A Review of Medical Image Classification Techniques

Smitha P.
Selection grade lecturer
Dept. of CSE, CE
Karunagapally

Shaji L.
Lecture in comp.Appln
Dept.of IT, CE,
Karunagapally

Dr. Mini M. G.
Asst. Prof. & HOD
Dept. of ECE, MEC
Karunagapally

ABSTRACT
Medical image Classification can play an important role in diagnostic and teaching purposes in medicine. For these purposes different imaging modalities are used. There are many classifications created for medical images using both grey-scale and color medical images. One way is to find the texture of the images and have the analysis. Texture classification is an image processing technique by which different regions of an image are identified based on texture properties[4]. Second way is by using neural network classification techniques and the final one is by using the data mining classification schemes. Neural networks play a vital role in classification, with the help of, supervised and unsupervised techniques. The word data mining refers to, extracting the knowledge from large amounts of data. It is one of the area, which uses statistical, machine learning, visualization and other data manipulation with knowledge extraction techniques[10]. This finds an insight into the relationship between the data and patterns hidden in the data. Using the digital data within the pictures actual communication systems creates a possibility for research enhancements.

Medical images form a vital component of a patient's health record and are associated with manipulation, processing and handling of data by computers. This makes the basis for the computer-assisted radiology development. Further developments are associated with the use of decision support systems which helps to decide, the relevant knowledge for diagnosis.

Keywords
Image classification, Texture classification, Data mining, Association rule mining.

1. INTRODUCTION
Effective medical images can play an important role in aiding in diagnosis and treatment; they can also be useful in the education domain for healthcare students by explaining with these images will help them in their studies. Advances in digital imaging technologies, created a large growth in the number of digital images taken, in recent years. In addition to the Picture Archival and Communication Systems (PACS) that are becoming omnipresent in hospital and clinics, there are numerous online collections of medical images[2]. On-line atlases of images can be found for many medical domains including dermatology, radiology, ophthalmology, cytopathology and gastroenterology. The sheer volume of medical image data is useful for numerous challenges and opportunities. Classification refers to as assigning a physical object or incident into one of a set of predefined categories. Medical image databases used for image classification or for teaching purposes often contain images of many different modalities, taken under varied conditions with variable accuracy of annotation. This can be true for images found in various on-line resources, including those that access the on-line content of journals. Approaches combining both visual and textual techniques for classification have shown some promise at medical image classification tasks. Figure 1 shows the overview of the medical image processing steps.

1.1 Texture Classification
In texture classification the goal is to assign an unknown sample image to one of a set of known texture classes. Texture classification is one of the main domains in the field of texture analysis. Texture analysis is important in many applications of computer image analysis for classification or segmentation of images based on local spatial variations of intensity or color. A successful classification or segmentation requires an efficient description of the image texture[1]. An important application area in texture classifications are industrial and biomedical surface inspection, for example finding the defects and disease, ground classification and segmentation of satellite or aerial imagery, segmentation of textured regions in document analysis, and content-based access to image databases etc. However, despite of many potential areas of application for texture analysis in industry, there are only a limited number of successful examples available. A wide variety of techniques for describing image texture have been proposed. Texture analysis methods can be divided into four categories called statistical, geometrical, model-based and signal processing [8].

1.2 Neural Network Classification
Neural networks have emerged as an important tool for the classification. The recent research activities in neural classification have established that neural networks are a promising alternative to different conventional classification methods. The advantage of neural networks resides in the following theoretical aspects[7]. First, neural networks are data driven self-adaptive methods in which they can adjust themselves to the data, without any explicit specification of functional or distributional form with the underlying model. Second, neural networks are universal functional approximations which can approximate any function with arbitrary accuracy. Since any classification procedure finds a functional relationship between the group membership with the attributes of the object, accurate identification of this underlying function is very important. Third, neural networks are nonlinear models, which makes them flexible in modeling complex real world applications. Finally, they are able to estimate the posterior probabilities, which provides the basis in establishing the classification rules and performing statistical analysis. Neural networks have been successfully applied to a wide variety of real world classification such as speech recognition, fault detection, medical diagnosis etc[8].

Neural networks play an important role in classifications by using its supervised and unsupervised techniques. Self-organizing maps (SOM) of neural networks are useful in cluster based classification of medical images. This methodology can be used in categorization and in computer-aided diagnostic decision making.

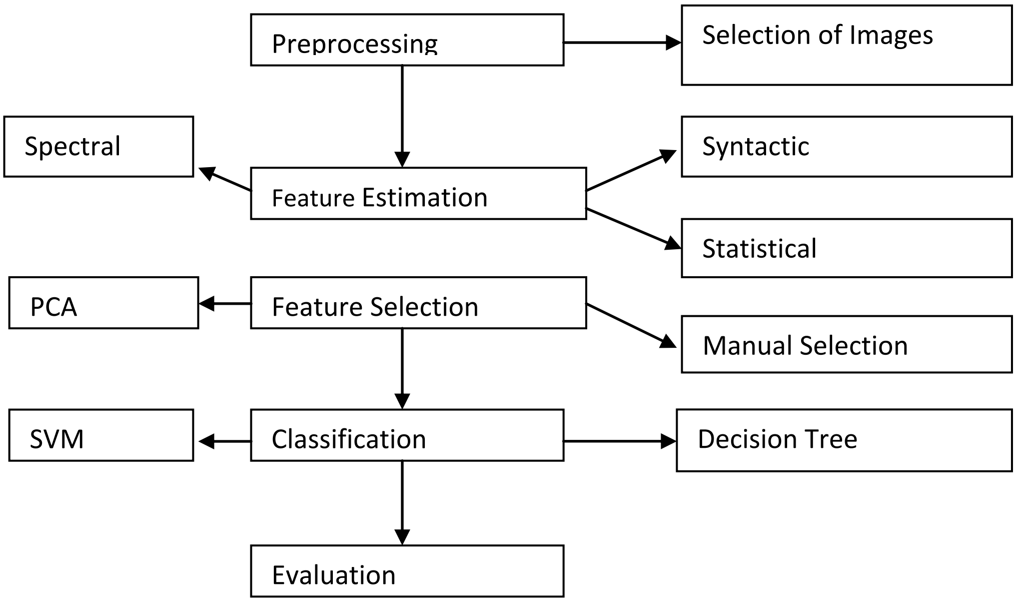

Figure 1. Overview of the steps involved in medical image processing.

1.3 Data Mining Techniques

Classification in data mining technique is used to predict group membership for data instances. Data mining involves the use of sophisticated data analysis tools to discover the relationships in large data set. Data mining never means a collection and managing data, it also includes analysis and prediction of data. Classification techniques are capable of processing a wider variety of data than regression and also growing in popularity.

People are often prone to making mistakes during analysis while establish relationships between multiple features. This makes it difficult for them to find solutions for certain problems. Data mining can often be successfully applied to these problems by improving the efficiency of the systems and the designs of the machines. Numerous data mining applications, involve tasks that can be set up as supervised.

2. TEXTURE CLASSIFICATIONS

Conventional radiologists produce diagnoses on the basis of combining their training, experience, and individual judgment. But this will create an inevitable degree of variability in image interpretation and relies primarily on human visual perception. Nowadays tools for automated pattern recognition and image analysis, can provide an objective information, to support clinical decision-making and may help to reduce the variability. Texture analysis describes a wide range of techniques that enable the spectral properties of an image. This technique is tremendously versatile and can be applied to virtually any modality of digital image[1]. If the spatial extent of the lesion is to be identified by an independent means, then the application of texture analysis is restricted to a set of predefined regions of interest. A verity of approaches to the quantification and characterization of image texture have been proposed, with the most textural features actually falling under three general categories: syntactic, statistical, and spectral. Syntactic texture analysis identifies fundamental or primitive elements of the images.

The first-order statistical features such as Mean, Standard Deviation, Energy and Variance are self explanatory. The second-order statistical features are extracted from the probabilistic matrix such as Angular second Movement, correlation, contrast, sum average, sum entropy, sum variance and entropy etc., In a similar manner, run-length features can be computed to evaluate the coarseness of a texture in a predetermined direction. A gray-level run consists of a set of consecutive collinear pixels in a given direction. The Spectral Features means the co-occurrence[12] or run-length[13] features may lack the sensitivity, to identify larger scale or more coarse changes in spatial frequency. Wavelet concept plays an important role in texture analysis. Wavelet functions can be designed to evaluate spatial frequencies at multiple scales[8]. Readers will understand the close relation of the wavelet transform, the Fourier transform, which can identify the spatial frequencies present in a signal intensity. But it cannot delineate temporal changes in frequency content and presumes, that all signals reflect a superposition of sinusoids. In Some cases localization can be imparted to Fourier analysis by means of the windowed or short-time method, which allows the Fourier transform to be performed on sequential portions of the entire signal intensity.

The wavelet transform provides more flexibility by enabling us to trade some degree of spatial-frequency resolutions for the ability to localize this frequency content in time. But unlike the Fourier Transform, the wavelet transform does not require sinusoidal basis functions. In fact, the word wavelet, refers to the generalized, basis function used to compute the transform, which involves scaling and translating a core wavelet function and comparing each resultant wavelet with the signal intensity and computing a coefficient reflecting the strength of this similarity as wavelet coefficients. As such, the wavelet transform is an inherently multi scale analysis method. Due to the computational complexity of the continuous wavelet transform, there is a substantial difficulties for implementing this type of analysis in the clinical setting. Other methods proposed such as the discrete ortho-normal Stockwell Transform(ST)[14], which can be computed in a clinically realistic timeframe.

The Feature Selection and Extraction[3,15] are the next steps used in texture analysis that is, select the subset of features most likely to distinguish one tissue class or patient diagnosis from another. The challenge here is that even a modest Gray Level Co-occurrence Matrix(GLCM) approach with 3 d and 4 theta, values can produce many more textural features than are suitable for the number of positive cases that will be ultimately subjected to classification. If measures are not taken, to reduce the number of features before classification, then the statistical model will better reflect the noise or random errors than the underlying data. Fortunately, there is a number of strategies available for dimensionality reduction, beginning with simple consolidation of a given feature over all directions. More systematic techniques can be implemented to search for feature subsets, such as those based on the Fisher criterion or by means of Principle Component Analysis, which are previously used to extract mutually orthogonal features from the larger consolidated feature space.

The classification of textural features is useful to a radiologist's clinical interpretation and involves partitioning the selected feature space according to tissue class or category. Classification is typically accomplished by using a decision or discriminate function. Textural features may complement the macro texture information already used by radiologists, such as the organization of lesions in normal brain parenchyma. One of the earliest neuro-MR imaging applications of statistical texture analysis deals the characterization of brain tumors. This remains a particularly important problem because there can be substantial intersection between the T1 and T2 of benign and malignant brain tumors, which complicates lesion characterization with conventional Magnetic Resonance imaging. MR imaging texture analysis procedure for identifying tumor constituents is proposed in Lerski (1993) feasibility study of some patients with intracranial tumors. The authors combined first-order histogram, gradient, and second-order GLCM features are extracted from T1- and T2-weighted spin-echo MR images into a 4- layer hierarchical decision tree with the stepwise Discreminent Analysis applied at each level, to identify the features of most capable of discriminating between and among, the tissue and tumor constituents. S ratifying the tumor constituents are identified by a separate texture class corresponding to peritumoral White Matter(WM) by Mahmoud-Ghoneim. These authors discovered that the specificity of GLCM features for the identification of tumor constituents could be improved by extending the evaluation of in-plane two dimensional inter pixel relationships to include through-plane 3 dimensional relationships (multiple-section volume).

While compared with their two dimensional equivalents, the 6 GLCM features extracted from the three dimensional noncontrast T1-weighted gradient-echo images were superior for discriminating between necrosis and solid tumor (sensitivity and specificity were each 100% for three dimensional versus 75% and 60% for two dimensional) and between solid tumor and edema (sensitivity and specificity were each 82% for three dimensional versus 60% and 55% for two dimensional). The objectivity of texture analysis depends on the assumption of that images which are acquired, processed, and analyzed under identical conditions. The author Herlidou-Meˆme acquired standard T1- and T2-weighted images from ten healthy volunteers and nearly sixty patients with confirmed intracranial tumors and data acquired by using three different scanners to test the robustness of the technique. The authors were able to use the same statistical textural features to segment the tumors, irrespective of the acquisition parameters, scanner, reconstruction, or processing used, and reported highly reproducible results in a head-to-head comparison of second-order features [5]. These features are computed from T1- and T2-weighted images of gliomas acquired with three different MR imaging scanners. With respect to multiscale or spectral features authors have suggested that the Stock well transform is sensitive enough to distinguish between tumor genotypes, such as in the identification of oligodendrogliomas with a genetic signature associated with better results. The spectral analysis of T2-weighted MR images with the Stockwell transform are more accurate than the visual assessment and with a sensitivity and specificity of 93.2% and 96.4%, respectively, for delineating these particular tumor genotypes.

3. New Applications of Texture Analysis

Techniques are needed for the early prediction of hemorrhagic complications. The findings reported for the Multiple Sclerosis(MS) studies and assuming that the blood-brain barrier disruption before hemorrhagic transformation(HT) is similar in degree to that occurring in acute MS. It is hypothesized that there would be differences in the complexity and homogeneity of HT-prone stroke infarcts. For this it is evaluated that, first-order (Mean Gray Level, Variance of Gray Levels) and four 2D-GLCM texture features are extracted from post contrast T1- weighted spin-echo images acquired from 34 patients with Accute Ischemic Stroke(AIS)[1]. Contrast and correlation factors were the only two features capable of predicting HT and were much more sensitive predictors than conventional visual assessment of post contrast T1-weighted images. Surprisingly, the addition of visual enhancement to either contrast and correlation factors did not significantly improve accuracy. Recently a combined 3D GLCM analysis with Dynamic Contrast Enhanced (DCE) imaging of breast lesions was proposed. By examining the time-evolution of contrast enhancement, with in which time constitutes the fourth dimension and the authors reported impressive differentiation of benign and malignant tumors, which may prove instructive in future neuro-MR imaging applications.

4. NEURAL BASED CLASSIFICATION

Bayesian decision theory is the basis of statistical classification methods and it provides the fundamental probability model for well-known classification procedures such as the statistical discriminant analysis. Supervised classification can be performed by using the Bayesian decision theory or linear discriminant analysis with a distance classifier. Supervised classification of nonparametric data can be accomplished by using decision trees, k-nearest neighbor, support vector machines[6], or neural network techniques. The accuracy or success of our feature classification strategy is generally evaluated by cross-validation. Beginning with, dividing the data into training and testing subsets, and then performing the classification process on the training set, and then validating the results of the classification on the testing set. If we then repartition the original cases and repeat the same procedure, we can ultimately use the average classification accuracy as our overall validation metric. We can also construct a logistic regression model by assigning the most discriminating features as predictors and we can get either the tissue class or the diagnosis, as the outcome measure. The classification accuracy is then calculated by measuring the area under the Receiver

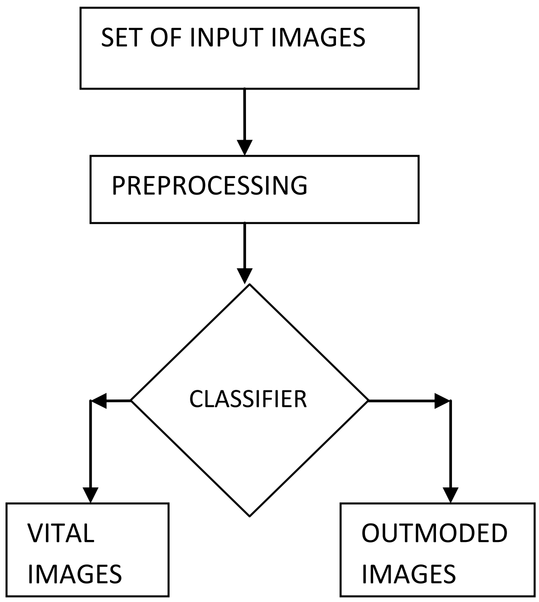

Figure 2. Neural network classifier

Operating Characteristic (ROC) curve. It is also noted that unsupervised classification techniques including the K-means or hierarchical clustering are suitable for this scenarios, in which there is no prior knowledge of how the feature space is organized; therefore, all cases belong to a single testing set. One such algorithm in this purpose is the Discriminant Analysis and the other one is the Support Vector Machine. The Discriminant Analysis (DA) algorithm is a simple algorithm which generally works well. Support Vector Machines (SVM) are generally produces better results, but it is more difficult to find the optimal parameters which give the best results. For this reason it is usually takes a longer time to produce a good SVM model than a DA model.

4.1 Support Vector Machines (SVM)

A Support Vector Machine is a binary classifier; it aims to classify two classes of instances by finding the maximum separating hyper plane between the two. For this reason SVM tends to generalize better. With the basic design of Support Vector Machines, it can only discriminate between two classes. In order to allow for the classification of more than two classes, one of the following methods can be employed. Figure 2 shows the general neural network classifier. One such method is the "one-vsone" method which creates one binary classifier for each pair of classes. If for our case, for example three classes are there then three binary classifiers are created. Support Vector Machines in their simple form, are called linear classifiers. It is possible however to create a nonlinear SVM by increasing the dimensionality of the feature space, and by using the so-called "kernel-trick". It is thus possible to find a separating the hyper plane in a higher dimensions where such a hyper plane would not exist in lower dimensions. There are many choices for which kernel, to use. The standard choices are the linear kernel (which is otherwise called as dot-product kernel), the polynomial kernel and the Gaussian Kernel. The Gaussian Kernel is the special case of Radial Basis Function (RBF) kernel. In the standard case, the distance used, is the Euclidean distance. In the RBF kernel, the parameters determine, the width of the kernel, and d(x, y) is the distance metric.

4.2 Discriminant Analysis

Discriminant Analysis (DA) is a parametric technique which uses the Maximum-Likelihood method, to make the predictions. A Baysian approach is taken to estimate the parameters of a normal distribution for all classes. Each class distribution is characterized by a mean value and a covariance matrix. Depending on how the covariance matrix is calculated for each class distribution will be resulting in discriminant function of the classifier, and can be linear or polynomial.

5. DATA MINING TECHNIQUES

Data mining is the process of discovering meaningful new correlations, patterns and trends by shifting through large amounts of data stored in databases, using pattern recognition techniques as well as statistical and mathematical techniques. There is another definitions for Data mining, which is the analysis of observational data sets to find unsuspected relationships and to summarize the data in many ways that are both understandable and useful to the data analyzer. Data mining is an interdisciplinary field bringing together the techniques from machine learning, pattern recognition, statistics, databases and visualization to address the issues of information extraction from a large data bases. This is an evolving and growing area of research and development. Researchers need to pay attention to the mining of different data types, including numeric and alphanumeric formats, video, voice, speech, graphics, text, images and also their combinations[10].

Data mining techniques can be used for preprocessing, extracting, analysis and segmentation areas. For preprocessing purposes cluster analysis can be used. Cluster analysis is a method of grouping data with similar characteristics into larger units for analysis. The C-MEANS algorithm is the most popular non-hierarchical iterative clustering algorithm. When it is applied on a set of data the c-means finds some of the most natural c-groups exiting in data. That is by (1) Choose c cluster centers to coincide with c randomly-chosen patterns or c randomly defined points inside the hyper volume containing the patterns. (2) Assign each pattern to the closest cluster center according to a certain pre-specified metric or dissimilarity measure. (3) Recalculate the cluster centers using the current cluster memberships. (4) If a convergence criterion is not met, go to step 2. Else the computation is over and the current clusters correspond to the c- groups are identified by the algorithm in the data. Since in fuzzy set theory which gave rise to the concept of partial membership which is based on membership functions and this fuzziness has received increasing attention. Fuzzy clustering which gives variations in clustering, can produce overlapping cluster partitions, and has been widely studied and applied in various areas. So far, there have been proposed a relatively small number of methods for testing the existence or inexistence of a natural grouping tendency in a data collection and most of them being based on arguments coming from mathematical statistics and heuristic graphical techniques[11].

The association rules play an important role in data mining techniques. Association rule mining has been used in most of the research for finding the rules for diagnosis in large and small databases. The frequent patterns from the CT scan images are generated by frequent pattern tree (FP-Tree) algorithm that mines the association rules. The decision tree method has been used to classify the medical images for diagnosis. These systems enhances the classification process to be more accurate. The hybrid methods actually improves the efficiency of the methods than the traditional image mining methods. The Watershed morphological transformation of images has been used for segmentation and removal of inconsistent data, from the given images. Two dimension median filter has been used prior to edge detection to reduce the amount of noise in the stack of images. This is preferred because it is a very simple smoothing filter but at the same time preserves the edges. The given input image is then segmented in to the number of objects using canny edge detection method. For each object, the texture features have been calculated, and stored in the some temporary or transaction database. Using these information databases one can go for association rule mining. In this it is possible to discover the hidden association relationship between the different item sets from the transaction databases. By using set theory the frequent item sets and maximum frequent item sets can be formed. All these information identified has been stored in the transactional database. The information stored in the transactional databases are given in the form of tree structure. Decision tree based classification methods are widely used in data mining for the decision support application. This type of systems use decisions that have to be made by the physicians whether the maximum frequent item set that are found in the transaction tree has been compared with the maximum frequent item of the test images. In this, the decision tree classification with association rule classification method provides a better option for the physicians, to classify the benign and malignant images. It is done by comparing the maximum frequent items generated by the association rules in the training image, have been compared with the maximum frequent items, of the test image. Hence the diagnosis can be made easily.

6. CONCLUSION

Texture analysis plays a supportive rather than a comprehensive role in the future of medical image interpretation. The robustness of texture analysis makes it particularly attractive for monitoring disease progression or treatment response with time, as demonstrated with MS. Support Vector Machines using the correlation kernel, Polinomial kernel, Gaussian kernel and RBF kernel all pay important role in each applications and give satisfactory results. Neural technicques are more useful in automatically analyzing the images by training the networks. Data Mining techniques are also playing equal role in keeping the image databases and making analysis on the set of images. Fuzzy based algorithms are giving better results. First two cases are more complicated techniques because of its tedious calculations, though giving good performance.

7. REFERRENCE

[1] A. Kassner, R.E. Thornhill, "*Texture Analysis: A Review of Neurologic MR Imaging Applications*", A Journal of NeuroRadiology 31:809, May 2010.

[2] A. Materka, M. Strzelecki, *Texture Analysis Methods – A Review*, Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels 1998

[3] E. L. Hall, Kruger RP, Dwyer SJ, Hall DL, Mclaren RW, Lodwick GS," *A survey of preprocessing and feature extraction techniques for radiographic images***.** IEEE Transactions on Computers 1971;20:1032–44

[4] Haralick RM, Shanmugam K, Dinstein I."*Textural features for image classification***."** IEEE Tran Syst Man Cybern 1973;3:610–21

[5] Julesz B, Gilbert EN, Shepp LA., "*Inability of humans to discriminate between visual textures that agree in second-order statistics: revisited".* Perception 1973;2:391–405

[6] D. Jegelevicius, A. Krisciukaitis, A. Lukosevicius, V. Marozas, A. Paunksnis, V. Barzdziukas, et al, *"Network Based Clinical Decision Support System"* Proceedings of the 9th International Conference on Info~ation Technology and Applications in Biomedicine, ITAB 2009, Larnaca, Cyprus, November 2009

[7] K.V.Ramana, Raghu.B.Korrapati, "*Neural Network Based Classification and Diagnosis of Brain Hemorrhages",* International Journal of Artificial Intelligence and Expert Systems (IJAE), Volume (1): Issue (2) 2009

[8] G.Wiselin Jiji, L.Ganesan, S.Sankar Ganesh**,** "*Unsupervised Texture Classification*" Journal of Theoretical and Applied Information Technology ; Vol.5 No.4 April 2009.

[9] P.S.Hiremath, S.Shivashankar , "*Texture Classification using Wavelet Packet Decomposition",* GVIP Journal, Volume 6, Issue 2, September, 2006.

[10] U. Rajendra Acharya, Wenwei Yu, "*Data Mining Techniques in Medical Informatics",* Open Medical Informatics Journal, Published online,May 2010.

[11] P. Rajendran, M.Madheswaran, "*Hybrid Medical Image Classification Using Association Rule Mining with Decision Tree Algorithm"*, Journal Of Computing, Volume 2, Issue 1, January 2010.

[12] Mahmoud-Ghoneim D, Alkaabi MK, de Certaines JD, *"The impact of image dynamic range on texture classification of brain white matter***",** MC Med Imaging 2008;8:18.

[13] Galloway MM, *"Texture analysis using gray level run lengths" Comp Graph Image Proc* 1975;4:172–79.

[14] Stockwell RG, " *A basis for efficient representation of the S-transform",* Digital Signal Processing 2007;17:371–93.

[15] J F Dale Addison, Stefan Wermter, Garen Z Arevian, "A Comparison of Feature Extraction and Selection Techniques", ND00909.