Rough Sets in Medical Informatics Applications

Aboul Ella Hassanien, Ajith Abraham, James F. Peters, and Gerald Schaefer

Abstract Rough sets offer an effective approach of managing uncertainties and can be employed for tasks such as data dependency analysis, feature identification, dimensionality reduction, and pattern classification. As these tasks are common in many medical applications it is only natural that rough sets, despite their relative 'youth' compared to other techniques, provide a suitable method in such applications. In this paper, we provide a short summary on the use of rough sets in the medical informatics domain, focussing on applications of medical image segmentation, pattern classification and computer assisted medical decision making.

1 Introduction

Rough set theory provides an approach to approximation of sets that leads to useful forms of granular computing. The underlying concept is to extract to what extent a given set of objects (e.g. extracted feature samples) approximate another set of objects of interest. Rough sets offer an effective approach of managing uncertainties

Aboul Ella Hassanien
Information Technology Department, FCI, Cairo University, and the System Department, CBA, Kuwait University, Kuwait, email: `abo@cba.edu.kw`

Ajith Abraham
Center for Quantifiable Quality of Service in Communication Systems, Norwegian University of Science and Technology, Trondheim, Norway, email: `ajith.abraham@ieee.org`

James F. Peters
Computational Intelligence Laboratory, Department of Electrical & Computer Engineering, University of Manitoba, Winnipeg, Canada, email: `jfpeters@ee.umanitoba.ca`

Gerald Schaefer
School of Engineering and Applied Science, Aston University Birmingham, U.K., email: `g.schaefer@aston.ac.uk`

and can be employed for tasks such as data dependency analysis, feature identification, dimensionality reduction, and pattern classification.

Based on rough set theory it is possible to construct a set of simple if-then rules from information tables. Often, these rules can reveal previously undiscovered patterns in sample data. Rough set methods can also be used to classify unknown data based on already gained knowledge. Unlike many other techniques, rough set analysis requires no external parameters and uses only the information present in the input data. Rough set theory can be utilised to determine whether sufficient data for a task is available respectively to extract a minimal sufficient set of features for classification which in turn effectively performs feature space dimensionality reduction.

Although, compared to other methods, a relatively recent technique, these characteristics have prompted various rough set approaches in the general domain of medical informatics. In the following we will therefore, after giving a brief introduction to basic rough set concepts, provide an overview of the use of rough sets in this area. In particular, we will show how rough sets have been used for medical image segmentation, classification, for mining medical data, and in medical decision support systems.

2 Rough Set Theory

Rough set theory [11, 14] is a fairly recent intelligent technique for managing uncertainty that is used for the discovery of data dependencies, to evaluate the importance of attributes, to discover patterns in data, to reduce redundancies, and to recognise and classify objects. Moreover, it is being used for the extraction of rules from databases where one advantage is the creation of readable if-then rules. Such rules have the potential to reveal previously undiscovered patterns in the data; furthermore, it also collectively functions as a classifier for unseen samples. Unlike other computational intelligence techniques, rough set analysis requires no external parameters and uses only the information presented in the given data. One of the useful features of rough set theory is that it can tell whether the data is complete or not based on the data itself. If the data is incomplete, it will suggest that more information about the objects is required. On the other hand, if the data is complete, rough sets are able to determine whether there are any redundancies and find the minimum data needed for classification. This property of rough sets is very important for applications where domain knowledge is very limited or data collection is expensive because it makes sure the data collected is just sufficient to build a good classification model without sacrificing accuracy [11, 14].

In rough set theory, sample objects of interest are usually represented by a table called an information table. Rows of an information table correspond to objects and columns correspond to object features. For a given set B of functions representing object features and a set of sample objects X, an indiscernibility relation $\sim_B$ is a set of pairs $(x, x') \in X \times X$ such that $f(x) = f(x')$ for all $f \in B$. The relation $\sim_B$ defines a quotient set $X / \sim_B$, i.e., a set of all classes in the partition of X defined by $\sim_B$. Rough

set theory identifies three approximation regions defined relative to $X / \sim_B$, namely, lower approximation, upper approximation and boundary. The lower approximation of a set X contains all classes that are subsets of X, the upper approximation contains all classes with non-empty intersections with X, and the boundary is the set difference between the upper and lower approximations.

Rough image processing can be defined as the collection of approaches and techniques that understand, represent and process images, their segments and features as rough sets [21]. In images boundaries between object regions are often ill-defined [10]. This uncertainty can be handled by describing the different objects as rough sets with upper (or outer) and lower (or inner) approximations.

3 Rough Sets in Medical Image Segmentation

One of the most important tasks in medical imaging is segmentation as it is often a pre-cursor to subsequent analysis, whether manual or automated. The basic idea behind segmentation-based rough sets is that while some cases may be clearly labelled as being in a set X (called positive region in rough sets theory), and some cases may be clearly labelled as not being in X (called negative region), limited information prevents us from labelling all possible cases clearly. The remaining cases cannot be distinguished and lie in what is known as the boundary region. Kobashi *et al.* [7] introduced rough sets to treat nominal data based on concepts of categorisation and approximation for medical image segmentation. The proposed clustering method extracts features of each pixel by using thresholding and labelling algorithms. Thus, the features are given by nominal data. The ability of the proposed method was evaluated by applying it to human brain MRI images. Peters *et al.* [12] presented a new form of indiscernibility relation based on k-means clustering of pixel values. The end result is a partitioning of a set of pixel values into bins that represent equivalence classes. The proposed approach allows to introduce a form of upper and lower approximation specialised relative to sets of pixel values.

An improved clustering algorithm based on rough sets and entropy theory was presented by Chena and Wang [1]. The method avoids the need to pre-specify the number of clusters which is a common problem in clustering based segmentation approaches. Clustering can be performed in both numerical and nominal feature spaces with a similarity introduced to replace the distance index. At the same time, rough sets are used to enhance the algorithm with the capability to deal with vagueness and uncertainty in data analysis. Shannon's entropy was used to refine the clustering results by assigning relative weights to the set of features according to the mutual entropy values. A novel measure of clustering quality was also presented to evaluate the clusters. The experimental results confirm that both efficiency and clustering quality of this algorithm are improved.

An interesting strategy for colour image segmentation using rough sets has been presented by Mohabey *et al.* [9]. They introduced a concept of encrustation of the histogram, called histon, for the visualisation of multi-dimensional colour informa-

tion in an integrated fashion and study its applicability in boundary region analysis. The histon correlates with the upper approximation of a set such that all elements belonging to this set are classified as possibly belonging to the same segment or segments showing similar colour value. The proposed encrustation provides a direct means of separating a pool of inhomogeneous regions into its components. This approach can then be extended to build a hybrid rough set theoretic approximations with fuzzy c-means based colour image segmentation. The technique extracts colour information regarding the number of segments and the segment centers of the image through rough set theoretic approximations which then serve as the input to a fuzzy c-means algorithm.

Widz *et al.* [20] introduced an automated multi-spectral MRI segmentation technique based on approximate reducts derived from the theory of rough sets. They utilised T1, T2 and PD MRI images from a simulated brain database as a gold standard to train and test their segmentation algorithm. The results suggest that approximate reducts, used alone or in combination with other classification methods, may provide a novel and efficient approach to the segmentation of volumetric MRI data sets. Segmentation accuracy reaches 96% for the highest resolution images and 89% for the noisiest image volume. They tested the resultant classifier on real clinical data, which yielded an accuracy of approximately 84%.

4 Rough Sets in Medical Classification

The computation of the core and reducts from a rough set decision table is a way of selecting relevant features [15]. It is a global method in the sense that the resultant reducts represent the minimal sets of features which are necessary to maintain the same classification power given by the original and complete set of features. A more direct manner for selecting relevant features is to assign a measure of relevance to each feature and choose the features with higher values. Based on the reduct system, we can generate the list of rules that will be used for building the classifier model for the new objects. Reduct is an important concept in rough set theory and data reduction is a main application of rough set theory in pattern recognition and data mining.

Wojcik [21] approached the nature of a feature recognition process through the description of image features in terms of rough sets. Since the basic condition for representing images must be satisfied by any recognition result, elementary features are defined as equivalence classes of possible occurrences of specific fragments existing in images. The names of the equivalence classes (defined through specific numbers of objects and numbers of background parts covered by a window) constitute the best lower approximation of the window contents (i.e., names of recognised features). The best upper approximation is formed by the best lower approximation, its features, and parameters, all referenced to the object fragments located within the window. The rough approximation of shapes is robust with respect to accidental changes in the width of contours and lines and to small discontinuities and, in gen-

eral, to possible positions or changes in shape of the same feature. Rough sets are also used for noiseless image quantisation.

Swiniarski and Skowron [16] presented applications of rough set methods for feature selection in pattern recognition. They emphasise the role of basic constructs of rough set approaches in feature selection, namely reducts and their approximations, including dynamic reducts. Their algorithm for feature selection is based on the application of a rough set method to the result of principal component analysis (PCA) used for feature projection and reduction. In their study, mammogram images were evaluated for recognition experiments. The database contains three types of images: normal, benign, and malignant. For each abnormal image the co-ordinates of centre of abnormality and proximate radius (in pixels) of a circle enclosing the abnormality, have been given. For classification the centre locations and radii apply to clusters rather than to the individual classifications. From the original mammograms, 64 x 64 pixel sub-images were extracted around the center of abnormality (or at the average co-ordinate for normal cases). They concluded that the rough set methods have shown ability to significantly reduce the pattern dimensionality and have proven to be viable image mining techniques as a front end of neural network classifiers.

Cyran and Mrzek [2] showed how rough sets can be applied to improve the classification ability of a hybrid pattern recognition system. Their system consists of a feature extractor based on a computer-generated hologram (CGH) where the extracted features are shift, rotation, and scale invariant. An original method of optimising the feature extraction abilities of a CGH was introduced which uses rough set concepts to measure the amount of essential information contained in the feature vector. This measure is used to define an objective function in the optimisation process. Since rough set based factors are not differentiable, they use a nongradient approach for a search in the space of possible solutions. Finally, rough sets are used to determine decision rules for the classification of feature vectors.

5 Rough Sets in Medical Data Mining

With increasing sizes of the amount of data stored in medical databases, efficient and effective techniques for medical data mining are highly sought after. Applications of rough sets in this domain include inducing propositional rules from databases using rough sets prior to using these rules in an expert system. Tsumoto [18] presented a knowledge discovery system based on rough sets and feature-oriented generalisation and its application to medicine. Diagnostic rules and information on features are extracted from clinical databases on diseases of congenital anomaly. Experimental results showed that the proposed method extracts expert knowledge correctly and also discovers that symptoms observed in six positions (eyes, noses, ears, lips, fingers, and feet) play important roles in differential diagnosis.

Hassanien *el al.* [4] presented a rough set approach to feature reduction and generation of classification rules from a set of medical datasets. They introduced a rough

set reduction technique to find all reducts of the data that contain the minimal subset of features associated with a class label for classification. To evaluate the validity of the rules based on the approximation quality of the features, a statistical test to evaluate the significance of the rules was introduced. A set of data samples of patients with suspected breast cancer were used and evaluated. The rough set classification accuracy was shown to compare favourably with the well-known ID3 classifier algorithm.

Huang and Zhang [5] presented a new application of rough sets to ECG recognition. First, the recognition rules for characteristic points in ECG are reduced using rough set theory. Then the reduced rules are used as restriction conditions of an eigenvalue determination arithmetic to recognise characteristic points in ECG. Several aspects of correlative arithmetic such as sizer method, difference method and how to choose difference parameters are discussed. They also adopted MIT-BIH data to verify R wave recognition and it is shown that the resulting detection rate is higher than those of conventional recognition methods.

Recently, Independent Component Analysis (ICA) [6] has gained popularity as an effective method for discovering statistically independent variables (sources) for blind source separation, as well as for feature extraction. Swiniarski *et al.* [17] studied several hybrid methods for feature extraction/reduction, feature selection, and classifier design for breast cancer recognition in mammograms. The methods included independent component analysis, principal component analysis (PCA) and rough set theory. Three classifiers were designed and tested: a rough sets rule-based classifier, an error back propagation neural network, and a Learning Vector Quantization neural network. Based on a comparative study on two different data sets of mammograms, rough sets rule-based classifier performed with a significantly better level of accuracy than the other classifiers. Therefore, the use of ICA or PCA as a feature extraction technique in combination with rough sets for feature selection and rule-based classification offers an improved solution for mammogram recognition in the detection of breast cancer.

6 Rough Sets in Medical Decision Support Systems

The medical diagnosis process can be interpreted as a decision-making process, during which the physician induces the diagnosis of a new and unknown case from an available set of clinical data and from clinical experience. This process can be computerised in order to present medical diagnostic procedures in a rational, objective, accurate and fast way. In fact, during the last two or three decades, diagnostic decision support systems have become a well-established component of medical technology.

Podraza *et. al* [13] presented an idea of complex data analysis and decision support system for medical staff based on rough set theory. The main aim of their system is to provide an easy to use, commonly available tool for efficiently diagnosing diseases, suggesting possible further treatment and deriving unknown dependencies

between different data coming from various patient's examinations. A blueprint of a possible architecture of such a system is presented including some example algorithms and suggested solutions, which may be applied during implementation. The unique feature of the system relies on removing some data through rough set decisions to enhance the quality of the generated rules. Usually such data is discarded, because it does not contribute to the knowledge acquisition task or even hinder it. In their approach, improper data (excluded from the data used for drawing conclusions) is carefully taken into considerations. This methodology can be very important in medical applications as a case not fitting to the general classification cannot be neglected, but should be examined with special care.

Mitra *et al.* [8] implemented a rule-based rough-set decision system for the development of a disease inference engine for ECG classification. ECG signals may be corrupted by various types of noise. Therefore, at first, the extracted signals are undergoing a noise removal stage. A QRS detector is also developed for the detection of R-R interval of ECG waves. After the detection of this R-R interval, the P and T waves are detected based on a syntactic approach. Isoelectric-level detection and base-line correction are also implemented for accurate computation of different features of P, QRS, and T waves. A knowledge base is developed from medical literature and feedback of reputed cardiologists regarding ECG interpretation and essential time-domain features of the ECG signal. Finally, a rule-based rough-set decision system is generated for the development of an inference engine for disease identification from these time-domain features.

Wakulicz-Deja and Paszek [19] implemented an application of rough set theory to decision making for diagnosing mitochondrial encephalomyopathies in children. The resulting decision support system maximally limits the indications for invasive diagnostic methods (puncture, muscle and/or nerve specimens). Moreover, it shortens the time necessary for making diagnosis. The system has been developed on the basis of data obtained from the Clinic Department of Pediatrics of the Silesian Academy of Medicine.

7 Conclusions

In this paper, we have provided a brief overview of rough sets and their use in various medical tasks. Although rough sets represent a relatively recent approach, a number of effective applications have demonstrated their potential and it is only to be expected that research will continue to improve upon and extend these techniques. Due to space constraints we were only able to highlight some of the work on medical imaging and medical decision making. A more comprehensive review of the literature on rough sets in medical imaging can be found in [3].

References

1. C-B. Chena and L-Y. Wang. Rough set-based clustering with refinement using shannon's entropy theory. *Computers and Mathematics with Applications*, 52(10–11):1563–1576, 2006.
2. K.A. Cyran and A Mrzek. Rough sets in hybrid methods for pattern recognition. *International Journal of Intelligent Systems*, 16(2):149–168, 2001.
3. A.E. Hassanien, A. Abraham, J.F. Peters, and G. Schaefer. Overview of rough-hybrid approaches in image processing. In *IEEE Conference on Fuzzy Systems*, pages 2135–2142, 2008.
4. A.E. Hassanien, J.M. Ali, and Hajime N. Detection of spiculated masses in mammograms based on fuzzy image processing. In *7th Int. Conference on Artificial Intelligence and Soft Computing*, volume 3070 of *Springer Lecture Notes in Artificial Intelligence*, pages 1002–1007, 2004.
5. X-M. Huang and Y-H. Zhang. A new application of rough set to ECG recognition. In *Int. Conference on Machine Learning and Cybernetics*, volume 3, pages 1729–1734, 2003.
6. A. Hyvärinen and E. Oja. Independent component analysis: A tutorial. Technical report, Laboratory of Computer and Information Science, Helsinki University of Technology, 1999.
7. S. Kobashi, K. Kondo, and Y. Hata. Rough sets based medical image segmentation with connectedness. In *5th Int. Forum on Multimedia and Image Processing*, pages 197–202, 2004.
8. S. Mitra, M. Mitra, and B.B. Chaudhuri. A rough-set-based inference engine for ECG classification. *IEEE Trans. on Instrumentation and Measurement*, 55(6):2198–2206, 2006.
9. A. Mohabey and A.K. Ray. Fusion of rough set theoretic approximations and FCM for color image segmentation. In *IEEE Int. Conference on Systems, Man, and Cybernetics*, volume 2, pages 1529–1534, 2000.
10. S.K. Pal, B. U. Pal, and P. Mitra. Granular computing, rough entropy and object extraction. *Pattern Recognition Letters*, 26(16):2509–2517, 2005.
11. Z. Pawlak. *Rough Sets. Theoretical Aspects of Reasoning About Data*. Kluwer, The Netherlands, 1991.
12. J.F. Peters and M. Borkowski. K-means indiscernibility relation over pixels. In *Int. Conference on Rough Sets and Current Trends in Computing*, pages 580–585, 2004.
13. R. Podraza, A. Dominik, and M. Walkiewicz. Decision support system for medical applications. In *Applied Simulation and Modelling*, 2003.
14. L. Polkowski. *Rough Sets. Mathematical Foundations*. Physica-Verlag, Heidelberg, 2003.
15. D. Ślęzak. Various approaches to reasoning with frequency-based decision reducts: a survey. In L. Polkowski, S. Tsumoto, and T.Y. Lin, editors, *Rough Sets in Soft Computing and Knowledge Discovery: New Developments*. Physica Verlag, 2000.
16. R. Swiniarski and A. Skowron. Rough set methods in feature selection and recognition. *Pattern Recognition Letters*, 24:833–849, 2003.
17. R.W. Swiniarski, H.J. Lim, Y.H. Shin, and A. Skowron. Independent component analysis, princpal component analysis and rough sets in hybrid mammogram classification. In *Int, Conference on Image Processing, Computer Vision, and Pattern Recognition*, pages 640–, 2006.
18. S. Tsumoto. Mining diagnostic rules from clinical databases using rough sets and medical diagnostic model. *Information Sciences: an International Journal*, 162(2):65–80, 2004.
19. A. Wakulicz-Deja and P. Paszek. Applying rough set theory to multi stage medical diagnosing. *Fundamenta Informaticae*, 54(4):387–408, 2003.
20. S. Widz, K. Revett, and Ślęzak D. Application of rough set based dynamic parameter optimization to mri segmentation. In *23rd Int. Conference of the North American Fuzzy Information Processing Society*, pages 440–445, 2004.
21. Z. Wojcik. Rough approximation of shapes in pattern recognition. *Computer Vision, Graphics, and Image Processing*, 40:228–249, 1987.