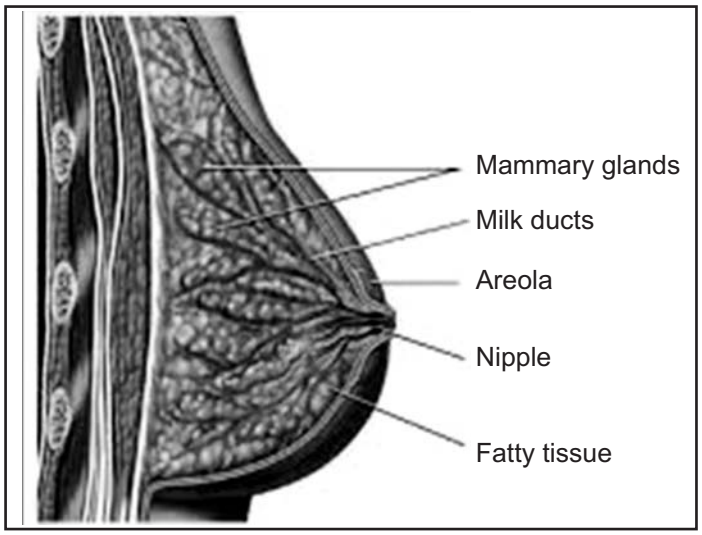

Figure 1: Anatomy of breast

The cancer in the breast can start from any parts of the breast. Mostly, it begins from the ducts which carry milk to the nipple. This is known as ductal cancer. The growth of cells inglands the produce milk form lobular cancer. The cancer growth in other tissues is known as sarcomas and lymphomas. Most of the cancer form lump in the breast. But, all the lumps need not be cancer. The lump is benign, that do not spread outside the breast. The lump that spreads to other areas is known as malignant. Mammography is the best screening method to detect the abnormality of breast at the early stage. It helps radiologist to find out the location, growth and shape of micro calcifications in the breast. It is an x-ray of low energy, around 30 kVp, penetrated into breast to capture the image. Mammograms image appears in the shades of gray, white and black that depends upon the density of the tissue.

In recent years, many techniques have been developed by many researchers to find out the presence of cancer and its growth. But still, the researchers progress their work to find out the better method to identify the cancer region in human. Combining more than one technique, also called as hybrid method to find out the cancer gives better performance. In this work, a new approach to segment affected regions in mammogram image using morphology, median filter, multilevel threshold and gravitational search algorithm is applied. The paper is organized as follows. Section II discusses about the related work of previous papers in the same topic. The methods and materials used for this research are discussed in section III. The experimental results of the work carried out for segmenting affected regions are explained in section IV. Finally, section V concludes the research work.

2. LITERATURE REVIEW

In recent years, many methods are proposed by researchers for segmenting image. The proposed methods are based on the combination of more than one technique. Researchers conclude that hybrid methods produce better results than traditional methods. The following section discusses about the various hybrid methods analyzed and applied by researchers. Rashedi et al.[1]proposed GSA algorithm and compared the performance with Particle Swarm optimization(PSO), Central Force Optimization(CFO) and Real Genetic Algorithm(RGA), GuptaChhavi and Sanjeev Jain[2] proposed threshold method based on entropy FCM and GSA. The results were assessed by computation time, entropy, peak signal to noise ration and

stability. The results showed that the performance of GSA was better than PSO. Tandanet al.[3] performed a study on PSO based image segmentation without knowing about the regions existing in the data.Ibrahim et al.[4] analyzed the applications of GSA in various fields to about the features of GSA. Chander et al.[5] presented a new variant of PSO for image segmentationusing optimal multi-level thresholding and also proposed an iterative scheme suitablefor obtaining initial values of candidate multilevel thresholds. The proposed PSO algorithmmakes a new role in adaptingsocial and momentum components of the velocity equationfor particle move updates. Majd Emadaldin Mozafari et al.[6] proposed hybrid method for segmentation of image. FCM is combined with GSA.

Li Linyi et al.[7] proposed fuzzy entropy image segmentation and fuzzy threshold by selecting the parameters using PSO. It showed efficient results. YeZhiwei et al.[8] developed aswarm intelligence based segmentation and 2-D Fisher criteria thresholding. The results showed improved search performance select optimum 2D thresholds with much less computation cost. Poli[9] analyzed the publications from IEEE Xplore based on applications of PSO algorithm. Das Swagatamet al.[10] proposed PSO based fuzzy clustering for image segmentation. Dheeba and S. Tamil Selvi [11] proposed swarm optimization neural network to detect the presence of micro calcification. Energy measures are extracted using Laws energy texture. SONN showed better performance than feedforward neural network and support vector machine. Dheeba J et al [12] proposed PSOWNN method to classify normal and abnormal breast tissues. The proposed method is compared with SONN and DEOWNN showed 94.167% of sensitivity, 92.105% specificity and 93.671% of accuracy.

Tan Khang Siang and Nor Ashidi Mat Isa[13] proposed a novel histogram – fuzzy c means hybrid for color image segmentation. The method showed low complexity but obtained better cluster quality and segmentation the ant colony optimization. Rashedi Esmatand Hossein Nezamabadi-pour [14] proposed a method to clustering image to segment using GSA.Sabri Norlina Mohd et al.[15] studied the performance of GSA. Shokouhifar Mohammad et al[16] Proposed FCM for segmentation of MRI image of brain. Initially noise probability is introduced and then the pixels are classified as normal or abnormal. The results show that the response time is reduced.Deepa and Aruna Devi [17] published a survey article on applications of artificial intelligence in medical image classification. ZyoutImad et al.[18] proposed a classification based on shape using PSO-kNN. The results showed better classification even shape is heterogeneous. Ramos-PollánRaúl et al[19] developed a machine learning classifier for breast cancer diagnosis. The features were extracted from CC and MLO views. The classifier was evaluated under ROC curve with 0.996.Venkatesan and T. Velmurugan[22] analysed the classification algorithm j48, Classification and Regression Trees (CART), Alternating Decision Tree (AD Tree) and Best First Tree (BF Tree) using data set of breast cancer. Performance result shows decision tree algorithm is better. Mahalakshmi, and T. Velmurugan[23] segmented MRI image of brain using particle swarm optimization. Different views of brain images are analyzed using different levels of segmentation. The result showed that the performance of PSO algorithm is better for the coronal plane than the axial plane.

3. MATERIALS AND METHODS

Segmentation is the process of fragmenting the image into various homogeneous parts based on common features. Preprocessing is the process of removing noise in the image which is present in image due to capturing device. Preprocessing or post processing is done on image based on the problem definition. The screening of breast at regular interval helps to identify the presence of cancer at the early stage. Mammography is the best method to identify the presence of abnormal regions in the breast. The image of the breast is captured at different views to obtain location of calcification with clear visual. Segmenting the abnormal regions through computer aided system assist the radiologist to confirm the presence of calcification and its growth.

A. Dataset

Segmentation technique applied in this work takes the input images from mini-MIAS database. MIAS contains 322 images with the combination of left and right breast. The images are categorized into three main groups as dense- glandular, fatty and fatty-glandular breasts based on the characteristics of breast tissue. The images are further classified into three major category based on the presence of calcification: malignant, benign and normal[21]. The images are digitized at 200 micron pixel edge and clipped (padded) in order to obtain all images with a size of 1024 × 1024 pixels in Portable Greymap (PGM) format. Each pixel in the images is represented as an 8-bit word, where the images are in grayscale format with a pixel intensity of range [0, 255].The mammogram image contains radio-opaque artifacts such as wedges and labels. In some cases it contains some patients personal information. The images are preprocessed to remove the wedges, labels and noise present in it during the acquisition.

B. Gravitational Search Algorithm(GSA)

Gravitational search algorithm (GSA) is a population based heuristic optimization method designed by Rashedi. It is theoretical substitution of Newton's law of Gravity and the law of motion. Newton's gravitational law states that every particle in the universe attracts other particles with force that is directly proportional to product of their masses and inversely proportional to distance between them. Newton's II law of motion which states that acceleration of an agent depends on its mass and total force applied on it. The searching agents interact with each other through gravity force. Attraction exists between all the masses, but the weaker ones are well pulled by heavy mass. Heavy mass, here represents efficient agent as it has better attraction capability which leads to better outcomes.

Each agent has gravitational mass which is calculated using fitness value given in Eq.1

$$\mathrm{M}_i(t) = \frac{m_i(t)}{\sum_{j=1}^{N} m_j(t)} \tag{1}$$

$$m_i(t) = \frac{fit_i(t) - worst(t)}{best(t) - worst(t)} \tag{2}$$

where fit *i* is fitness value of agent *i* and worst and best are the worst and best fitness value among all the agents. The position of agent *i* is defined by Eq.3

$$\mathrm{X}_i = (x_i^1 .... x_i^d .... x_i^n) \tag{3}$$

where *d* is the dimension of the agent. The velocity and position of an agent for next iteration is calculated by Eq.4 and Eq. 5

$$v_i^d(t+1) = r_i v_i^d(t) + a_i^d \tag{4}$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \tag{5}$$

where *r* is uniform random number which lies between [0-1]. The acceleration of the agent *i* by Newtons II law of motion is obtained by Eq.6

$$a_i^d(t) = \frac{\mathrm{F}_i^d(t)}{\mathrm{M}_i(t)} \tag{6}$$

The total force Fi applied on agent i by other agents is given by Eq.7

$$\mathrm{F}_i^d(t) = \sum_{j \in \mathrm{Kbest},\, j \neq i}^{\mathrm{N}} r_j \mathrm{F}_{ij}^d(t) \tag{7}$$

where r_j is a random number which lies between[0-1]. where N represents number of agents. Kbest represents agents which are heavier. Initially Kbest will have all agents, after execution of each step the value decreases and at the end Kbest will have only one value which is optimal solution. Force acting on mass M_i from mass M_j is defined as

$$F_{ij}^{d}(t) = G(t)\frac{M_i(t)M_j(t)}{R_{ij}(t)+\in}(x_i^d - x_j^d) \quad (8)$$

where G(t) is a gravitational constant, $R_{ij}(t)$ is the Euclidean distance between the ith agent and jth agent. G(t) is calculated using following

$$G(t) = G_0(t)X\left(\frac{t}{T}\right)^{\beta} \quad (9)$$

$\beta < 1$ and T is total number of iteration

C. Multilevel Threshold

Multilevel threshold is the process that segments a gray level image into several discrete regions. This method determines more than one threshold for the input image and partitions the image. The pixels in each discrete region belong to same class having gray levels within a specific range defined by multiple thresholds. It computes the probability occurrence of gray level. This technique works well for colored image.

Identification of affected areas in mammogram is a challenging task. Suspicious regions are extracted from other areas by applying some techniques. Abnormal regions has high intensity pixel[20]. GSA clusters the pixel which has high intensity value using the following steps

1. The input image is preprocessed using morphology and median filter. Morphology removes the unwanted information present in the input image such as patients name, age,etc. median filter is applied to remove the noise in the image.
2. Histogram is calculated from the preprocessed image.
3. Calculate the probability of gray level occurrence Prob(i) for i = 0 to level-1
4. Initialize the population, velocity, dimension. $G_0 = 100 = 100$, $\beta = 0.2$
5. Evaluate the fitness for each agent using equation 2
6. Update the best and worst of the population
7. Calculate the mass and acceleration for each agent using equation 1 and 6
8. Update the velocity and position using equation 4 and 5
9. Repeat the steps 5 to 8 until the maximum iteration is reached.
10. Each level of threshold computes the best values. These values are clustered.

Initially, level in step 3 is considered as 2. This will perform two level thresholds, as the feature of threshold is to separate the foreground and background objects in the image. Level one is not applicable as it will not show any significance change in image. The above steps are applied for many abnormal images that are taken from dataset.

4. RESULTS AND DISCUSSION

Segmenting image to extract hidden information is a crucial task. In biomedical image the abnormal region is segmented based on the intensity level of pixel. The high intensity regions represent the affected portion. Those regions are separated from other region to improve the visualization of affected region. By this method, the preprocessed image is taken as input data from our previous work. The original image is preprocessed by applying morphology and median filter. Different abnormal images are taken as input from MIAS database and preprocessed. The different levels of threshold are applied to preprocessed image along with GSA. Initial value of level is considered as 2. At level 2, two thresholds are applied to the input image. The image is segmented with two levels: one is background and the breast region as foreground. It does not show any significance result in the output image. The breast region has pixels of different intensity, but it is not visible at level 2. At level 3, the output image showed clusters of high

intensity regions in the input image. All the clusters are not affected regions as compared with input image. At level 4, the output image showed the significant result. The affected region alone segmented separately. Then the method is applied by assigning different values to levels such as 6, 7, 8 and 9. The output image obtained at different levels does not showed the considerable output. Figure 1 shows the input image and preprocessed image. In the preprocessed image the irrelevant labels are removed.

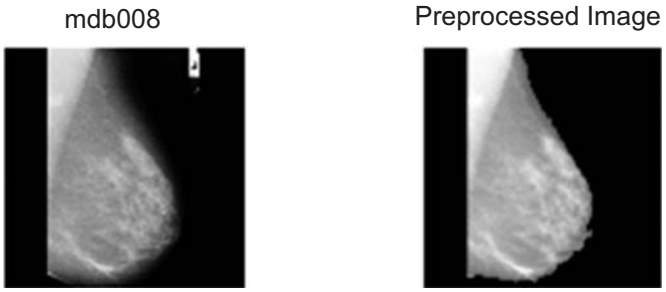

Figure 1: Input image mdb008 and preprocessed image

Figure 2 shows the segmentation of preprocessed image at different levels of threshold. At Level 4, output image shows the affected region better than other levels. Table 1 shows the execution time of input image mdb008 at different levels of threshold. Figure 3 is the graphical representation of execution time of mdb008 in seconds.

Table 1

Execution time for mdb008 image

Threshold level	L = 2	L = 3	L = 4	L = 5	L = 6	L = 7	L = 8
Run time(seconds)	8.479	9.473	11.01	12.35	13.44	16.04	16.37

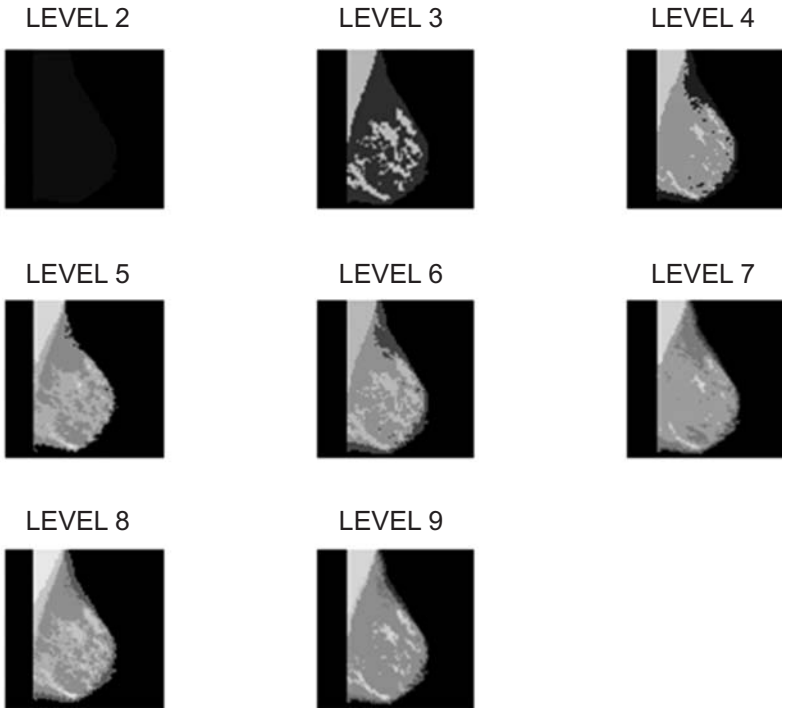

Figure 2: Segmentation at different levels of threshold

From the figure 2, the expected output is obtained at level 4. Different abnormal images are taken as input and GSA is applied on them with multilevel threshold up to level 5. The levels, execution time and average execution time are tabulated in table 2. The average execution time of all the images is shown as graph in figure 4. In table 2, the execution time increases as the levels of threshold increase. The average

runtime graph shows that image mdb015 has high execution time than other images. At level 4 and 5 the execution time of mdb015 is higher than other image.

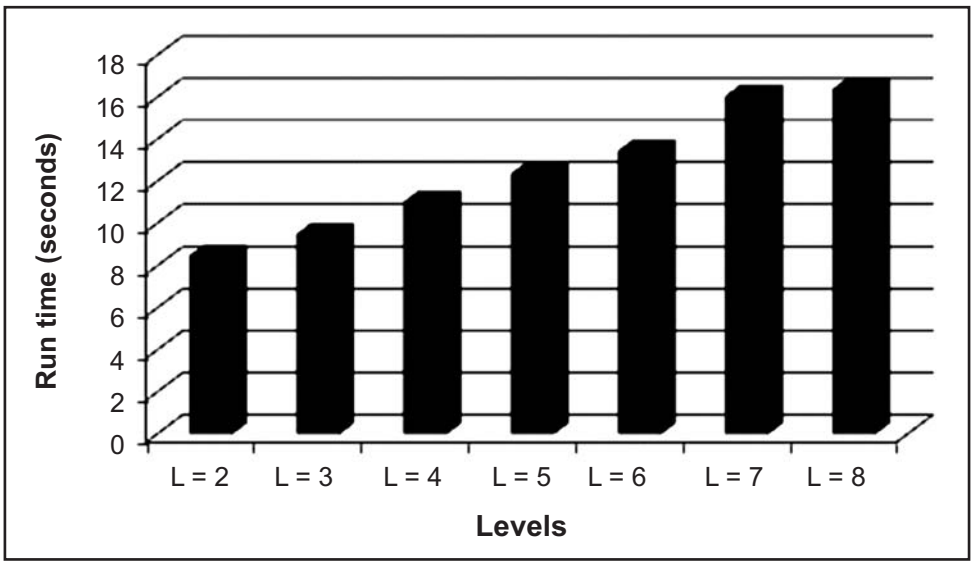

Figure 3: Runtime of GSA for mdb008

Table 2

Execution time of different images

Images	*Runtime at level 2*	*Runtime at Level 3*	*Runtime at Level 4*	*Runtime at Level 5*	*Average run time(s)*
Mdb001	8.346	9.687	10.88	12.31	10.305
Mdb005	8.221	9.703	10.96	12.27	10.288
Mdb015	8.284	9.750	11.02	13.47	10.631
Mdb021	8.268	9.751	10.90	12.32	10.309
Mdb320	8.393	9.828	11.02	12.34	10.395

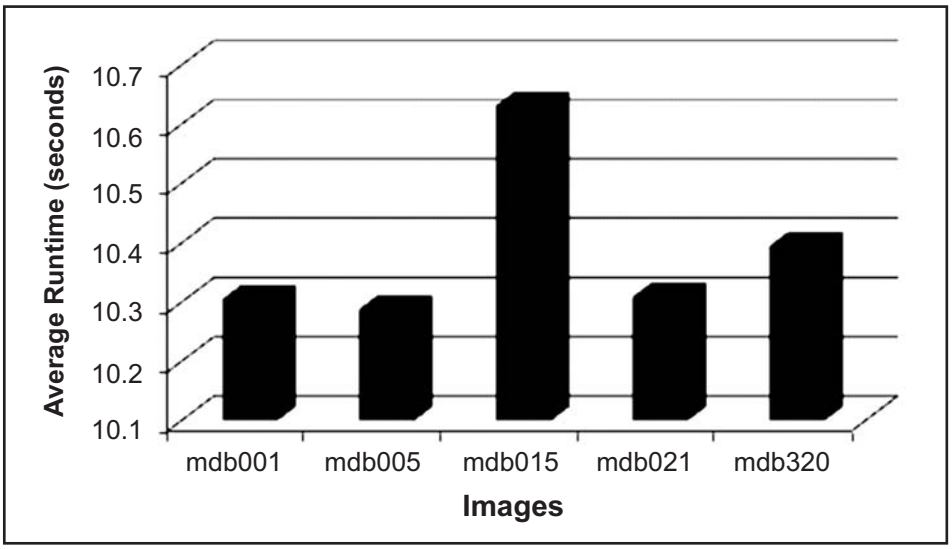

Figure 4: Runtime of GSA

5. CONCLUSION

Image processing is carried out to extract unknown information hidden in it. Images are processed according to the requirement of the problem domain. Processing image differs from image to image. The color images are processed based on the RGB component. Gray level images are processed on the intensity level of pixels. In the field of biomedical, processing the image plays an important role. Each step of processing explores some important information. Biomedical image is a gray level image. The high intensity values represent the affected regions in the image. Mammogram assists the radiologist to identify the suspicious regions of the breast. Preprocessing and segmenting the anomalous portions in the mammogram image, improvise the visualization effect of those regions. This research work is carried out to segment the unusual regions in mammogram by applying gravitational search algorithm at different levels of threshold. Two level threshold separated breast region from background. The affected regions are not visible as segment method extracted only the breast region from the background. In the levels from 3 to 5, most of the images are grouped the pixels that fall under the specified range in the threshold level. The affected regions with high intensities are well grouped. The segmentation at level 4 gives better visual result than the other levels. The execution time at each level is calculated and tabulated. This work concludes that segmentation at level 4 find the abnormal area better than other levels.

6. REFERENCES

1. Rashedi Esmat, HosseinNezamabadi-Pour and SaeidSaryazdi, “GSA: a gravitational search algorithm”, Information sciences 179, Vol. 13, pp. 2232-2248, 2009.
2. Gupta Chhavi and Sanjeev Jain, “Multilevel thresholding based on fuzzy C partition and gravitational search algorithm”, INFOCOMP Journal of Computer Science 13,Vol. 1,pp. 1-11, 2014.
3. Tandan Anita, Rohit Raja and YaminiChouhan, “Image Segmentation Based on Particle Swarm Optimization Technique”, International Journal of Science, Engineering and Technology Research (IJSETR) ,Vol. 2, 2014.
4. Ibrahim Tengku, TengkuNadzion, MarapanThiyagarajan, SaipolHadiHasim, Amar FaizZainalAbidin, Norhaizat Omar, NurAnisNordin et al,”A Brief Analysis of Gravitational Search Algorithm (GSA) Publication from 2009 to May 2013”,pp.16-24, 2014.
5. ChanderAkhilesh, AmitavaChatterjee and Patrick Siarry, “A new social and momentum component adaptive PSO algorithm for image segmentation”, Expert Systems with Applications, Vol. 5, pp. 4998-5004, 2011.
6. MajdEmadaldinMozafari, M. A. As’ ari, U. U. Sheikh and S. A. R. Abu-Bakar, “Hybrid image segmentation using fuzzy c-means and gravitational search algorithm”, In Fourth International Conference on Digital Image Processing (ICDIP 2012), International Society for Optics and Photonics, pp. 83342-83342, 2012.
7. Li Linyi and Deren Li, “Fuzzy entropy image segmentation based on particle swarm optimization”, Progress in Natural Science, Vol. 9, pp. 1167-1171, 2008.
8. Ye Zhiwei, Zhengbing Hu, Xudong Lai, and Hongwei Chen, “Image segmentation using thresholding and swarm intelligence”, Journal of Software, Vol. 5, pp. 1074-1082, 2012.
9. Poli Riccardo, “Analysis of the publications on the applications of particle swarm optimisation”, Journal of Artificial Evolution and Applications, 2008.
10. DasSwagatam, Ajith Abraham and AmitKonar, “Spatial information based image segmentation using a modified particle swarm optimization algorithm,” In Sixth International Conference on Intelligent Systems Design and Applications, Vol. 2, pp. 438-444, IEEE, 2006.
11. Dheeba J. and S. Tamil Selvi, “A swarm optimized neural network system for classification of microcalcification in mammograms”, Journal of medical systems , Vol. 5, pp. 3051-3061, 2012.
12. Dheeba, J., N. Albert Singh and S. Tamil Selvi. “Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach.” Journal of biomedical informatics, Vol. 49, pp. 45-52, 2014.
13. Tan Khang Siang and Nor Ashidi Mat Isa, “Color image segmentation using histogram thresholding–Fuzzy C-means hybrid approach”, Pattern Recognition ,Vol.1, pp. 1-15, 2011.
14. Rashedi Esmat and HosseinNezamabadi-Pour, “A stochastic gravitational approach to feature based color image segmentation”, Engineering Applications of Artificial Intelligence, Vol. 4, pp. 1322-1332, 2013.

15. Sabri Norlina Mohd, MazidahPuteh, and MohamadRusopMahmood, "A review of gravitational search algorithm", Int. J. Advance. Soft Computing, Vol. 3, pp. 1-39, 2013.

16. Shokouhifar Mohammad and Gholamhasan Sajedy Abkenar, "An artificial bee colony optimization for mri fuzzy segmentation of brain tissue", In 2011 International Conference on Management and Artificial Intelligence IPEDR, Vol. 6, 2011.

17. Deepa S. N. and B. Aruna Devi, "A survey on artificial intelligence approaches for medical image classification", Indian Journal of Science and Technology,Vol. 11,pp. 1583-1595, 2011.

18. Ramos-Pollán, Raúl, Miguel Angel Guevara-López, Cesar Suárez-Ortega, Guillermo Díaz-Herrero, Jose Miguel Franco-Valiente, Manuel Rubio-del-Solar, Naimy González-de-Posada, Mario Augusto PiresVaz, Joana Loureiro and Isabel Ramos, "Discovering mammography-based machine learning classifiers for breast cancer diagnosis", Journal of medical systems,Vol. 4, pp. 2259-2269, 2012.

19. ZyoutImad, Ikhlas Abdel-Qader and Christina Jacobs, "Embedded feature selection using PSO-kNN: Shape-based diagnosis of microcalcification clusters inmammography", JUSPN ,Vol.1, pp. 7-11, 2011.

20. Gonzalez Rafael C, Digital image processing, Pearson Education India, Third Edition, 2009.

21. Mammographic Image Analysis Society, http://www.wiau.man.ac.uk/services/MIAS/MIAS web.html

22. Venkatesan E and T. Velmurugan, "Performance analysis of decision tree algorithms for breast cancer classification", Indian Journal of Science and Technology, Vol. 8, pp. 1-8,2015.

23. Mahalakshmi and T. Velmurugan, "Detection of brain tumor by particle swarm optimization using image segmentation", Indian Journal of Science and Technology, Vol.8, pp. 13-19, 2015.

24. Sukassini and T. Velmurugan, "A Survey on the Analysis of Segmentation Techniques in Mammogram Images", Indian Journal of Science and Technology, Vol. 8, pp. 1-8,2015..