International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

A Hybrid of Aco-Ffa Algorithm for Feature Selection in Digital Mammogram

J. B. Jona

Associate Professor, Department of Computer Applications, CIT, Coimbatore, Tamil Nadu India.

ABSTRACT
Digital mammogram is the only effective screening method to detect the breast cancer. Gray level co-occurrence matrix (GLCM) textural features are extracted from the mammogram. All the features are not essential to detect the mammogram. Therefore identifying the relevant feature is the aim of this work. Feature selection improves the classification rate and accuracy of any classifier. In this paper a new hybrid metaheuristic named ACO-FFA a hybrid of Ant Colony Optimization (ACO) and Firefly Algorithm (FFA) is proposed for feature selection in Digital Mammogram. ACO is a good metaheuristic optimization technique but the drawback of this algorithm is that the ant will walk through the path where the pheromone density is high which makes the whole process slow hence FFA is employed to carry out the local search of ACO. Support Vector Machine (SVM) classifier with Radial Basis Kernel Function (RBF) is done along with the ACO to classify the normal mammogram from the abnormal mammogram. Experiments are conducted in mini-MIAS database. The performance of the new hybrid algorithm is compared with the ACO and PSO algorithm. The results show that the hybrid ACO-FFA algorithm is more accurate than the other techniques.

Keywords: Firefly Algorithm, Ant Colony Optimization, ROC curve, Support Vector Machine.

1. INTRODUCTION

Breast cancer is one of the leading causes of mortality in women. Breast cancer is caused due to uncontrolled growth of cells in the breast. So far mammography is the only effective screening method for detection of breast cancer in early stage. Due to wrong interpretation of the radiologist or the limitation of human visualization system certain errors like false negative errors may arise. To overcome such limitation of mammography the researchers developed Computer Aided Diagnosis (CADx) which evaluate or assess mammographic abnormality by automating the segmentation, detection, feature extraction and classification processes.

During the past two decades the focus of researchers falls on the nature inspired metaheuristic algorithms. They concentrated more on the Nature-Inspired Computation (NIC). The NIC refer to algorithms that are derived by mimicking natural phenomena and biological models to solve a problem. Many more nature inspired algorithms had evolved in the past few decades. The most well-known NIC are the artificial neural networks, Genetic Algorithm (GA), Particle Swarm Optimization (PSO) and Artificial Immune System (AIS).

ACO is a population based optimization technique which was first introduced by Marco Dorigo, as Ant System (Dorigo et al., 1996, Dorigo and Gambardella, 1997). In 1999 it was redefined as Ant Colony Optimization and used first to solve the travelling salesman problem (Dorigo and Stuetzle, 2004). It is

inspired by the behavior of ants in finding shortest paths from the colony to its food. Al-Ani used ACO for feature selection (Al-Ani, 2005). Firefly Algorithm (FFA) is an optimization technique developed by Xin-She Yang at Cambridge University in 2007 which was inspired by the flashing characteristic of the fireflies[Yang., 2009 and Yang., 2010]. Firefly with greater flashing light intensity will attract the other fireflies to move toward it. Firefly algorithm has some advantages such as simplicity and intrinsic capability of finding local optimums. The major disadvantage of ACO is that the local search it performs is not much faster. So Firefly Algorithm is proposed to carry out the local search of ACO. In this paper a hybrid of Ant Colony Optimization (ACO) and Firefly Algorithm is proposed.
This paper is organized as follows: Section 2 describes the overview of Firefly algorithm; Section 3 describes the methodology of the proposed system which includes feature extraction, feature selection by the proposed hybrid technique and the Support Vector Machine (SVM) classification. Section 4 and section 5 describes the experimental results, conclusion and the future enhancements respectively.

2. AN OVERVIEW OF FIREFLY ALGORITHM (FFA)

Firefly algorithm is population based metaheuristic optimization technique. There are about two thousand firefly species and most fireflies produce short and rhythmic flashes. The pattern of flashes is often unique for a particular species. The flashing light is produced by a process of bioluminescence, and the true functions of such signalling systems are still being debated. However, two fundamental functions of such flashes are to attract mating partners (communication), and to attract potential prey. In addition, flashing may also serve as a protective warning mechanism to remind potential predators of the bitter taste of fireflies. Thus the whole process is induced by a brighter firefly.
Firefly optimization algorithm is based on three simple rules stated:

- All fireflies are unisex so that one firefly will be attracted to other fireflies regardless of their sex
- Attractiveness is proportional to the brightness or light intensity, thus for any two flashing fireflies, a firefly with lesser brightness will move towards the brightest one. The attractiveness is proportional to the brightness and they both decrease as their distance increases. If there is no brighter firefly then it will move randomly
- The brightness of a firefly is affected or determined by the landscape of the objective function

2.1. Firefly Algorithm for Feature Selection

In the FFA, there are two important issues:

- Variation of light intensity
- Formulation of the attractiveness.

For simplicity, it is always assumed that the attractiveness of a firefly is determined by its light intensity which in turn is associated with the encoded objective function. In the simplest case for maximum optimization problems, the light intensity I of a firefly at a particular location x can be chosen as $I(x) \propto f(x)$. In order to improve own solution, the firefly needs to advance towards the fireflies that have brighter light emission than is his own.
In this algorithm, each firefly has a location $X = (x_1, x_2 \dots x_d)^T$ in a d-dimensional space and a light intensity I(x) or attractiveness β(x) which are proportional to objective function f(x). Attractiveness β(x) and light intensity I(x) are relative and these should be judged by the other fireflies. Thus, it will vary with the distance r_{ij} between firefly i and firefly j. So attractiveness

$$\beta = \beta_0 e^{-\gamma r^2} \qquad (1)$$

In which β_0 is the attractiveness in distance r=0 and γ is light absorption coefficient in the range $[0,\infty]$The distance r between firefly i and j at and is defined as Cartesian distance:

$$r = r_{ij} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^{d}(x_{i,k} - x_{j,k})^2} \qquad (2)$$

where $x_{i,k}$ is the k^{th} component of the spatial coordinate x_i of the i^{th} firefly and d is the number of dimensions. Moreover, the movement of firefly i which is attracted by a more attractive or brighter firefly j is given by the following equation:

$$x_i = x_i + \beta_0 e^{-\gamma r x_{i,j}^2}(x_j - x_i) + \alpha(\epsilon - 0.5) \qquad (3)$$

where the second term is due to the attraction. The third term is randomization with α being the randomization parameter such that $\alpha \in [0, 1]$, and ∞ is a vector of random numbers drawn from a Gaussian distribution or uniform distribution in the range [0, 1]. Furthermore, for most problems, one can take $\beta_0=1$.

The attracted firefly move towards the attractive one and the location gets updated and the final step is checking the location of the firefly whether it is inside the range within which it is specified. Then based on the rank the best firefly is computed.

Algorithm 1: FFA algorithm

1. Generate initial population of fireflies x_i (i = 1, 2... n)
2. Compute the fitness light intensity I_i at x_i is determined by f (x_i)
3. Define light absorption coefficient γ
4. while (t <MaxGeneration)

4.1 Move each firefly *i* towards other brighter fireflies, and if there is no other brighter firefly, move it randomly.

4.2 Attractiveness varies with distance r through exp[-γr]

4.3 Evaluate new solutions and update light intensity

4.4 If maximum iterations reached, then stop; otherwise go to step (4).

4.5 Rank the fireflies and find the current global best

end while

FFA has the advantage that it can find the global optima as well as the local optima simultaneously and effectively. A further advantage of FFA is that different fireflies will work almost independently, it is thus particular suitable for parallel implementation.

3. METHODOLOGY

The mammographic images from Mini-MIAS database is used in this research. 78 Gray level co-occurrence matrix (GLCM) textural features are extracted. The extracted feature set is reduced by the proposed hybrid ACO-FFA optimization technique. SVM classifier is used along with ACO to classify the normal mammogram from an abnormal mammogram. The performances of all the proposed techniques are compared by Receiver Operating Characteristic curve (ROC).

4.1.Feature Extraction

Feature extraction is the process of reducing the original mammogram image into a set of features, by measuring certain properties or features that distinguish one input pattern from another pattern. GLCM features are extracted from the mammogram. GLCM features are calculated based on the haralick's texture feature. The haralick features are namely energy, correlation, inertia, entropy, inverse difference moment, sum average, sum variance, sum entropy, difference average, difference variance, difference entropy, information measure of correlation 1 and information measure of correlation 2 are extracted at four directions (0°, 45°, 90°, 135°). The mean and variance of each of the thirteen haralick feature at four directions are extracted making a total of 78 features (Haralick et al., 1973).

4.2.Feature Selection using ACO-FFA

ACO is inspired by the foraging behavior of real ants. While walking from food sources to the nest and vice versa, ants deposit a chemical substance called pheromone on the ground. When they decide about a direction to go, they choose probabilistically paths marked by strong pheromone concentrations. This behavior is the basis for a cooperative interaction which leads to the emergence of shortest paths between food sources and their nest. In ACO algorithms, the artificial ants incrementally construct a solution by adding appropriately defined solution components to the current partial solution. Each of the construction steps is a probabilistic decision based on local information, which is represented by the pheromone information.

The pseudo code of the ACO-FFA algorithm is shown in Algorithm 2. The algorithm is initialized with the no. of features, no. of iteration, no. of ants, no. of selected features, trail intensity (T_i), evaporation rate, etc., In the first iteration, each ant will randomly choose a feature subset of m features. Only the best k subsets, ($k < na$), will be used to update the pheromone trial and influence the feature subsets of the next iteration. In the second and following iterations, each ant will start with $m - p$ features that are randomly chosen from the previously selected k-best subsets, where p is an integer that ranges between 1 and $m - 1$. This process of feature selection is a sequential process, which may delay the entire system of feature selection. So the Firefly algorithm of Algorithm 1 selects the optimal feature. The feature selected by the firefly algorithm is now carried to the next generation. In this way, the features that constitute the best k subsets will have more chance to be present in the subsets of the next iteration. However, it will still be possible for each ant to consider other features as well. For a given ant j, those features are the ones that achieve the best compromise between previous knowledge, i.e., pheromone trails, and the current best of the firefly algorithm.

The problem of feature selection can be stated as follows: given the feature set, F, of n features, find the feature subset S, which consists of m features where $m < n$, and $S \subset F$), such that the classification accuracy is maximized. The feature selection problem representation exploited by the artificial ants includes the following:

- n features that constitute the original set,
  $F = \{f_1, \ldots, f_n\}$.
- na, the number of artificial ants to search through the feature space.
- T_i, the intensity of pheromone trail associated with feature f_i.
- For each ant j, a list that contains the selected feature subset, $S_j = \{s_1, \ldots, s_m\}$.

Algorithm2:ACO-FFA algorithm for feature selection

1. Initialization:

- Set T_i = cc where cc is a constant
- $\Delta T_i = 0$, where $i = 1, \ldots, n$, and T_i is the amount of change of pheromone trial quantity for feature f_i.
- Define the maximum number of iterations.
- Define k, where the k-best subsets will influence the subsets of the next iteration.
- Define m, the number of features to select

- Define na, number of ants
- Define p, where m – p is the number of features each ant will start with in the second and following iterations.

2. First iteration,

2.1 For j = 1 to na,

- Randomly assign a subset of m features to S_j.

2.2 Go to step 4.

3. Select the remaining p features for each ant:

3.1 For mm = m – p + 1 to m,

- For j = 1 to na,
- Given subset S_j, Choose feature fi from the current best of Firefly Algorithm of algorithm 1.
- $S_j = S_j \cup \{f_i\}$.

3.2 Replace the duplicated subsets, if any, with randomly chosen subsets.

4. Evaluate the selected subset of each ant using a SVM classification algorithm:

- For j = 1 to na,
- Estimate the Mean Square Error (MSE_j) of the classification results obtained by classifying the features of S_j.
- Sort the subsets according to their MSE. Update the minimum MSE (if achieved by any ant), and store the corresponding subset of features.

5. Using the feature subsets of the best k ant:

- For j = 1 to k,

$$\Delta T_i = \begin{cases} \frac{\max\limits_{g=1:k}(MSE_g) - MSE_j}{\max\limits_{h=1:k}(\max\limits_{g=1:k}(MSE_g) - MSE_j)} & if\ f_i \in S_j \\ 0 & Otherwise \end{cases} \qquad (4)$$

$$T_i = \rho . T_i + \Delta T_i \qquad (5)$$

where ρ is a constant such that (1 - ρ) represents the evaporation of pheromone trails.

- For j = 1 to na,
- Randomly produce m – p feature subset for ant j, to be used in the next iteration, and store it in S_j.

6.If the number of iterations is less than the maximum number of iterations, go to step 3.

4.3.SVM Classifier

SVM is a classification techniques based on statistical learning theory [Bazzani et al., 2000 and Issam et al., 2002). Support vector machines (SVM) are statistical learning theory (SLT) problems used for classification. The SVM algorithm constructs a separating hypersurface in the input space by transforming the input space into a high dimensional feature space through some nonlinear mapping chosen a priori (Kernel). It constructs the maximal margin hyperplane in the feature space and the support vectors that lies on this hyperplane.

4. RESULTS

4.4.Image Database

In this research, mammograms from the Mammographic Image Analysis Society (MIAS), a Mini Mammographic Database [J. Suckling] is used. Each mammogram image has a spatial resolution of

1024x1024 pixels. This database is chosen since it contains various types of abnormalities such as calcification, well-defined, circumscribed masses, spiculated masses, ill-defined masses, architectural distortion, asymmetry and normal. Each of these abnormalities has been diagnosed and confirmed by a biopsy.

4.5.Experimental Setup

The experiments implemented in MATLAB. These techniques are experimented on 100 mammogram images with various abnormalities, 50 abnormal images with microcalcification, spiculation, circumscribed and 50 normal mammogram images. The following table (Table 1) shows the parameters used in the algorithms. The parameters of the ACO-FFA algorithm are listed below.

Table 1. Table type styles

FFA parameters	ACO parameters
Randomness α :0.2 Absorption co-efficient γ: 1.0 Randomness reduc-tion delta:0.97 Number of fireflies n:12 Max. generation = 50.	Population size : 100 Number of ant : 78 Number of Itera-tions : 200 Evaporation rate ρ : 0.75 Heuristic value η : 1 $\alpha = 1.0$ $\beta = 1.0$.

4.6.Experimental Results

The features selected by ACO, PSO and the proposed ACO-FFA are listed in Table 2. The 78 GLCM features are considered for the experiment and only the best 5 features are extracted.

Table 2. Features Selected by ACO, PSO and ACO-FFA

Techniques	Selected GLCM Feature
ACO	Correlation, Difference Average, Energy, Inertia, Sum Variance
PSO	Correlation, Energy, Inverse Difference, Information measure of correlation2, Sum Variance
ACO-FFA	Correlation, Difference Variance, Information measure of correlation2, Inertia, Sum Variance

With the selected features the testing is performed on the same set of 100 mammograms. The SVM classification results are pictorially depicted in Fig. 1.

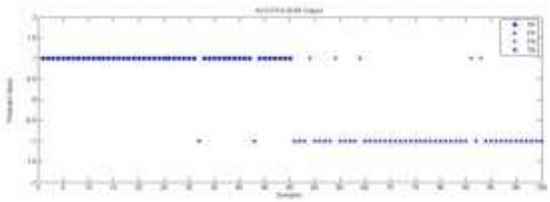

Fig. 1. Result of ACO-FFA-SVM

Receiver operating characteristic (ROC) is a statistical tool used in medical decision making and is graphical plot of Sensitivity Vs 1-Specificity for a classifier. Sensitivity is the ratio of malignant samples which have a positive test result. Specificity is ratio of benign samples which have a negative test result. The true positive fraction (or true positive ratio or Sensitivity), false positive fraction (or false positive ratio or 1-Specificity) and accuracy is defined as:

$$TPF = \frac{TP}{TP+FN} \qquad (6)$$

$$FPF = \frac{FP}{FP+TN} \qquad (7)$$

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \qquad (8)$$

where,
TP (true positives) is the no. of mammograms correctly classified as abnormal mammogram
FP (false positive) is the no. of mammograms incorrectly classified as abnormal mammogram
TN (true negative) is the no. of mammograms correctly classified as normal mammogram
FN (false negatives) is the no. of mammograms incorrectly classified as normal mammogram.

Table 3. Confusion matrix

Technique	**Actual**	**Predicted**	
		Abnormal	**Normal**
ACO	**Abnormal**	(TP) 45	(FP) 3
	Normal	(FN) 5	(TN) 47
PSO	**Abnormal**	44	4
	Normal	6	46
ACO-FFA	**Abnormal**	48	5
	Normal	2	45

The confusion matrix is constructed from the obtained results is shown in Table 3. The accuracy of the hybrid technique is tabulated in Table 4. The results are compared with the traditional ACO and PSO. Fig. 2 plots the excel graph to compare the performances of the algorithms. ROC curve plotted in Fig. 3 compares the accuracy percentage of all the three classifiers.

Table 4. Performance of the proposed techniques

Classifier	**TPF**	**TNF**	**Accuracy**
ACO	0.9	0.94	92%
PSO	0.88	0.92	90%
ACO-FFA	0.96	0.9	93%

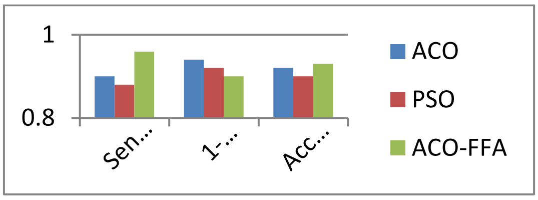

Fig. 2. Performance measure of the proposed techniques

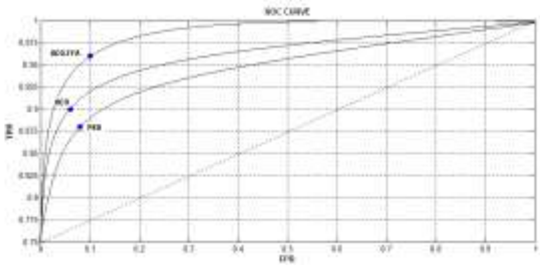

Fig. 3. ROC curves of ACO, PSO and ACO-FFA

From Table 2, it is inferred that the features correlation and sum variance are selected by all the techniques. Energy, inertia and information measures of correlation 2 are the features selected by atmost two of the techniques. The overall performance of the hybrid techniques is better than both PSO and ACO. ACO-FFA shows 2% better accuracy than ACO and 4% better accuracy than PSO.

5. CONCLUSION

This paper presents the hybrid of ACO-FFA to select the best feature of digital mammogram. The local search of ACO, stimulated by the firefly algorithm works well and gives an accuracy of 93%. The SVM classifier is the added feature of the proposed system. The SVM classifier plays dual role; it goes along with the ACO-FFA optimization and also used in testing the performance of the total system. The hybrid algorithm shows promising accuracy and is better than the ACO and PSO algorithm.

REFERENCES

1. Al-Ani, A., 2005. Feature subset selection using ant colony optimization. Int. J. Computational Int., 2: 53-58. http://www.citeulike.org/user/mass845/article/10466718
2. Bazzani, A., R. Bevilacqua, D. Bollini, R.B. Renato and R. Campanini *et al*., 2000. Automatic detection of clustered microcalcifications in digital mammograms using an SVM Classifier. ESANN'2000 Proceedings-European Symposium on Artificial Neural Networks, Bruges (Belgium), pp: 195-200. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.6592
3. Dorigo, M., V. Maniezzo and A. Colorni, 1996. Ant system: Optimization by a colony of cooperating agents. IEEE Transactions Syst. Man Cybernetics-Part B 26: 29-41. DOI: 10.1109/3477.484436
4. Dorigo, M. and L.M. Gambardella, 1997. Ant Colony System: A cooperative learning approach to the traveling salesman problem. IEEE Transactions Evolutionary Computation, 1: 53-66. DOI: 10.1109/4235.585892
5. Dorigo, M. and T. Stuetzle, 2004. Ant Colony Optimization, MIT Press/Bradford Books, Cambridge, MA. DOI: 10.1007/b99492

6. Haralick, R., M. Shanmugam and K. Its'hak Dinstein, 1973. Textural features for image classification. IEEE Transaction Syst. Man Cybernetics, 3: 610-621. DOI: 10.1109/TSMC.1973.4309314
7. El-Naqa, I., Y. Yang, M.N. Wernick, N. Galatsanos and P. Nishikawa, 2002. A Support Vector Machine Approach for Detection of Microcalcifications. IEEE Trans. Med. Imag., 21: 1552-1563. DOI: 10.1109/TMI.2002.806569.
8. Yang, X. S., 2009. Firefly algorithms for multimodal optimisation, Stochastic Algorithms Foundations and Applications 169-178. DOI: 10.1007/978-3-642-04944-6_14
9. Yang, X. S., 2011. Nature-Inspired Metaheuristic Algorithm. 2 Edn., Luniver Press, Frome, ISBN-10: 1905986289, pp: 160.
10. Yang, X. S., 2010. A Discrete Firefly Meta-heuristic with Local Search for Make span Minimization in Permutation Flow Shop Scheduling Problems. Int. J. Industrial Eng. Computations. 1: 1-10. DOI: 10.1109/ICCIE.2010.5668209