International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-9 Issue-1, November 2019

Image Feature Selection using Ant Colony Optimization

Richa Sharma, Anuradha Purohit

Abstract: With the enhancement in imaging technology, huge amount of information is being created whose classification is a challenging task among researchers. The classification performance is solely dependent on the quality of the features defining the respective images. Hence, feature extraction from images and feature selection from this huge data is required. Image Feature Selection is a vital step which can significantly affect the performance of image classification system. This paper proposes an efficient combination of image feature extraction technique and feature selection strategy using Ant Colony Optimization (ACO). ACO utilizes overall subsets performance and local feature importance to fetch problem domain finding prime solutions. For ACO based feature selection in images most of the researchers use priori information of features. However, in this work, the features are extracted separately through the detailed analytical process which carries the best suited features for different classes of images. The performance in terms of classification accuracy has been enhanced through the optimization of features in the dataset. For experimentation and evaluating the performance of proposed work, Corel and Caltech datasets are used. Satisfactory results have been obtained with the proposed approach.

Keyword : Image Feature Selection; Feature Extraction; Ant Colony Optimization; MATLAB; Grey Level Co-occurrence Matrix.

I. INTRODUCTION

In several pattern recognition algorithms, like image classification feature selection plays a significant role. Feature selection is performed to select the most relevant features in order to reduce the dimensionality of feature space [18]. Considering all features does not always leads to better performance, thus feature selection eventually improves the classification accuracy and helps in reducing the time consumption.

Feature Selection can be performed using combinatorial optimization problem with objective of finding the optimal solution from the feasible solution set [17]. Ant colony optimization (ACO) is a combinatorial optimization problem solving technique which utilizes collective behaviour of real ants. It is an iterative and probabilistic meta-heuristic method whose behaviour is simulated by the natural behaviour of real ants consisting of mechanisms of adaptation and cooperation.

Revised Manuscript Received on November 05, 2019.

Richa Sharma*, Department of Computer Engineering, Shri G.S. Institute of Technology and Science, Indore, Madhya Pradesh, India.Email:rruchi.64@gmail.com

Anuradha Purohit, Department of Computer Engineering, Shri G.S. Institute of Technology and Science, Indore, Madhya Pradesh, India.Email:anuradhapurohit78@gmail.com.

A. *Image Feature Selection*

Image classification is one of the important pattern recognition methods. High dimensional space contains redundant and irrelevant features, thus performance of image classification is degraded. Therefore, feature selection in images, from all the feature extracted is required. By reducing unessential features, Image Feature Selection improves the quality of data set. Figure 1 demonstrates the general process of feature subset selection using search technique.

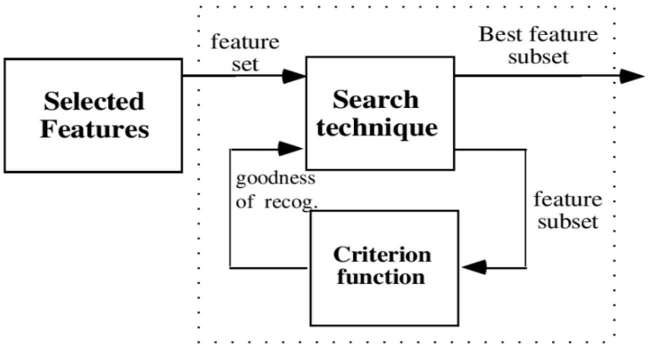

Figure 1 : Feature Selection

B. Ant Colony Optimization

ACO is metaheuristic search algorithm, based on foraging behaviour of ants and population-based approach for finding the shortest path to obtain food sources. ACO can be used for finding solutions to various combinatorial optimization problems.

Ant Colony Optimization uses a chemical substance called pheromone and heuristic information to guide its search. Thus, ACO can explore more solutions than greedy heuristics. Pheromone is a source of communication in ants which they deposits while travelling, other ants follows the path containing largest amount of pheromone. As it gets evaporated over time, the amount of pheromone in shorter route will be more than larger one and will subsequently attract a greater number of ants. This is a positive feedback loop, in which the probability of an ant to choose a particular path will depends upon the number of times same path is used by other ants. In this way, ACO optimizes its path to the most optimal solution.

ACO being a metaheuristic method, because of its random nature, can perform faster than heuristic and complete search methods. Thus, using ACO for selection of features in images will result in larger space exploration with lesser time complexity. Figure 2 demonstrates the general behaviour of ants for the selection of shorter routes.

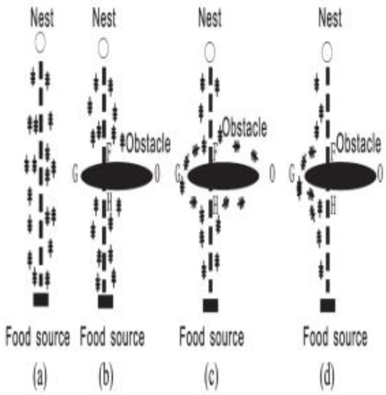

Figure 2: Route choice behavior of ants

The rest of the paper is organized as follows: Section I contains introduction of feature selection along with details of ant colony optimization. Section II reviews the related work done. Section III includes the problem statement for the proposed approach and steps involved for methodology are discussed in Section IV. Section V contains experimental analysis for the proposed work along with results obtained. Section VI concludes the paper along with the future research work.

II. RELATED WORK

Goal of feature selection is described in [1] using three folds that include: enhancing the predictive accuracy of the predictors, providing cost-effective and faster predictors, and providing the good knowledge of the underlying process that have generated the data. Researchers in [2] portrays the existing strategies that are accessible for dealing with various classes of problems as there is a plenty of work in this field. Search systems that can be utilized with multivariate data are classified as sequential, exponential and randomized algorithms [19].

The categorization for feature selection methods is described as filter methods, wrapper methods, hybrid and embedded methods [2]. In image processing, features selection method for processing of medical images are widely used for data mining and knowledge discovery and allows elimination of irrelevant and redundant features[16].

Ant Colony Optimization (ACO) metaheuristic was proposed to solve hard combinatorial optimization problems [20]. There are several variants of ACO algorithm like Ant Colony System (ACS) or Min Max Ant System (MMAS). ACO was first used to solve TSP problem and succeeded in finding better solutions [3]. Later on it was applied on various optimization problems including classification and data mining [9],[10].

ACO based feature subset search procedure is presented in [4] where local importance of a feature and overall performance of various subsets is determined using evaluation functions and classifier algorithms. In [15] a filter based strategy is described in which a classifier is combined with ACO to upgrade the performance. Proposed technique has remarkable capacity to create subsets with lesser features and accomplished higher classification accuracy. ACO can be used for selecting relevant features in images with lesser time in computing and reducing memory requirement [7]. Band selection for hyperspectral images is also an optimization problem having the objective function as highest total distance. In [11] Unsupervised based band selection algorithm is presented. To solve this problem ACO algorithm including the transition probability and pheromone update is used.

In [13], ACO algorithm has been used for decreasing the input space dimensions based on selecting feature subsets in images. It has strong capability in search space and will be able to find optimal subset of feature for hyper-spectral imagery. ACO outperforms GA based feature selection methodology. An ACO-based approach is proposed in [7] for choosing subsets from feature set in images to decrease the memory necessity and calculation time. This algorithm utilizes the classifier performance and list of feature set size to direct the search, and advances the list of features considering its size.

III. PROBLEM STATEMENT

Various feature selection techniques have been proposed by the researcher over the last decade. All these techniques were based on either filtering or wrapping strategies. The methodologies proposed for the feature selection were mostly based on the statistical measures to evaluate the features and the respective goodness of feature subset. However, these measures were subject to its biases towards attributes with large number of distinct values. These methods used the already available features of the database and its classes. The features utilized for the analytical study of the feature selection algorithm also played an important role in the effectiveness of the feature selection method. So the best suited features should be selected through proper feature extraction algorithm and it should be effectively augmented with the feature selection methodology.

ACO being a subset of Swarm Intelligence works on collaborative behavior of agents, hence used for solving various types of optimization problems. Also, the classification accuracy and computation complexity has been compromised in these algorithms. So, a combination of image feature extraction algorithm for the extraction of best suited features and feature selection algorithm for providing better classification accuracy at less computational complexity is the rationale behind this research work.

IV. METHODOLOGY

The proposed methodology in this work includes augmentation of the feature extraction and feature selection strategy using ant colony optimization algorithm. To attain the accuracy at low computation complexity, the overall work is divided in two parts: Feature Extraction and Feature Selection. The overall system architecture is shown in Figure 3. The features of the images of various classes of the database are extracted using Grey Level Co-occurrence Matrix (GLCM) and other statistical analysis. This feature encompasses the best possible characteristics of images which are required in most of the real time application like face recognition, biomedical signal analysis, cyber security etc.

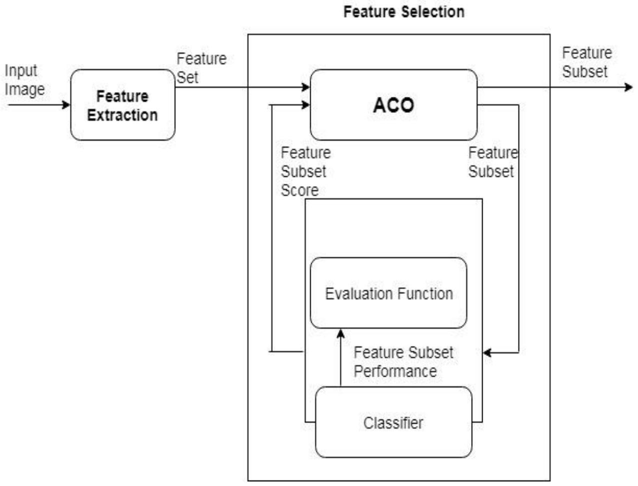

Figure 3: System Architecture with ACO

I.Feature Extraction

Relevant features needs to be extracted from the images to represent features in feature vector form. In this process, firstly data is fetched in form of images which will be converted into numerical data representing several extracted features from images including entropy, correlation etc. Feature Extraction is very essential step of the system. It can be described using following steps:

1. Convert rgb(color) image to gray scale image.
2. Using this gray scale image, calculate GLCMs (Gray Level Co-occurrence Matrix) for feature calculation.
3. Using GLCMs, compute features like entropy, moments, contrast etc. using respective formulas.

Grey Level Co-occurrence Matrix uses image as input, then convert that image in matrix form. The features considered in this work are listed in TABLE I.

TABLE I. LIST OF FEATURES EXTRACTED

S.No.	Feature
1.	Uniformity / Energy / Angular Second Moment
2.	Entropy
3.	Dissimilarity
4.	Contrast / Inertia
5.	Inverse difference
6.	Correlation
7.	Homogeneity / Inverse difference moment
8.	Autocorrelation
9.	Cluster Shade
10.	Cluster Prominence
11.	Maximum probability
12.	Sum of Squares
13.	Sum Average
14.	Sum Variance
15.	Sum Entropy
16.	Difference variance
17.	Difference entropy
18.	Information measures of correlation (1)
19.	Information measures of correlation (2)
20.	Maximal correlation coefficient
21.	Inverse difference normalized (INN)
22.	Inverse difference moment normalized (IDN)
23.	Mean
24.	Variance
25.	Skewness
26.	Kurtosis
27.	4 Moments

After extracting the features from the images with different orientation and characteristics, the dataset of features is prepared and passed to the feature selection algorithm for the classification and optimization of classes.

II. Feature Selection

Ant Colony optimization is used in this work owing to its superior performance in terms of classification accuracy and computation complexity as compared to the conventional statistical methods. As ACO takes decision on the probabilistic domain through artificial pheromone trails and local heuristic information, it can explore larger number of solutions than conventional one. Also the phenomenon of trail evaporation and intensity reduction in pheromone trail with time makes it very adaptive and dynamic for most of the applications as it helps in the convergence of the algorithm towards the desired solution. Table II describes the parameter initialized for ACO algorithm.

TABLE II. PARAMETER INITIALIZATION

Alpha	1
Beta	1
Rho	0.75
No. of ants	200
No. of iteration	10
Initial pheromone intensity	1

The input required for ACO includes: number of features(n), number of ants(na), pheromone intensity (Ti), list of selected features (S_j) for ant j, next iterations depends on k best subset (k), initial pheromone(p), features to be used in first iteration(m). The output of the algorithm will be global best solution. The steps involved in applying ACO for selecting feature subsets to solve any optimization problem are as given:

1. Initialize value of pheromone and other parameters.
2. If in the first iteration, for all ants randomly assigns subsets of features to S_j(Initial Subset). Goto step 4.
3. For every ant, select the remaining features such that: The feature f_j that maximizes the SM (Selection Measure) value will be selected.
4. Replace randomly chosen subsets from duplicate subsets.
5. Using classification algorithm, evaluate fitness of chosen subsets. Evaluate MSE_j (Mean Square Error) and sort subsets accordingly.
6. Update pheromone information from subsets obtained by best k ants.
7. Produce m-p feature subsets randomly for ant j, from subsets of best k ants. Store it in S_j.
8. If number of iterations is less than maximum iteration or desired MSE not achieved goto step 3.

Mathematical representations of formulas used for the algorithm are as follows:

a) Selection Measure (SM):

$$SM_i^{Sj} = \frac{(\tau i)^{n}(LI_i^{S_j})^{k})}{\sum g \in Sj\ (\tau g)^{n}(LI_g^{S_j})^{k}} \quad (1)$$

where LI for feature f_i represents its local importance, effect of local pheromone and trail intensity controlled by n and k.

b) Local Importance (LI):

$$LI_i^{S_j} = I(C; f_i) \times \left[\frac{2}{1 + \exp(-\alpha D_i^{S_j})} - 1\right] \qquad (2)$$

Where,

$$D_i^{S_j} = \min\left[\frac{H(f_i) - I(f_i, f_s)}{H(f_i)}\right] * \frac{1}{|S_j|}\sum_{f_s \in S_j}\left[\beta\left(\frac{I(C;\{f_i, f_s\})}{I(C;f_i) + I(C;f_s)}\right)^{\gamma}\right] \qquad (3)$$

c) Pheromone Update:

$$\Delta\tau_i = \frac{\max(MSEg) - MSEj}{\max h = 1{:}k(\max g = 1{:}k(MSEg) - MSEh)} \qquad (4)$$

$$\tau_i = \rho{:}\ \tau_i + \Delta I \qquad (5)$$

α, β, and γ are constant parameters, Entropy of f_i is represented by $H(f_i)$, $I(f_i, f_s)$ represents value of mutual information between fs and f_i and mutual information between class labels and feature is represented by $I(C; f_i)$.

III. Evaluation Measure using KNN Classifier

The KNN classifier will be used to test accuracy of the data obtained. KNN works for both classification and regression. It helps to find whether a given image is classified accurately or not. Steps involved for classification are as follows:

1. Load trained model formed using subset of selected features calculated in above step.
2. Train the classifier using training data.
3. Use KNN (K-Nearest Neighbor) for classification.

The complete algorithm of proposed approach in detail is shown in the form of a flow chart in Figure 4. It reflects the complete process of ACO implemented over the features database. After extracting the features from the images with different orientation and characteristics, the dataset of features is prepared and passed to the feature selection algorithm for the classification and optimization of classes. On the basis of iterative analysis of the pheromone performance and regressive stopping criteria analysis, the best suited subclasses of the features are generated in the output.

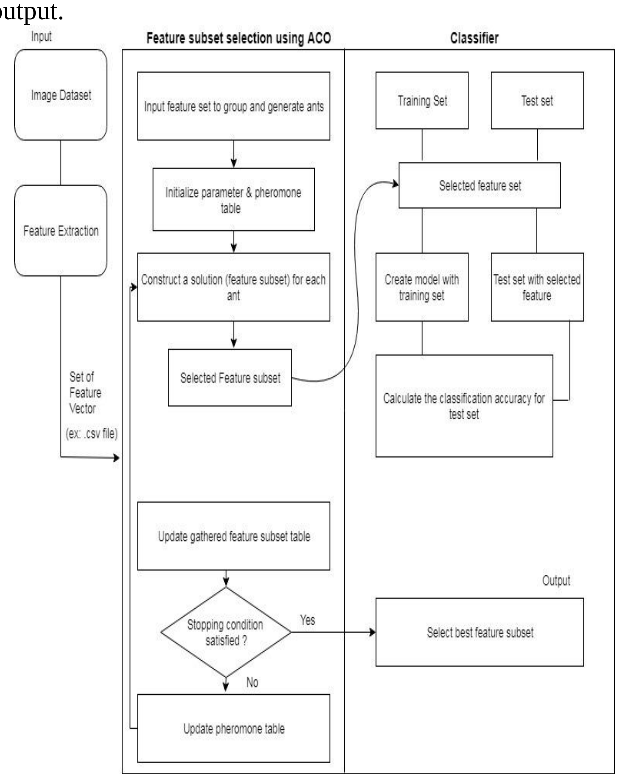

Figure 4 : Flow Diagram of proposed solution

V. EXPERIMENTAL ANALYSIS AND RESULT

For experimentation of proposed approach dataset of images were tested to demonstrate the classification accuracy and determine whether the proposed algorithm can correctly select the relevant features. Following datasets are used for experimentation:

1. Corel image dataset containing 100 categories, and there are 10,000 images of size 192×128 or 128×192 from diverse contents in the JPEG format. The dataset tested contains 500 images in 20 classes of dogs, girl, elephant, building etc.
2. Caltech dataset containing images from 101 categories where each category contains about 40 to 800 images per category. The size of each image is roughly 300 x 200 pixels. The dataset tested contains 300 images in 15 classes of chair, laptop, cup, butterfly etc.

From these image datasets 30 features were extracted including correlation, entropy, contrast, sum average etc.

Proposed approach is implemented using Matlab version 2015a. The performance of the proposed approach is tested by applying kNN classification without feature selection and with features selected from ACO algorithm using K-fold Cross Validation. Cross-validation is a re-sampling technique used for the performance evaluation of machine learning models. In this approach training set is divided into K-partition, K-1 partitions will be used for model building and on K^{th} partition test error will be computed. 10 fold cross validation is used for different values of K nearest neighbours to measure the average classification accuracy of the selected feature subsets.

The parameters of feature selection based on ACO are assigned with following values: Parameters α(pheromone's relative importance), β(importance of heuristic information), ρ(rate of evaporation) have important influence on result. The best values obtained by experimentations are:

- $\rho = 0.75$, $\alpha = 1$, $\beta = 0.5$.
- Initial pheromone trail intensity, cc = 1.
- The number of ants, na = 200.
- Max number of iteration = 20.

For evaluation of results, three parameters are used namely accuracy, recall and precision.

1. Accuracy is fraction of correct predicted samples to total number of samples.
2. Precision is a measure of exactness. It can be defined as, from the samples predicted as positive, how many actually belongs to positive class.
3. Recall is measure of completeness. It tells about how many positive class samples are correctly classified.

Analysis of performance parameters is performed on datasets and comparison is made by applying KNN classifier without using feature selection with proposed model (feature subsets generated using ACO) to validate the efficiency of the proposed model.

The values of performance parameters for different values of K is displayed in Table III and Table IV for proposed approach (kNN+ACO) and kNN without using ACO on Corel Dataset and Caltech Dataset respectively. The proposed algorithm has selected 14 subclasses of features out of 30 features for Caltech Dataset and 16 features for Corel Dataset.

TABLE III. Classification Performance on Corel Dataset

Parameters	K=3		K=5		K=7	
	KNN	KNN+ ACO	KNN	KNN+ ACO	KNN	KNN+ ACO
Accuracy	0.73	0.81	0.68	0.75	0.63	0.68
Precision	0.75	0.79	0.70	0.65	0.65	0.70
Recall	0.78	0.84	0.70	0.79	0.75	0.65

TABLE IV. Classification Performance on Caltech Dataset

Param eters	K=3		K=5		K=7	
	Knn	Knn+A co	Kn n	Knn+A co	Kn n	Knn+A co
Accur acy	0.7	0.741	0. 68	0.725	0. 64	0.69
Precisi on	0.71	0.8	0. 67	0.72	0. 6	0.65
Recall	0.82	0.765	0. 56	0.7	0. 58	0.62

From results on Image Datasets, it is observed that for various values of k in KNN classifier accuracy is more for k as 3 as shown in Table III and Table IV. Thus, from the results, it can be concluded that for lesser of k better values are observed whereas as value of k increases, the accuracy gets affected.

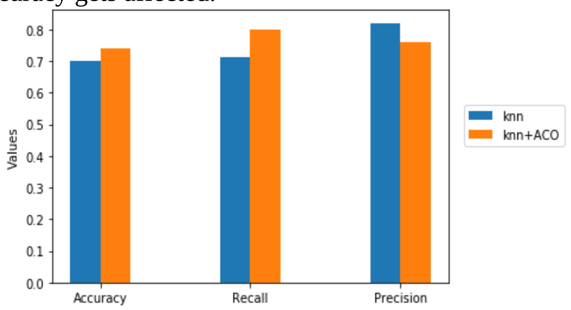

Figure 5: Comparision of performance parameters for Proposed (kNN+ACO)Model and Base (kNN) Model for Corel Dataset .

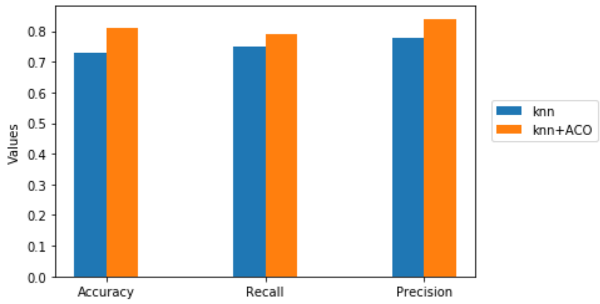

Figure 6: Comparision of performance parameters for Proposed (kNN+ACO)Model and Base (kNN) Model for Caltech Dataset.

In Figure 5 and Figure 6 results of evaluation of the values of accuracy, recall and precision has been plotted to graphically illustrate the performance analysis of techniques for Corel Dataset and Caltech dataset respectively. This shows that the overall performance of the proposed method (kNN+ACO) for feature selection is better than that of the base method (kNN) without using ACO with accuracy of 81% and 73% respectively for Caltech Dataset. Also, for Corel Dataset accuracy is 74% for proposed method and 70% for kNN without using ACO. From the tables, it is observed that overall performance is better for the proposed model (Feature Selection Using Ant Colony Optimization) compared to base model (Without Feature Selection).

VI. CONCLUSION AND FUTURE SCOPE

In this paper, an approach for Feature Subset Selection using ACO has been proposed. The quality of the feature selection is governed by the efficient GLCM based feature extraction technique. The complete process is carried out in three steps namely feature extraction, feature selection using ACO and performance evaluation using kNN Classifier. Simulation analysis has been carried out to evaluate the performance of the approach in terms of the classification accuracy for the datasets taken. Future work involves implementing augmented technique on various image datasets. Further research can be carried out by considering ensemble of classifiers.

REFERENCES

1. Isabelle Guyon, Andre Elisseeff, "An Introduction to Variable and Feature Selection", Journal of Machine Learning Research 3, 2003, 1157-1182.
2. Jović, K. Brkić and N. Bogunovic, "A review of feature selection methods with applications" 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO, 2015).
3. Marco Dorigo1 and Krzysztof Socha2," An Introduction to Ant Colony Optimization",:IRIDIA April 2006 ,Published as a chapter in Approximation Algorithms and Metaheuristics, a book edited by T. F. Gonzalez.
4. Ahmed Al-Ani, "Feature Subset Selection Using Ant Colony Optimization", World Academy of Science, Engineering and Technology 17, 2008.
5. Anam Naseer, Waseem Shahzad, Arslan Ellahi, "A Hybrid Approach for Feature Subset Selection using Ant Colony Optimization and Multi-Classifier Ensemble". (IJACSA) International Journal of Advanced Computer Science and Applications, 2018,Vol. 9, No. 1.
6. Avrim L. Bluma "Artificial Intelligence Selection of relevant features and examples in machine learning". Amficial Intelligence 97 School of Computer Science, Carnegie Mellon University Pittsburgh, 1997, pp. 15213-3891.
7. Bolun Chen,Ling Chen and Yixin Chen, "Efficient ant colony optimization for image feature selection" Signal Processing 93, 2013, Vol. 93, Issue 6, pp. 1566-1576.
8. Farhad Samadzadegan and Tahmineh Partovi, "Feature selection based on ant colony algorithm for hyperspectral remote sensing images" workshop on Hyperspectral Image and Signal processing: Evolution in Remote Sensing, IEEE, 2010.
9. David Martens, Manu De Backer, Raf Haesen," Classification With Ant Colony Optimization",IEEE Transactions on evolutionary computation, 2007, Vol. 11, No. 5, 651.
10. Rafael S. Parpinelli, Heitor S. Lopes, "Data Mining With an Ant Colony Optimization Algorithm" IEEE Transactions on Evolutionary Computing, Vol. 6, No. 4, August 2002 321, Member, IEEE.
11. Xu Sun, Lina Yang, Qian Shen and Bing Zhang, "A new ant colony optimization algorithm based band selection method", Hyperspectral Image and Signal Processing: (WHISPERS) Lausanne, Switzerland, 2014.
12. Andrea Paoli, Farid Melgani, Edoardo Pasolli, "Swarm Intelligence for Unsupervised Classification of Hyperspectral Images", IEEE International Geoscience and Remote Sensing Symposium 12-17 July 2009.

13. Mohammad-Reza Feizi-Derakhshi, and Manizheh Ghaemi, "Classifying Different Feature Selection Algorithms Based on the Search Strategies", in International Conference on Machine Learning, Electrical and Mechanical Engineering (ICMLEME'2014) Jan. 8-9, 2014 Dubai (UAE).
14. D. Asir, S. Appavu alias balamurugan, E. Jebamalar leavline, "A novel feature selection method for image classification", Optoelectronics And Advanced Materials – Rapid Communications Vol. 9, No. 11-12, Nov. – Dec. 2015,pp. 1362 – 1368.
15. Kamilia Menghour and Labiba Souici-Meslati, "Hybrid ACO-PSO Based Approaches for Feature Selection", International Journal of Intelligent Engineering and Systems, Vol.9, No.3, 2016.
16. Adegoke and Omotayo, "Review of Feature Selection Methods in Medical Image Processing", IOSR Journal of Engineering (IOSRJEN), 2014, Vol. 04, Issue 01, pp. 2278-8719.
17. Anuradha Purohit, NS Chaudhari, A Tiwari, "Construction of classifier with feature selection based on genetic programming", IEEE Congress on Evolutionary Computation, 2010, pp.1-5.
18. Khushboo Jain, Anuradha Purohit ," Feature selection using modified particle swarm optimization", International Journal of Computer Applications, 2017, Vol. 161, Issue 7, pp.8-12.
19. Anuradha Purohit, Priyansh Jain," Feature Subset Selection Using Agglomerative Clustering", International Journal of Research in Electronics and Computer Engineering, 2018, Vol. 6, Issue 2, pp. 726-729.
20. Aruna Tiwari Swati Jain, Anuradha Purohit," Multiclass Classifier with Feature Selection using Genetic Programming", International Journal on Computer Engineering and Information Technology (IJCEIT), 2010, Vol. 8, Issue 13, pp.26-30.

AUTHORS PROFILE

Ms. Richa Sharma has completed her Bachelor of Engineering in Computer Science and Engineering from Rajiv Gandhi Prodyogiki Vishvavidyalaya, Bhopal in 2017. She is currently pursuing Master of Engineering in Department of Computer Engineering, in Shri Govindram Seksaria Institute of Technology and Science, Indore.

Dr. Anuradha Purohit pursued her Bachelor of Engineering in Computer Enginering from Vikram University, Ujjain. She has completed her Master of Engineering and Ph.D. in Computer Science and Engineering from Rajiv Gandhi Prodyogiki Vishvavidyalaya, Bhopal. She is currently working as Associate Professor in Department of Computer Engineering in Shri Govindram Seksaria Institute of Technology and Science, Indore. She has 19 years of teaching experience.