Segmentation of Mammogram Abnormalities Using Ant System based Contour Clustering Algorithm

Sudha Subramanian
Department of Computer Applications, RVS College of Engineering, India
drssudhasasikumar@gmail.com

Ganesan Rasu Thevar
Department of Electrical and Electronics Engineering, E.G.S. Pillay Engineering College, India
ganesanhod@gmail.com

Abstract: *Breast cancer is the most widespread cancer that affects females all over the world. The Computer-aided Detection Systems (CADs) could assist radiologists' in locating and classifying the breast tissues into normal and abnormal, however the absolute decisions are still made by the radiologist. In general, CAD system consists of four stages: Pre-processing, segmentation, feature extraction, and classification. This research work focuses on the segmentation step, where the abnormal tissues are segmented from the normal tissues. There are numerous approaches presented in the literature for mammogram segmentation. The major limitation of these methods is that they have to test each and every pixel of the image at least once, which is computationally expensive. This research work focuses on detection of microcalcifications from the digital mammograms using a novel segmentation approach based on novel Ant Clustering approach called Ant System based Contour Clustering (ASCC) that simulates the ants' foraging behavior. The performance of the ASCC based segmentation algorithm is investigated with the mammogram images received from Mammographic Image Analysis Society (MIAS) database.*

Keywords: *Mammogram, ant colony optimization, segmentation, ant clustering.*

Received June 6, 2021; accepted April 13, 2022
https://doi.org/10.34028/iajit/20/3/4

1. Introduction

Breast cancer causes highest incidence rate among women in most of the countries. Initially it seems to be an asymptomatic lesion of the breast, and then it may extend to the entire organ if untreated. Worldwide breast cancer statistics report that the incidence rate is kept on rising among all other cancers in women. In India, breast cancer is very common in urban areas, which contributes about 25-35% of all cancers in women. And most of them are either in 3rd and 4th stages, the most complicated stages, those who are in these stages are undeniably face the survival problem [4]. Initially breast cancer starts from milk ducts or the lobules and spreads across the breast tissues. The specific causes of breast cancer are yet to identify, however, there are some risk factors which might be the reason for developing the breast cancer. Some of the risk factors are menopause delay, heritage, hormone therapy and dietary factors. There are five stages of breast cancer, at every stage the severity of the disease increases while the survival rate decreases.

Mammography is the key screening method to identify the breast cancer at an early stage, however, the mammographic image consists of non-pathological structures. Literature studies shown that, 10%-25% of the masses are overlooked by the radiologists [5]. Mammographic abnormality exists in different types, in general, they are called as microcalcifications or masses. Microcalcifications are small granule like calcium deposits that are found either individual or clusters, and they are arbitrary in shape. Dixon [10] reported that mass detection is more challenging than locating microcalcifications, due to size and shape variation as well as poor contrast.

The practical difficulties like false-negative screening, visual tiredness and consistency of the radiologists have increased a great demand in developing Computer Aided Detection (CAD) system for mammographic analysis. The shape and size of the tumors are irregular in nature, which encounters the CAD systems difficult in mammogram analysis. The major objectives of the CAD system are to improve the classification accuracy as well as to reduce false positive rate. A typical CAD system consists of five stages: preprocessing, segmentation, feature extraction and selection and classification [1]. This research work focuses on segmentation step. The primary objective of mammogram image segmentation phase is to extract one or more regions of abnormalities from the background tissues [28]. Agrawal *et al*. [3] reported the following challenges for developing an efficient mass segmentation algorithm:

1. More efficient preprocessing algorithms are required to preprocess the images to remove noises.
2. It is necessary to develop robust segmentation algorithm which doesn't require eliminating the pectoral region as they share similar gray-level similar to mass regions.
3. More effective segmentation is required to reduce the false positive rate, which should not partition the

normal tissues as masses.

4. Numerous works have been proposed in the literature to improve the mammogram segmentation methods. However, it is still expected to explore hybrid intelligence to overcome the drawbacks of the reported segmentation methods. The clustering and contour based methods are integrated in this research work to efficiently detect the mammogram abnormalities.

The rest of the paper is organized as follows: the following section presents a comprehensive review on image segmentation methods. Section 3 presents an ant clustering based image segmentation approach. Section 4 discusses the proposed Ant System based Contour Clustering (ASCC) method. Section 5 presents the experimental setup and discusses the investigation results. The paper is concluded at section 6.

2. Related Work

An image segmentation problem could be viewed as an optimization problem, where it is expected to receive connected and well-separated regions. The metaheuristic algorithms are more suitable for optimization problems [18] and this research work utilizes an Ant System based Clustering Algorithm (ASCA) for image segmentation as it achieved more significant results as summarized in the following text. There are two types of Ant System (AS) based image segmentation algorithms were proposed in the literature. The first one is based on ants' foraging behavior, known as Ant Colony Optimization (ACO) algorithm, introduced by Dorigo *et al*. [11]. As ants forage for food, they deposit a chemical substance called pheromone on their path that connects the food and the nest, this real time behavior is simulated in ACO. The optimum path is constructed when more ants follow the shortest path where the concentration of the pheromone would be more. In ACO based image segmentation, the ants are allowed to move on 2D image grid, and assign each pixel to desired cluster based on the pheromone density. The second class of AS based segmentation simulates the cemetery building or brood sorting behavior of the real ants. In real time, the ants are used to clean their nest by collecting the dead ants, this process is known as cemetery building. Sometimes, they used to sort their larvae according to their size, known brood sorting. Deneubourg *et al*. [9] introduced the ant clustering method. Here the ant agents are allowed to randomly move around a 2D space, where an unloaded agent can pick an isolated object and carry it with random movement, or a loaded agent can drop its object to a location where it finds more similar objects. The agents are simulated to repeat this process of picking and dropping the objects would result in group of similar items or connected clusters.

2.1. Ant Colony Optimization Based Image Segmentation

Image segmentation through clustering is the method of partitioning the given pixels into number of clusters based on a similarity measure. This could be viewed as an optimization problem that has a couple of difficulties like, the high resolution images have larger search space, and the non-convex objective function may lead to huge number of local minima. Ouadfel *et al*. [31] proposed an Ant Colony System (ACS) hybrid with Markov Random Field (MRF) for image segmentation. This method is further extended with improved local search in Ouadfel and Batouche [30]. Here the segmentation results are compared with other stochastic optimization methods like Simulated Annealing (SA), and Genetic Algorithm (GA). The experimental results reported that ACS-MRF based image segmentation is more efficient than SA and GA based segmentation. Feng [13] presented a Maximum-Minimum (MAX-MIN) ant system for image segmentation and reported that ACS is outperforming GA and SA based segmentation. The pheromone update procedure used in conventional ACO algorithm is cost expensive. Wang *et al*. [37] and Yuanjing *et al*. [42] proposed a Finite Grade ACO (FGACO) to reduce the time complexity of ACO algorithm. Here the pheromones are ranked with three grades such as finite, higher and more, and the pheromone update is achieved through changing the grade rather than updating its value. FGACO is applied for active contour model based image segmentation, and the investigation results shown that FGACO achieved better segmentation results with medical images. Han and Shi [15] proposed an (ACO) based fuzzy clustering for image segmentation. Here the fuzzy membership function is estimated based on pixels' gray value, gradient and adjacency. The cluster centers are improved heuristically that improves the searching process. The simulation results indicate that ACO integrated with fuzzy based image clustering is significantly improving the segmentation results.

In general, ACO based segmentation is performed based on pheromone update, however, Ghosh *et al*. [14] proposed a new concept to pheromone aggregation that is stimulated from the behavior of ants to gather around points with denser pheromone trail. Laptik and Navakauskas [22] applied an ACO model for two-dimensional electrophoresis gel image segmentation. Here the time complexity is reduced by tuning the ACO parameters and an image pre-processing steps. This application is able to reach a maximum of 66% of accurate segmentation. Saatchi and Hung [34] proposed an ACO conjunction with Self-Organizing Map (SOM) based image segmentation. In addition to the conventional ACO algorithm, here the objective function and pheromone values are normalized at each step. The performance is

compared with conventional ACO and Simple Cooperative Learning (SCL) based image segmentation methods. The experimental results report that the normalization procedure enhances the segmentation performance.

Snake model based image segmentation procedure suffers with the issues like local minima, convergence speed, and concave boundaries. Li *et al*. [23] proposed an ACO with Gradient Vector Flow (GVF) snake model for addressing the local minima issue in medical image segmentation problem. The optimization function is added with crowded degree function that improves the ants' traversal while avoiding the local minimum. Also the reported results indicate that the ACO-GVF model based segmentation resolves the entire problem of snake model and achieved superior segmentation results. Zou *et al*. [46] proposed an ACO algorithm for image segmentation. Initially the image is partitioned based on binary thresholding, and for each binary region a different Ant model is applied to receive the segmentation results. The simulation results show that the binary image based ACO segmentation significantly improves the segmentation accuracy while reduce the time complexity. Hung and Sun [17] improved the k-means clustering based image segmentation using an ACO algorithm. Two different pheromone updating strategy: one with the spatial distance and other without using the spatial distance are proposed to enhance the ACO algorithm. The experimental results show that the segmentation performance is more acceptable. Abdullah and Jasim [2] applied an ACO for document image segmentation and reported a greater accuracy of 96.95% comparatively. Khorram and Yazdi [21] proposed an optimized thresholding method for brain image segmentation. Here, ACO is used to estimate the optimum threshold, where the texture features are adopted as heuristic information. The segmentation performance is compared with other stochastic optimization methods like Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC) optimization, and k-means algorithm. The results indicate the superior performance of the ACO algorithm.

2.2. Ant based Clustering Algorithms

Zhao *et al*. [44] proposed an improved ACO based image segmentation. The pheromone initialization and updates are performed based on k-means clustering which reduces the computation of ACO. The promising segmentation results indicate the significance of the ACO hybrid with k-means clustering based image segmentation. Yang *et al*. [39] proposed an Ant-Tree based fuzzy clustering method for image segmentation. In addition to pixel intensity, the gradient and neighborhood features are used as feature vectors to achieve better segmentation results. The clusters are initialized based on image histogram. The simulation results prove the efficiency of the Ant-Tree based fuzzy clustering method. Hao *et al*. [16] proposed an adaptive ACO hybrid with hierarchical clustering based image segmentation. The experimental results indicate that the hierarchical clustering based seeds and pheromone update improves the ACO based segmentation performance significantly. Jevtić *et al*. [20] proposed an ACSA based image segmentation with mammogram images. The segmentation performance of ACSA method is compared with SOM, k-means, Fuzzy C-Means (FCM), and Possibilistic fuzzy C-Means (PCM). The simulation results indicate the superior segmentation performance of the ACSA. Yu *et al*. [41] improved FCM clustering with PCM to make FCM method as noise insensitive. However, the quality of PCM clustering depends on initial parameters, so the PCM is hybrid with ACO to solve image segmentation problem. The greater segmentation accuracy reported in the results shown that ACOPCM is superior to FCM and PCM. Liu *et al*. [24] proposed an improved ant clustering for color image segmentation. Here the initial segmentation is received from Mean Shift (MS) algorithm and the result is represented as a graph. Further the problem is modeled as a graph partitioning problem and solved by using a Similarity Carrying Ant Model (SCAM). Simulation results indicate that SCAM based color image segmentation outperforms the conventional methods in terms of complexity and quality. Yan [38] proposed an ACO based FCM for remote sensing image segmentation. Here, the cluster initialization is achieved using ACO, then FCM is proceeded to extract the final segmentation. The experimental results show that ACO-FCM improves the segmentation performance significantly. The major limitations of FCM clustering algorithm are the time complexity and noise insensitive. Inkaya *et al*. [18] proposed a multi-objective ACO for a clustering problem. The simulation results are outperforming, however, the method is cost expensive as it involves an additional preprocessing step called neighborhood construction. Zou [45] focused to solve these issues by using an ACS. Here the clusters are initialized with ACO algorithm to improve the fuzzy clustering. The investigation results with Magnetic Resonance (MR) brain images reported that the ACO-FCM reduces the time complexity of FCM clustering and improves the segmentation quality through optimum seed points.

3. Ant Clustering Model for Image Segmentation

The basic idea of Deneuborug's *et al*. [9] ant clustering model is to pick up the isolated objects and drop them at some other location where more objects of that kind are present. For an illustration, consider that there is

only one type of object in the search space, the pick probability P_{pick} to pick up an object is computed as:

$$p_{pick} = \left(\frac{k_1}{k_1+f}\right)^2 \quad (1)$$

Where, f is the perceived fraction of objects in the adjacent of the ant and k_1 is a threshold factor between 0 and 1. Similarly, the drop probability p_{drop} for a loaded agent to drop the object is defined as:

$$p_{drop} = \left(\frac{f}{k_2+f}\right)^2 \quad (2)$$

Where, k_2 is another threshold factor between 0 and 1. Lumer and Faieta [25] extended Deneuborug's *et al*. [9] ant clustering model, where the similarity between the objects are estimated by using a neighborhood function is defined as:

$$f(i) = max\left(0, \frac{1}{\sigma^2}\sum_j\left(1-\frac{\delta(i,j)}{\alpha}\right)\right) \quad (3)$$

Where *a* is a threshold factor for the distance metric $\delta(i,j)$ between a picked object *i* and all the other adjacent objects *j*. It is suggested to restrict the neighborhood size as either 3×3 or 5×5. The pick and drop probabilities are defined based on neighborhood function, given as:

$$p_{pick}(i) = \left(\frac{k_1}{k_1+f(i)}\right)^2 \quad (4)$$

And

$$p_{drop}(i) = \begin{cases} 2f(i), & iff(i) < k_2 \\ 1, & otherwise \end{cases} \quad (5)$$

The k_1 and k_2 parameters are used to represent the amount of influence the f has over pick up and drop probabilities. For an object *i* when its similarity value is low with respect to k_1, then an unloaded ant is likely to pick up that object. Similarly, when the similarity value is high with respective to k_2, represents that the agent reaches the location where it finds more similar objects. Hence it is likely to drop the carrying objects. Subsequently, this procedure constructs the clusters in 2D space. In Lumer and Faieta [25] ant clustering model, each ant has a heap to store a list of recently visited objects. With that list, when an ant picks up an object, it can compare with the recently visited objects and choose to move towards the most similar item from the list. This process is known as matching search. The Deneuborug *et al*. [9] and Lumer and Faieta [25] models have formed the basic framework for more subsequent work. Once such work reported in Ouadfel and Batouche [29] is modified here in this research work. Ouadfel and Batouche [29] proposed AntClust, a novel ant-based clustering algorithm for image segmentation, which modifies the basic ant clustering model to improve the clustering quality and to reduce the time complexity.

In the basic ant clustering algorithms, the ants are allowed to walk around on a two-dimensional grid. The dimension of the grid is based on the size of the data. In general, with the smaller grid size, it is difficult for the ants to find a location to drop the object. In the other case, if the grid size is too huge, then the ants will be idle for a long time to pick up an object. Hence, the grid size should be optimum enough to avoid such time complexity. Ouadfel and Batouche [29] addressed this issue by converting the 2D space into one-dimensional space, here the data points are arranged into 1D vectors and each cell is connected to the nest of colony, so it is simple for the ants to move from one cell to another. With this 1D data points, the ants are allowed to create, build or destroy the clusters of pixels. Each cell in the 1D space could represent a cluster of two or more pixels, unlike the basic models which represents the clusters as spatial patterns. In the AntClust algorithm, initially the image pixels are transformed from 2D to 1D array, where each cell can contain only one pixel. At the first step, from K number of ants, each ant is allowed to choose a pixel at random and back to its nest. Then the clustering procedure starts, at this time the ants are allowed to move between their nest and the cells of the array. While visiting a cell, the ant can decided whether or not to drop the pixel to the current cell based on the probability, p_{drop}. Suppose, an ant drops the pixel, then the ant is free now, and it start searching for the other pixel to pick up from the list of unvisited or unloaded pixels. This iterative process could be terminated by fixing the maximum number of iterations. The similarity function f, for picking or dropping is pixel p_i from a cell c_k is defined as:

$$f(p_i, c_k) = \frac{1}{n_k}\sum_{p_j \in c_k}\frac{\alpha^2}{\alpha^2+\delta(p_i,p_j)^2} \quad (6)$$

Where $\delta(p_i,p_j)$ represents the distance measure, which estimates the intensity variation between the two pixels p_i and p_j in terms of gray level and is computed as:

$$\delta(p_i, p_j) = \frac{|g_i - g_j|}{NG} \quad (7)$$

Where g_i and g_i represents the intensity value of the pixels p_i and p_i respectively, and Number of Gray Levels (NG) denotes the maximum number of gray levels presents in the image. The symbol α represents the average distance between all pixels and is given as:

$$\alpha = \frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{j=1}^{N}\delta(p_i, p_j) \quad (8)$$

This could be calculated before initiating the clustering procedure. The similarity function f returns a maximum value when the distance is zero. The pickup and dropping procedure re explained below.

Picking up a pixel-when the ant is unloaded, it searches for a free pixel based on an index table which consists all unloaded pixels. For an unloaded ant, the following three situations are considered to pick up a pixel:

1. A cell contains only one pixel, and then the ant

picks it.
2. A cell which contains two dissimilar pixels, then the ant destroy this cluster and picks up a pixel based on probability q, a random number between [0, 1].
3. A cell of more pixels, where an isolated pixel is found with lower similarity to all other pixels, then the ant chooses that pixel.

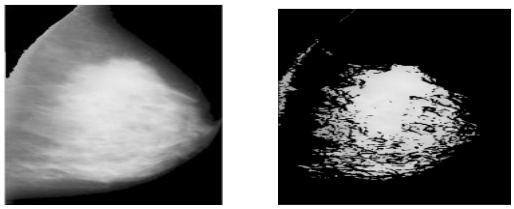
These three cases are mathematically represented as pick probability,

$$p_{pick}(p_i, c_k) = \begin{cases} 1 & if|c_k| = 1 \\ q & if|c_k| = 2 \\ cos^2\left(\frac{\pi}{2} f(p_i, c_k\right) & otherwise \end{cases} \quad (9)$$

Dropping a pixel-for a loaded ant, it searches for the similar cell where it can drop the pixel. As discussed in the basic ant clustering model, in AntClust algorithm also, a small list of recently visited cells are maintained for each ant. For dropping a pixel, the cells in the memory are considered and choose the most similar cell to drop the pixel. The drop probability is computed as:

$$p_{drop}(p_i, c_k) = 1 - cos^2(\frac{\pi}{2} f(p_i, c_k)) \quad (10)$$

Figure 1 depicts a typical segmentation output from AntClust algorithm.



a) Original mammogram image. b) Segmented image

Figure 1. Results from ant clustering.

4. Ant System Based Contour Clustering Model For Image Segmentation

The ant clustering algorithm is achieving significant results for image segmentation problem. However, the following characteristics show the inefficiency of them.

1. All the pixels of the image have to be either picked or dropped at least once, which increases the time complexity of the algorithm.
2. There is no optimal value to initialize the ant system parameters like number of ants, maximum number of iterations and thresholds employed in the similarity functions.
3. In specific, the AntClust algorithm, converts 2D image pixels into 1D, seems to be an additional step, which could be avoided as the segmentation algorithm has to detect the suspicious the region in the image grid.

The proposed ASCC based image segmentation simulates an interesting behavior of real time ants. In real-time, once an ant find the food source, then the other ants from the same colony are start surrounding the food source as shown in Figure 2. This point motivates this research work to extract the boundary of the surrounding region as a contour of the mass region to be segmented. The following text explains the segmentation procedure based on the proposed ASCC algorithm.

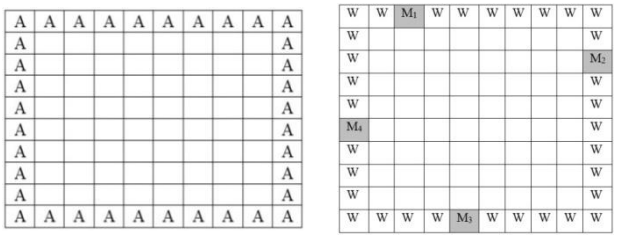
a) Ants surrounds the food source. b) Densed ants over the food source.

Figure 2. Ants Foraging for food

The proposed ASCC algorithm assumes that there is exactly only one abnormality region means that there are two kinds of tissues exist in the image, the background (normal) and the foreground (abnormal). ASCC algorithm, initially place one ant at the first and last row as well column of the image grid as shown in Figure 3-a), with the grid size as 10×10, each 'A' represents an ant agent. Hence, for an image dimension m×n, the ASCC algorithm initializes the number of ants (k) as:

$$k = 2m + 2n - 4 \quad (11)$$

It is noted that most of the ant based algorithms initialize the number of ants with some real number in random. In ASCC, the number of ants is computed based on the image dimension, which resolves the problem of optimizing the 'number of ants' parameter. From the initialized ants, 5% of ants from each row and column are chosen for foraging and the rest of the ants are kept in wait state. This is to avoid collision and the time complexity of the algorithm. The ants chosen for foraging are called as Marker ants (M) and the rest of the ants are called as Walker ants (W).



a) Sample image grid with initial ant position. b) Sample image grid with Marker and Walker ants' position.

Figure 3. Placement of ants.

Figure 3-b) illustrates a typical initialization of the image grid with marker and walking ants, for the demonstration, one ant from each row and column is chosen as a marker ant and rest of them are in idle state. The segmentation phase starts with the marker

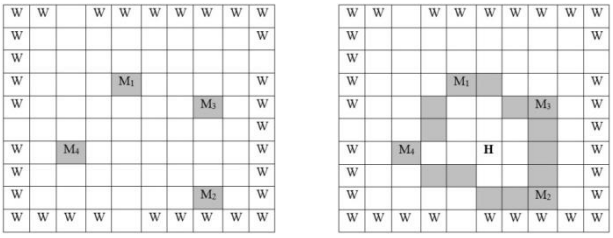
ants; they are allowed to search for a suspicious (abnormal) pixel in the image grid with random walk. Initially a marker ant is loaded with a pixel from its neighbor, the pixels' similarity with its neighbor is estimated as given in Equation (9), and the decision to drop the pixel is decided based on the drop probability as defined in Equation (10). If the drop probability is smaller than a random number, then the marker ant is allowed to search for the next pixel. Otherwise, the mean value of the surrounding neighbors are estimated and compared with a threshold value. For a pixel p, located at (x, y), the mean of its neighbor surrounded in 3×3 window of 9 pixels is computed as:

$$\mu = \frac{1}{9}\sum_{i=x-1}^{x+1}\sum_{j=y-1}^{y+1} p(i,j)) \quad (12)$$

If it is greater than the pixel will be marked as contour pixel, then the same procedure is repeated for the other ants in the marker set. If the mean value of the surround pixel is smaller than the threshold, then the marker ant drops the current pixel and chooses another pixel from its neighbor.

$$contour\_pixel = \begin{cases} 1 & if \mu \geq threshold \\ 0 & Otherwise \end{cases} \quad (13)$$

For example, after several iterations, the marker ants could be placed as shown in Figure 4-a).

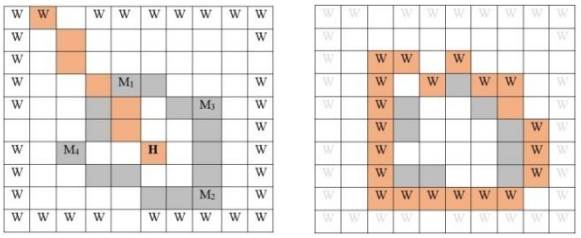


a) A typical output of selected contour pixels from marker ants.

b) A typical convex hull and its center.

Figure 4. Typical Ants' movement.

Once all the marker ants end up with the abnormal pixel, a convex hull is constructed with their current position, and the center of the hull is computed ($H_{x,y}$) as illustrated in Figure 4-b).The convex hull generated with the marker ants might generate a rough boundary over the mass region of the mammogram. The boundary constructed from the convex hull is going to be refined with the walker ants. At the second stage, the walker ants are resumed from the idle state and start moving towards the convex hull center $H_{x,}$. This movement of walker ant is a directed one, as they are restricted to travel through straight line that is established between the current spatial location of the walker ant and the hull center. The straight line is computed using the basic Bresenham's line drawing algorithm. While walking along the straight line, for every encountering pixel, the drop probability and the neighboring mean is estimated as similar to the procedure followed for the marker ants. The walk on the straight line is continued either they found a contour pixel or they reached the hull center. Once all the walker ants have completed their walk, then the convex hull is reconstructed with the current spatial locations of all the ants. And the contour of the convex hull would be the refined boundary defines the abnormal region of a mammogram image. Figure 5-a) depicts a typical path of a walker ant and their final position in Figure 5-b). It is possible that more than one walker ant might ends up at a same pixel location, hence there is a chance for overlapping.

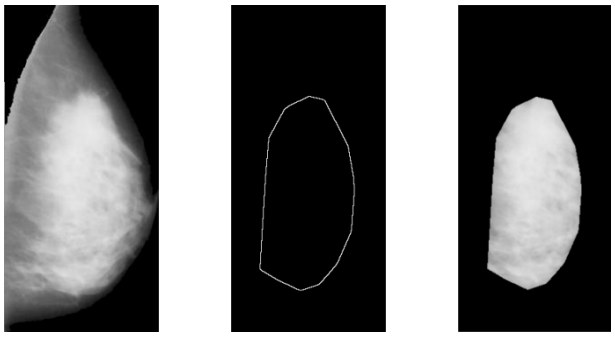


a) A straight line path for a walker ant.

b) Refined contour with the location of all walker ants.

Figure 5. Construction of segmentation boundary.

.



a) Breast region.

b) Contour of the suspicious region.

c) Segmented image.

Figure 6. Ant system-based contour clustering.

Figure 6 illustrates the proposed ASCC based mammogram segmentation result. The advantages of the proposed ASCC based image segmentation is summarized below.

1. The number of ant agent is estimated rather than initializing with the random number.
2. The termination condition for each ant is to end up with a contour pixel rather than fixing it with maximum number of iterations.
3. Only the drop probability is estimated for the pixels, the pick is based on the neighboring mean.
4. Definitely the ASCC algorithm doesn't evaluate each and every pixel in the image, and moreover the pixels are retained at their original position rather than moving them to form the clusters.
5. It is not necessary to assign a heap memory for each agent to remember their recent visits as specified in the basic ant clustering model from [9, 25] and the other models.

The following pseudocode summarizes the ASCC Algorithm (1) based mammogram segmentation.

Algorithm 1: Ant System based Contour Clustering

Input – Digital Mammogram Image
Output – Segmented Mammogram Image
/ Ant Initialization */*
Place each ant a_i at every border pixel p_i
Categorize the ants as Marker (M) and Walker (W) ants
For each marker ant
  Compute $f(p_i, c_k)$ and $p_{drop}(p_i, c_k)$
Select a random number R between 0 and 1
If$\left(R \leq p_{drop}(p_i, c_k))\right)$ then
    Move a_i to the next adjacent pixel p_i
Else
    Find the mean (μ) of the neighbors,
    If $\mu >$ th
    $p_c \leftarrow (x, y)\ of\ p_i$, contour pixel
      break;
    End If
  End If
End For
Construct the Convex Hull with the positions of M
Computer the Convex Hull center (H)
For each walker ant
  Draw a straight line from (p_i) to H
    / using Bresenham's line drawing algorithm */*
For each pixel in the straight line path
    Compute $f(p_i, c_k)$ and $p_{drop}(p_i, c_k)$
    Select a random number R between 0 and 1
    If$\left(R \leq p_{drop}(p_i, c_k))\right)$ then
      Move a_i to the next pixel in the path, p_i
    Else
    Find the mean (μ) of the neighbors,
    If$\mu >$ th
    $p_c \leftarrow (x, y)\ of\ p_i$, add p_i to the contour
      break;
    End If
End If
  End For
End For
Refine the convex hull with the current positions of W
Extract the segmented region surrounded by the contour pixels.

5. Experimental Setup and Results

The proposed segmentation algorithm is evaluated with the mammogram images received from the Mammographic Image Analysis Society (MIAS). This dataset consists of 312 mammograms acquired from 161 patients, for each of them the left and right breasts are screened [28]. The digitized mammograms are 8-bit gray-scaled images, captured with 200μm per pixel, with the dimension of 1024×1024 pixels. The segmentation performance of the proposed ASCC algorithm is evaluated with various performance measures, and compared against with the recent segmentation methods. The measures are grouped under three categories: area-, pixel, and edge-based measures.

5.1. Area based Metrics

The area based methods generally overlaps the segmented and the reference region and estimate the size of the common area between them. Jaccard index [19] is a metric used to measure the overlap between the segmented region (*S*) and the ground truth region (*R*). It is computed as:

$$JI = \frac{|S \cap R|}{|S \cup R|} \tag{14}$$

Sezgin and Sankur [36] proposed a Relative Foreground Area Error (RAE) to compare the shape and area between the reference and segmented regions. It is estimated as:

$$RAE = \begin{cases} \frac{|R|-|S|}{|R|} & if\ |S| < |R| \\ \frac{|S|-|R|}{|R|} & if\ |S| \geq |R| \end{cases} \tag{15}$$

Where $|R|$ and $|S|$ denotes the total number of pixels availabe in the reference and segmented regions respectively. A zero indicates the accurate segmentation, and the 1 denote the inexact segmentation. Also, the Goodness based on intra-region Uniformity (GU), and Goodness based on inter-region Contrast (GC) measures [43] are estimated to quantify the segmentation performance.

5.2. Pixel based Metrics

The pixel-based segmentation evaluation measures quantify the performance based on the number of correctly segmented and mis-segmented pixels from the background and foreground of the image. The pixel-based measures are presented below. Yasnoff *et al*. [40] proposed an area based metric, Segmentation Error (SE), that indicate the proportion of misclassified pixels to the whole ROI. SE is computed as:

$$SE = 1 - \frac{|B_S \cap B_R| + |M_S \cap M_R|}{|B_S + M_R|} \tag{16}$$

Where B and M denotes the background and the mass pixels, subscript S and R represents the segmented and reference (ground truth) regions respectively. The lower the error indicates better segmentation performance. But, this measure fails when the actual mass region is very small, though the segmentation would not be able to locate any of the mass pixels. To overcome these issues, Yasnoff *et al*. [40] proposed Distance Error (DE) which considers the spatial distance between the misclassified pixels and the actual location. In addition to that, Overlay Index (OI), Precision (Pr), Recall (Re), F-measures and Specificity (Sp) metrics are also used to analyze the segmentation performance [32, 35].

5.3. Edge based Metrics

The edge-based segmentation evaluation measures compare the goodness of the boundaries between the segmented and the reference regions. Edge based measures are used to evaluate the boundary between the reference and segmented region. Initially the edge pixels of both the regions are stored in two sets *A*

$=\{a_1,a_2, \ldots, a_n\}$ and $B=\{b_1, b_2,\ldots, b_3\}$, where *ai* and *bi* are the edge points. The minimum distance from the edge pixels of *A* set to *B* is computed as:

$$d(a_i, B) = \min_j \|b_j - a_i\| \quad (17)$$

The distance between two edges could be measured with Hausdorff Distance (*H*) measure [6] defined as:

$$H(A, B) = max\left(\max_i\{d(a_i, B\}, \max_j\{d(b_j, A)\}\right) \quad (18)$$

The Hausdorff distance measure quantifies the common edge dissimilarity between two boundaries. The under-segmented (UDI), over-segmented (ODI) pixel rates, Non-Uniformity (NU), Edge Mismatch (EM), Figure of Merit (FOM) metrics are also used to investigate the segmentation performance [27, 36, 43].

On the MIAS Dataset, the proposed ASCC method's segmentation performance is tested using area, pixel, and edge based metrics. On the same dataset as the proposed ASCC algorithm, ACO based segmentation [15, 21], Active Contour (AC) based segmentation [12], Level Set (LS) segmentation [33], k-means clustering [26], Threshold based Segmentation (TS) [8], and FCM clustering [7] are evaluated and compared.

To know the overlaps between the segmented and the reference region and to estimate the size of the common area between them, the area based metrics such as Jaccard Index (JI), RAE, GU, and GC are evaluated for the proposed ASCC algorithm along with other existing algorithms. Further, to quantify the segmentation performance based on the number of correctly segmented and missegmented pixels from the background and foreground of the image, the pixel based metrics such as SE, DE, OI, Pr, Re, F-Measure, and Sp are evaluated on the proposed and existing algorithms. Furthermore, to compare the goodness of the boundaries between the segmented and the reference regions, the proposed and existing algorithms are evaluated with edge based metrics such as Hausdorff Distance (H), ODI, UDI, NU, EM, and FOM. The following Table 1, quantifies the performance of segmentation with area based evaluation measures.

Table 1. Segmentation performance with area based metrics.

Methods	JI	GC	RAE	GU
ASCC[A]	0.6247	0.6659	0.1668	0.2557
ACO[B]	0.6237	0.6583	0.2592	0.3128
FCM[C]	0.6224	0.6015	0.3364	0.3586
LS[D]	0.6223	0.5857	0.3814	0.3689
AC[E]	0.5484	0.5746	0.3841	0.3712
TS[F]	0.5374	0.5669	0.4222	0.3859
k-means[G]	0.4784	0.5648	0.4366	0.3994

The segmentation performance comparison of the proposed ASCC algorithm with the other existing algorithms with respect to the area based metrics is depicted graphically in the Figure 7. The proposed ASCC algorithm outperforms the other segmentation methods with better quantitative rates for all the areas based measures compared to other algorithms mentioned.

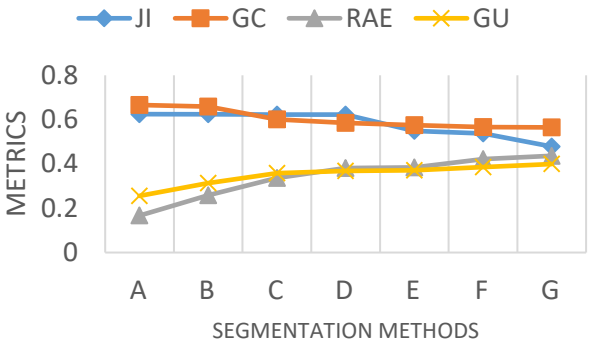


Figure 7. Performance comparison of segmentation methods with area based metrics.

The following Table 2, quantifies the performance of segmentation with pixel based evaluation measures. The segmentation performance comparison of the proposed ASCC algorithm with the other existing algorithms with respect to the pixel based metrics is depicted graphically in the following Figure 8. The proposed ASCC algorithm outperforms the other segmentation methods with better quantitative rates for all the pixel based measures compared to other algorithms mentioned.

Table 2. Segmentation performance with pixel based metrics.

Methods	SE	DE	OI	Pr	Re	F	Sp
ASCC[A]	9.51	0.23	0.75	93.67	90.90	0.84	89.63
ACO[B]	16.61	0.31	0.72	86.79	89.05	0.77	86.62
FCM[C]	18.23	0.33	0.71	81.31	83.88	0.75	86.19
LS[D]	19.67	0.38	0.64	78.70	79.77	0.41	85.70
AC[E]	26.10	0.40	0.55	73.23	71.95	0.34	84.63
TS[F]	29.33	0.42	0.50	70.42	69.03	0.33	82.92
k-means[G]	32.51	0.42	0.48	69.39	68.74	0.30	81.52

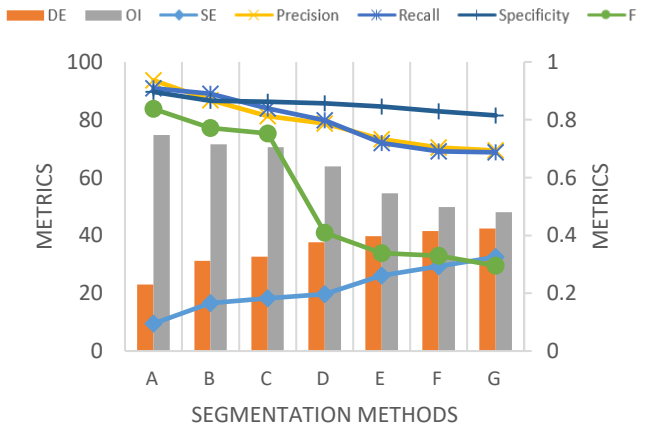


Figure 8. Performance comparison of segmentation methods with pixel based metrics.

The following Table 3, quantifies the performance of segmentation with edge based evaluation measures. The segmentation performance comparison of the proposed ASCC algorithm with the other existing algorithms with respect to the edge based metrics is depicted graphically in the following Figure 9. The

proposed ASCC algorithm outperforms the other segmentation methods with better quantitative rates for all the edge based measures compared to other algorithms mentioned.

Table 3. Segmentation performance with edge based metrics.

Methods	H	ODI	UDI	NU	EM	FOM
ASCC[A]	35.5640	0.0043	0.0026	0.0234	0.1864	0.8549
ACO[B]	75.8811	0.0104	0.0357	0.0413	0.2020	0.7933
FCM[C]	94.0456	0.0376	0.0364	0.0601	0.2181	0.7594
LS[D]	130.0218	0.0425	0.0685	0.0752	0.2831	0.7314
AC[E]	165.7160	0.0455	0.0742	0.0804	0.3434	0.5122
TS[F]	166.3824	0.0623	0.1163	0.0839	0.3618	0.2904
k-means[G]	197.1012	0.0747	0.1747	0.0898	0.3792	0.2231

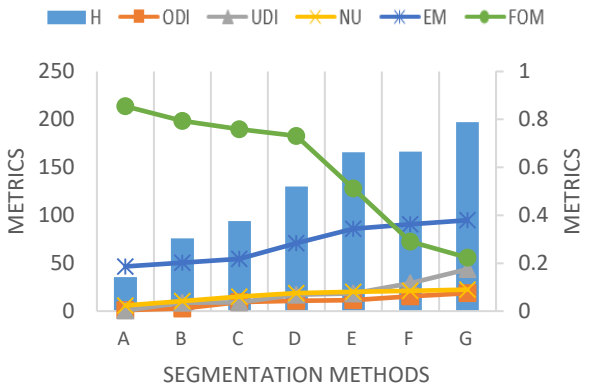
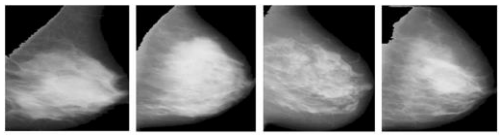
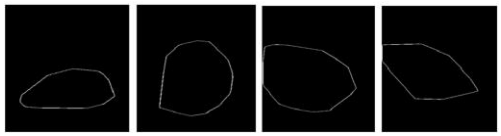


Figure 9. Performance comparison of segmentation methods with edge based metrics.

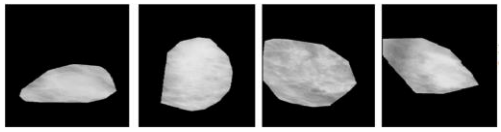
The qualitative typical segmentation outputs for the Ant Clustering algorithm are shown in the Figure 10. When it comes to the proposed ASCC algorithm, only the drop probability is estimated for the pixels, the pick is based on the neighboring mean and also the number ant agent is estimated rather than initializing with the random number.



a) Row-1: Original mammogram images.



b) Row-2: Contour of the suspicious region.



c) Row-3: Segmented output images.

Figure 10. Sample segmentation results from ASCC algorithm.

The segmentation performance of the proposed ASCC method is evaluated with the area, pixel and edge based metrics on the MIAS dataset. Table 1, shows the area based metrics with JI of 62.47%, Goodness based on intra-region Contrast (GC) of 66.59%, RAE of 16.68% and GU of 25.57%. All these edge based metrics showing superiority of the proposed ASCC algorithm compared to other existing methods. Table 2 shows the evaluation of the pixel based metrics for the proposed ASCC algorithm with SE of 9.51%, DE of 0.22%, OI of 74.70%, Pr of 93.67%, Re of 90.90%, F-Measure of 83.77% and Sp of 89.63%. These entire pixel-based metrics showing superiority of the proposed ASCC algorithm compared to other existing methods. Table 3 shows the performance of segmentation of the proposed ASCC method with the help of edge based measures. For proposed algorithm, the edge based metrics such as Hausdorff Distance (H) value of 35.56%, ODI of 0.43%, UDI of 0.26%, NU of 2.34%, and EM of 18.64% and FOM of 85.49% are obtained. These entire edge based metrics show the superiority of the proposed ASCC algorithm compared to other existing methods.

Existing approaches must process every pixel of an image, which is expensive. ACO can detect microcalcifications in digital mammograms. Ant clustering achieves remarkable results for image segmentation. Picking or dropping every pixel increases computing time. There is no best value for initializing ant system parameters such as number of ants, maximum iterations, and similarity function thresholds. The Ant Clustering technique, which converts 2-D pixels to 1-D, seems superfluous for image segmentation. For this study, the surrounding region's boundary will be employed as a contour to segment the mass region to be studied. The advantages of the proposed ASCC based image segmentation are: the number ant agent is estimated rather than initializing with the random number, the termination condition for each ant is to end up with a contour pixel rather than fixing it with maximum number of iterations, only the drop probability is estimated for the pixels, the pick is based on the neighboring mean. Definitely the ASCC algorithm doesn't evaluate each and every pixel in the image, and moreover the pixels are retained at their original position rather than moving them to form the clusters.

6. Conclusions

Detecting the abnormalities in digital mammogram is the key preprocessing technique for the accurate diagnosis. A novel Ant System based segmentation is proposed in this paper for efficient mammogram segmentation. The proposed ASCC is a hybridization of clustering and contour based segmentation approaches. Here the AS parameters like number of ants and the maximum number of iterations are computed rather than initializing them. And, the time complexity of the clustering process is reduced by investigating limited number of pixels rather than testing the entire pixels in the image grid as

conventional clustering methods do. The proposed ASCC method doesn't require any heap memory storage for the ants to remember their recently visited locations. With all these merits, the segmentation performance of the proposed ASCC is evaluated with the mammogram images received from MIAS database. The performance metrics and the comparative study show that the proposed ASCC based segmentation is more efficient than the other recently reported methods. The proposed segmentation algorithm is evaluated with only one benchmark dataset which could be considered as a limitation. It is planned to evaluate the same with more datasets and real-time images to improve the robustness of the proposed segmentation.

References

[1] Abbas Q., Fondo'n I., and Celebi E., "A Computerized System for Detection of Spiculated Margins Based on Mammography," *The International Arab Journal of Information Technology*, vol. 12, no. 6, pp. 582-587, 2015.

[2] Abdullah H. and Jasim A., "Improved Ant Colony Optimization for Document Image Segmentation," *International Journal of Computer Science and Information Security*, vol. 14, no. 11, pp. 775-785, 2016.

[3] Agrawal P., Vatsa M., and Singh R., "Saliency Based Mass Detection from Screening Mammograms," *Signal Processing*, vol. 99, pp. 29-47, 2014.

[4] Angayarkanni S., Kamal N., and Thangaiya R., "Dynamic Graph Cut Based Segmentation of Mammogram," *SpringerPlus*, vol. 4, no. 1, pp. 1-9, 2015.

[5] Cascio D., Fauci F., Magro R., Raso G., Bellotti R., De Carlo F., Tangaro S., De Nunzio G., Quarta M., Forni G., and Lauria A., "Mammogram Segmentation by Contour Searching and Mass Lesions Classification with Neural Network," *IEEE Transactions on Nuclear Science*, vol. 53, no. 5, pp. 2827-2833, 2006.

[6] Chalana V. and Kim Y., "A Methodology for Evaluation of Boundary Detection Algorithms on Medical Images," *IEEE Transactions on Medical Imaging*, vol. 16, no. 5, pp. 642-652, 1997.

[7] Chowdhary C. and Acharjya D., "Segmentation of Mammograms Using a Novel Intuitionistic Possibilistic Fuzzy C-Mean Clustering Algorithm," *in Proceedings of the Nature Inspired Computing*, Singapore, pp. 75-82, 2018.

[8] De Nazaré Silva J., De Carvalho Filho A., Silva A., De Paiva A., and Gattass M., "Automatic Detection of Masses in Mammograms Using Quality Threshold Clustering, Correlogram Function, and SVM," *Journal of Digital Imaging*, vol. 28, no. 3, pp. 323-337, 2015.

[9] Deneubourg J., Goss S., Franks N., Sendova-Franks A., Detrain C., and Chrétien L., "The Dynamics of Collective Sorting Robot-like Ants and Ant-like Robots," *in Proceedings of the 1st International Conference on Simulation of Adaptive Behavior on from Animals to Animats*, Cambridge, pp. 356-365, 1991.

[10] Dixon A., "Diagnostic Breast Imaging: Mammography, Sonography, Magnetic Resonance Imaging, and Interventional Procedures," *Ultrasound: Journal of the British Medical Ultrasound Society*, vol. 22, no. 3, pp. 182-183, 2014.

[11] Dorigo M., Maniezzo V., and Colorni A., "Ant System: Optimization by A Colony of Cooperating Agents," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 26, no. 1, pp. 29-41, 1996.

[12] Duarte M., Alvarenga A., Azevedo C., Calas M., Infantosi A., and Pereira W., "Evaluating Geodesic Active Contours in Microcalcifications Segmentation on Mammograms," *Computer Methods and Programs in Biomedicine*, vol. 122, no. 3, pp. 304-315, 2015.

[13] Feng Y., "Ant Colony Cooperative Optimization and its Application in Image Segmentation," Ph. D. Theses, Xi'an Jiaotong University, 2005.

[14] Ghosh S., Kothari M., Halder A., and Ghosh A., "Use of Aggregation Pheromone Density for Image Segmentation," *Pattern Recognition Letters*, vol. 30, no. 10, pp. 939-949, 2009.

[15] Han Y. and Shi P., "An Improved Ant Colony Algorithm for Fuzzy Clustering in Image Segmentation," *Neurocomputing*, vol. 70, no. 4-6, pp. 665-671, 2007.

[16] Hao Y., Yang H., Long B., and Liu J., "Image Segmentation Based on a New Self-Adaptive Ant Clustering Algorithm," *in Proceedings of the IEEE International Conference on Apperceiving Computing and Intelligence Analysis Proceeding*, Chengdu, pp. 258-261, 2010.

[17] Hung C. and Sun M., "Ant Colony Optimization for the K-Means Algorithm in Image Segmentation," *in Proceedings of the 48th Annual Southeast Regional Conference*, Mississippi, pp. 1-4, 2010.

[18] İnkaya T., Kayalıgil S., and Özdemirel N., "Ant Colony Optimization Based Clustering Methodology," *Applied Soft Computing*, vol. 28, pp. 301-311, 2015.

[19] Jaccard P., "Nouvelles Recherches Sur La Distribution Florale," *Bulletin Société Vaudoise des Sciences Naturelles*, vol. 44, no. 163, pp. 223-270, 1908.

[20] Jevtić A., Quintanilla-Domínguez J., Barrón-Adame J., and Andina D., "Image Segmentation Using Ant System-Based Clustering Algorithm," *in Proceedings of the 6th International*

Conference on Soft Computing Models in Industrial and Environmental Applications, Berlin, pp. 35-45, 2011.

[21] Khorram B. and Yazdi M., "A New Optimized Thresholding Method Using Ant Colony Algorithm for MR Brain Image Segmentation," *Journal of Digital Imaging*, vol. 32, no. 1, pp. 162-174, 2019.

[22] Laptik R. and Navakauskas D., "Application of Ant Colony Optimization for Image Segmentation," *Elektronika Ir Elektrotechnika*, vol. 80, no. 8, pp. 13-18, 2007.

[23] Li L., Ren Y., and Gong X., "Medical Image Segmentation Based on Modified Ant Colony Algorithm with GVF Snake Model," *in Proceedings of the IEEE International Seminar on Future BioMedical Information Engineering*, Wuhan, pp. 11-14, 2008.

[24] Liu L., Tan G., and Soliman M., "Color Image Segmentation Using Mean Shift and Improved Ant Clustering," *Journal of Central South University*, vol. 19, no. 4, pp. 1040-1048, 2012.

[25] Lumer E. and Faieta B., "Diversity and Adaptation in Populations of Clustering Ants," *in Proceedings of the 3rd International Conference on Simulation of Adaptive Behavior: From Animals to Animats 3*, Cambridge, pp. 501-508, 1994.

[26] Moftah H., Azar A., Al-Shammari E., Ghali N., Hassanien A., and Shoman M., "Adaptive K-Means Clustering Algorithm for MR Breast Image Segmentation," *Neural Computing and Applications*, vol. 24, no. 7, pp. 1917-1928, 2014.

[27] Odet C., Belaroussi B., and Benoit-Cattin H., "Scalable Discrepancy Measures for Segmentation Evaluation," *in Proceedings of the IEEE International Conference on Image Processing*, Rochester, pp. I-I, 2002.

[28] Oliver A., Freixenet J., Marti J., Perez E., Pont J., Denton E., and Zwiggelaar R., "A Review of Automatic Mass Detection and Segmentation in Mammographic Images," *Medical Image Analysis*, vol. 14, no. 2, pp. 87-110, 2010.

[29] Ouadfel S. and Batouche M., "An Efficient Ant Algorithm for Swarm-Based Image Clustering," *Journal of Computer Science*, vol. 3, no. 3, pp. 162-167, 2007.

[30] Ouadfel S. and Batouche M., "MRF-based Image Segmentation Using Ant Colony System," *ELCVIA Electronic Letters on Computer Vision and Image Analysis*, vol. 2, no. 1, pp. 12-24, 2003.

[31] Ouadfel S., Batouche M., and Garbay C., "Ant Colony System for Image Segmentation Using Markov Random Field," *in Proceedings of the International Workshop on Ant Algorithms*, Berlin, pp. 294-295, 2002.

[32] Rosa B., Mozer P., and Szewczyk J., "An Algorithm for Calculi Segmentation on Ureteroscopic Images," *International Journal of Computer Assisted Radiology and Surgery*, vol. 6, no. 2, pp. 237-246, 2011.

[33] Rouhi R. and Jafari M., "Classification of Benign and Malignant Breast Tumors Based on Hybrid Level Set Segmentation," *Expert Systems with Applications*, vol. 46, pp. 45-59, 2016.

[34] Saatchi S. and Hung C., "Using Ant Colony Optimization and Self-Organizing Map for Image Segmentation," *in Proceedings of the Mexican International Conference on Artificial Intelligence*, Berlin, pp. 570-579, 2007.

[35] Sampaio W., Diniz E., Silva A., De Paiva A., and Gattass M., "Detection of Masses in Mammogram Images Using CNN, Geostatistic Functions and SVM," *Computers in Biology and Medicine*, vol. 41, no. 8, pp. 653-664, 2011.

[36] Sezgin M. and Sankur B., "Survey Over Image Thresholding Techniques and Quantitative Performance Evaluation," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146-165, 2004.

[37] Wang X., Feng Y., and Feng Z., "Ant Colony Optimization for Image Segmentation," *in Proceedings of the IEEE International Conference on Machine Learning and Cybernetics*, Guangzhou, pp. 5355-5360, 2005.

[38] Yan J., "Remote Sensing Image Segmentation Based on Ant Colony Optimized Fuzzy C-Means Clustering," *Journal of Chemical and Pharmaceutical Research*, vol. 6, no. 6, pp. 2675-2679, 2014.

[39] Yang X., Zhao W., Chen Y., and Fang X., "Image Segmentation with a Fuzzy Clustering Algorithm Based on Ant-Tree," *Signal Processing*, vol. 88, no. 10, pp. 2453-2462, 2008.

[40] Yasnoff W., Mui J., and Bacus J., "Error Measures for Scene Segmentation," *Pattern Recognition*, vol. 9, no. 4, pp. 217-231, 1977.

[41] Yu J., Lee S., and Jeon M., "Medical Image Segmentation by Hybridizing Ant Colony Optimization and Fuzzy Clustering Algorithm," *in Proceedings of the 13th Annual Conference Companion on Genetic and Evolutionary Computation*, Dublin, pp. 217-218, 2011.

[42] Yuanjing F., Li Y., and Liangjun K., "Finite Grade Pheromone Ant Colony Optimization for Image Segmentation," *Opto-Electronics Review*, vol. 16, no. 2, pp. 163-171, 2008.

[43] Zhang Y., "A Survey on Evaluation Methods for Image Segmentation," *Pattern Recognition*, vol. 29, no. 8, pp. 1335-1346, 1996.

[44] Zhao B., Zhu Z., Mao E., and Song Z., "Image Segmentation Based on Ant Colony Optimization and K-Means Clustering," *in Proceedings of the IEEE International Conference on Automation and Logistics*, Jinan, pp. 459-463, 2007.

[45] Zou G., "Ant Colony Clustering Algorithm and Improved Markov Random Fusion Algorithm in Image Segmentation of Brain Images," *International Journal Bioautomation*, vol. 20, no. 4, pp. 505-514, 2016.

[46] Zou R., Yu W., Yu Z., and Yu X., "Image Segmentation Based on Local Ant Colony Optimization," *in Proceedings of the IEEE 5th International Conference on Natural Computation*, Tianjian, pp. 35-39, 2009.

Sudha Subramanian received Bachelor's degree in Computer Science from Madurai Kamaraj University, Madurai in 2002, and Master in Computer Applications from Madurai Kamaraj University, Madurai in 2005, and Master of Engineering in Computer Science Engineering from University College of Engineering, BIT Campus, Tiruchirappalli in 2013. Her research interests include Medical Image processing, Data Mining, Machine Learning and its applications.

Ganesan Rasu Thevar received his B.E, M.E., and Ph.D in the years 1991, 2000 and 2010 respectively. He designed and developed a frame work in Bio Medical area for his Ph.D research under the title of "Design of CT Scan Brain Image Analysis System". He possesses about 25 years of Teaching experience and 3 years of Industrial experience in Tuticorin Thermal Power Station. He is a Senior Member of ISTE, IETE and MIE. He convened and organized many Workshops, Seminars, International Conferences and Symposia. He produced 5 Ph.D holders. Received R and D project fund from TNSCST, I.E, AICTE and DST. His Research areas are Medical Imaging, Image Processing, Video Processing, Neural Network, Instrumentation Systems and Genetic Algorithm.