ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

Mammographic Cluster detection using Markov Random Field with Ant Colony and Support Vector based Classification

Venmathi A R[1], Dr E N Ganesh[2], Dr.N Kumaratharan [3]

Research Scholar, Department of ECE, Kings Engineering College Anna University, Chennai, India. [1]

Professor, Department of ECE, Saveetha Engineering College, Chennai, India[2]

Professor, Department of CSE, Sri Venkateswara Engineering College, Chennai, India[3]

ABSTRACT: Breast cancer is one of the deadly diseases found among female population. Mammography a screening method plays a vital role for the detection of breast cancer. Detection of microcalcification based on textural image segmentation and classification is the most effective early-diagnosis of breast cancer. Early detection of micro calcification in breast tissue, which is an indication of developing breast cancer. It is found to be difficult for the radiologist to pinpoint effectively to identify the affected area in the regions of the breast. In this paper, a proposed technique, Markov Random Field method involved with Ant Colony System, to find the affected region of microcalcification in digital mammograms. 161 pairs of digitized mammograms from MIAS database are tested for proposed algorithm and the techniques and intern these images are subjected to Ant colony optimization and Markov Random Fields approaches , to find spatial relation for different neighboring pixels via an energy function, and optimized using ant colony to find its complete region coverage with the technique called region coverage with the technique called Euclidian distance , so that the complete affected area can be identified. 161 pairs of mammograms taken from MIAS database to evaluate the performance of the proposed method.

KEYWORDS: Mammogram, MIAS, Database, Microcalcification, Ant Colony Optimization, MRF, SVM

I.INTRODUCTION

Breast cancer continues to be a significant public health problem in the world. Breast cancer is one of the major causes for the increase in mortality among women, especially in developed and underdeveloped countries. The World Health Organization's International agency for Research on Cancer in Lyon, France, estimates that more than 150 000 women worldwide die of breast cancer each year. In India, breast cancer accounts for 23% of all the female cancers followed by cervical cancers (17.5%) in metropolitan cities such as Mumbai, Calcutta, and Bangalore. Early detection is the key to improve breast cancer prognosis and sutle signs of breast cancer requires high-quality images and skilled mammographic interpretation [1]. There are some limitations for the radiologist to give accurate and better evaluation. The accumulation of calcification forms a cluster and this presence is an important sign for the detection of breast cancer. The dense tissues, and especially in younger women, cause the suspicious region to be almost invisible and may be easily misinterpreted as calcifications and yield a high False Positive (FP) rate that is a major problem with most of the existing algorithms [2]. The main goal of the research is designing a system for the automatic classification of mammogram images using CAD (Computerize Aided Diagnosis)[3]. The texture energy is extracted as features from the ROI of the digital mammogram for the identification of microcalcification clusters. The extracted region of interest can be given to Markov Random Field approach and Ant Colony Optimization to identify the particular mass separately [4].

Markov Random Field approach is a graphical model of a joint probability distribution. It consists of a unidirectional graph in which the nodes represent random variables. This model aims at classifying the images as relevant or

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering
(An ISO 3297: 2007 Certified Organization)
Website: www.ijircce.com
Vol. 5, Special Issue 3, April 2017

irrelevant, and the output of MRF is used to generate a new list of ranked images. The MRF takes into account the following three processes [5].

1. The rank information provided by the initial retrieval system
2. Similarities among images in the list.
3. Relevance feedback information.

By using the above processes satisfies the Marko property which can make predictions for the failure of the process based solely on its present state just as well as one could knowing the process's full history. The goal of this method is to find out the optimum label of the image that minimizes the posterior energy function value[6].

Ant Colony Optimization is a computational method and the algorithm uses the virtual ants. These ants can create a solution which can be represented by paths on a graph[7][8]. This algorithm finds one dimensional manifolds on the spectral data cloud, and then links the similarity pixels to create the path. The low dimensional manifold estimated by the best fit "Ant Path". The path plotted by the ants with the help of phorems will be aligned to form a graph and is analyzed by the use of Markov[8].

Segmentation is the initial step for any image analysis. The MRF based image segmentation of mammogram images is used to assist radiologists to classify the abnormalities of the breast into benign or malignant. The segmented image is given to the SVM classifier for the better classification [9]. The segmented image is given to the SVM classifier for the better classification.

II. LITERATURE STUDY

Each family of algorithms has its own implementation for the data to be analyzed. W Hai-ningand[3]describe the standard models in developing quantitative algorithms for processing. The graph based models have been developed that leverage the data with minimal assumptions. Beckers [7]describe how graph-based methods can be used to develop animal detection. Colorni [8] describes how graph can be used to describe a spectral image, to manipulate through edge cutting to produce a cluster map. Karaboga[10] describe the spectral image clustering with the application of commute time distance transformation to the spectral image. This method can preserves and enhances the connectivity of the graph.

III. PROPOSED METHOD

In this paper, meta-heuristic algorithms such as ACO and MRF are implemented to extract the suspicious region. The textural features can be extracted from the suspicious region to classify the microcalcification into benign or malign using SVM[11]. The mammogram image is stored in a two-dimensional matrix. A unique label assigned to the kernels having similar patterns. The optimum label is the one, which minimizes the MAP estimate. A labeling process consists of assigning the same label to the kernels having similar patterns. [Kernel is a 3×3 window of neighborhood pixels]. Ant Colony Algorithm which is described in the proposed system comes from the family of swarm intelligence algorithms which can model the behavior of animals to solve problems. Ant Colony Optimization (ACO) is a population based approach first designed by Marco Dorigo and coworkers, inspired by the foraging behavior of ant colonies [12]. An ant colony is the basic family unit around which ants organize their lifecycle [13]. Ant colonies are eusocial, and are very much like those found in other social Hymenoptera, though the various groups of these developed sociality independently through convergent evolution.

ACO optimize virtual ants on a mathematical graph which is formed by MRF [14], Individuals ants are simple insects with limited memory and capable of performing simple actions. However, the united behavior of ants yields intelligent way out to problems such as finding the shortest paths from the nest to a food source. Ants foraging for food dispose quantities of a volatile chemical substance named pheromone, marking their path that it follows [2]. Ants smell pheromone and decide to follow the path with a high probability and thereby bolster up it with a further amount of pheromone. The probability that an ant selects a path increases with the number of ants selecting the path at previous times and with the strength of the pheromone concentration.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

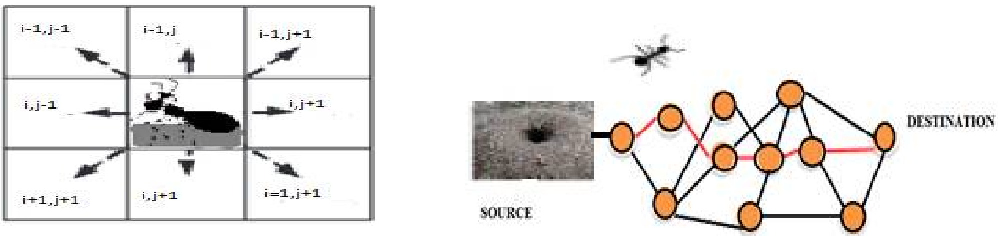
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

ACO also can be used to effectively divide an image into clusters [15]. Because of these reasons, ACO can be accepted as a viable method in analyzing spherical data



(a) (b)

Fig .1(a)Defined Matrix for the Tour of the ant (b) Paths of the ant

A. Markov Random Filed (MRF)

The mammogram image is stored in a two dimensional matrix. A unique label assigned to the kernels having similar patterns. For each kernel in the image, calculate the posterior energy function value U(x).

$U(x)=\{\sum[(y-\mu)2/(2*\sigma 2)]+\sum\log(\sigma)+\sum V(x)\}$ (1)

where, y is the intensity value of pixels in the kernel

μ is the mean value of the kernel; σ is the standard deviation of the kernel

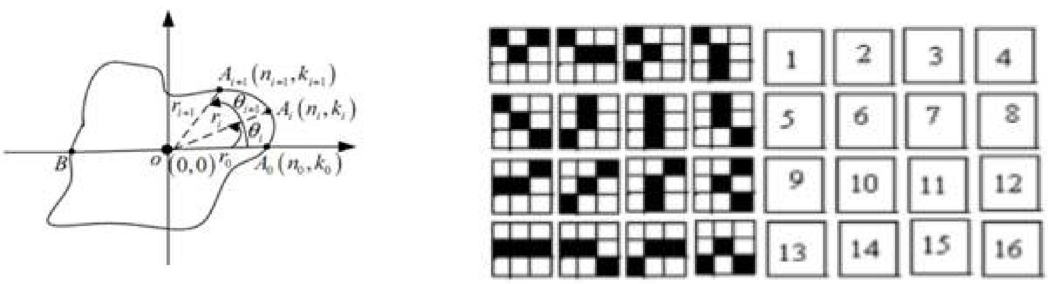
V is the potential function of the kernel, and x is the label of the pixel.

If x1 is equal to x2 in a kernel, then $V(x) = \beta$, otherwise 0, where β is visibility relative parameter ($\beta \geq 0$).

The maximizing the a posteriori probability (MAP) estimate can be written as:

$P(x|y) = \exp(-U(x))$, the challenge of finding the MAP estimate of the segmentation is search for the optimum label which minimizes the posterior energy function U(x).

This paper describes a new effective approach for the minimization of the energy function, the concept of Ant Colony Optimization.



(a) (b) (c)

Fig. 2(a) Markov model of Microcalcification boundary (b) Image edge modes, (c)The codes of image edge patterns

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

B. Ant Colony Algorithms

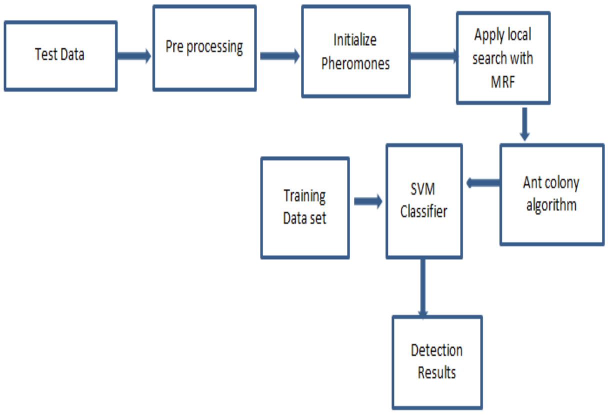


Fig. 3 Flow diagram for the algorithm

In ACO, a set of software agents called artificial ants search for good solutions optimization for this problem.

Differences with Ant System:

1. Decision Rule - Pseudorandom proportional rule
2. Local Pheromone Update
3. Best only offline Pheromone Update

C. Initialize Pheromone Values

At the beginning of the algorithm all the pheromone values are initialized to the numerical value. These are two measures called pheromone and heuristic level which are stored in an array for every edge on the graph. The pheromone level gives information on the ant's movement α that the past ants fared on the edge and the heuristic level is the predetermined measure which helps the ants to proceed in the right direction.

While the ants proceed they construct the graph following the decision rule and this rule purely depend on the pheromone and the heuristic values. Using two parameters α and β ,

$$P^{K}_{xy} = \frac{\gamma_{xy}^{\alpha}\eta_{xy}^{\beta}}{\Sigma\gamma_{xy}^{\alpha}\ \eta_{xy}^{\beta}}$$

Where P_{xy} is the probability in x and y directions. It gives the probability of moving from point x to point y and is the pheromone and is the heuristic value. This can be considered as simply the length of the edge. The summation on the denominator gives an average values for all the edges.

The end point is defined by the user and the ant path ends at this point. After all the iteration of the algorithm according to the decision the pheromones are created and updated. The pheromones update rule is based on the ant's solution. The update rule is as follows

$$\Delta\gamma_{xy}^{K} = \begin{cases} \frac{Q}{Lk} & if\ ant\ uses\ curve\ path \\ 0 & OW \end{cases}$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

Where K is the number of ants

Q is the constant

Lk is the quality of the solution.

For every iteration j the pheromones should be evaporated in order to put more emphasis on current path. The evaporation rule is as follows

$$\gamma_{xy}^{j} = (1-\rho)\gamma_{xy}^{(j-1)}$$

where is the evaporation rule parameter.

This value is small (i.e., in [0,1]) as in most ACO algorithms.

The graph is very much needed for the optimization problem. The solution for the optimization can be given by the paths on the graph. When ants go on adding pheromones, it is very important to concentrate on edges and sections so that it leads the falsies ants to follow the right direction. The familiar way to construct a graph is to build a mode on each mixed so that they can be interconnected based on similarities. To give remedy for cracking unconnected graphs, the dense points in the graph can be used to define the Euclidean distance the dense point and the radius used to calculate density are considered as parameters in this algorithms the ACO algorithms ants can more from darker pixel while constructing the path on a graph.

In the construction phase an ant incrementally constructs a solution by adding solution components to the partial solution constructed so far. The probabilistic selection of the next way out component to be added is carried out by using the transition probabilities. The transition probabilities are determined by a function that is called the state transition rule, which is a function of the pheromone values and possibly of other information. The objective of this amend rule is to enlarge the pheromone values of solution components that have been found in high-quality solutions.

The Goal of this Algorithm is to find paths along a point in the cloud where data are desire. The start of the program focuses the dearest point to calculate the Euclidean distance, fill the ending point. Once the Euclidean distance is found, then, the points are labeled and continued fill no then, the points left in the image.

IV.CLASSIFICATION TOOL SUPPORT VECTOR MACHINES

A Support Vector Machine (SVM) is a discriminated classifier officially defined by a separating hyper plane. SVM fits a hyper plane/function between 2 different classes given a maximum margin parameter.

This hyper plane attempts to separate the classes so that each falls on either side of the plane, and by a specified margin. There is a specific cost function for this kind of model which adjusts the plane until error is minimized.

In the below picture you can see that there exists multiple lines that offer a solution to the problem. Is any of them better than the others? We can intuitively define a criterion to estimate the worth of the lines:

A line is bad if it passes too close to the points because it will be noise sensitive and it will not generalize correctly. Here, our aim is to find the line passing as far as possible from all points.

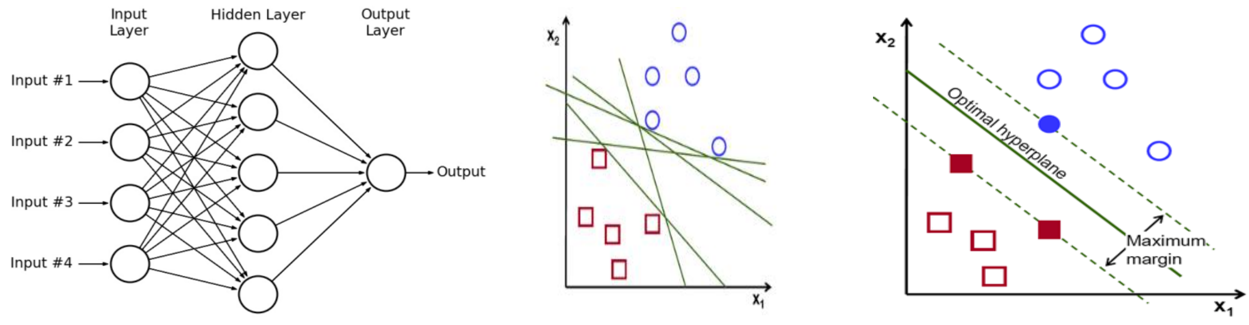
ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017



(a) (b) (c)

Fig.4 Support Vector Machine (a) Architecture (b) Hyper plane construction (c) Separated Outputs

Here the performance of the SVM algorithm is rested on finding the hyper plane that provides the largest minimum distance to the training examples. Twice, this distance receives the important name of margin within SVM's theory. SVMs are being developed in the reverse order to the buildout neural networks (NNs). SVMs evolved from the sound theory to the implementation and experiments, while the NNs followed more heuristic path, from applications and extensive experimentation to the theory. Therefore, the optimal separating hyper plane *maximizes* the margin of the training data as well as this machine operates with the principles of Ex- OR gates and identified as the best for classification also it reduces the delay in operation.

The optimal hyper plane can be represented in an infinite number of different ways by scaling of and . For a matter of convention, amid all the
possible representations of the hyperplane, the one choice is

$$|\beta 0 + \beta^T x| = 1$$

where symbolizes the training examples closest to the hyper plane. Normally, the training examples that are nearest to the hyper plane are called support vectors. This representation is known as the canonical hyper plane. Now, we use the result of geometry that gives the distance between a point and a hyper plane $(\beta, \beta o)$

distance =

In particular, for the canonical hyper plane, the numerator is equal to one and the distance to the support vectors is

$$\text{Distance}_{\text{support vectors}} = \left| \frac{|\beta 0 + \beta^T x|}{\|\beta\|} \right| = \frac{1}{\|\beta\|}$$

Recall that the margin introduced in the previous section, here denoted as M, is twice the distance to the closest examples

$$M = \frac{2}{\|\beta\|}$$

Finally, the problem of maximizing M is equivalent to the problem of minimizing a function L(subject to some constraints. The constraints model the requirement for the hyper plane to classify correctly all the training examples xi. Formally,

$$\min_{\beta, \beta o} L(\beta) = \frac{1}{2} \|\beta\| 2 \; subject \; to \; yi \; (\beta T \; xi + \beta 0) \geq 1 \forall i,$$

where yi represents each of the labels of the training examples

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

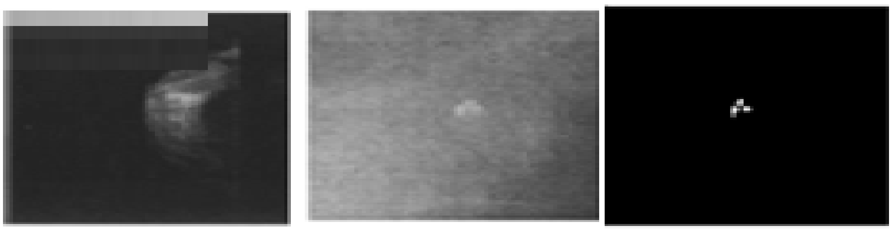
(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

A.MAMMOGRAM CLASSIFICATION WITH SVM

SVM is not able to perform the classification tasks efficiently in the nonlinear case.To overcome this limitation on SVM, kernel approaches are developed. The kernel function in an SVM plays the central role of implicitly mapping the input vector into a high-dimensional feature space. Typical choices for kernel function a Gaussian Radial Basis Function (RBF), Polynomial, Sigmoid, Inverse multi-quadratic, etc. The Gaussian RBF and polynomial kernels are
Polynomial kernel KP
Where P is Gaussian kernel
K = exp.
Both of these kernels satisfy the conditions and are among the most commonly used in SVM



(a) (b) (c)

Fig.5.Classified Outputs for Benign Image a)Original image(b)Segmented using MRF (c)Classified Benign with SVM

(a) (b) (c)

Fig.6 Classified Outputs for Malignant Image (a)Original image (b)Segmented image using MRF (c)Classified Malignant with SVM

V. RESUTS AND DISSCUSSIONS

Classification performance can be measured by plotting the probability curves using the positive examples against the rate of incorrectly classifying true negative examples. To evaluate the performance of SVM classifier, the classification performances of the textural features extracted by texture-analysis method are evaluated. The analysis is based on statistical decision theory and has been applied extensively to the evaluation of clinical diagnosis. ACO algorithms have been successfully applied to diverse combinational optimization problems. A good review of this optimization technique is studied and it is observed that ACO is extensively used in medical image processing.

VI. CONCLUSION

In this paper, a novel approach is applied for mammogram image segmentation and classification based on the combination of Markov Random Field, Ant Colony System., and SVM Neural Network. In MRF the image pixels are labeled and their posterior function values are computed. Ant Colony Optimization (ACO) have been used to find out the optimum label that minimizes the Maximizing a Posterior estimate to segment the image. The ACO search is inspired by the foraging behavior of real ants. Each ant constructs a solution using the pheromone information accumulated by the other ants. In each iteration, a local minimum value is selected from the ants' solution and the

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

pheromones are updated locally to find the global minimum. The update has done the pheromone of the ant that generates the global minimum. At the final iteration, global minimum returns the optimum label for image segmentation. Then an SVM classifier is considered for the task of classifying the regions as malignant or benign MCs. As a result, the average accuracy was approximately 97.5% (sensitivity: 95.7%, specificity: 100%).

REFERENCES

[1] J. Siddiqui, M. Anand, K Mehrotr, R.Sarangi and N.Mathur. Biomonitoring of Organochlorines inWomen with Benign and Malignant Breast Disease. Environmental Research, 1-8, 2004.

[2] L. YunyueandS.Zhifeng,"Cognitive Radio Decision Engine Based on Adaptive Ant Colony Optimization",Computer Science, vol. 38, no. 8,(2011), pp.253-256.

[3] WHai-ningandS.Shou-qian, "Adaptive Ant Colony OptimizationAlgorithm Based on Maturity ",Microelectronics Computer,vol. 27, no.8,2010), pp.140-44.

[4] A. Colorni, M.Dorigo, and V. Maniezzo. InF.J. Varela, P. Bourgine.Distributed optimization byant colonies. Proceedings of the First EuropeanConferenceonArtificialLife, Paris, 1991.

[5] C. Kervrann and F. Heitz. A Markov Random Field Model based Approach to Unsupervised Texture Segmentation Using Local and Global IEEE Transactions on Image Processing, 4(6): 856 –862, 1995.

[6] A. Speis and G. Healey. An Analytical and Experimental Study of the Performance of Markov Random Fields Applied toTextured Images UsingSmall Samples. IEEE Transactions on ImageProcessing 5(3):447—458,1996

[7] R. Beckers, J. L. Deneubourg, and S. Goss. Trails and U-turns in the selection of the shortest path bythe antLasius. niger. J.Theor.Bio., 159:397–415, 1992.

[8] Y. Ge, Q. C. Meng, C. J. Yan and J. Xu. A Hybrid Ant Colony Algorithm for Global Optimization of Continuous Multi –Extreme Functions.Processiing of the Third International Conference on Machine Learning and Cybermetics W. Feng, Q. Zhng, Shanghai, 2427-2432,

[9] K. Polat, S. Guneş, "Breast cancer diagnosis using least square support vector machine", *Digital Signal Process.* "in press", 2006.

[10] N. Karaboga, A. Kalinli and D. Karaboga. Designing Digital IIR filters using ant colony optimization algorithm Engineering applications of Artificial Intelligence 17:301-309,2004

[11] A. Bazzani, A.D. Bevilacqua, "Automatic detection of clustered microcalcifications in digital mammograms using an SV classifier", ESANN'2000 proceedings, European Symposi- um on Artificial Neural Networks, Bruges, Belgium, 2000.

[12] G. Hu, J Xiangji Huang," Mining network data for intrusion Detection through combining SVMs with ant colony networks "Future Generation Computer Systems, 2013.

[13] S XiaoyaandLYan"Improved Artificial Bee Colony Algorithm for Assignment Problem"Microelectronics & Computer, vol.29, no. 1,(2012).

[14] WZhenqing,"Study on the Method of Image Segmentation Based on Ant Colony Algorithm", University of Petroleum, msd,Beijing,China, (2011)

[15] J. He, D. Long, C. Chen, "An improved ant-based classifier for intrusion detection", in: Proceedings of the 3rd International Conference on Natural Computation, Vol. 4, ICNC'07, IEEE Computer Society, 2007, pp. 819–823

BIOGRAPHY

VENMATHI A R is an Associate Professor in the Department of Electronics and Communication Engineering, Kings Engineering College, Anna University. She is currently a PhD Scholar pursuing PhD degree from Anna University chennai, Tamilnadu, India. Her Research Domain is medical image processing.