

FEATURE DIMENSION REDUCTION FOR EFFICIENT MEDICAL IMAGE RETRIEVAL SYSTEM USING UNIFIED FRAMEWORK

¹Yogapriya Jaganathan and ²Ila Vennila

¹Department of Computer Science and Engineering, Paavai Engineering College, Namakkal, Tamil Nadu, India

²Department of Electrical and Electronics Engineering, PSG College of Technology, Coimbatore, Tamil Nadu, India

Received 2013-06-23, Revised 2013-09-04; Accepted 2013-09-24

ABSTRACT

Feature dimensionality reduction problem is a major issue in Content Based Medical Image Retrieval (CBMIR) for the effective management of medical images with the support of visual features for the purpose of diagnosis and educational research field. The proposed CBMIR is used a unified approach based on extraction of visual features, optimized feature selection, classification of optimized features and similarity measurements. The Texture features are selected using Gray Level Co-occurrence Matrix (GLCM), Tamura Features (TF) and Gabor Filter (GF) in which pull out of features are formed a feature vector database. Fuzzy based PSO (FPSO) is applied for Feature selection to overcome the difficulty of feature vectors being surrounded in local optima of original PSO. This procedure also integrates a smart policymaking structure of ACO procedure into the novel FPSO where the global optimum position to be exclusive for every feature particle. The Fuzzy based Particle Swarm Optimization and Ant Colony Optimization (FPSO-ACO) technique is used to trim down the feature vector dimensionality and classification is accomplished using an extensive Fuzzy based Relevance Vector Machine (FRVM) to form collections of relevant image features that would provide an accepted way to classify dimensionally concentrated feature vectors of images. The Euclidean Distance (ED) is recognized as finest for similarity measurement between the medical query image and the medical image database. This proposed approach can acquire the query from the user and had retrieved the desired images from the database. The retrieval performance would be assessed based on precision and recall. This proposed CBMIR is used to provide comfort to the physician to obtain more assurance in their decisions for diagnosis and research students of medicine are keenness to get the crucial images fruitfully for additional investigation of their exploration.

Keywords: Medical Image Retrieval, Texture Features, Feature Optimization, Dimensionality Reduction, Feature Classification, Similarity Measurements

1. INTRODUCTION

Due to rapid growth of medical images in the hospitals, there would be a growing demand for medical image management tools and techniques to physicians for instantaneous diagnosis and to the medical research students for effective examination of their research. Content Based

Medical Image Retrieval (CBMIR) (Ghosh *et al.*, 2011; Kyung *et al.*, 2012) is receiving a vigorous research to automate medical image database. CBMIR systems have increased attractiveness because of their neutral means of considering medical image content such as texture, shape, color. Text based medical image retrieval could not precisely capture the pictorial features of an image in

Corresponding Author: Yogapriya Jaganathan, Department of Computer Science and Engineering, Paavai Engineering College, Namakkal, Tamil Nadu, India

the user's mind also it could beset by discrepancies of the annotator. Medical image retrieval system shows a dynamic role that would trace a preferred image from a large wide-ranging collection of image database. CBMIR has not made significant growth due to feature dimensionality reduction problem. Many researchers are looking to overwhelmed the Visual Feature dimensionality reduction problem and to create CBMIR as a extremely proficient system (Yogapriya and Vennila, 2013; Esmat *et al.*, 2013).

A unified approach is needed to build an efficient CBMIR which would retrieve the images quickly by selecting proficient methods and parameters in the following four levels in which feature dimensionality problem is reduced and convergence speed could be raised to acquire a optimal solutions:

- Level I : Extraction of Visual Features
- Level II: Optimized Feature Selection
- Level III: Classification of Optimized Features
- Level IV: Similarity measurements

1.1. Level 1: Extraction of Visual Features

The first level of CBMIR is Feature extraction (Akgul *et al.*, 2011; Suganya and Rajaram, 2012) which would extract the visual features and are designed as Feature Vector Database. Related to shape and color features, texture features have periodicity and scale results to the possessions of catching semantic features in images since it consumes a complete association of innumerable grey levels inside the images. The proposed system uses texture feature extraction methods such as Gray Level Co-occurrence Matrix, Tamura Features, Gabor Filter Features. From GLCM, texture features are extracted which would process the confined variations, joint probability incidence of quantified pixel pairs, the sum of squared pixel features, the companionship of distribution of elements. Tamura extracted six texture features in respect to human visual perception such as coarseness, contrast, directionality, line-likeness, regularity and roughness. Gabor filters have used sets of filters to investigate the image spectrum. The texture feature vectors are established by calculating the mean and standard deviation based on appropriate dilations and a rotation of Gabor Function. These features would be formed as entire feature vector database for further process of retrieval.

1.2. Level 2: Optimized Feature Selection

The second level of CBMIR is Feature selection which would be defined as selecting the amalgamation of features between a larger feature vector database that defines a specific feature set is finest. Researchers have addressed the dimensionality reduction problem by applying various algorithms such as Principal Component Analysis, Weighted Multi-Dimensional Scaling, Tabu Search Method (Wu *et al.*, 2009) and Evolutionary Algorithms for optimizing the features are Particle Swarm Optimization (PSO) (Ye *et al.*, 2009), Ant Colony Optimization (ACO) (Patrik and Izquierdo, 2009), Genetic Algorithms (GA) (Silva *et al.*, 2011), Gravitational Search Algorithm (GSA) (Rashedi *et al.*, 2009). The individual dimensional reduction approach based CBMIR is used to recognize the probable optimal resolutions within a sensible amount of time but the convergence rapidity is declined. PSO is one of best heuristic algorithms but the basic drawback of classic PSO is that the selection of parameters that are impulsive convergence whenever the particle and cluster finest solutions are narrowed into local minimums through the search process. A Fuzzy based PSO (FPSO) (Yong-Feng and Shu-Ling, 2009) approach is applied to astonish the impulsive convergence and besides to increase the speed of the penetrating process (Yogapriya and Vennila, 2013). The parameters of inertia weight and learning factors of PSO are robustly adapted using fuzzy IF/THEN rules. FPSO changes its behaviour during the optimization process based on information gathered at each iteration. Hence FPSO procedure would integrate a smart policymaking structure of ACO algorithm called FPSO-ACO (Nafar *et al.*, 2012) where the global optimum feature position is exclusive for every particle also it would disperse different global optimum feature positions to every distinct particle agent.

1.3. Level 3: Classification of Optimized Features

The third level of CBMIR is Feature Classification which could train and classify the optimized features into separate classes based on classification decision rules. The most extensively used classification algorithms (Rajendran and Madheswaran, 2010) are K-nearest neighbor, Fuzzy C-Means clustering, Decision Tree, Bayesian Classification also machine learning algorithms are enabled to train and classify medical images and to enrich the evidence by using Support Vector Machine (SVM), Relevance Vector Machine (RVM) (Yogapriya and Vennila, 2012).

SVM has major drawbacks such as predictions are not probabilistic also kernel functions have to satisfy Mercer conditions. Relevance Vector Machine (RVM) is used to overcome SVMs drawback also it is a standard machine learning techniques centered on statistical learning theory and having dazzling features and attentive functional performance. The shortcomings of RVM (Hong and Chen, 2011) are to delight the training points homogeneously during training; however, the consequence of the training points is different. Confidently, a fuzzy membership is required to each input point so that unrelated input points can create altered effects in learning. Fuzzy Relevance Vector Machine (FRVM) is deliberated to overcome this training struggle in which a fuzzy membership is dispensed to each training input point such that dissimilar input points can make diverse impacts in learning progression.

1.4. Level 4: Similarity Measurements

The fourth level of CBMIR is Similarity Measurements which could retrieve the relevant images from the classified image feature vector database. Various similarity measurement distance metrics are available such as Manhattan Distance (L1 metric), Euclidean Distance (L2 metric), Vector Cosine Angle Distance (VCAD), Chord Distance, Pearson’s Correlation Coefficient, Spearman Rank Coefficient (Zhu *et al.*, 2012). The Euclidean Distance (ED) is recognized as finest for similarity measurement between the medical query image and the medical image database.

2. MATERIALS AND METHODS

2.1. Methods

The proposed CBMIR system is depicted in **Fig. 1**:

- The Feature vector database containing x number of feature vectors using texture extraction methods such as GLCM, Tamura and Gabor, $x > 100$
- Execute Fuzzy Based PSO-ACO on the high dimensional feature dataset to choose the lower dimensional feature set
- Fuzzy RVM is performed for Classification in each of the optimized feature set.
- Euclidean Distance is applied as Similarity Measurements to retrieve the feature classified image databases and query image

2.2. A Unified approach for CBMIR: Texture-FPSO-ACO-FRVM

The following steps should be taken and repeated by using an unified approach as Texture-FPSO-ACO-FRVM in which an efficient CBMIR system would be developed.

Step 1: A Texture Feature Vector Database Containing a x Number of Vectors using GLCM, Tamura and Gabor, $x > 100$ and Defined in Table 1

Step 2: Generate the Preliminary Population of N Features of Swarm and Preliminary Velocity

The preliminary feature population and preliminary feature velocity of each particle would be randomly generated and defined in Equation 1 and 2:

$$\text{FeaturePopulation} = \begin{bmatrix} FC_1 \\ FC_2 \\ \dots \\ FC_{N_{\text{swarm}}} \end{bmatrix}$$

$$FC_i = [\text{FeatureCenter}_1, \text{FeatureCenter}_2, \dots, \text{FeatureCenter}_k], i = 1, 2, 3, \dots, N_{\text{swarm}}$$

$$\text{FeatureCenter}_j = [Fc_1, Fc_2, \dots, Fc_d] \tag{1}$$

$$Fc_i^{\min} < Fc_i < Fc_i^{\max}$$

$$\text{FeatureVelocity} = \begin{bmatrix} FV_1 \\ FV_2 \\ \dots \\ FV_{N_{\text{swarm}}} \end{bmatrix}$$

$$FV_i = [\text{FeatureCenter_V}_1, \text{FeatureCenter_V}_2, \dots, \text{FeatureCenter_V}_k], i = 1, 2, 3, \dots, N_{\text{swarm}}$$

$$\text{FeatureCenter_V}_j = [Fv_1, Fv_2, \dots, Fv_d] \tag{2}$$

$$Fv_i^{\min} < Fv_i < Fv_i^{\max}$$

where, FeatureCenter_j is the j^{th} feature center for the i^{th} specific features. FeatureCenter_V_j is the feature velocity of the j^{th} feature center for the i^{th} specific feature. FV_i and FC_i are the feature velocity and feature position of the i^{th} specific features respectively. d is the dimension of each feature center. Fv_i^{\max} and Fv_i^{\min} are the maximum and minimum value of the feature velocity of each point belonging to the j^{th} feature centre, respectively. Fc_i^{\max} and Fc_i^{\min} are the maximum and minimum value of each feature point fitting to the j^{th} feature center, respectively.

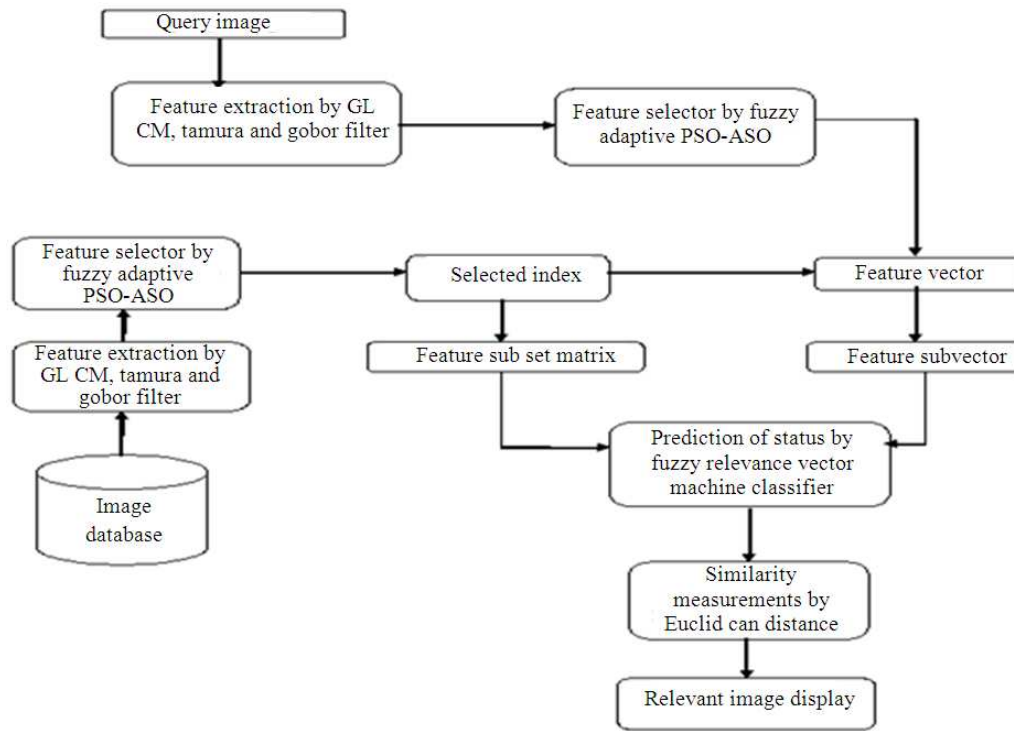


Fig. 1. The proposed unified approach based on FPSO-ACO-FRVM for medical image retrieval system

Table 1. Features extracted by GLCM, Tamura, Gabor

Features extracted by GLCM, Tamura, Gabor
Angular Second Moment, Contrast, Correlation, Information Measure of Correlation 1, Information Measure of Correlation 2, Maximum Correlation Coefficient, Inverse Difference Moment, Sum Average, Variation, Sum Variance, Difference Variance, Sum Entrophy, Entrophy, Difference Entrophy, Coarseness directionality, line likeness, regularity roughness, Mean, Standard Deviation

Step 3: Create the Preliminary Trail Intensity

The trail intensity between each feature pair of swarms is the identical and would be generated and defined in Equation 3:

$$\text{Feature_Trail_Intensity} = [ft_{ij}]_{N_{\text{swarm}} \times N_{\text{swarm}}} \quad ft_{ij} = ft_0 \quad (3)$$

Step 4: Estimate the Objective Function Value

The objective function is estimated for each specific population of N Features.

Step 5: Categorize the Preliminary Feature Population based on Objective Function Values

The preliminary feature population is arising with respect to objective function values.

Step 6: Choose the Optimum Feature Global Position (G_{optimum})

The specific features that have the minimum objective function is nominated as the optimum feature global position.

Step 7: Select the Optimum feature Local Position (P_{optimum})

The optimum feature local position is nominated for each specific features.

Step 8: Select the ith Specific Features

The ith specific feature is selected and neighbours of this feature particle should be well-defined vigorously as defined in Equation 4:

$$FS_i = \left\{ FC_j \parallel FC_i - FC_j \parallel \leq 2R_o \left(\frac{1}{1 - \exp(-pt/t_{\text{max}})} \right) \mid i \neq j \right\} \quad (4)$$

where, R_0 is the preliminary neighborhood radius, p is a parameter used to tune the neighbourhood radius over the iteration and $\|FC_i - FC_j\|$ is the Euclidean distance operator.

Step 9: Estimate the Next Feature Position for the i^{th} Specific Feature

There are two methods to Calculate the next feature position as follows.

Method 1:

If $FS_i \neq \{\}$ where $\{\}$ shows null set, then the transition probabilities between FC_i and each specific feature in FS_i are determined and defined in Equation 5:

$$[Feature_Probability] = [fP_{i1}, fP_{i2}, \dots, fP_{iM}]_{1 \times M}$$

$$FP_{ij} = \frac{(f\tau_{ij})^{\gamma_2} (1/Q_{ij})^{\gamma_1}}{\sum_{j=1}^M (f\tau_{ij})^{\gamma_2} (1/Q_{ij})^{\gamma_1}} \quad Q_{ij} = \frac{1}{|J(FC_i) - J(FC_j)|} \quad (5)$$

where, FP_{ij} is the Feature state transition probability between FC_i and the j^{th} individual in FS_i . M is the number of supporters in FS_i .

Then the cumulative feature probabilities are determined and defined in Equation 6:

$$[Cumulative_Feature_Probability] = [CfP_1, CfP_2, \dots, CfP_M]_{1 \times M} \quad (6)$$

Where $CfP_1 = fP_{i1}, CfP_2 = CfP_1 + fP_{i2}, \dots$
 $CfP_M = CfP_{M-1} + fP_{iM}$

The randomly generated number between 0 and 1 is compared with the determined cumulative feature probabilities. The first term of the cumulative feature probabilities, which is higher than the generated number, would be selected and the related position is identified as the optimum feature global position.

The i^{th} particle is then stimulated based on the following rules if FC_j is selected as the optimum feature and defined in Equation 7:

$$FV_i^{t+1} = \omega * FV_i^t + Fc_1 * rand_1(0...1) * (P_{optimumj} - FC_i^t) + Fc_2 * rand_2(0...1) * (FC_j - FC_i^t) \quad (7)$$

$$FC_i^{t+1} = FC_i^t + FV_i^{t+1}$$

The recognized pheromone level between FC_i and FC_j is modernized at the next stage and defined in Equation 8:

$$f\tau_{ij}(t+1) = \rho.f\tau_{ij}(t) + FP_{ij} \quad (8)$$

Method 2:

If $FS_i = \{\}$, then it defines that there is not any specific feature in particle's neighborhood.

The i^{th} particle is then stimulated based on the following rules and defined in Equation 9:

$$FV_i^{t+1} = \omega * FV_i^t + Fc_1 * rand_1(0...1) * (P_{optimumj} - FC_i^t) + Fc_2 * rand_2(0...1) * (G_{optimumj} - FC_i^t) \quad (9)$$

$$FC_i^{t+1} = FC_i^t + FV_i^{t+1}$$

where, FV_i^{t+1} represents the rate of position change feature velocity of the j th particle in the d th dimension and t denotes the iteration number, $rand(0...1)$ is a random number between 0 and 1. $P_{optimumj}$ is the optimum position found so far by the feature particle j . $G_{optimumj}$ is the optimum feature position among all particles. The three parameters C_1 , C_2 and ω of PSO are prominently manipulating the performance of an algorithm via Fuzzy Rules and specified in Step 10.

Next, the trail intensity would be stimulated in which the index j represents the global optimum feature particle index in the group and defined in Equation 10:

$$f\tau_{ij}(t+1) = \rho.f\tau_{ij}(t) + r; 0.1 \leq r \leq 0.5 \quad (10)$$

The altered position for the i^{th} specific feature is checked with its limit.

Step 10: Fuzzy Rules for Inertia Weight and Learning Factors

The suitable choice of inertia weight ω would provide a balance between global and local optimum points and the learning factors $c1$ and $c2$ determines the influence of personal optimum $P_{optimum}$ and global optimum $G_{optimumj}$ via Fuzzy Rules. The Input and Output variables and Linguistic values for inertia weight ω and learning factors $c1$, $c2$ are shown in **Table 2**. The Fuzzy rules for the change of inertia weight ω and learning factors $c1$, $c2$ are mentioned in **Table 3 and 4**.

Table 2. Input and output variables and Linguistic values for inertia weight ω and learning factors $c1, c2$

Input variable	Input linguistic values	Output variable	Output linguistic values
SFFV & ω	S-Small M-Medium L-Large	Change in $\omega(\Delta\omega)$	Zero-ZR OA-Optimistically accepted NA-Not Accepted
STFF & UTF	OS-Optimistically Small OM -Optimistically Medium OB -Optimistically Big OR- Optimistically Bigger	Learning Factors $c1, c2$	OS-Optimistically small OM-Optimistically medium OB-Optimistically big OR-Optimistically bigger

Table 3. Fuzzy rules for the change of inertia weight ω

		ω		
		S	M	L
$\Delta\omega$	S	ZR	NA	NA
	M	OA	ZR	NA
	L	OA	ZR	NA

Table 4. Fuzzy rules for the learning factors $c1, c2$

		UTF			
		OS	OM	OB	OR
$c1, c2$	OS	OR	OB	OB	OM
	OM	OB	OM	OM	OS
	OB	OB	OM	OS	OS
	OR	OM	OM	OS	OS

The Standardized Feature Fitness Value (SFFV) and current inertia weight ω are used as an input variable between 0 and 1 and is defined in Equation 11:

$$SFFV = \frac{FFV - FFV_{\min}}{FFV_{\max} - FFV_{\min}} \tag{11}$$

The Standardized Top Feature Fitness (STFF) is shown in Equation 12 and no. of divisions for Unchanged Top Feature Fitness (UTF) are used as input variable:

$$STFF = \frac{TFF - TFF_{\min}}{TFF_{\max} - TFF_{\min}} \tag{12}$$

The triangular membership functions of inputs and outputs are used and Mamdani-type fuzzy rule base has been used to formulate the conditional statements which involve fuzzification process. The Member functions of SFFV and STFF are represented in **Fig. 2 and 3**.

Step 11: If All of the Feature Individuals are Chosen then go to the Step 12, Otherwise go Back to Step 8 and Set as $i=i+1$

Step 12: Check the Termination Criteria

If the present iteration number touches the prearranged maximum iteration number, the search procedures would go to Step 13; otherwise the preliminary feature population would be replaced by the new population of swarms and then the algorithm goes back to Step 4. The last G_{optimum} is the solution of the problem.

Step 13: Consider the Last G Optimum Value as the Initial Solution for the Fuzzy RVM Algorithm

The G_{optimum} is considered as an initial solution of the FRVM classification problem.

The RVM is based on Bayesian probabilistic learning approach that would obtain a relevance vectors and weights by maximizing a marginal likelihood. RVM makes predictions by the sum of product of weights and kernel functions and defined in Equation 13:

$$y(x, w) = \sum_{i=1}^n w_i K(ox, ox_n) \tag{13}$$

where, $K(ox, ox_i)$ is a kernel function for optimized features and $\{w_i\}$ are the model weights. RVM use hyper planes in order to separate the two parts of the image classes such as relevant and irrelevant.

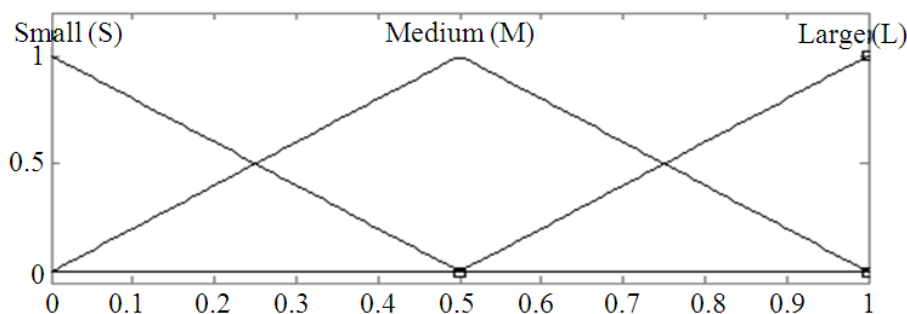


Fig. 2. Membership functions of SFFV

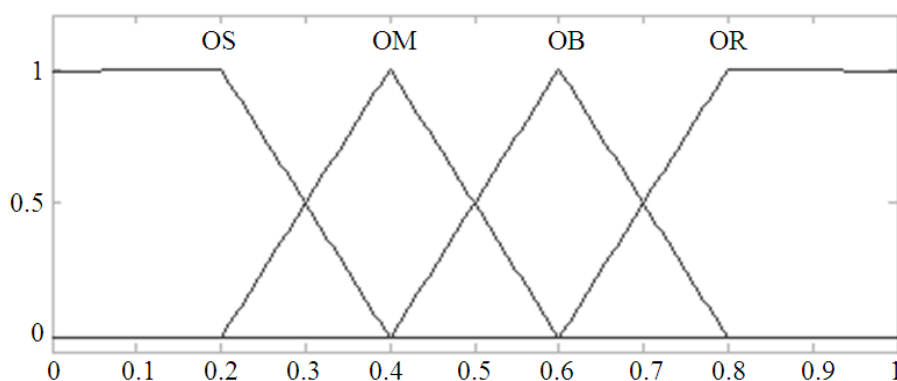


Fig. 3. Membership functions of STFF or UTTF

The fuzzy membership is introduced in RVM and represented in Equation 14:

$$-\log\{p(t|\omega)p(\omega|\alpha)\} = -\sum F_{s_n}[t_n \log y_n + (1 - t_n) \times \log(1 - y_n)] + \frac{1}{2} \omega^T A \omega \tag{14}$$

$p(t|\omega)$ -Likelihood of the optimized feature training data set $p(\omega|\alpha)$ -Gaussian distribution over w with variance α to control over fitting.

Based on the different values of FS_i , there would be a control over the transaction of the respective training points (x_i, t_i) during the classification stage. A negligible value of s_i marks the corresponding point (x_i, t_i) less significant in training. So RVM is the separate case of FRVM if we set all $FS_i = 1$ and if FS_i is positive then conclude that the optimized feature set could be correctly classified otherwise it is not under the class and defined in Equation 15:

$$FS_i = \begin{cases} FS_+, y_i = 1 \\ FS_-, y_i = -1 \end{cases} \tag{15}$$

Rules for FRVM:

R1: If ox_1 is $K(ox_1, ox_{11})$ and $\dots ox_D$ is $K(ox_D, ox_{1D})$

Then $f_1 = c_{10} + c_{11} ox_1 + \dots + c_{1D} ox_D$

R2: If ox_1 is $K(ox_2, ox_{21})$ and $\dots ox_D$ is $K(ox_D, ox_{2D})$

Then $f_1 = c_{20} + c_{21} ox_1 + \dots + c_{2D} ox_D$

⋮

⋮

⋮

Rn: If ox_1 is $K(ox_1, ox_{n1})$ and $\dots ox_D$ is $K(ox_D, ox_{nD})$

Then $f_1 = c_{n0} + c_{n1} ox_1 + \dots + c_{nD} ox_D$

D-Dimension of optimized feature input.

ox_j is an input variable; $j = 1, 2, \dots, D$.

f_i is i th local output feature variable.

$K(ox_j, ox_{ij})$ is a fuzzy set $i = 1, 2, \dots, n; j = 1, 2, \dots, D$.

c_{ij} is a consequent parameter $i = 1, 2, \dots, n; j = 1, 2, \dots, D$.

The number of rule would be selected based on number of relevance vectors and membership function is represented in Fig. 4.

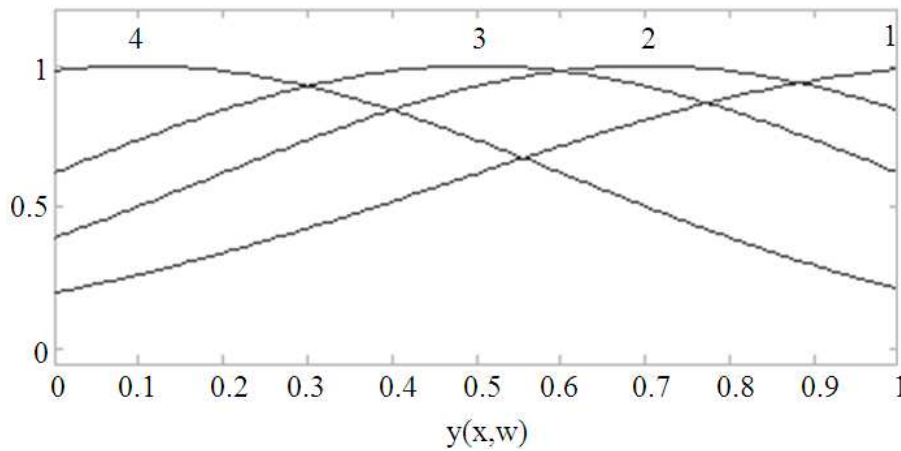


Fig. 4. Membership Functions of $y(x,w)$ relevance vectors

Step 14: Similarity Measurements

Euclidean distance is used to find the similarity between the query image features and categorized image features in the database which would retrieve an effective images based on the principle of smaller the distance also better the similarity defined in Equation 16:

$$FS_{ED} = \sqrt{\sum_{i=1}^N (Q_F[i] - DB_{CF}[i])^2} \quad (16)$$

where, $Q_F[i]$ the i^{th} query image features and $DB_{CF}[i]$ is the matching feature in the classified feature vector database and N refers to the total number of images in the database.

3. RESULTS

This proposed CBMIR is implemented with the image database of 1000 images which are gray level images and basically related images includes some parts of human body like lung, liver, kidney, brain. A vast number of features are extracted from the image and the sample lung feature extraction is represented in **Table 5**. This would upturn the complexity of the system. So FPSO-ACO would extract the features that are appropriate to retrieve images from the database. This study streamlines the system, rises the accuracy and reduces the complexity of the system. **Table 6** shows the ten optimization results of FPSO-FRVM and FPSO-ACO-FRVM in respect to Relevance vectors and Particle Fitness Value.

This results shows that approximately flawless fit with the optimal value with great probability also which

proves reliable relevance vector number and close by sparse level. For each image, selected texture features are given as input to the classification phase. The total number of feature classes are 58. The number of relevance vectors would be less due to feature optimization. The proposed FPSO-ACO-FRVM method parameters, optimum and average solutions are represented in **Table 7**.

3.1. Comparisons on Retrieval Performance

The retrieval performance would be assessed based on precision and recall. Precision is the ratio of retrieved images that could be relevant and measures the quality, while recall is the ratio of relevant images that could be retrieved and measures the completeness. The Precision (P) and Recall (R) are then defined as:

$$p = \frac{r}{n_1} = \frac{\text{Number of relevant images}}{\text{Number of retrieved images}}$$

$$R = \frac{r}{n_2} = \frac{\text{Number of relevant images}}{\text{Total number of retrieved images}}$$

in the Database

The **Fig. 5-8** represent the sample query image and output of retrieved images of the brain and lung images and **Fig. 9-12** represents brain, lung, liver and spine image retrieval performance analysis based on Precision and Recall of the proposed approach and remaining three methods and shows that the proposed method has more exactness than the remaining methods.

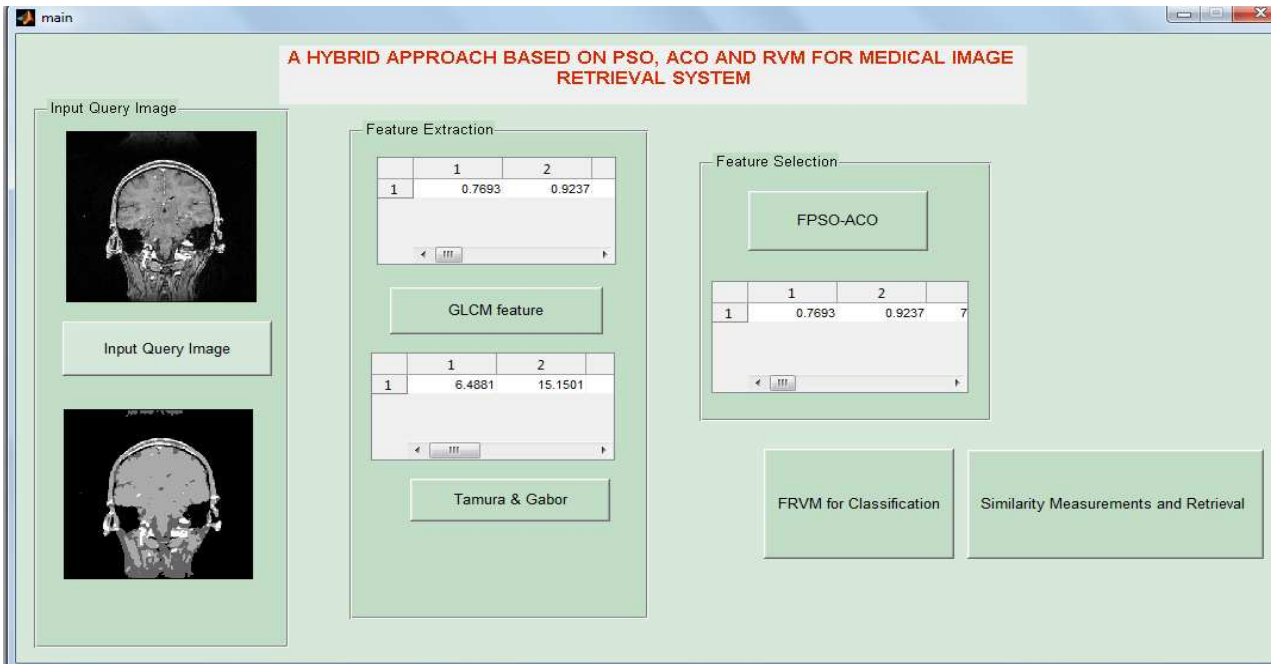


Fig. 5. Sample query image "brain.jpg"

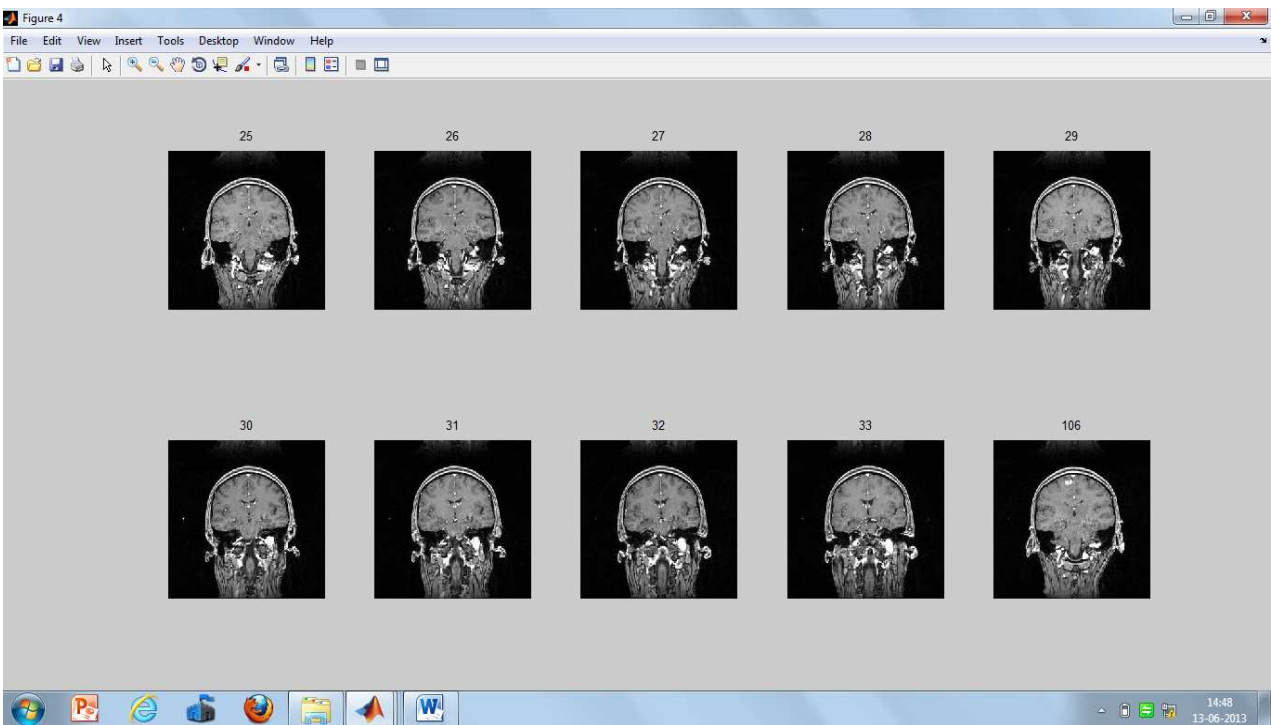


Fig. 6. Output of retrieved images of "brain.jpg"

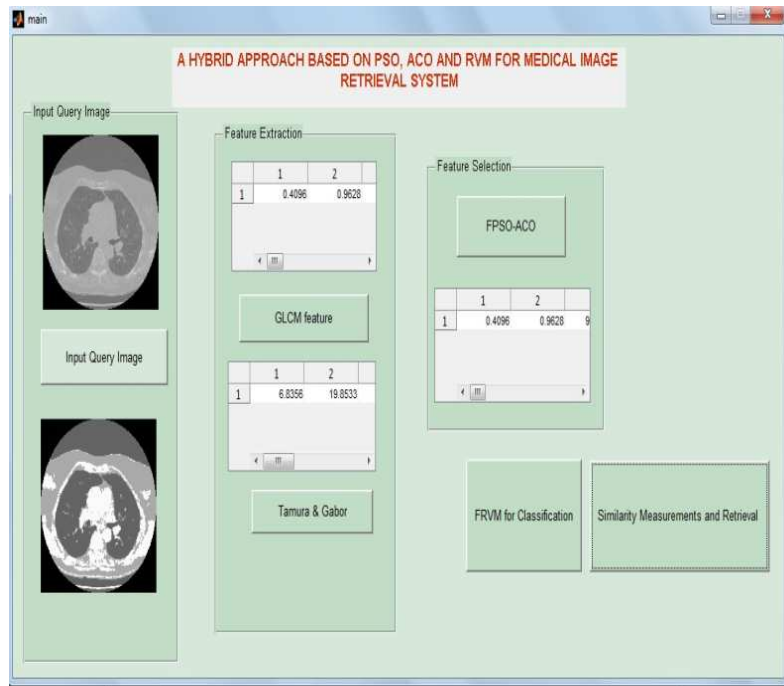


Fig. 7. Sample query image “lung.jpg”

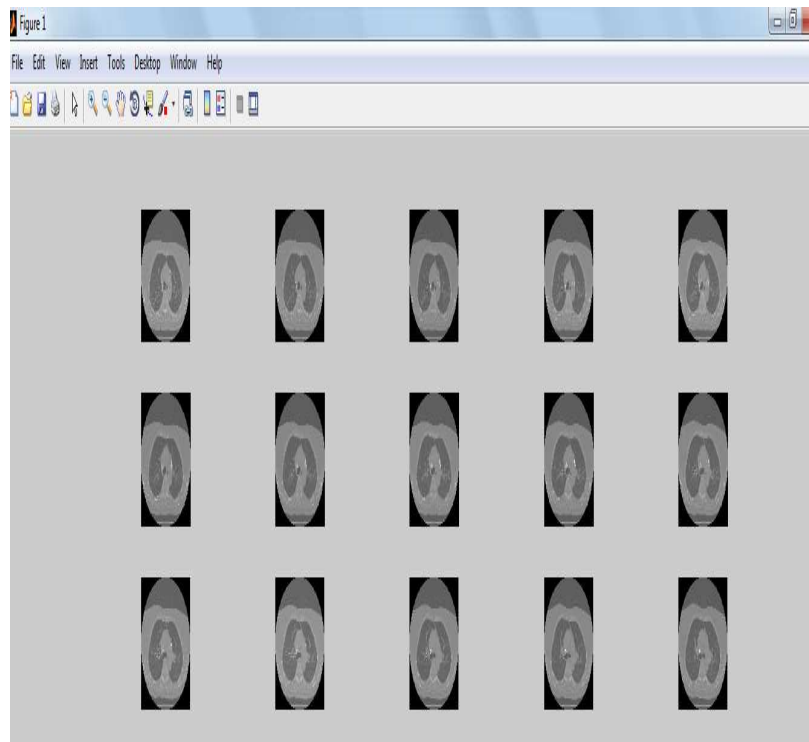


Fig. 8. Output of retrieved images of “lung.jpg”

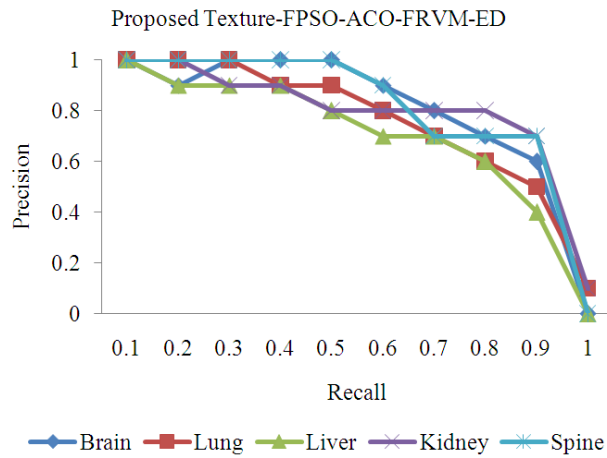


Fig. 9. Precision-Recall of Texture-FPSO-ACO-FRVM-ED

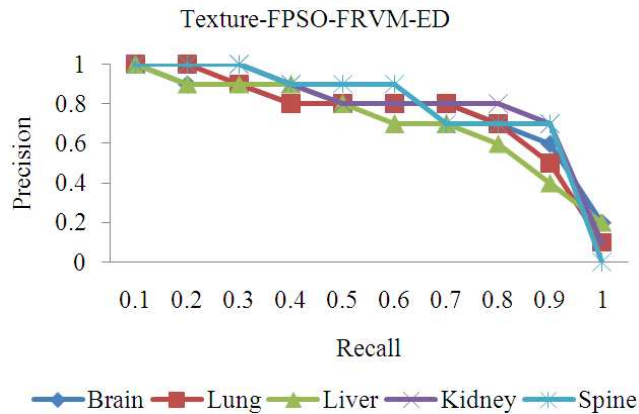


Fig. 10. Precision-Recall of Textures FPSO-FRVM-ED

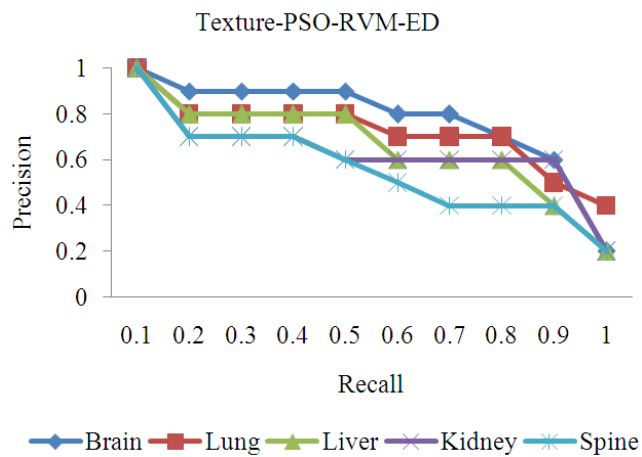


Fig. 11. Precision-Recall of Texture PSO-RVM-ED

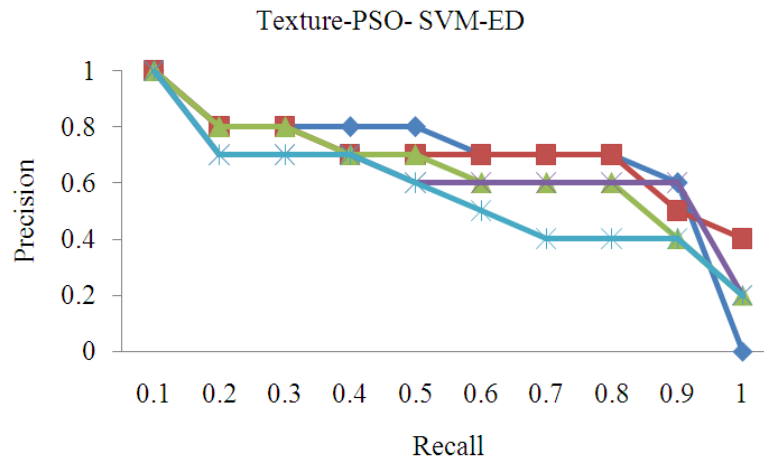


Fig. 12. Precision-Recall of Texture-PSO-SVM-ED

Table 5. Feature extraction for 15 lung images

ID	Image	Autocorrelation	Contrast	Correlation	Cluster pemi	Cluster shade	Dissimilarity	Energy
1	Data/lung/1JPG	16.49044884	0.1168811	0.974434180	215.5401972	28.80818928	0.080407475	0.4211946
2	Data/lung/2JPG	16.99867494	0.1163450	0.975797141	237.1443339	29.45932528	0.071430760	0.3013322
3	Data/lung/3JPG	16.89834594	0.1021063	0.978038160	223.1869941	29.55379165	0.064016544	0.3104339
4	Data/lung/4JPG	16.94276195	0.1141736	0.975990500	232.7653248	29.55379165	0.070009957	0.3059237
5	Data/lung/5JPG	16.97571998	0.1150314	0.975849798	233.2700141	30.10903766	0.070216759	0.3064252
6	Data/lung/6JPG	17.42662377	0.1495672	0.970202633	254.0640583	30.37509347	0.099360447	0.2770359
7	Data/lung/7JPG	17.47815181	0.1435011	0.971509917	255.9109313	30.51999637	0.093478094	0.2780954
8	Data/lung/8JPG	17.38241039	0.1346775	0.972897020	250.5308642	30.62789637	0.098493030	0.2875938
9	Data/lung/9JPG	17.64131817	0.1616613	0.968404297	261.7255266	30.62789637	0.110359222	0.2663671
10	Data/lung/10JPG	17.18958716	0.1211818	0.974951698	239.8068901	31.02157492	0.074077053	0.3036289
11	Data/lung/11JPG	17.72218137	0.1626838	0.968374900	264.2013553	31.38325475	0.110891544	0.2642610
12	Data/lung/12JPG	17.25053615	0.1175513	0.975770723	241.4548161	31.36176062	0.070515472	0.3065470
13	Data/lung/13JPG	17.75861673	0.1516965	0.970503263	265.0095427	31.65757232	0.100547641	0.2676540
14	Data/lung/14JPG	18.09853707	0.1774050	0.966752542	280.1805949	32.05722478	0.125337010	0.2454860
15	Data/lung/15JPG	18.61542969	0.2049786	0.963559882	303.5662378	32.67733637	0.152136949	0.2190821
		Entropy	Homogeneity	Coarseness	Directionality	Line likeness	Mean	Standard deviation
1	Data/lung/1JPG	1.42119460	0.964383681	1920.989200	0.000553	0.079632	0.007025736	16.49044884
2	Data/lung/2JPG	1.42541700	0.969734796	1508.143100	0.000686	0.165142	0.038862222	16.99867494
3	Data/lung/3JPG	1.37397670	0.972703674	601.661290	0.000499	0.191917	0.024487982	16.89834594
4	Data/lung/4JPG	1.14033479	0.970391136	1056.338500	0.000469	0.236568	0.030158262	16.94276195
5	Data/lung/5JPG	1.30999500	0.970356988	1254.091200	0.000330	0.145695	0.025487982	16.89994594
6	Data/lung/6JPG	1.52892200	0.956349507	891.842340	0.000743	0.169151	0.079286820	16.65276195
7	Data/lung/7JPG	1.52155870	0.959229282	535.302950	8.95E-05	0.260812	0.028307035	17.42662377
8	Data/lung/8JPG	1.48060780	0.963456201	362.980090	0.000629	0.182369	0.029307035	17.47815181
9	Data/lung/9JPG	1.56689170	0.950923649	437.278580	0.000212	0.245956	0.027248695	17.38241039
10	Data/lung/10JPG	1.41194280	0.968622409	490.452300	0.000647	0.225941	0.123231752	17.64131817
11	Data/lung/11JPG	1.57181830	0.950706189	181.496400	0.000870	0.229542	0.041653468	17.18958716
12	Data/lung/12JPG	1.39860130	0.970354754	74.437341	0.000942	0.210794	0.073491856	17.72218137
13	Data/lung/13JPG	1.48860130	0.955773208	784.192740	0.000419	0.245446	0.041776773	17.25053615
14	Data/lung/14JPG	1.63693200	0.965577321	939.429980	0.000462	0.154800	0.019832619	17.75861673
15	Data/lung/15JPG	1.73264650	0.930152457	424.143320	0.000284	0.232140	0.054368663	18.09853707

Table 6. Sample optimization results of FPSO-FRVM

Optimization: FPSO	Relevance vectors FRVM	Particle fitness value	Optimization FPSO-ACO	Relevance vectorsFRVM	Particle fitness value
3.8222	9	0.1056	4.9111	7	0.1054
3.8658	9	0.1042	4.9432	7	0.1049
3.5672	9	0.1062	4.9334	7	0.1043
3.4682	9	0.1063	4.8968	7	0.1056
3.8598	9	0.1009	4.6745	7	0.1052
3.5056	9	0.1064	4.6989	7	0.1054
3.8216	9	0.1051	4.8832	7	0.1058
3.4345	9	0.1026	4.7896	7	0.1048
3.8277	9	0.1044	4.6856	7	0.1047
3.4402	9	0.1041	4.8234	7	0.1053

Table 7. No. of relevance vectors for No. of training data and parameter value

No. of training data	FPSO-ACO-FRVM:							Optimum Solutions	Average Solutions
	No.of relvance vectors	c1,c2	ω	a	r	R_0	ρ		
50	7	1,1	0.4	16	0.5	10	0.99	95.564	95.555
56	7	1,1	0.4	15	0.5	10	0.99	95.567	95.565
75	7	1,1	0.4	14	0.5	11	0.99	96.556	95.554
100	7	1,1	0.4	15	0.5	11	0.99	96.558	95.546
125	9	2,2	0.2	15	0.5	10	0.99	96.556	95.453
150	9	2,2	0.2	15	0.5	10	0.99	96.456	95.521
175	9	2,2	0.2	15	0.5	10	0.99	96.501	95.499
200	9	2,2	0.2	15	0.5	10	0.99	96.558	95.554

Table 8. Comparison of proposed method

Methods	Overall features	Corrected features	Model error	Accuracy
PSO-SVM	58	45	0.089	77.58
PSO-RVM	58	49	0.032	84.48
FPSO-FRVM	58	53	0.019	91.37
FPSO-ACO-FRVM	58	56	0.009	96.55

4. DISCUSSION

The proposed system has addressed the feature dimensionality reduction problem and effectively retrieves images based on Feature Extraction, Feature Selection, Feature Classification and Similarity measurements. This study includes analysis of proposed Unified algorithm FPSO-ACO-FRVM with three methods such as FPSO-FRVM, PSO-RVM and PSO-SVM.

In feature extraction phase, the results were reasonably good and it has given about 95% of the performance of texture feature extraction. In feature selection, the parameters are impulsive convergence whenever the particle and cluster finest solutions are narrowed into local minimums through the search process. Initially, a Fuzzy PSO is applied for feature selection but it changes the behaviour during the optimization process based on information gathered at each iteration. Hence, A fuzzy PSO is combined with ACO and used to astonish the impulsive convergence

and besides to increase the speed of the penetrating process and it gives about 96% of the performance of this phase. The parameter selection to be taken in elegant way in future. In Classification phase, approximately all the images were classified using a fuzzy membership assigned to input point using FRVM in which unrelated optimized feature points can create altered effects in learning and produces 96% of classification performance and conventional Euclidean distance method was used for retrieval and it has given better performance result.

The combined approach FPSO-ACO-FRVM gives the better performance than FPSO-FRVM, PSO-RVM and PSO-SVM. This unified environment is offer more confidence to the physician and medical research students. The traditional parameters such as precision and recall measurements were used for measuring the retrieval performance. The effectiveness of the optimization, classification testing, potential performance and accuracy is analyzed with 160 training feature samples and 58 testing feature samples of lung images. For 58 testing samples the

overall features, corrected features, model error and accuracy of PSO-SVM, PSO-RVM, FPSO-FRVM and FPSO-ACO-FRVM is represented in **Table 8**, indicate that the proposed method has more accuracy than the remaining three methods. The results shows that FPSO-ACO-FRVM converges to the global optimum of nearly 96% in all training data. Hence this approach is able to find the optimum feature selection and classification in all runs in which it could speed up the retrieval process of CBMIR.

5. CONCLUSION

A unified approach based on FPSO-ACO-FRVM is proposed for effective retrieval of medical images, it comprises numerous benefits when compared to existing CBMIR. Texture Features are extracted using GLCM, Tamura and Gabor features and stored in a feature vector database. FPSO-ACO is applied for feature selection to overcome the difficulty of feature vectors being surrounded in local optima of PSO and to identify the global optimum position to be exclusive for every feature particle. FPSO-ACO is used to diminish the feature vector dimensionality concerns while choosing the significant features in the feature vector database in which computational complexity is reduced. FRVM is used for feature classification in which the speed and response rate of the retrieval procedure is increased. Euclidean distance has been used to recognize the similarity between classified features and query image features. This proposed framework is used to help the physician to obtain more assurance in their decisions for diagnosis and medical research students are keenness to get the crucial images fruitfully for additional investigation of their exploration.

6. ACKNOWLEDGEMENT

We would like to thank Dr. P. Senthil Kumar, MDRD, DNB, Radiologist, Saravana Scans, Salem, for providing suggestions and continuous encouragement to make this CBMIR as a fruitful system. We would like to thank Prof. R. Arangasamy and Prof. K. Devaki, who have worked with us in our dataset collection and their valuable support to our work.

7. REFERENCES

- Akgul, C.B., D.L. Rubin, S. Napel, C.F. Beaulieu and H. Greenspan *et al.*, 2011. Content-based image retrieval in radiology: Current status and future directions. *J. Digital Imag.*, 24: 208-222. DOI: 10.1007/s10278-010-9290-9
- Esmat, R., H. Nezamabadi-pour and S. Saryazdi, 2013. A simultaneous feature adaptation and feature selection method for content-based image retrieval systems. *J. Knowel. Based Syst.*, 39: 85-94. DOI: 10.1016/j.knosys.2012.10.011
- Ghosh, P., S. Antani, R. Long and G.R. Thoma, 2011. Review of medical image retrieval systems and future directions. *Proceedings of the 24th International Symposium on Computer Based Medical Systems*, Jun. 27-30, IEEE Xplore Press, Bristol, pp: 1-6. DOI: 10.1109/CBMS.2011.5999142
- Hong, L.I. and G. Chen, 2011. Image classification technique based on fuzzy-relevance vector machine. *J. Inform. Comput. Sci.*, 8: 3989-3995.
- Kyung, H.H., H. Lee and D. Choi, 2012. Medical image retrieval: Past and present. *J. Healthc. Inform. Res.*, 18: 3-9. DOI: 10.4258/hir.2012.18.1.3
- Nafar, M., G.B. Gharehpetian and T. Niknam, 2012. Using modified fuzzy particle swarm optimization algorithm for parameter estimation of surge arresters models. *Int. J. Innov. Comput. Inform. Control*, 8: 567-581.
- Piatrik, T. and E. Izquierdo, 2009. Subspace clustering of images using Ant colony optimization. *Proceedings of the International Conference on Image Processing*, Nov. 7-10, IEEE Xplore Press, Cairo, pp: 229-232. DOI: 10.1109/ICIP.2009.5414503
- Rajendran, P. and M. Madheswaran, 2010. Hybrid medical image classification using association rule mining with decision tree algorithm. *J. Comput.*, 2: 127-136.
- Rashedi, E., H. Nezamabadi-Pour and S. Saryazdi, 2009. GSA: A gravitational search algorithm. *Inform. Sci.*, 179: 2232-2248. DOI: 10.1016/j.ins.2009.03.004
- Silva, S.F.D., M.X. Ribeiro, J.D.E.S.B. Neto, C. Traina-Jr. and A.J.M. Traina, 2011. Improving the ranking quality of medical image retrieval using a genetic feature selection method. *Decision Support Syst.*, 51: 810-820. DOI: 10.1016/j.dss.2011.01.015
- Suganya, R. and S. Rajaram, 2012. Content based imageretrieval of ultrasound liver diseases based on hybridapproach. *Am. J. Applied Sci.*, 9: 938-945. DOI: 10.3844/ajassp.2012.938.945
- Wu, J., X. Wang, D. Yan, C. Wei and Y. Zhang, 2009. Method for medical image retrieval based on frequency layer feature. *Proceedings of the International Conference on Information Technology and Computer Science*, Kiev, Jul. 25-26, pp: 622-624. DOI: 10.1109/ITCS.2009.321

- Ye, Z., B. Xia and D. Wang and X. Zhou, 2009. Weight optimization of image retrieval based on particle swarm optimization algorithm. Proceedings of International Symposium on Computer Network and Multimedia Technology, Jan. 18-20, IEEE Xplore Press, Wuhan, pp: 1-3. DOI: 10.1109/CNMT.2009.5374753
- Yogapriya, J. and I. Vennila, 2012. Medical image retrieval system using GGRE framework. Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering, Mar. 21-23, IEEE Xplore Press, Salem, Tamilnadu, pp: 246-251. DOI: 10.1109/ICPRIME.2012.6208352
- Yogapriya, J. and I. Vennila, 2013. A hybrid approach based medical image retrieval system using feature optimized classification similarity framework. Am. J. Applied Sci., 9: 549-562. DOI: 10.3844/ajassp.2013.549.562
- Yong-Feng, X. and Z. Shu-Ling, 2009. Fuzzy particle swarm clustering of infrared images. Proceedings of International Conference on Information and Computing Science, May 21-22, Manchester, pp: 122-124. DOI: 10.1109/ICIC.2009.139
- Zhu, Z., C. Zhao and Y. Hou, 2012. Research on similarity measurement for texture image retrieval. PLoS ONE, 7: 1-14. DOI: 10.1371/journal.pone.0045302