IRACST – Engineering Science and Technology: An International Journal (ESTIJ), ISSN: 2250-3498, Vol.3, No.1, February 2013

A Comparative study of Medical Images by Feature Extraction and Classification

P. Gnanasekar
Department of Computer Science & Engg.
Vinayaka Mission's Research Foundation University
Salem, India

A. Nagappan
Principal
V.M.K.V. Engineering College
Salem, India

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 $F = (f1, f2, \dots, fn)$ where n is the number of AVOIs. The features used for brain classification are extracted from 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

For professional applications in biometric authentication, industrial automation, biomedicine, social security, and crime detection and prevention a number of systems using image content feature extraction technologies have been proved reliable [7]. The features are cluster shade, entropy, homogeneity, are autocorrelation, contrast, cross correlation, cluster prominence, energy, and homogeneity. We select only a subset of the features that are useful for the image recognition. Here we use Ant Colony Optimization. Then we use the Euclidean Distance to measure images which are retrieved from the database.

Ant Colony Optimization is the technique we use for feature selection. ACO, inspired by the behavior of ants, is a population based meta heuristic. Here a colony of ants work in tandem to look for solutions for a specific problem. A solution is build incrementally by adding components to a partial solution under construction by artificial ants. Thus the Feature extraction for ACO applied image data can be classified as follows.

Histogram :

Histogram provides a good description of the plaque structure despite its simplicity.

Shape :

To investigate the size and complexity of the shape of the segmented plaque whether it has any diagnostic value.

Correlogram:

Correlograms measure not only statistics about the features of the image, but also take into account the spatial distribution of these features [7].

Correlograms can be differentiated,

(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.

Morphology:

Using Morphological image processing the presence of specific patterns viz., structural elements are detected. Two different morphological feature sets can be computed.

1. The mean cumulative distribution functions (CDF) and
2. The mean probability density functions (PDF)

Fig. 1 shows the overall process of ACO Feature selection. First a number of ants, say k are generated. Place these ants randomly on the graph. We can even equate the number of ants on the graph to the number of features in the data. The ants traverse along edges probabilistically till they satisfy a traversal stopping criterion.

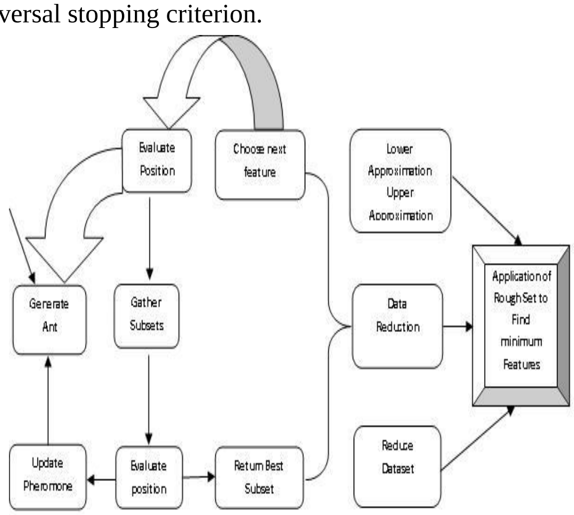

Fig. 1 Hybridized Model using ACO and RST

The subsets thus resulting are evaluated before gathering. The best feature subset is taken as output if an optional subset is found or the algorithm executes a number of times. The process iterates once more if neither condition holds.

III. Genetic Programming and Computer Vision

We can apply genetic programming to a variety of image classification problems. To classify military targets in infrared images (Fig. 2), Tackett [4] used statistical features such as mean and standard deviation of grey levels within a window.

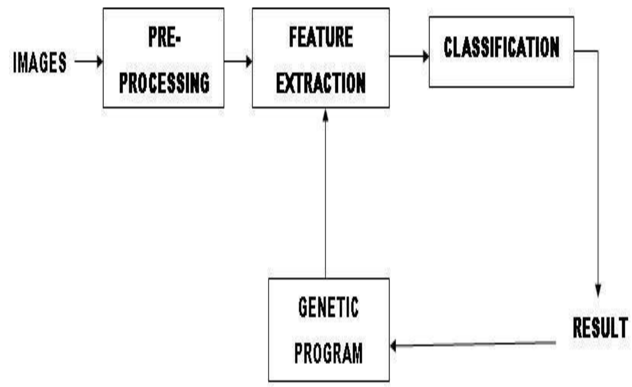

Fig. 2 Feature Extraction Discovery using Genetic Programming

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

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

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

(ii) A new mathematical approach to imprecision, vagueness and uncertainty is Rough set theory (RST) [6],[9],[16]. Every object of the universe is associated with some information in an information system. Objects characterized by the same information are indiscernible with respect to the available information about them. An elementary set is any set of indiscernible objects. Any union of elementary sets is referred to as a crisp set- otherwise a set is rough (imprecise, vague). In terms of information about the elements vague concepts cannot be characterized. A rough set is the approximation of a vague concept by a pair of precise concepts, called lower and upper approximations. The upper approximation is a description of the objects that possibly belong to the subset whereas the lower approximation is a description of the domain objects, which are known with certainty to belong to the subset of interest. A set is rough if its lower and upper approximations are not equal, relative to a given set of attributes.

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 ΔTi is the amount of change of pheromone trial quantity for feature *fi* and *cc* is a constant.

- Define *p*, where b – *p* is the number of features each ant will start with in the second and following iterations.
- Define *k*, where the *k*-best subsets will influence the subsets of the next iteration.
- Define the maximum number of iterations.

2. If in the first iteration,

- For *j* = 1 to *na*,
- Randomly assign a subset of *m* features to S*j*.

  Go to step 4.

3. Select the remaining *p* features for each ant:

- For bb = b – *p* + 1 to b,
- For *j* = 1 to *na*,
  - Given subset S*j*, Choose feature *fi* that maximizes *USMi* S*j*.
  - □S*j* = S*j* Ụ {*fi*}.

4. If number of iterations < maximum number of iterations,

   Go to step 3.

The main focus is on how to improve the time efficiency of a heuristic feature subset selection algorithm. Here a positive approximation is employed for a new rough set framework hybridized with ACO. This framework is able to characterize the granulation structure of a rough set using a granulation order. Based on the positive approximation, 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, we develop a common strategy.

We can get a new algorithm combining ACO method with Rough Set theory. As there seems to be no heuristic that can guide search to the optimal minimal subset every time, ACO is particularly attractive for feature selection. Additionally, it can be the case that ants discover the best feature combinations as they proceed throughout the search space. One of the effective methods for dealing with incomplete information, however is Rough Set Theory, 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 as far as it is possible to define any combinatorial problem:

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

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

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

4) 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.

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

Algorithmic Framework for a Rough ACO

The following steps are followed

1: The input is taken as a decision table S= (U, C, D)
2: Calculate the POSc (D), keeping Core=φ
3: For ∀a C, calculate
   POS(c-{a}) (D).
   If POS(c-{a}) (D) ≠ POSc(D),
   then ORE=CORE ∪{a};
   Else C=C-{a}
4: Step 2 is executed iteratively until all attributes among C are calculated.
5: If POScore (D) = POSc (D), Stop algorithm
   Return CORE as the result of feature selection; Otherwise go to next step

6: The pheromone of each arc (i,j) is assigned to an constant, i.e. $\tau ij(0) = c$

7: Distribute the ants say b numbers to each core attribute node to conduct feature selection

8: Next feature node is selected by each ant

9: Calculate POScore (D), a Є C- CORE,
if POScore (D) = POSc (D)
Stop algorithm
Return FS= CORE ∪ai as the result of feature selection; else go to next step

10: For each path link update value of pheromone τij
Go to step 8

Ant Colony Optimization

To solve several discrete optimization problems [2] in Ant colony optimization (ACO) Meta heuristic, a novel population-based approach was recently proposed. The ACO mimics the way real ants find the shortest route between a food source and their nest. The ants communicate with each one by means of pheromone trails. They 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. This is then followed by other ants by depositing their own pheromone. The establishment of the shortest route results in this auto catalytic and collective behavior. With the help of pheromone trail, ants find the shortest path from their nest to the food source This characteristic of ants is adapted on ant colony optimization algorithms [15] to solve real problems by using exactly some characteristics of ants and some new additions. The method improved by modeling real ants use exactly the same specifications taken from real ants and is explained below:

- Through pheromone trail the communication is established with ants.
- Paths deposited with more amount of pheromone is preferred.
- Pheromone trail on short paths increases more rapidly and steadily.

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

- In the environment they live in, the time is discrete.
- They will reach the details about the problem. They will not be completely blind.
- With some memory, they will keep information formed for the solution of the problem.

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 in ant colony optimization algorithms. This situation is named as *stagnation behavior*, since after a point, algorithm finishes to generate alternative solutions. The reason of this situation is, that ants generate continuously the same solutions because pheromone amount intensifies in some points and the difference between pheromone concentrations on paths become very huge after a certain number of iterations.

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

Initiate the parameters which determine the pheromone trail
While *(until result conditions supplied)* ***do***
Generate Solutions
Apply Local Search continually
Update Pheromone Trail
End

VI. Comparisons

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

TABLE I. FEATURES SELECTED

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, which increases the complexity of the system. So using ACO only the features that are useful in retrieving images from the database are extracted. This simplifies the system and increases its accuracy.

TABLE II. PERFORMANCE

No. of images	Images retrieved	Accuracy
140	125	93.4%

A database of 140 images was used to experiment. An accuracy of 93.4% is obtained (Table II) which is sufficiently higher for an image retrieval system.

The experimental result is the general information of selected data set as shown in Table III for a Rough Set ACO.

TABLE III. INFORMATION OF SELECTED DATASET

Dataset	No. of features	Condition features	Decision features	No. of records
SPECTF Heart	27	26	1	135
Brain Tumor	43	42	1	41

The performance of feature selection of Rough Set based algorithms is showed in Table IV. 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.

It describes how the image data are reduced according to the features of the respective image with the model of classification technology.

TABLE IV. PERFORMANCE OF FEATURE SELECTION BASED ON ROUGH ACO

Data set	Average iteration	Number of reduction	Reduction features
SPECTF	>=20	7	0,1,2,3,4,5,6
Brain Tumor	>=75	3	0,4,8

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.

a. Original **b. ACO Method** **c. RST ACO**

Fig. 3. Brain CT Image Sample 1

a. Original **b. ACO Method** **c. RST ACO**

Fig. 4. Brain CT Image Sample 2

TABLE V. COMPARISON OF ACO AND RTS ACO

Ant colony Optimization (ACO)	Rough-Set ACO
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.
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.
ACO mainly used for remote sensing image classifications.	It implements new mathematical approach to imprecision, vagueness and uncertainty.
ACO mainly uses Feature Extraction only through **Co-Occurrence Matrix (CCM)**.	An elementary set is defined as any set of indiscernible objects whereas any union of elementary set is referred to as a **crisp-set** otherwise a set is rough.
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.
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].
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.

VII. CONCLUSION

The combined method will deliver the compatible image classification and feature extraction with high efficiency. The Accuracy can be further improved over 93.4% when we use a parallel combination of single classifier, based on scaled representations and global texture features.

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

REFERENCES

[1] P. Gnanasekar, A. Nagappan, S. Sharavanan, O. Saravanan, D. Vinodkumar, T. Elayabharathi and G. Karthik, "Investigation on Feature Extraction and Classification of Medical Images" World Academy of Science, Engineering and Technology Issue 60, 2011, pp. 327-332

[2] 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.

[3] 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.

[4] 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.

[5] 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.

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

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

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

[9] 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.

[10] 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.

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

[12] 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

[13] 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

[14] 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.

[15] 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.

[16] 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.

[17] 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.