AN IMAGE INFORMATION MINING FRAMEWORK FOR RAPID ASSESSMENT AND RESPONSE IN COASTAL DISASTER EVENTS

Surya S. Durbha, [1, 2], Roger L. King [1, 2], Nicolas H. Younan [1, 2], David R. Shaw [2] Ioana Banicescu [3]

[1] *Department of Electrical and Computer Engineering (USA),*
E-mail: suryad@gri.msstate.edu ,rking@engr.msstate.edu, younan@ece.msstate.edu
[2] *GeoResources Institute (GRI) (USA), E-mail:dshaw@gri.msstate.edu*
[3] *Department of Computer Science and Engineering (USA), E-mail:ioana@cse.msstate.edu*
Mississippi State University, Mississippi State, MS 39762-9571, USA

ABSTRACT

In this paper, we propose a framework that focuses on the need for rapid image information mining in a coastal disaster event where it is necessary to explore vast amounts of data from multiple remote sensing sensors in real or near real time. The proposed system; Rapid Image Information Mining (RIIM) is a region based approach where, in *lieu* of prevalent pixel based methods, it localizes interesting zones and extracts information from them that are stored in a knowledge base. A set of primitive features are extracted from the regions, whose relevance for a particular land cover class or a combination of classes is then assessed based on a wrapper-based genetic algorithm (GA) approach. In this, we use an induction algorithm along with the GA to arrive at an optimal set of features. We investigate feature selection and feature generation using this wrapper approach. A support vector machines-based classification is applied for generating predictive models for those land cover classes that are important in coastal disaster events. In RIIM, searching for a particular land cover type (e.g. flooded agriculture) is based on the actual meaning and content of the image instead of just the metadata.

1. INTRODUCTION

Observations from satellites comprise a significant portion of the data used by coastal zone monitoring systems. The data are particularly valuable because of the synoptic coverage and because they can provide a variety of measurements that are not otherwise available or affordable. However, the use of such valuable information in a rapid assessment scenario is hindered by the fact that it is cumbersome to explore huge image databases through manual or semi automated methods. In a coastal disaster event, it is necessary to obtain information about water level (depth), winds, currents, waves, temperature-salinity stratification in real time and predictions of water level (12-24 hrs), storm surge (48-72 hours) in advance. It has been estimated that better preparation, response, and mitigation will reduce average costs of storm-related disasters by 10%. The dissemination of information that is time critical calls for systems that will facilitate quick assessment of the scenario from multiple perspectives. While some of the information is available only through *in situ* sensors, this paper focuses on the rapid retrieval of the status of different land covers using satellite remote sensing data. Such information is normally made available after a lengthy process of manually identifying the affected areas, classifying the image data and then updating the GIS (maps) for the area. This process provides planners and responders with critical information, but it is time critical.

Image information mining provides an attractive alternative to traditional exploration of imagery through the provision of state of the art methods for content based extraction of knowledge from vast amounts of imagery. The proposed system; Rapid Image Information Mining (RIIM) is a region based approach that localizes interesting zones and extracts characteristic information from them and stores this information in a database for later use during the disaster. This content is then available for a variety of queries based on the image content for searching relevant imagery. This paper deals with the followings:
i) The application of image information mining in coastal disaster events, with particular emphasis on the image mining of post hurricane events and present results using imagery from Landsat ETM+ of post Katrina hurricane.
ii) We propose an image information mining system that is fast and reliable and able to perform the tasks of identifying affected regions with minimal expert supervision.

Previous efforts in image information mining areas have focused mainly on the reduction of features using clustering approaches, but little has been reported on the selection of best feature subsets. In our view, this is of more importance than clustering of the data features, as feature data reduction irrespective of understanding which features are optimal for the prediction of a particular semantic class or a set of classes, does not enable maximum utilization of the hypothesis space. Hence, predictive model development should go in conjunction with feature selection and feature generation approaches. Towards this end, we pursue this strategy of development of semantic land cover classes that could be used to retrieve knowledge about the image content that is relevant to the user's conjecture.

1.1. Feature Selection and Generation for Image Information Mining Applications

Feature extraction is an important and integral part of image mining systems; however, feature extraction is computationally intensive and produces huge amounts of data that is difficult to manage in practice. In this paper, we focus on the selection of a minimal subset of features that can give optimal performance for the identification of a particular land cover class. We propose a genetic algorithm- (GA) based wrapper approach for this purpose. In addition to feature selection, models for feature generation are developed to enrich the hypothesis language with additional constructed and derived features [1]. Hence, we pursue a hybrid approach for both feature selection and feature generation using GAs.

Feature selection is defined as the selection of a subset of features to describe a phenomenon from a larger set that may contain irrelevant or redundant features [2]. Feature selection techniques usually involve a criterion function and a search algorithm. The former aims at evaluating the separability of classes for a given subset of features. And the latter identifies the subset of features that maximize the adopted criterion [3]. Several separability indexes have been proposed in the remote sensing literature [4-8]. These indexes are generally based on the average distance among the classes and are computed by using the statistical distance between a pair of classes and are dependent upon the set of features considered. Lorenzo Bruzzone [3] argues that criterion functions based on the average pair wise distances without taking into consideration the costs associated with classes are not appropriate for selecting features that minimizes the total classification cost. A criterion function based on the Bayes rule for minimum cost (BRMC) has been proposed [3] and uses a neural network as the induction algorithm. However, recently, it has been noted that the feature selection stage and classification stage are not independent because the goal is correct classification with a corresponding feature pattern extracted with the intermediate step of feature extraction and dimensionality reduction [9]. Hence, it is recommended to couple feature selection with effective classification techniques. Wrappers-based feature selection is a methodology that has had a long history within the literature on statistics and pattern recognition [10], but its use within machine learning is relatively recent, and to the best of our knowledge no studies have been made for their applications in the remote sensing domain. In a wrapper approach, the feature subset selection algorithm exists as a wrapper around the induction algorithm. The feature subset selection algorithm conducts a search for a good subset using the induction algorithm itself as a part of the function evaluating feature subsets [11] [12] [13]. The RIIM system adopts a Genetic algorithm-based wrapper approach for feature selection and generation. Genetic algorithms (GAs) are randomized search and optimization techniques guided by the principles of evolution and natural genetics. They are efficient, adaptive, and robust search processes, producing near optimal solutions and have a large amount of implicit parallelism. The utility of GAs in solving problems that are large, multimodal, and highly complex has been demonstrated in several areas [14]. GAs have been used to search for feature subsets in conjunction with several classification methods such as neural networks [15, 16], decision trees [17], k-nearest neighbors [18-21], and Naïve Bayes [22, 23].

2. FRAMEWORK FOR RAPID IMAGE INFORMATION MINING (RIIM)

The RIIM system uses machine learning to address the task of exploring remote sensing imagery based on its content. The process of knowledge extraction from the imagery starts with the creation of tiles of the full scenes of the images and then performs an unsupervised segmentation on each tile. We use a hierarchical segmentation algorithm [24] to delineate regions of interest which are then used for primitive features extraction based on color, texture, and shape [25].

2.1. Features Selection using Wrapper-based Approach

The task of concept learning could be divided into two subtasks [26]:

- Deciding which features to use in describing the concept and
- Deciding how to combine those features.

Thus, the selection of relevant features, and the elimination of irrelevant ones, is one of the central problems in machine learning, and many induction algorithms incorporate some approaches to address this issue. Numerous search algorithms have been used to search for feature subsets [27]. The application of evolutionary learning algorithms to pattern recognition is becoming increasingly common. A variety of researchers have used evolutionary algorithms to perform feature selection [28]. The majority of these approaches begin with a large pool of potential features and an evolutionary process is used to evolve a population of feature subsets drawn from the pool. The subsets are evaluated using a standard classifier.

In this paper, we pursue a wrapper approach that uses a modified genetic algorithm for the incremental selection and generation of new features [29]. It uses an attribute-based induction algorithm for the evaluation of the features at hand. Inductive learning involves the process of *learning by example* (i.e., where a system tries to induce a general rule from a set of observed instances). Computational studies of Darwinian evolution and natural selection have led to numerous models for solving optimization [30-34]. GAs comprise of a subset of these evolution-based optimization problems techniques focusing on the application of selection, mutation, and recombination to a population of competing problem solutions [35, 36]. The population is usually formed from a constant number of individuals representing samples from the search space, including:

i). Chromosome

In the RIIM system, a chromosome is defined as an individual whose length is the same as the total number of features corresponding to each segmented region in the image. Each bit of the chromosome is initialized by a randomly selected value of 0 or 1. The fitness of the chromosome during the evolutionary process is calculated by considering only features that have 1s.

New individuals for the next generation are formed by applying two genetic operators; crossover and mutation to the individuals from the current generation. In each generation, half of those individuals with higher fitness values survive, and the others are extinguished. Two parents are selected from the survived individuals, and children are generated by a two-point crossover.

ii). Cross over

The crossover process defines how genes (chromosomes) from the parents have been passed to the offspring. In each generation, once two individuals are selected as the parents, a gene from each parent is broken into several segments and recombined with gene segments from the other parent based on a predefined crossover probability. After the crossover operation, every two parents will produce two children. The above selection and crossover processes will continue to run in each generation until the number of children equals the population size. At the end of each generation, it is useful to pass a certain number of the best individuals directly to the next generation, which is called elitism. In elitism, the best individual from the current generation is copied directly to the next generation, and is used for fast convergence.

iii). Mutation

The mutation process simulates the natural disturbance during crossover. It is a bit-by-bit operation based on the mutation probability (mutation rate). The mutation rate is generally selected based on the population size and other factors, such as the selection method and with or without an elitism policy. The mutation operation follows immediately after the crossover operation. Figure 1 depicts a two-point crossover and mutation processes used in GA. Parents that will produce new individuals are chosen according to their fitness. Better individuals are more likely to pass their genes to the next generation. Therefore, each generation should have a better overall fitness.

Original	1 0 0 0 0 1
Parent 1	1 0 0 1 1 0 0
Parent 2	0 1 1 1 0 1 0
Child 1	1 0 1 1 0 0 0
Child 2	0 1 0 1 1 1 0
Mutated	1 0 0 0 1 1 1

Fig. 1. Crossover and mutation

2.2. Hybrid Wrapper-based Genetic Algorithm Approach for Feature Selection and Generation

This algorithm combines the positive search properties of conventional genetic algorithms with the option to adapt the search space incrementally. In the wrapper approach, the feature subset selection algorithm exists as a wrapper around the induction algorithm. The feature subset selection algorithm conducts a search for a good subset using the induction algorithm itself as part of the function evaluating the feature subsets.

As shown in Figure 2, the outer cross validation wrapper randomly splits the original data set into n equally sized parts. For each run, the i^{th} part is kept as a test set while the remaining n-1 parts are passed to the genetic algorithm and subsequently to the final SVM learner whose learned model is tested on the i^{th} part of the dataset.

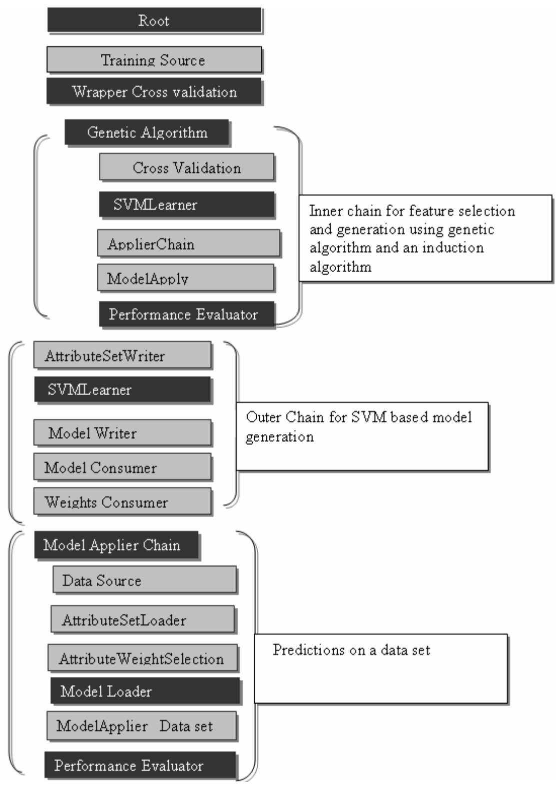

Fig. 2. Algorithm for wrapper-based approach for feature selection, generation, model creation, and performance evaluation

The inner cross validation trains the model on the training data training set and optimizes the choice of an attribute set using the disjunct evaluation data test set to avoid a bias in the selection of the attribute set. For reliable performance estimation of the complete operator chain for feature selection and classification learning, the training evaluation and test data sets need to be disjunct, which is guaranteed here by the nested cross validations [29].

The combined feature selection and generation of new features using a wrapper-based approach employs an attribute-based induction algorithm for the evaluation of the feature sets. This is achieved by a modified genetic algorithm as proposed by [29], which is based on a canonical genetic algorithm [36] using a bit string representation, fitness proportional selection, one-point crossover and standard mutation. In the combined feature selection and generation case, in addition to the standard mutation operator, the crossover operator is modified to make it a variable length crossover operator, which is based on the variable length genetic algorithms [37]. A new operator that produces new features is also introduced. This operator uses a type restriction on the generator (e.g. Boolean, mathematical, etc, operator) to combine a given set of compatible features and the resulting new features are generated. For instance, the suitable features from *Colordescriptor1*, *uniformity*, *entropy*, *eccentricity* could be combined with an addition operator which produces a new feature and added to the original feature set. The set of the compatible features is not limited to the original features, but can contain compound features that have already been created by the generator [29]. We adopt the above methodology for feature selection and generation in our study.

2.3. Support Vector Machines

Support Vector Machines (SVMs), as originally introduced by Vapnik within the area of statistical learning theory and structural risk minimization [38], have proven to work successfully on many applications of nonlinear classification and function estimation. The problems are formulated as convex optimization problems, usually quadratic programs, for which the dual problem is solved. Within the models and the formulation, one makes use of the kernel trick, which is based on the Mercer theorem related to positive definite kernels [39]. One can plug in any positive definite kernel (e.g., linear, polynomial, or Radial Basis Function (RBF)) for a support vector machine classifier. The conceptual idea of generalizing an existing linear technique to a nonlinear version by applying the kernel trick is an area of active research.

We try to find an optimal hyperplane that separates two classes. In order to find an optimal hyperplane, we need to minimize the norm of the vector **w**, which defines the separating hyperplane. This is equivalent to maximizing the margin between two classes. This concept is illustrated in Figure 3 for the nonlinearly separable case; note that only support vectors (dark circled) are required to define the optimally defined hyperplane. Linear decision hyperplanes in nonlinearly separable data can be handled by including slack variables to allow classification errors.

Normally the training data is separated into two parts; one is used for training and the other is used for testing. An improved version of handling the training sets is cross validation. In V-fold cross validation, we first divide the training set into V subsets of equal size. Sequentially, one subset is tested using the classifier trained on the remaining V-1 subsets. Thus, each instance of the whole training set is predicted once and the cross-validation accuracy is the percentage of data that are correctly classified. This cross validation method is used in conjunction with the wrapper-based approach presented in this study. The cross validation procedure can help alleviate over fitting the data.

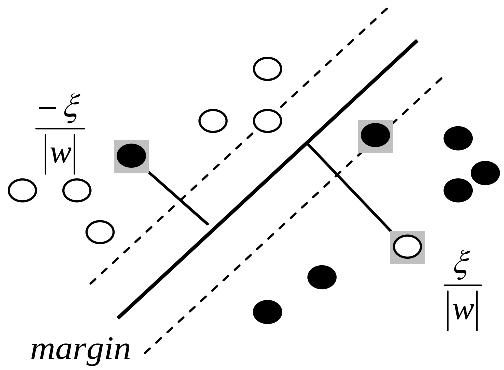

Fig. 3. A nonlinearly separable case

3. EXPERIMENTAL RESULTS

The system was implemented in a Java-based framework that provides an easy to use interface for interactivity and exploration of the imagery. This facilitated evaluating the presented wrapper-based approach on hurricane related events and in testing the system's ability to discover affected land cover regions. The system consists of offline and online components. In the former, at time of data ingestion, images are tiled in subimages, indexed, and stored in a repository. From these images, primitive features are extracted, such as color, textural, and geometrical attributes and stored in a RDBMS system, which forms the basis for the development of semantic models. The RIIM system optimizes the number of features required for identifying a certain semantic class; this is particularly helpful in real time assessment of the affected regions. Thus, once a semantic model was built based on an optimal subset of features, the RIIM system uses this knowledge to extract only those features or combination of features for a particular land cover class that the user is interested to explore on a new image which is ingested into the system. It automatically identifies the optimal subset of features required and, after the segmentation process, uses this knowledge about the selected subset to extract the primitive features and stores them in the database. This results in considerable reduction in the time gap between image ingestion, features extraction, and querying.

3.1. Data-set Description

To evaluate the presented wrapper-based approach on hurricane related events, data from Landsat ETM+ data (path 22, row 39, Aug 30, 2005) corresponding to post Katrina hurricane, and Landsat ETM + data (path 23, row 36, Sep 22,1999) that is not related to the hurricane (pre hurricane) were used in this study. This would help to identify training sites of different land covers that are specific to post hurricane areas (e.g. flooded fastlands) and training sites that are present in general during other times (agriculture, fallow, etc). Such a strategy helps to develop predictive models, which have the capability for image information mining from affected areas and compare the land cover classes in the pre and post hurricane events and also the evaluation over a period of time.

The database consists of primitive features from 7117 segmented regions extracted from 60 tiles (each of 967 x 915 dimension) corresponding to post-hurricane Landsat ETM+ data and primitive features from 4592 segmented regions extracted from 60 tiles (each of 719 x 575 dimensions) corresponding to pre-hurricane data. The bands 4, 3, 2 corresponding to near infrared, red, and green were selected and the false color composites (FCC) were derived from these bands. The flooded area classes selected for this study were based on the USGS-NWRC classification system (Figure 4), which provides specific land cover types that need to be assessed in a post hurricane scenario.

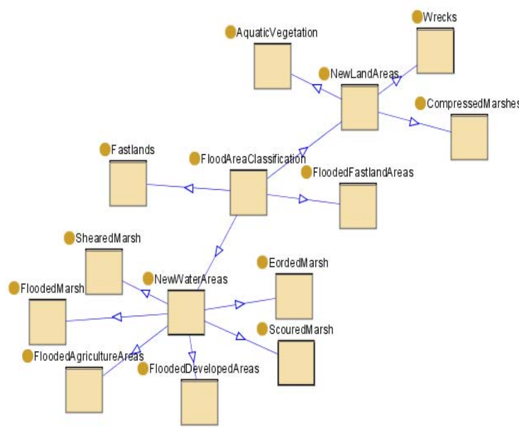

Fig. 4. Flooded area classes based on the USGS-NWRC classification system

3.2. Results and Discussion

Several experiments were conducted to formulate the semantic models. In the wrapper-based approach of feature selection/generation, the performance evaluation was done using two nested cross validations. The inner cross validation was used to find an optimal subset of features and the other outer cross validation was used to evaluate the performance of this subset of features.

The genetic algorithm parameters were set to 100 for the number of generations, 15 for population size, 0.5 for the crossover probability, 0.2 for the mutation (feature selection probability), and 0.5 for the feature generation probability. The induction algorithm used for classification was a SVM with complexity C=1000, epsilon=0.1, and using a RBF kernel.

3.3. Precision, Recall, and F-measure:

In content-based image retrieval systems (CBIR), *recall* and *precision* measurements are most often used to visualize how many relevant (target) and irrelevant (misdirected) images are contained in the highest ranked images [40]. In region-based image information mining, instead of accounting for the number of images retrieved, it is more relevant to account for the number of correct regions that are retrieved, which indirectly would correspond to the images retrieved. Hence, *precision* in this case is the proportion of relevant regions to all the regions retrieved. *Recall* is the proportion of relevant regions that are retrieved, out of all relevant regions. In general, high *precision* indicates that most of the items you retrieve are relevant**.** High *recall* indicates that you have retrieved most of the available relevant regions in a repository. The *F-measure* is the weighted harmonic mean of precision.

We performed the GA-based wrapper approach for feature selection for the semantic classes. Table 1 depicts the feature selection results; *precision*, *recall* and *F-measure* values using the reduced feature set as obtained from the GA. The accuracy of the retrieval is measured in terms of the relative number of correctly classified examples. From an initial set of 15 features, the GA algorithm produced an optimal feature subset specific to each semantic class.

Table 1: Accuracy, precision, recall, and F-measure obtained using feature selection by GA

Class	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
Water bodies	98.80	93.80	87.76	90.53
Agriculture	97.63	98.83	98.18	98.50
Fallow Land	96.49	98.38	94.81	96.56
Forest	98.53	97.84	94.44	96.11
Flooded vegetation	91.04	89.83	100	94.64
Flooded fastlands	96.92	93.75	83.33	88.24

The number of features reduced is more than 50% in most of the cases while providing good accuracy. These selected features were then used to generate a semantic model for each class. This helps to rapidly extract a limited number of features that are highly relevant to a semantic class from images in a hurricane disaster event and begin the process of knowledge discovery.

The retrieval from a semantic query relevant to the hurricane event is depicted in Figure 5. Several tiles images have been retrieved by the system that matches the query. The user then has the option to further look into the details of the system derived knowledge by clicking on an image, which brings up a window that provides a detailed view of the actual regions that matched the user's semantic query.

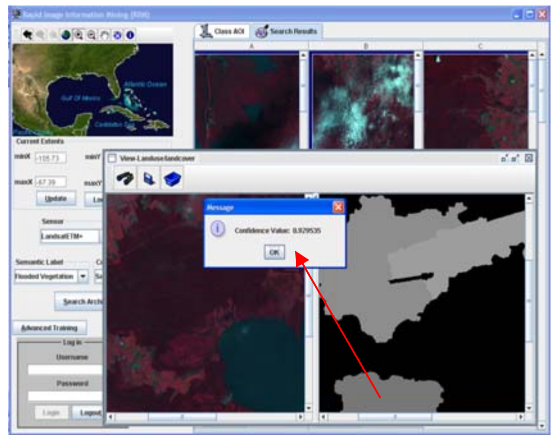

Fig. 5. Results of a semantic query (flooded vegetation)

It is also possible to know the confidence level of each of the retrieved regions; this is helpful to understand how probable the region of a particular type. Accordingly, the RIIM system automatically calculates these confidence values for all the regions and stores them in the database. This further facilitates the retrieval of regions which are above a particular confidence threshold that satisfies the user's needs.

4. CONCLUSIONS

The RIIM prototype is reliable and fast and is focused on image exploration for hurricane affected regions in near real time scenarios. The computationally intensive tasks of feature extraction and model generation are considerably reduced by a wrapper-based approach for feature selection proposed in this research. This is vital for emergency response activities. The RIIM system provides capabilities for a first assessment of the disaster situation through the querying of the actual content in the remote sensing images, which is currently limited by queries only at the image metadata level. The proposed system currently uses imagery from only one sensor, but can be easily scaled up to be used with a variety of sensors.

5. REFERENCES

1. R. S. Micahalski. "A theory and methodology of inductive learning". R.S. Michalski, J.G. Carbonell, and T. M. Mitchell, editors, *Machine learning: An artificial Intelligence approach*, volume 1, chapter 4, pp 83-134. Morgan Kaufmann, Palo Alto, CA, USA, 1983.
2. John Loughery, Padraig Cunningham,"Overfitting in wrapper-Based Feature subsets Selection: The harder you try the worse it gets" In Proceedings of the Twenty-fourth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, 2004.
3. Lorenzon Bruzzone, "An approach to feature selection and classification of remote sensing images based on the bayes rule for minimum cost. *IEEE Trans. on Geosci. Remote Sensing.*, vol.38, no.1, January 2000.
4. J. A Richards, Remote sensing Digital image analysis, 2nd Ed. New York: Springer-Verlag, 1993.
5. P. H. Swain and S. M. Davis, *Remote Sensing: The Quantitative Approach*. New York: McGraw-Hill, 1978.
6. I. L. Thomas, N. P. Ching, "A review of multi-channel indices of class separability," *Int. J. Remote Sens.*, vol. 8, pp. 331–350, 1987.
7. P.W. Mausel, W. J. Kramber, and J. K. Lee, "Optimum band selection for supervised classification of multispectral data", *Photogram. Eng. Remote Sens.*, vol. 56, pp.55-60, Jan.1990.
8. L. Bruzzone, F. Roli, and S. B. Serpico, "An extension of the Jeffreys-Matusita distance to multiclass cases of feature selection, "*IEEE Trans. on Geosci. Remote Sensing*, vol. 33, pp. 1318-1321, Nov.1995
9. M. Pei, E.D.GoodNam, W.F. Punch " feature extraction using genetic algorithms", *Proceeding of International Symposium on Intelligent Data Engineering and Learning'98 (IDEAL'98)*, Hong Kong, Oct. 1998.
10. Deviver, P.A., & Kittler, J. "Pattern recognition: A statistical approach". *New York: Prentice-Hall*. 1982.
11. Ron Kohavi, George H. John, "Wrappers for Feature subset selection", *Artificial intelligence journal special issue on relevance*, 1997.
12. Eric Cantu-Paz, "Feature subset selection, class separability and genetic algorithms", submitted to Genetic and evolutionary computing conference , WA, 2006
13. Aha, D., Bankert, R. A "Comparative evaluation of sequential feature selection algorithms", *Artificial intelligence and Statistics*, D. Fisher and J.H. Lenz, New York, 1996.
14. S.K. Pal and P.P. Wang Eds., "Genetic algorithms for Pattern Recognition", CRC Press, Boca Raton, FL, 1996.
15. Brill, F.Z., Brown, D.E., Martin, W.N., "Genetic algorithms for feature selection for counter propagation networks". Tech. Rep. No. IPC-TR-90-004, University of Virginia, Institute of Parallel Computation, Charlottesville, 1990.
16. Brotherton, T.W., Simpson, and P.K., "Dynamic feature set training of neural nets for classification", In McDonnell, J.R., Reynolds, R.G., Fogel, D.B., eds.: *Evolutionary Programming IV*, Cambridge, MA, MIT Press, 1995
17. Bala, J., De Jong, K., Huang, J., Vafaie, H., Wechsler, H. "Using learning to facilitate the evolution of features for recognizing visual concepts", *Evolutionary Computation*, 4, pp. 297-311, 1996.
18. Kelly, J.D., Davis, L., "Hybridizing the genetic algorithm and the K-nearest neighbors classification algorithm". In Belew, R.K., Booker, L.B., eds.: *Proceedings of the Fourth International Conference on Genetic Algorithms*, San Mateo, CA, Morgan.
19. Punch, W.F., Goodman, E.D., Pei, M., Chia-Shun, L., Hovland, P., Enbody, "Further research on feature selection and classification using genetic algorithms", *Proceedings of the Fifth International Conference on Genetic Algorithms*, San Mateo, CA, Morgan Kaufmann, pp. 557-564, 1993.
20. Raymer, M.L., Punch, W.F., Goodman, E.D., Sanschagrin, P.C., Kuhn, L.A, "Simultaneous feature scaling and selection using a genetic algorithm", *Proceedings of the Seventh International Conference on Genetic Algorithms*, San Francisco, Morgan Kaufmann, pp. 561-567, 1997.
21. Kudo, M., Sklansky, K. "Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition*, vol. 33, pp. 25-41, 2000.
22. Inza, I., Larra~naga, P., Etxeberria, R., Sierra, B., "Feature subset selection by Bayesian networks based optimization", *Artificial Intelligence*, Vol. 123, pp. 157-184, 1999.
23. Cantu-Paz, E., "Feature subset selection by estimation of distribution algorithms" In Langdon, W.B., Cantu-Paz, E., Mathias, K., Roy, R., Davis, D., Poli, R., Balakrishnan, K., Honavar, V., Rudolph, G., Wegener, J., Bull, L., Potter, M.A., Schultz, A.C., Miller, J.F., Burke, E., Jonoska, N., eds., *Proceedings of the Genetic and Evolutionary Computation Conference*, San Francisco, CA, Morgan Kaufmann Publishers, pp. 303-310, 2002.
24. Y.Deng, and B.S.Manjunath, "Unsupervised segmentation of color-texture regions in images and video," *IEEE Transactions on Pattern Analysis and*

25. *Machine Intell.*, vol. 23, no. 8, pp. 800-810, Aug. 2001.
26. S. S. Durbha and R. L. King, "Semantics enabled framework for knowledge discovery from earth observation data archives", *IEEE Transactions on Geosciences and Remote Sensing,* Vol.43, NO.11, Nov 2005.
27. A. Blum and Pat Langley, "Selection of Relevant features and examples in machine learning", *Artificial Intelligence on relevance*, R. Greiner and D. Subramanian (eds.), 1997.
28. Jain A., Zongker, D, "Feature selection: evaluation, application and small sample performance". *IEEE transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, pp. 153-158, 1997.
29. M. M. Rizki, M. A. Zmuda and L. A. Tamburino, "Evolving pattern recognition systems", *IEEE trans. on evolutionary computation*, vol. 6, no. 6. Dec, 2002.
30. Oliver Ritthoff, Ralf Klinkenberg, Simon Fischer, Ingo Mierswa, "A Hybrid Approach to Feature Selection and Generation Using an Evolutionary Algorithm**",** Proceedings of the 2002 U.K. Workshop on Computational Intelligence (UKCI-02), 2002.
31. A.S. Fraser," Simulation of genetic systems by automatic digital computers-I: Introduction," *Australian J. Bio.Sci*, vol.10, pp.484-491, 1957.
32. J.L. Crosby, "Computers in the study of evolution", Sci.Prog.Oxf., vol.55, pp.279-292, 1967.
33. H. J. Bremermann, M. Rogson, and S. Salaff, "Global properties of evolution processes, "Natural automata and useful simulation, H. H. Pattee, E. E. Edlsack, L. Fein, and A.B. Callahan, Eds. Washington, DC: Spartan, 1966, pp.3-41.
34. J. Reed, R. Toombs, and N. A. Barracelli, "Simulation of biological evolution and machine learning: Selection of self-reproducing numeric patterns by data processing machines, effects of hereditary control, mutation type and crossing, " J. theorist. Biol., vol. 17, pp. 319-342, 1967.
35. D. B. Fogel, Ed., Evolutionary Computing: The Fossil Record. New York: IEEE Press, 1998.
36. J. H. Holland, Adaptation in natural and Artificial Systems. Ann Arbor, MI:Univ. Michigan Press, 1975.
37. D. Goldberg, Genetic Algorithms in search and optimization, and Machine learning, Reasoing, MA: Addison-Wesley, 1989.
38. R. Poli, J. E. Rowe, and N. McPhee. Markov chain models for gp and variable-length gas with homologous crossover. In Lee Spector et al., editor, *proceedings of the genetic and evolutionary computation conference (GECCO-2001)*, pp 112-119, San Francisco, CA, USA, 2001.
39. V. Vapnik. *The Nature of Statistical Learning Theory*. New york: Springer-Verlag, 1995.
40. R. Courant and D. Hilbert, Methods of Mathematical Physics. New York: Wiley, 1953.
41. M. Datcu, H. Daschiel, A. Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P. G. Machetti, and S D'Elia, "Information mining in remote sensing image archives: system concepts," IEEE Trans. on Geosci. Remote Sensing, vol. 41, no.12, pp. 2923-2936, 2003.