Investigation on Feature Extraction and Classification of Medical Images

P. Gnanasekar, A. Nagappan, S. Sharavanan, O. Saravanan, D. Vinodkumar, T. Elayabharathi and G. Karthik

Abstract—In this paper we present the deep study about the Bio-Medical Images and tag it with some basic extracting features (e.g. color, pixel value etc). The classification is done by using a nearest neighbor classifier with various distance measures as well as the automatic combination of classifier results. This process selects a subset of relevant features from a group of features of the image. It also helps to acquire better understanding about the image by describing which the important features are. The accuracy can be improved by increasing the number of features selected. Various types of classifications were evolved for the medical images like Support Vector Machine (SVM) which is used for classifying the Bacterial types. Ant Colony Optimization method is used for optimal results. It has high approximation capability and much faster convergence, Texture feature extraction method based on Gabor wavelets etc..

Keywords—ACO Ant Colony Optimization, Correlogram, CCM Co-Occurrence Matrix, RTS Rough-Set theory

I. INTRODUCTION

A key function in different image applications is feature extraction. Feature extraction is part of the dimension reduction, in a typical classification task, if the number of relevant features (voxels) is N, the feature extraction problem is defined as obtaining the $n < N$ features that enable the construction of the best classifier. For the Brain images the features are extracted by masking the pre-processed PET images with the brain mask. This leads to the extraction of the anatomical volumes of interest (AVOI). Then, each AVOI is represented by the mean value of the intensities inside this AVOI. At the end, each image will be represented by a feature vector $F = (f1, f2, …, fn)$ where n is the number of AVOIs. The features used for brain classification are extracted from

P. Gnanasekar is a research Scholar in the Department of Computer Science & Engineering at the Vinayaka Missions Research Foundation Deemed University, Salem, Tamilnadu, India. e-mail: peegeeyes@gmail.com.

A. Nagappan is the Principal of VMKV Engineering College, Salem, Tamilnadu, India. e-mail: nags_slm@yahoo.com.

S. Sharavanan is a research Scholar in the Department of Computer Science & Engineering at the Vinayaka Missions Research Foundation Deemed University, Salem, Tamilnadu, India. e-mail: sharavanan33@gmail.com.

O. Saravanan is a research Scholar in the Department of Computer Science & Engineering at the Vinayaka Missions Research Foundation Deemed University, Salem, Tamilnadu, India. e-mail: osaravanan.it@gmail.com

D. Vinodkumar is a research Scholar in the Department of Computer Science & Engineering at the Vinayaka Missions Research Foundation Deemed University, Salem, Tamilnadu, India. e-mail: vino.kd@gmail.com

T. Elayabarathi is a research Scholar in the Department of Computer Science & Engineering at the Vinayaka Missions Research Foundation Deemed University, Salem, Tamilnadu, India. e-mail: bharathieee2k6@gmail.com

G. Karthik is a research Scholar in the Department of Computer Science & Engineering at the Vinayaka Missions Research Foundation Deemed University, Salem, Tamilnadu, India. e-mail: gkarthikme@gmail.com

automatically generated regions, which are determined from the training data. Several issues are taken into consideration here. First, morphological changes of brain structures resulting from pathological processes usually do not occur in isolated regions or in regions necessarily having regular shapes. The implicit knowledge extraction, image data relationship and other patterns are not explicitly stored in the images. This technique is an extension of data mining to image domain.

It is an inter disciplinary field that combines techniques like computer vision, image processing, data mining, machine learning, data base and artificial intelligence [1]. Features are used in different applications such as image processing, remote sensing and content-based image retrieval. These features can be extracted in several ways. The most common way is using a Gray Level Co-occurrence Matrix (GLCM). GLCM contains the second - order statistical information of neighboring pixels of an image. Textural properties can be calculated from GLCM to understand the details about the image content. A feature is a characteristic that can capture a certain visual property of an image either globally for the whole image, or locally for objects or regions. Different features such as color, shape, and texture can be extricated from an image. Texture is the variation of data at different scales. A number of methods have been developed for texture feature extraction.

It is the process of generating features to be used in the selection and classification tasks. Feature selection reduces the number of features provided to the classification task. Those features which are likely to assist in discrimination are selected and used in the classification task. Features which are not selected are discarded [5]. In these three activities, feature extraction is most critical because the particular features made available for discrimination directly influence the efficacy of the classification task [13]. The end result of the extraction task is a set of features, commonly called a feature vector, which constitutes a representation of the image.

II. RELATED WORK

In the last few years, a number of systems using image content feature extraction technologies proved reliable enough for professional applications in industrial automation, biomedicine, social security, biometric authentication and crime prevention [7]. The features are autocorrelation, contrast, cross correlation, cluster prominence, cluster shade, energy, entropy and homogeneity. From these features we select only a subset of the features that are useful for the image recognition. This is done using Ant Colony Optimization. Then using the Euclidean Distance measure images are retrieved from the database.ACO is the technique for feature selection. ACO is a population based meta heuristic

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:12, 2011

which is inspired by the behavior of ants. It is used to find solutions for various optimization problems. In ACO, a colony of ants cooperates to look for solutions for the problem. Artificial ants incrementally build a solution by adding components to a partial solution under construction. Feature extraction for ACO applied image data is classified as following methods.

A. Shape Features

The following shape features were calculated from the plaque images [9]: 1) X - coordinate maximum length of the plaque image frame, 2) Y - coordinate maximum length of the plaque image frame, 3) Area of ROI, 4) Perimeter of ROI, and 5) Perimeter/Area. The idea was to investigate whether the size and complexity of the shape of the segmented plaque had any diagnostic value.

B. Morphology Features

Morphological image processing allows the detection of the presence of specific patterns, called structural elements, at different scales. The simplest structural element for near isotropic detection is the cross '+' consisting of five image pixels. Using the cross '+' as a structural element, pattern spectra were computed for each plaque image. The mean cumulative distribution functions (CDF) and the mean probability density functions (PDF) were computed as two different morphological feature sets.

C. Histogram Features

(i) Histogram: The grey level histogram of the ROI of the plaque image was computed for 32 equal width bins and was used as an additional feature set [14]. Histogram despite its simplicity provides a good description of the plaque structure.

(ii) Multi-region Histogram: Three equidistant ROIs were identified by eroding the plaque image outline by a factor based on the plaque size. The histogram was computed for each one of the three regions as described above and the 96 values comprised the new feature vector. This feature was computed in order to investigate whether the distribution of the plaque structure in equidistant ROIs has a diagnostic value and more specifically if the structure of the outer region of the plaque is critical whether the plaque will rupture or not.

D. Correlogram Features

Correlograms are histograms, which measure not only statistics about the features of the image, but also take into account the spatial distribution of these features [7]. In this work two correlograms were implemented for the ROI of the plaque image:

(i) based on the distance of the distribution of the pixels' gray level values from the center of the image, and (ii) based on their angle of distribution.

For each pixel the distance and the angle from the image center was calculated and for all pixels with the same distance or angle their histograms were computed. In order to make the comparison between images of different sizes feasible, the distance correlograms were normalized into 32 possible distances from the center by dividing the calculated distances with *maximum_ distance/32*. The angle of the correlogram was allowed to vary among 32 possible values starting from the left middle of the image and moving clockwise. The resulting correlograms were matrices 32x32 (gray level values over 32 were set to be the white area surrounding the region of interest and were not considered for the calculation of the features).
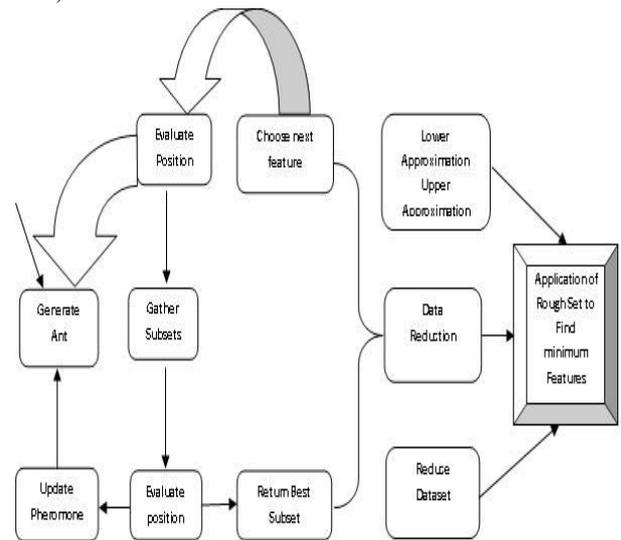

Fig. 1 Proposed Hybridized Model using ACO and RST (Rough ACO)

The overall process of ACO feature selection can be seen in Fig. 1. It begins by generating a number of ants, *k*, which are then placed randomly on the graph (i.e. each ant starts with one random feature). Alternatively, the number of ants to place on the graph may be set equal to the number of features within the data; each ant starts path construction at a different feature. From these initial positions, they traverse edges probabilistically until a traversal-stopping criterion is satisfied. The resulting subsets are gathered and then evaluated. If an optimal subset has been found or the algorithm has executed a certain number of times, then the process halts and outputs the best feature subset encountered. If neither condition holds, then the pheromone is updated, a new set of ants are created and the process iterates once more.

III. GENETIC PROGRAMMING AND COMPUTER VISION

Genetic programming has been applied to a variety of image classification problems. The work so far can be grouped into three approaches. The first approach involves pre-processing images with low-level feature extraction algorithms followed by genetic programming designed to discover the classification rule. Tackett [4] used statistical features such as mean and standard deviation of grey levels

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:12, 2011

within a window to classify military targets in infrared images (Fig. 2).

Existing work on pathological medical image database retrieval deals mostly with 2D images and has not paid much attention to a systematic approach for image feature weighting. In the work of dental radiography image database retrieval [8], the authors use a deformable shape contour selected by the system designer as the primary feature for image indexing. By selecting different modes in the Finite element representation and eigen-decomposition of the contours (hand drawn by an expert dentist), the authors achieve classification rates between 87% (for normals) and 62% (for pathologies). What is missing in the image retrieval practice is an objective, quantitative evaluation of the extracted image features before they are used for image retrieval.
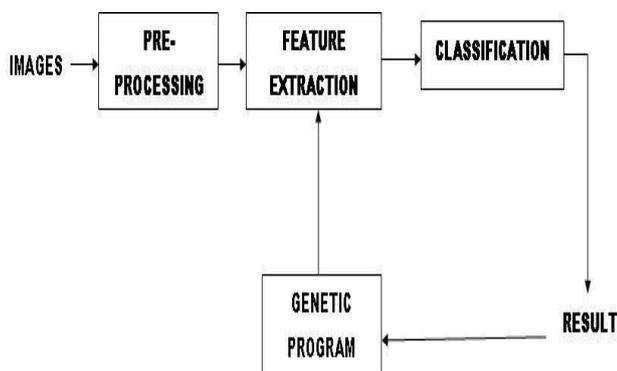

Fig. 2 Feature Extraction Discovery using Genetic Programming

Before applying the feature-extraction and classification for the images we have to perform the following methodology for image data.

(i) An ant colony optimization (ACO) as the technique for feature selection which is inspired by the behavior of ants and has numerous applications in various fields such as image processing, networking and neural networks. ACO was used for image edge detection [2],[12]. The proposed edge detection method takes advantage of the improvements introduced in ant colony system. With suitable parameter values, the algorithm was able to successfully identify edges in the canonical test images. Parallel ACO [3],[6] is used for the segmentation of MR brain tumor. The proposed method has advantage that it can effectively segment the fine details of the image. It has higher accuracy as compared to the existing algorithms. ACO is used for remote sensing image classification as in [10].

 (ii) Rough set theory [6],[9],[16] is a new mathematical approach to imprecision, vagueness and uncertainty. In an information system, every object of the universe is associated with some information. Objects characterized by the same information are indiscernible with respect to the available information about them. Any set of indiscernible objects is called an elementary set. Any union of elementary sets is referred to as a crisp set- otherwise a set is rough (imprecise, vague). Vague concepts cannot be characterized in terms of

information about their elements. A rough set is the approximation of a vague concept by a pair of precise concepts, called lower and upper approximations. The lower approximation is a description of the domain objects, which are known with certainty to belong to the subset of interest, whereas the upper approximation is a description of the objects that possibly belong to the subset. Relative to a given set of attributes, a set is rough if its lower and upper approximations are not equal.

IV. PROPOSED METHOD

The proposed section covers the algorithmic implementation for the image data which could undergo the process of feature extraction and image classification in optimal way.

ACO Algorithm

The steps of the ACO algorithm are as follows:

1. Initialization: Set $Ti = cc$ and $\Delta Ti = 0$, $(i = 1, \ldots, n)$, where cc is a constant and ΔTi is the amount of change of pheromone trial quantity for feature fi.
- Define the maximum number of iterations.
- Define k, where the k-best subsets will influence the subsets of the next iteration.
- Define p, where $m - p$ is the number of features each ant will start with in the second and following iterations.

2. If in the first iteration,
- For $j = 1$ to na,
- Randomly assign a subset of m features to Sj. Go to step 4.

3. Select the remaining p features for each ant:
- For $mm = m - p + 1$ to m,
- For $j = 1$ to na,
  ➤ Given subset Sj, Choose feature fi that maximizes $USMi$ Sj.
  ➤ . S$j = $ Sj U {fi}.

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

Main objective of this paper is to focus on how to improve the time efficiency of a heuristic feature subset selection algorithm. We employ a new rough set framework hybridized with ACO, which is called positive approximation. The main advantage of this approach stems from the fact that this framework is able to characterize the granulation structure of a rough set using a granulation order. Based on the positive approximation, we develop a common strategy for improving the time efficiency of a heuristic feature selection, which provides a vehicle of making algorithms of rough set based feature selection techniques faster.ACO method can be combined with rough set (RST) theory to get a new algorithm.

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:12, 2011

ACO is particularly attractive for feature selection as there seems to be no heuristic that can guide search to the optimal minimal subset every time. Additionally, it can be the case that ants discover the best feature combinations as they proceed throughout the search space. However Rough Set Theory is one of effective methods for dealing with incomplete information, which can reduce decision-making and classification rules so as to establish knowledge model through data analysis and knowledge reduction under the condition of maintaining the ability of classification unchangeable.

V. ACO FRAMEWORK

An ACO algorithm can be applied to any combinatorial problem as far as it is possible to define:

1) Appropriate problem representation: The problem can be described as a graph with a set of nodes and edges between nodes.

2) Heuristic desirability *(h)* of edges: A suitable heuristic measure of the "goodness" of paths from one node to every other connected node in the graph.

3) Construction of feasible solutions: A mechanism must be in place whereby possible solutions are efficiently created. This requires the definition of a suitable traversal stopping criterion to stop path construction when a solution has been reached.

4) Pheromone updating rule: A suitable method of updating the pheromone levels on edges is required with a corresponding evaporation rule, typically involving the selection of the n best ants and updating the paths they chose.

5) Probabilistic transition rule: The rule that determines the probability of an ant traversing from one node in the graph to the next.

Rough ACO: Algorithmic Framework
Step1: Take the input as a decision table S= (U, C, D)
Step2: Let Core=φ and Calculate the POSc (D)
Step3: For ∀a C, calculate POS(c-{a}) (D).If POS(c-{a}) (D) ≠ POSc(D), then ORE=CORE ∪{a}; Else C=C-{a}
Step4: Execute iteratively step 2 until all attributes among C are calculated.
Step5: If POScore (D) = POSc (D), algorithm stops and return CORE as the result of feature selection; otherwise go to step 6
Step6: The pheromone of each arc (i,j) is assigned to an constant, i.e. $\tau_{ij}(0) = c$
Step7: Some ants (assume the number of ants is m) are distributed to each core attribute node to conduct feature selection
Step8: Each ant selects next feature node
Step9: Calculate POScore (D), a Є C- CORE, if POScore (D) = POSc (D) algorithm stops and return FS= CORE ∪ai as the result of feature selection; else go to step 10
Step10: Update value of pheromone τ_{ij} for each path link and go to step 8
Ant Colony Optimization

Ant colony optimization (ACO) Meta heuristic, a novel population-based approach was recently proposed to solve several discrete optimization problems [2]. The ACO mimics the way real ants find the shortest route between a food source and their nest. The ants communicate with one another by means of pheromone trails and exchange information about which path should be followed. The more the number of ants traces a given path, the more attractive this path (trail) becomes and is followed by other ants by depositing their own pheromone. This auto catalytic and collective behavior results in the establishment of the shortest route. Ants find the shortest path from their nest to the food source with the help of pheromone trail. This characteristic of ants is adapted on ant colony optimization algorithms [15] to solve real problems with using exactly some characteristics of ants and some new addition.

The method improved by modeling real ants use exactly the same specifications taken from real ants are below:

• The communication established with ants through pheromone trail

• Paths deposited more pheromone preferred previously

• Pheromone trail on short paths increase more rapidly.

Addition of new specifications to this new technique as is below:

• They live in an environment where time is discrete

• They will not be completely blind, they will reach the details about the problem

• They will keep information formed for the solution of the problem with has some memory.

In ant colony optimization algorithms, operations described above are iterated in main loop until a certain number of iterations are completed or all ants begin to generate the same result. This situation is named as *stagnation behavior*, because after a point, algorithm finishes to generate alternative solutions. The reason of this situation is, after a certain number of iterations, ants generate continuously the same solutions because pheromone amount intensifies in some points and the difference between pheromone concentrations on paths become very huge.

Most ant colony optimization algorithms use this algorithmic diagram demonstrated below:

Initiation of the parameters which determines the pheromone trail

While *(until result conditions supplied)* ***do***
Generate Solutions
Apply Local Search
Update Pheromone Trail
End

VI. COMPARISONS

Ant colony optimization (ACO) adopts the texture features of the image for the image retrieval. To analyze the effect of the proposed method, experiments were conducted on image databases having more than 100 images. The performance of the proposed method was found to be superior when compared to other methods (see Table I).

TABLE I
FEATURES SELECTED USING ACO

Feature Extraction	Feature Selection
Autocorrelation, Contrast, Cross correlation, Cluster prominence, Cluster shade, Energy, Entropy and Homogeneity	Autocorrelation, Contrast, Cluster prominence Energy, Homogeneity

A large number of features are extracted from the image. This increases the complexity of the system. So using ACO we are extracting only the features that are useful in retrieving images from the database. This simplifies the system and increases its accuracy.

TABLE II
PERFORMANCE OF THE SYSTEM

No. of images in the database	No. of images retrieved	Accuracy
120	110	91.6%

The experiments were conducted on a database of 120 images. An accuracy of 91.6% is obtained (see Table II) which is sufficiently higher for an image retrieval system.

For rough-set ACO the experimental result is the general information of selected data set as shown in Table III. The performance of feature selection of Rough Set based algorithms is showed in Table IV. The experimental result shows that Feature selection based on Rough ACO algorithm achieves better results than the traditional algorithms. The Rough ACO algorithm has higher speed convergence and has better capacity of optimization.

TABLE III
GENERAL INFORMATION OF SELECTED DATASET

Dataset	No. of features	Condition features	Decision features	No. of records
SPECTF Heart	23	22	1	188
Lung cancer	57	56	1	33

It describes how the image data are reduced according to the features of the respective image with the model of classification technology. Table V presents the comparison of ACO with Rough set methodology. We collaborate the ACO with Rough set methodology to develop a new improved Rough-set ACO for the feature extraction of images and its classification. After the process is completed the images are reset with corresponding data for further tagging techniques.

TABLE IV
FEATURE SELECTION BASED ON ROUGH ACO

Data set	Average iteration	Number of reduction	Reduction features
SPECTF	>=30	8	0,1,2,3,4,5,6,7
Lung cancer	>=80	4	0,3,6,11

TABLE V
COMPARISON OF ACO AND RTS ACO

Ant colony Optimization (ACO)	Rough-Set ACO
Ant colony optimization successfully identifies edges in canonical test images.	Effective method for dealing with in complete information, which can reduce decision-making and classification rules.
It has higher accuracy as compared to the existing algorithms.	So as to establish knowledge model through data analysis and knowledge reduction, under the condition of maintaining the ability of classification unchangeable.
ACO mainly used for remote sensing image classifications.	It implements new mathematical approach to imprecision, vagueness and uncertainty.
It is a population based met-heuristic which is inspired by the behavior of ants. Used to find solutions for various optimization problems.	Information of universe objects characterized by the same information are indiscernible with respect to available information about them.
ACO mainly uses Feature Extraction only through **Co-Occurrence Matrix (CCM)**.	Any set of indiscernible object is called an elementary set. Any union of elementary set is referred to as a **crisp-set** otherwise a set is rough.
CCM is a tabulation of how a different combination of pixel occurs in an image.	Rough set requires no additional knowledge except for the supplied data.
CCM is used to get more useful features of image. These features include contrast, correlation, energy and homogeneity.	It performs feature select using only the granularity structure of the data [11].

VII. CONCLUSION

The combined method will deliver the compatible image classification and feature extraction with high efficiency. The categorization rate can be further improved to 85.5% when a parallel combination of single classifier based on scaled representations and global texture features is used.

Considering image categorization as initial step for image retrieval based on local features, the correct image category should be within the five or ten nearest neighbors. These methods can be integrated in one methodology that is capable of significantly improve the classification problem at hand. The proposed approach is computationally less expensive and yields good result. The classification accuracy can be improved by extracting more features and increasing the training data set.

REFERENCES

[1] C. Ordonez, E. Omiecinski ,"Image mining: A new approach for data mining" *Technical Report GITCC*-98-12, Georgia Institute of Technology, College of Computing, 1998, pp. 1-21.

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:12, 2011

[2]   Anna Veronica Baterina, Carlos Oppus, 2010. "Image Edge Detection Using Ant Colony Optimization", *International Journal of Circuits, Systems and Signal Processing*, Issue 2, Vol.4.

[3]   J.Jaya, K. Thanushkodi,2010. "Segmentation of MR Brain Tumor Using Parallel ACO", *International Journal of Computer and Network Security*, Vol.2, No.6, pp. 150-153.

[4]   Tackkett WA , "Genetic Programming for Feature Discovery and Image Discrimination", in *Proceedings of the 5th International Conference on genetic Algorithms, ICGA-93*, University of Illinois at Urbana-Champaign, 17-21 July 1993, pp. 303-309.

[5]   *IEEE Computer, Special issue on Content Based Image Retrieval*, 28, 9, 1995.

[6]   Pawlak, Z. (1991) "Rough Sets: Theoretical Aspects of Reasoning about Data". *Kluwer Academic Publishing*, Dordrecht.

[7]   A.K. Jain, R.M. Bolle and S. Pankanti (eds.), (1999) *Biometrics: Personal Identification in Networked Society*, Norwell, MA: Kluwer, 1999.

[8]   W. Zhang, S. Sclaro_, S. Dickinson, J. Feldman, and S. Dunn. Shape-based indexing in a medical image database. *Workshop on Biomedical Image Analysis*, pp. 221-230, June 1998.

[9]   Polkowski, L., Lin, T. Y., & Tsumoto, S. (Eds). (2000),"Rough Set Methods and Applications: New Developments in Knowledge Discovery in Information Systems", Vol.56. *Studies in Fuzziness and Soft Computing*, Physica-Verlag, Heidelberg.

[10] Fizazi Hadria, Hannane Amir Mokhtar, "Remote Sensing Image Classification using Ant Colony Optimization"

[11] Hossein Nezamabadi-pour , Saeid Saryazdi , Esmat Rashedi, Edge detection using ant algorithms, Soft Computing - *A Fusion of Foundations, Methodologies and Applications*, v.10 n.7, pp. 623-628, May 2006

[12] X. Zhuang , N. E. Mastorakis, "Edge detection based on the collective intelligence of artificial swarms", *Proceedings of the 4th WSEAS International Conference on Electronic, Signal Processing and Control*, pp. 1-7, April 25-27, 2005, Rio de Janeiro, Brazil

[13] X. Zhuang, "Edge Feature Extraction in Digital Images with the Ant Colony System", *IEEE International Conference in Computational Intelligence for Measurement Systems and Applications*, 2004.

[14] N. Otsu, "A Threshold Selection Method from Gray-level Histograms", *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.

[15] M. Dorigo and T. Stützle, "An Experimental Study of the Simple Ant Colony Optimization Algorithm", *Proceedings of the WSES International Conference on Evolutionary Computation*, 2001.

[16] Jensen, R., Shen, Q., 2003 "Finding Rough Set Reduces with Ant Colony Optimization", *Proceedings of the 2003 UK Workshop on Computational Intelligence*, pp. 15-22.