Medical Image Diagnosis Based on Rough Sets Theory

A.L. Ion and S. Udristoiu

University of Craiova/Software Engineering Department, Craiova, Romania

Abstract— **This paper proposes the utilization of rough set theory for modeling the medical images to help physicians in diagnosing. The rough set theory is a powerful approach that permits the searching for patterns in medical images using the minimal length principles. Searching for models with small size is performed by means of many different kinds of reducts that generate the decision rules capable for identifying the medical diagnosis.**

Keywords— **medical image diagnosis, rough sets, image colour, image texture, image shape.**

I. INTRODUCTION

The rough set theory was discovered by Zdzislaw Pawlak [1], [2] and is a powerful mathematical tool for modeling the imperfect and incomplete knowledge, which is an issue debated for a long period of time, by logicians, mathematicians, philosophers and computer scientists [3].

Among methods proposed for modeling the imperfect knowledge [4], the rough set theory is an interesting attempt to solve this problem. This theory is based on an assumption that objects are recognized by partial information about them and some objects can be indiscernible. From this fact it follows that some sets cannot be exactly described by the available information about objects [1], [3].

The methods based on rough set theory have an important utilization in many real life applications. Among the rough set based software systems are ROSETTA [5], RSES [6], and LERS [7], which have been applied to knowledge discover problems.

In this paper we use the rough set approach to discover patterns from medical images for establishing their diagnosis. In the medical domain, a lot of researches were developed to investigate automated techniques for extracting the low–level features that could generate semantic descriptions of the medical image content. Among these techniques are the methods based on machine learning that manually annotate the test image datasets. In the medical domain, algorithms that recognize specific organs with different structures of the medical images are studied in [9]. FIRE [8] application and IRMA [10] use with good results the sub-symbolic processing of images. Though, the actual methodologies of medical image analysis are not generically sufficient for interpreting different diseases. Their major problems are:

1. The description of semantic concepts and the problem understanding-the relationships between the low-level features and semantic concepts are unclear in the actual developed methods. So detailed tests and analysis have to be realized to ensure which combinations of low-level features capture the best the semantic concepts.
2. The generality of the application-in some of the previous researches, only certain semantic concepts could be learned, or the rule were generated of a fixed set of visual features.

The medical applications with automatic diagnosis capacity imply unique challenges, but at the same time new opportunities. Unless we are not physicians, it is a lot harder to understand a medical image than an image taken from nature. On the other hand, there are a lot of formal representations of the medical knowledge that could be exploited to realize the automation of the medical diagnosis in any medical domain.

II. IMAGE REPRESENTATION AND DISCRETIZATIONS

The diagnosis of medical images is directly related to the visual features (colour, texture, shape, position, dimension, etc.), because these attributes capture the information about the semantic meaning. A set of dominant colour regions is obtained from each image by segmentation after the colour characteristic [12]. The HSV colour space quantized to 166 colours is used to represent the colour information [12], [15]. The extraction of colour regions is realized by the colour set back projection algorithm [11]. The specialist selects the representative colour set *C* for the sick regions of medical images from the digestive domain. The algorithm detects the regions having the colour in the colour set *C*. The results of the segmentation algorithm applied to an image diagnosed with gastric ulcer can be visualized in Figure 1.

The visual features of a sick region are represented by 14 parameters [12]:

- The colour, which is represented in the HSV colour space quantized at 166 colours.
- The spatial coherency, which measures the spatial compactness of the pixels of the same colour.

S. Vlad and R.V. Ciupa (Eds.): MEDITECH 2011, IFMBE Proceedings 36, pp. 206–211, 2011.
www.springerlink.com

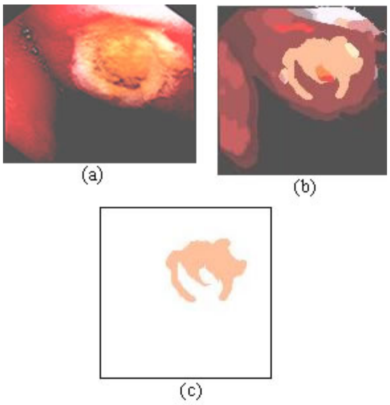

Fig. 1 Segmentation results from an image diagnosed with gastric ulcer: (a) The original image; (b) The quantized image; (c) The sick region.

- A seven-dimension vector (maximum probability, energy, entropy, contrast, cluster shade, cluster prominence, correlation), which represents the texture characteristics.
- The region dimension descriptor, which represents the number of pixels from region.
- The spatial information which is represented by the centroid coordinates of the region and by minimum bounded rectangle.
- A two-dimensional vector (eccentricity and compactness), which represents the shape feature.

The visual features of sick regions were discretized over intervals, using the concept of semantic indicators, which are visual elements: the colour (colour-light-red, etc.), spatial coherency (spatial coherency-weak, spatial coherency-medium, spatial coherency-strong), texture (energy-small, energy-medium, energy-big), dimension (dimension-small, dimension-medium, dimension-big), position (vertical-upper, vertical-center, vertical-bottom, horizontal-upper, etc.), shape (eccentricity- small, compactness-small, etc.).

The values of each semantic descriptor are mapped to a value domain, which corresponds to the mathematical descriptor [12]. At the end of the mapping process, a medical image is represented by means of the terms *figure(ListofRegions)*, where *ListofRegions* is a list of images' sick regions.

III. Modeling image diagnosis using rough sets

A. Rough Sets Foundations

Rough sets theory is an intelligent mathematical tool and it is based on the concept of approximation space [1], [2], [13].

In rough sets theory, the notion of information system determines the knowledge representation system. In this section, we recall some basic definitions from literature [1], [2], [3], [13].

Let U denote a finite non-empty set of objects (sick image regions) called the universe. Further, let A denote a finite non-empty set of attributes. Every attribute $a \in A$, there is a function $a: U \to V_a$, where V_a is the set of all possible values of a, to be called the domain of a. *A pair $IS = (U, A)$ is an information system.* Usually, the specification of an information system can be presented in tabular form. Each subset of attributes $B \subseteq A$ determines a binary *B–indiscernibility* relation *IND(B)* consisting of pairs of objects indiscernible with respect to attributes from B like in (1):

$$IND(B) = \{(x, y) \in U \times U : \forall a \in B,\ a(x) = a(y)\} \quad (1)$$

IND(B) is an equivalence relation and determines a partition of U, which is denoted by *U/IND(B)*. The set of objects indiscernible with an object $x \in U$ with respect to the attribute set, B, is denoted by $I_B(x)$ and is called *B-indiscernibility* class.

Thus,

$$I_B(x) = \{y \in U : (x, y) \in IND(B)\} \quad (2)$$

$$U / IND(B) = \{I_B(x) : x \in U\} \quad (3)$$

Table 1 Medical Information System

U	Colour	Texture-entropy	Diagnosis
R_1	light-red	Small	gastric-ulcer
R_2	light-red	Small	gastric-ulcer
R_3	light-red	Small	gastric-ulcer
R_4	light-red	Big	gastric-ulcer
R_5	light-yellow	Big	gastric-ulcer
R_6	light-yellow	Medium	duodenal-ulcer
R_7	light-yellow	Medium	duodenal-ulcer
R_8	medium-yellow	Small	duodenal-ulcer
R_9	medium-yellow	Small	duodenal-ulcer
R_{10}	dark-yellow	Small	duodenal-ulcer
R_{11}	dark-yellow	Small	duodenal-ulcer

Table 2 Partitions Defined by Indiscernibility Relations

IND(B)	Partitions U/IND(B)
IND({Colour})	$\{R_1, R_2, R_3, R_4\}$, $\{R_5, R_6, R_7\}$, $\{R_8, R_9\}$, $\{R_{10}, R_{11}\}$
IND({Colour,Texture-entropy})	$\{R_1, R_2, R_3\}$, $\{R_4\}$, $\{R_5\}$, $\{R_6, R_7\}$, $\{R_8, R_9\}$, $\{R_{10}, R_{11}\}$

It is said that a pair $AS_B = (U, IND(B))$ is an approximation space for the information system $IS=(U, A)$, where $B \subseteq A$.

The information system from Table 1 represents the sick regions of images from different diagnoses represented in terms of semantic indicators values, as described in Section II. For simplicity we consider only two semantic indicators as attributes, namely the colour and texture-entropy.

So, our information systems is IS = (U, B), where U = $\{R_1, R_2, R_3, R_4, R_5, R_6, R_7, R_8, R_9, R_{10}, R_{11}\}$ *and B = {colour, texture-entropy}.* Some examples of partitions defined by indiscernibility relations for the information system in Table 1 are given in Table 2.

In rough sets theory, the approximations of sets are introduced to deal with inconsistency. A rough set approximates traditional sets using a pair of sets named the lower and upper approximations of the set.

Let $W = \{w_1, \ldots, w_n\}$ be the elements of the approximation space $AS_B=(U, IND(B))$. We want to represent *X*, a subset of U, using attribute subset *B*. In general, *X* cannot be expressed exactly, because the set may include and exclude objects which are indistinguishable on the basis of attributes *B,* so we could define *X* using the lower and upper approximation.

The *B-lower approximation X,* $\underline{B}X$, is the union of all equivalence classes in *IND(B)* which are contained by the target set *X*. The lower approximation of *X* is called the positive region of *X* and is noted *POS(X)*.

$$\underline{B}X = \bigcup \{ w_i \mid w_i \subseteq X \} \quad (4)$$

The *B-upper approximation* $\overline{B}X$ is the union of all equivalence classes in *IND(B)* which have non-empty intersection with the target set *X*.

$$\overline{B}X = \bigcup \{ w_i \mid w_i \cap X \neq \varnothing \} \quad (5)$$

Example: Let $X = \{R_1, R_2, R_3, R_4, R_5, R_6, R_7, R_8\}$ be the subset of *U* that we wish to be represented by the attributes set *B={colour, texture-entropy}*. We can approximate *X*, by computing its *B-lower approximation,* $\underline{B}X$ and *B-upper approximation,* $\overline{B}X$.

So, $\underline{B}X = \{\{R_1, R_2, R_3\}, \{R_4\}, \{R_5\}, \{R_6, R_7\}\}$ and $\overline{B}X = \{\{R_1, R_2, R_3\}, \{R_4\}, \{R_5\}, \{R_6, R_7\}, \{R_8, R_9\}\}$.

The tuple ($\underline{B}X$, $\overline{B}X$) composed of the lower and upper approximation is called a rough set; thus, a rough set is composed of two crisp sets, one representing a *lower boundary* of the target set *X*, and the other representing an *upper boundary* of the target set *X*.

The accuracy of a rough set is defined as: cardinality($\underline{B}X$)/cardinality($\overline{B}X$). If the accuracy is equal to 1, then the approximation is perfect.

B. Dispensable Features, Reducts and Core

An important notion used in rough set theory is the decision table. Pawlak [1], [2] gives also a formal definition of a decision table: an information system with distinguished conditional attributes and decision attribute is called a decision table. So, a *tuple DT = (U, C, D), is a decision table. The attributes C = {colour, texture-entropy} are called conditional attributes, instead D = {diagnosis} is called decision attribute.*

The classes *U/IND(C)* and *U/IND(D)* are called condition and decision classes, respectively.

The *C-Positive* region of *D* is given by:

$$POS_C(D) = \bigcup_{X \in IND(D)} \underline{C}X \quad (6)$$

Let $c \in C$ a feature. It is said that *c* is dispensable in the decision table *DT*, if $POS_{C-\{c\}}(D) = POS_C(D)$; otherwise the feature *c* is called indispensable in *DT*. If *c* is an indispensable feature, deleting it from *DT* makes it to be inconsistent.

A set of features *R* in *C* is called a reduct, if *DT'= (U, R, D)* is independent and $POS_R(D)=POS_C(D)$. In other words, a reduct is the minimal feature subset preserving the above condition.

The set of all features indispensable in *C* is denoted by *CORE(C)*. In other words, *CORE(C)* is the set of all reducts of *C*.

C. Producing Rules by Discernibility Matrix

We transform the decision table into discernibility matrix to compute the reducts. Let *DT = (U, C, D)* be the decision table, with $U = \{R_1, R_2, R_3, R_4, R_5, R_6, R_7, R_8, R_9, R_{10}, R_{11}\}$. By a discernibility matrix of *DT*, denoted *DM(T),* we will mean *nxn* matrix defined as:

$$m_{ij}^{a(R_i)} = \{(a \in C : a(R_i) \neq a(R_j)) \text{ and } (d(R_i) \neq d(R_j))\} \quad (7)$$

where *i, j=1,2,…,11*.

We construct the discernibility matrix, *DM(DT)* as in Table 3, where the colour and texture-entropy are denoted by *C*, respectively *T*. The items within each cell are aggregated disjunctively, and the individual cells are then aggregated conjunctively.

To compute the reducts of the discernibility matrix we use the following theorems that demonstrate equivalence between reducts and prime implicants of suitable Boolean functions [3], [13].

For every object $R_i \in U$, the following Boolean function is defined:

$$g_{R_i}(Colour, Texture) = \bigwedge_{R_j \in U} \left(\bigvee_{a \in m_{ij}} a \right) \quad (8)$$

The following conditions are equivalent [3]:

1. $\{a_{i1}, \ldots, a_{in}\}$ *is a reduct for the object* R_i, *i=1..n.*
2. $a_{i1} \wedge a_{i2} \wedge .. \wedge a_{in}$ *is a prime implicant of the Boolean function* g_{R_i}.

Table 3 Discernibility Matrix

	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}
R_1	-	-	-	-	-	$C^{light\text{-}red}$ T^{small}	$C^{light\text{-}red}$ T^{small}	$C^{light\text{-}red}$	$C^{light\text{-}red}$	$C^{light\text{-}red}$	$C^{light\text{-}red}$
R_2	-	-	-	-	-	$C^{light\text{-}red}$ T^{small}	$C^{light\text{-}red}$ T^{small}	$C^{light\text{-}red}$	$C^{light\text{-}red}$	$C^{light\text{-}red}$	$C^{light\text{-}red}$
R_3	-	-	-	-	-	$C^{light\text{-}red}$ T^{small}	$C^{light\text{-}red}$ T^{small}	$C^{light\text{-}red}$	$C^{light\text{-}red}$	$C^{light\text{-}red}$	$C^{light\text{-}red}$
R_4	-	-	-	-	-	$C^{light\text{-}red}$ T^{big}	$C^{light\text{-}red}$ T^{big}	$C^{light\text{-}red}$ T^{big}	$C^{light\text{-}red}$ T^{big}	$C^{light\text{-}red}$ T^{big}	$C^{light\text{-}red}$ T^{big}
R_5	-	-	-	-	-	T^{big}	T^{big}	$C^{light\text{-}yellow}$ T^{big}	$C^{light\text{-}yellow}$ T^{big}	$C^{light\text{-}yellow}$ T^{big}	$C^{light\text{-}yellow}$ T^{big}
R_6	$C^{light\text{-}yellow}$ T^{medium}	$C^{light\text{-}yellow}$ T^{medium}	$C^{light\text{-}yellow}$ T^{medium}	$C^{light\text{-}yellow}$ T^{medium}	T^{medium}	-	-	-	-	-	-
R_7	$C^{light\text{-}yellow}$ T^{medium}	$C^{light\text{-}yellow}$ T^{medium}	$C^{light\text{-}yellow}$ T^{medium}	$C^{light\text{-}yellow}$ T^{medium}	T^{medium}	-	-	-	-	-	-
R_8	$C^{medium\text{-}yellow}$	$C^{medium\text{-}yellow}$	$C^{medium\text{-}yellow}$	$C^{medium\text{-}yellow}$ T^{small}	$C^{medium\text{-}yellow}$ T^{small}	-	-	-	-	-	-
R_9	$C^{medium\text{-}yellow}$	$C^{medium\text{-}yellow}$	$C^{medium\text{-}yellow}$	$C^{medium\text{-}yellow}$ T^{small}	$C^{medium\text{-}yellow}$ T^{small}	-	-	-	-	-	-
R_{10}	$C^{dark\text{-}yellow}$	$C^{dark\text{-}yellow}$	$C^{dark\text{-}yellow}$	$C^{dark\text{-}yellow}$ T^{small}	$C^{dark\text{-}yellow}$ T^{small}	-	-	-	-	-	-
R_{11}	$C^{dark\text{-}yellow}$	$C^{dark\text{-}yellow}$	$C^{dark\text{-}yellow}$	$C^{dark\text{-}yellow}$ T^{small}	$C^{dark\text{-}yellow}$ T^{small}	-	-	-	-	-	-

Next, from each decision matrix we form a set of Boolean expressions, one expression for each row of the matrix.

For the gastric ulcer we obtain the following rules based on the table reducts:

1. $(C^{light\text{-}red} \vee T^{small}) \wedge (C^{light\text{-}red})$
2. $(C^{light\text{-}red} \vee T^{small}) \wedge (C^{light\text{-}red})$
3. $(C^{light\text{-}red} \vee T^{small}) \wedge (C^{light\text{-}red})$
4. $(C^{light\text{-}red} \vee T^{big}) \wedge ((C^{light\text{-}red}))$
5. $(T^{big}) \wedge (C^{light\text{-}yellow} \vee T^{big})$

For the duodenal ulcer we obtain the following rules based on the table reducts:

1. $(C^{light\text{-}yellow} \vee T^{medium}) \wedge T^{medium}$
2. $(C^{medium\text{-}yellow}) \wedge (C^{medium\text{-}yellow} \vee T^{small})$
3. $(C^{dark\text{-}yellow}) \wedge (C^{dark\text{-}yellow} \vee T^{small})$

On Boolean expression the absorption Boolean algebra rule is applied. The absorption law is an identity linking a pair of binary operations.

For example: $a \vee (a \wedge b) = a \wedge (a \vee b) = a$

By applying the absorption rule on the prime implicants, the following rules are generated:

1. Rule 1: (Colour = light-red) → gastric ulcer;
2. Rule 2: (Texture-entropy = big) → gastric ulcer;
3. Rule 3: (Texture-entropy = medium) → duodenal ulcer;
4. Rule 4: (Colour = dark-yellow) → duodenal ulcer.

D. *Evaluation of Decision Rules*

Decision rules can be evaluated along at least two dimensions: performance (prediction) and explanatory features (description). The performance estimates how well the rules classify new images. The explanatory feature estimates how interpretable the rules are [3].

Let be our decision table *DT = (U, C, D)*. We use the set-theoretical interpretation of rules. It relates a rule to data sets from which the rule is discovered [3]. Using the cardinalities of sets, we obtain the 2×2 contingency table representing the quantitative information about the rule *if features then diagnosis*.

In table 4 the number of images that have a certain feature set and a certain diagnosis is computing. Using the elements of the contingency table, we may define the support (s) and accuracy (a) of a decision rule by:

$$s(rule) = cardinality(featureSet \cap diagnosisSet) \tag{9}$$

$$a(rule) = \frac{cardinality(featureSet \cap diagnosisSet)}{cardinality(featureSet)} \tag{10}$$

where the set $featureSet \cap diagnosisSet$ is composed from image regions which have a certain *featureSet* and a certain *diagnosis*. In term of set theory, the accuracy is the degree in which the set of features rule is included in the set of diagnosis rule.

The coverage(c) of a rule is defined by:

$$c(rule) = \frac{cardinality(featureSet \cap diagnosisSet)}{cardinality(diagnosisSet)} \tag{11}$$

The coverage of a rule is the degree in which the set of diagnosis rule is included in the features set of rule. For the generated Rule 1, the contingency table Table 5 is obtained. For the Rule 1, the support is 4, accuracy is 4/4 and coverage is 4/5. Ryszard et al [14] suggests that high accuracy and coverage are requirements of decision rules.

IV. DECISION RULE EXTRACTION USING ROUGH SETS MODELS AND EXPERIMENTS

In this paper we present the application of rough set to discover the medical diagnosis of images from digestive apparatus. To establish the medical diagnosis the following tasks are carried out:

- selection of the most relevant condition attributes in our case 14 image visual semantic indicators,
- application of rough set based on reduced data,
- discovery of decision rules characterizing the dependency between values of condition attributes and decision attribute.

A rule has the form:

if (colour is red and texture-entropy is small) then the diagnosis is ulcer.

Decision rules are generated from reducts. So in order to compute decision rules, reducts have to be computed first.

This method finds all reducts by computing prime implicants of a Boolean function, as described in Section III.

The rule generation algorithm can be resumed as:

- construct the decision table and discernibility matrix,
- obtain the discernibility function and the prime implicants,
- apply the Boolean algebra rules,
- compute the reducts,
- produce the rules using the reducts.

The image classification algorithm can be resumed as:

- collect all the decision rules in a classifier,
- compute for each rule the support, accuracy and coverage,
- eliminate the rules with the support less than the minimum defined support,
- order the rules by accuracy, than by coverage,
- if an image matches more rules select the first one: an image matches a rule, if all the semantic indicators, which appear in the body of the rule, are included in the characteristics of the image regions.

The image collections used in our experiments were taken from free repositories on the Internet [16], [17]. Two image databases are used for learning and diagnosing process. The database used to learn the correlations between images and digestive diagnoses, contains 200 images. The learning database is categorized into the following diagnoses: duodenal ulcer, gastric ulcer, gastric cancer, esophagitis, and rectocolitis. The system learns each concept by submitting about 20 images per diagnosis. For example, we analyze the performance of the proposed method for colon cancer diagnosis. The rule generation algorithm produces 12 semantic rules that recognize this diagnosis. The test database contains 450 images, from which 67 are relevant for duodenal ulcer diagnosis. A part of images used for learning the diagnoses can be analyzed in Figure 2.

Table 4 General Contingency Table Representing the Quantitative Information about the Rule

	Diagnosis	not(Diagnosis)	
Features	cardinality(features and diagnosis)	cardinality(features and not(diagnosis))	card(features)
not(Features)	cardinality((not)features and diagnosis)	cardinality((not)features and (not)diagnosis)	card(not(features))
	cardinality(diagnosis)	card(not(diagnosis))	card(U)

Table 5 Contingency Table Representing the Quantitative Information about the Rule 1

	diagnosis = gastric ulcer	not(diagnosis = gastric ulcer)	
colour = light-red	card*(colour = light-red and diagnosis = gastric ulcer)* = 4	card*(colour=light-red and not(diagnosis=gastric ulcer))* = 0	card*(colour=light-red)*= 4
not(colour= light-red)	card*(not(colour=light-red) and diagnosis= gastric ulcer)*=1	card*(not(colour=light-red) and not(diagnosis= gastric ulcer))*=6	card*(not(colour=light-red)*)=7
	card(diagnosis = gastric ulcer) =5	*card(not(diagnosis = gastric ulcer))*= 6	card(U)=11

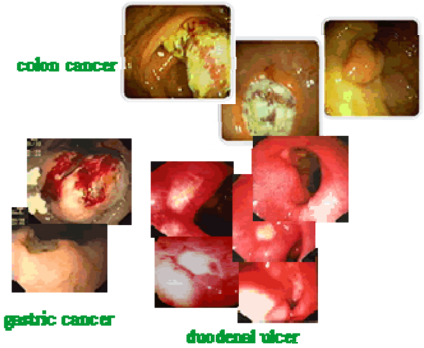

Fig. 2 Medical diagnosed images

After classification, we counted: the number of true positives (images correctly diagnosed with the colon cancer diagnosis) and we found 56 images; the number of false positives (images incorrectly diagnosed with the colon cancer diagnosis) and we found 9 images; the number of true negatives (images correctly diagnosed with a different diagnosis) and we found 377 images; the number of false negatives (images incorrectly diagnosed with a different diagnosis) and we found 8 images. The accuracy, which measures the proportion of true results, is 96.2%. The specificity, which measures the capability of colon cancer rules not to miss the colon cancer images, and not to diagnose images with a different diagnosis, is 97.6%. In our case, this set of rules is very specific.

For the other diagnoses, the counted results are presented in Table 6.

Table 6 Results recorded for different diagnoses

Diagnosis	Accuracy(%)	Specificity(%)
Duodenal Ulcer	96.3	95
Gastric Ulcer	96.7	95.1
Gastric Cancer	95.9	93
Rectocolitis	96.3	95.2

V. Conclusion

Methods proposed and developed in this study could assist physicians by doing automatic diagnose based on visual content of medical images. An important improvement of this paper is in the generation of rules with very high specificity using the rough set theory. The language used for rules representation is Prolog. The advantages of using Prolog are its flexibility and simplicity in representation of rules. The results of the presented method are very promising, being influenced by the complexity and number of endoscopic images.

Acknowledgment

This work was supported by the strategic grant POSDRU/89/1.5/S/61968, Project ID 61968 (2009), co-financed by the European Social Fund within the Sectorial Operational Program Human Resources Development 2007 - 2013.

References

1. Pawlak Z (1986) Rough Relations. Bulletin of the Polish Academy of Sciences, Technical Sciences 34(9-10): 587–590
2. Pawlak Z, Skowron A (1994) Rough Membership Functions. Advances in the Dempster-Shafer Theory of Evidence pp.251-271. John Wiley and Sons, New York
3. Stepaniuk J (2008) Rough Granular Computing in Knowledge Discovery and Data Mining. Springer-Verlag, Germany
4. Zadeh LA, Kacprzyk J (eds.) (1999) Computing with Words in Information/Intelligent Systems 2. Physica-Verlag, Heidelberg
5. Ohrn A, Komorowski J, Skowron A, et al. (1998) The Design and Implementation of a Knowledge Discovery Toolkit Based on Rough Sets - The Rosetta System. Polkowski, L., Skowron, A. (eds.) Rough Sets in Knowledge Discovery 1, Methodology and Applications, pp. 376–399. Physica-Verlag, Heidelberg
6. Bazan J, Szczuk M (2005) The Rough Set Exploration System. Transactions on Rough Sets III, LNCS 3400: 37–56
7. Grzymala-Busse JW (2005) LERS-A Data Mining System. Springer, US
8. Deselaers T, Keysers D, Ney H (2004) FIRE – flexible image retrieval engine: ImageCLEF 2004 evaluation. Multilingual Information Access for Text, Speech and Images 3491:688–698
9. Hong W, Georgescu B, Zhou XS, et al. (2006) Database-guided simultaneous multi-slice 3D segmentation for volumetric data, Proceedings of 9th European Conference on Computer Vision, Graz, Austria; 2006, pp. 397-409
10. Lehmann T, Güld M, Thies C, et al. (2004) Content-based image retrieval in medical applications. Methods Inf Med.43(4): 354-361
11. Smith JR, Chang SF. VisualSEEk: a fully auto-mated content-based image query system. The Fourth ACM International Multimedia Conference and Exhibition, Boston, MA, USA, 1996, pp. 87-98
12. Ion AL, Udristoiu S (2010) Image Mining for discovering medical diagnosis. Information Technology and Control 39(1): 123-129
13. Hassanien AE, Abraham A, Peters JF et al. (2008). Rough Sets in Medical Imaging: Foundations and Trends. Computational Intelligence in Medical Imaging: Techniques and Applications, pp. 47-87. CRC Press, USA
14. Ryszard A, Michalski S (1993) A Theory and Methodology of Inductive Learning. Readings in knowledge acquisition and learning: 323-348
15. Stanescu L, Burdescu D, Ion AL, et al. (2008) Imagistic Database for Medical E-learning, Proceedings of 21st IEEE International Symposium on Computer-Based Medical Systems, Jyvaskyla, Finland; 2008, pp.427-429.
16. Jackson Siegelbaum Gastroenterology at http://gicare.com/Endoscopy-Center/Endoscopy-images.aspx.
17. The Gatrolab Image Library at http://www.gastrolab.net/