ISSN 1392 – 124X INFORMATION TECHNOLOGY AND CONTROL, 2010, Vol. 39, No. 2

IMAGE MINING FOR ESTABLISHING MEDICAL DIAGNOSIS

Anca Loredana Ion, Stefan Cristinel Udristoiu

University of Craiova, Bvd. Decebal, Nr. 107
200440, Craiova, Romania
e-mail: anca_soimu@yahoo.com

Abstract. In this paper, a method based on association rules that support medical image diagnosis is developed. The proposed method has important characteristics that make it different from other computer assisted diagnosis methods: the process is completely automatic, having the possibility to define a great number of diagnoses; the method could be applied to any medical domain, because the visual features, the semantic indicators remain unchangeable, and the semantic rules are generated by learning from labelled images-examples; the selection of the visual characteristics set is based on their retrieval accuracy; the spatial information of the regions is considered, offering important medical information of the relationships between sick regions. Although we present the results achieved in endoscopic images analysis, our method can be used to analyze other types of medical images. The prototype system was applied to real datasets and the results show high accuracy.

Keywords: medical image diagnosis, image colour, image texture, image shape, image mining.

1. Introduction

Many researches were developed to support query by high-level concepts and to provide full support in bridging the 'semantic gap' between numerical image features and the richness of human semantics [9]. In the recent past, applications like FIRE [10] showed that image retrieval based only on sub-symbolic interpretation of the images works well for a number of applications. In the IRMA project [11], this technology was applied to the medical domain. A number of researches use ontology-based image retrieval to emphasize the necessity to combine sub-symbolic object recognition and abstract domain knowledge. In [12] an integration of spatial context and semantic concepts into the feature extraction and retrieval process using a relevance feedback procedure was proposed. In [13] a system that aims at applying a knowledge-based approach to interpret X-ray images of bones and to identify the fractured regions was proposed. In [14] a hybrid method which combines symbolic and sub-symbolic techniques for the annotation of brain Magnetic Resonance images was developed. In [5, 6] relevant methods using association rules for radiology image analysis were developed.

In this paper, we propose a method that employs semantic rules to support computer assisted diagnosis systems in digestive endoscopy domain. The rules reflect how physicians analyze and combine the visual features to determine the image diagnoses.

This work presents a different algorithm for finding semantic association rules than the algorithm and system that we developed in [2], [3]. The novelty is in the representation of sick regions, as well as in the algorithm used to discover the semantic association rules. The comparison between the presented algorithm and the previous ones [2], [3] will be made in a next paper.

The remainder of this paper is structured as follows. Section 2 presents the visual representation of medical images; Section 3 presents the vocabulary and syntax used to represent medical images; Section 4 presents the generation of semantic rules and details the process of medical diagnosis based on semantic rules. Finally, Section 5 discusses the experiments and summarizes the conclusions of this study.

2. Visual representation of images

The feature extraction phase is needed in order to create the transactional database to be mined. The features that were extracted are organized in a database, which is the input to the mining phase of the classification system.

The selection of the visual feature set and the image segmentation algorithm is the definitive stage for establishing the diagnosis of medical colour images [4]. 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 [8]. The HSV colour space quantized to 166 colours is used to represent the colour information. Before segmentation, the images are transformed from RGB to HSV colour space and quantized at a colour set of 166 colours [8]. The extraction of colour regions is realized by the colour set back projection algorithm [7]. 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 implementation of this algorithm is described in pseudo-code:

```
Algorithm 1. Segmentation of an image in
regions of a single colour.
Input: the selected colour set C and a
stack S.
Output: the set M of single colour
regions, and the spatial coherencies,
Coh of each region.
Method:
1.InitStack(S)
2.Visited = ∅
3.foreach pixel P do
4. if P unvisited and colour(P) in C then{
5. PUSH(P)
6. R = ∅
7. Coh = 0
8.   PointsofRegion = 0
9.     while not Empty(S) do{
10.      Crt = POP()
11.       if CrtPoint is unvisited then{
12.         Add Crt at R
13.         Visited=Visited∪{Crt}
14.         Connectivity8=0
15.           foreach neighbour N of Crt
16.         do
17.          if colour(N)=colour(Crt)
              then{
18.            Increment Connectivity8
19.            PUSH(N)
20.          }
21.         }
22.         Coh=Coh+Connectivity8
23.         Increment PointsofRegion
24.       }
25.     }
26.     Coh=Coh/PointsofRegion
27.     Add region R with Coh to M
28. }
```

For traversing the image, a stack, *S*, is used as data structure to store the unvisited pixels. First, the algorithm initializes the stack and the set *Visited* to null. Each pixel, which is unvisited and has the colour in the set, *C,* is added to the top of the stack (lines 4-5). The region, *R*, is initialized to null, the spatial coherency and the number of region pixels are initialized to zero (lines 6-7).

Next, while the stack *S* contains pixels, a pixel is extracted from it and, if the pixel was not visited before, the pixel is added to the current extracted region, *R*, then all pixel's neighbours having the same colour are pushed to the stack *S* to be visited later. The pixel's connectivity is computed using the variable *Connectivity8*, which is added to the coherency counter, *Coh*. When the stack *S* is empty, then a region has been extracted having the spatial coherency equal to the average connectivity. The extraction of a region starts with an unvisited pixel having a colour in the *C* set, meaning that there is a new region, and continues recursively to visit all new neighbours with the same colour until all of them have been visited (lines 9-28).

The results of the segmentation algorithm applied to two images diagnosed with gastric ulcer, respectively with duodenal ulcer, can be visualized in Figure 1 and Figure 2.

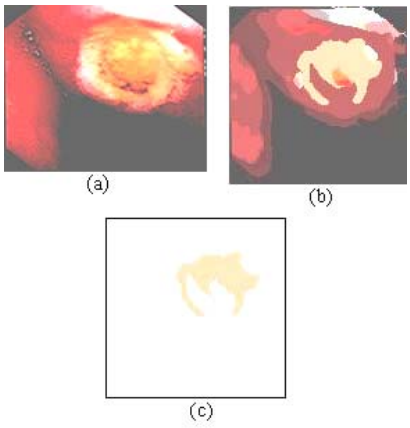

Figure 1. Segmentation results from an image diagnosed with gastric ulcer: (a) The original image; (b) The quantized image; (c) The sick region

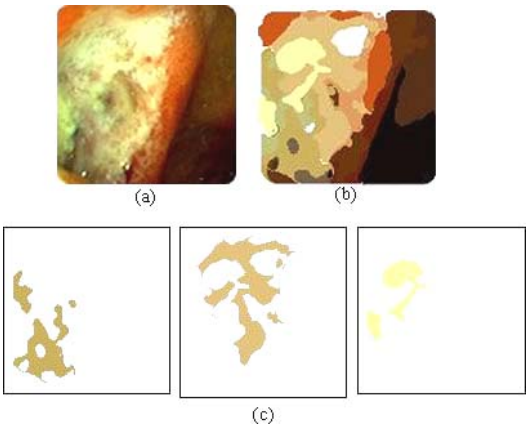

Figure 2. Segmentation results from an image diagnosed with duodenal ulcer: (a) The original image; (b) The quantized image; (c)Three sick regions

The visual features of sick region are represented by 14 parameters [8]:

- The colour, which is represented in the HSV colour space quantized at 166 colours. A region is represented by a colour index which is an integer number between 0...165.
- The spatial coherency, which measures the spatial compactness of the pixels of the same colour. It is a natural number associated to each colour 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.

3. Visual vocabulary and syntax

The visual features of sick regions were discretized over intervals, using the concept of semantic indicators. So, a simple vocabulary based on semantic indicators is used. The syntax captures the basic models about patterns and diagnosis. The proposed representation language is simple, because the syntax and vocabulary are elementary. Being visual elements, the semantic indicators are, by example, 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 syntax is represented by the model, which describes the images in terms of semantic indicators values. The values of each semantic descriptor are mapped to a value domain, which corresponds to the mathematical descriptor [8].

At the end of the mapping process, a medical image is represented in Prolog by means of the terms *figure(ListofRegions)*, where *ListofRegions* is a list of images' sick regions.

The term *region(ListofDescriptors)* is used for region representation, where the argument is a list of terms used to specify the semantic indicators. The term used to specify the semantic indicators is of the form:

descriptor(DescriptorName, DescriptorValue).

The model representation of image from Figure 1 can be observed in the following example:

```
figure([
  region([descriptor(colour,light-red),
   descriptor(dimension, medium),
   descriptor(horizontal-position,
      right),
   descriptor(vertical-position, upper),
   descriptor(shape-eccentricity,
      medium),
   descriptor(texture-probability,
       medium),
   descriptor(texture-inversedifference,
       medium),
   descriptor(texture-entropy,small),
   descriptor(texture-energy,big),
   descriptor(texture-contrast, small),
   descriptor(texture-correlation,
     small)]).
```

4. Associative classifiers

In image discovery, to label each object that appears in an image is a complex task, because it implies to find associations. Association rule mining is one of the most important tasks in Data Mining and initially, methods applied to market basket analysis, were developed. So, the generation of association rules was introduced in [15] and the algorithm AIS was proposed. In [16], the algorithm called SETM was proposed to discover association rules using relational operation. In [17], the algorithms called Apriori and AprioriTid were proposed, bringing important improvements to older methods.

In this paper, the associative classifiers are used to define rules that convert the visual features of medical images to semantic features of diagnoses.

The proposed method considers that the dataset contains N cases, described by 14 categorical parameters. These N cases have been classified into M diagnosis. Let D be the dataset (transactional database). Let I be the set of all items in D and Y be the set of diagnoses. A data case $d \in D$ contains $X \subseteq I$, a subset of items, if $X \subseteq d$. A diagnosis association rule (DAR) is an implication of the form: $X \rightarrow y$, where $y \in Y$. The rule has the confidence c, if c% of cases in D that contain X are labelled with diagnosis y. The rule has the support s, if s% of the cases in D contain X and are labelled with diagnosis y.

The main objective is to generate a set of DARs that satisfy the specified minimum support(*minsup*) and minimum confidence (*minconf*), and to annotate medical images using the DARs.

Associative classifiers are a two-stage approach to classification, in which a set of association rules between the semantic indicator values and diagnoses is first discovered and then a compact classifier is created by selecting the most important rules for classification, as in Figure 3.

4.1 Semantic rule generation

In our system, the learning of semantic rules is continuously made, because when a diagnosed image is added in the learning database, the system continues the process of rules generation.

The proposed method uses a modified version of CBA algorithm [1], for discovering the semantic rules between the images' sick regions and diagnoses.

A semantic rule is of form:

$SI \rightarrow d$,

where *SI,* the body of the rule, is composed by conjunctions of semantic indicators, while *d*, the head of the rule, is the diagnosis.

The image modelling in terms of itemsets and transactions is the following:

- the set of sick regions of the training images represent the transaction set, D.
- the itemsets are formed by semantic indicators of transactional regions, so an item is represented by a pair (semantic indicator, value).
- the frequent itemsets represent the itemsets with the support greater than or equal to the minimum support defined (minsup).
- the itemsets of cardinality between 1 and k are iteratively found, where k represents the maximum length of an itemset; in our case, k is the number of semantic indicators, namely fourteen.
- for a rule of the form SI → d, SICount represents the number of cases in the transactional set, D, that contain the semantic indicator set, SI.
- for a rule of the form SI → d, ruleCount represents the number of cases in the transactional set, D, that contain the semantic indicator set, SI and are labelled with the diagnosis d.
- the support is (ruleCount/|D|)*100%.
- the confidence is (ruleCount/SICount)*100%.
- the frequent itemsets are used for rules generation.

A rule is represented in the algorithm, in the form:

<(SI, support), (d, confidence)>.

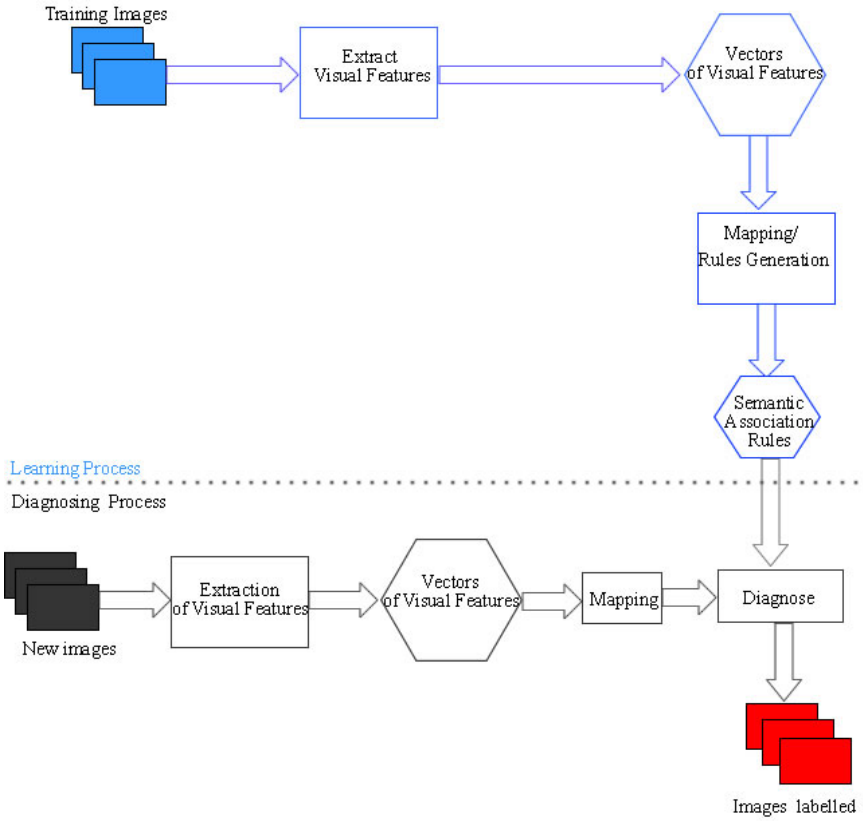

Figure 3. Description of the image diagnosis process

Algorithm 2. Rule generation on the training set of the transactional database.

Input: the transactional set, *D*, containing the sick regions of images with various diagnoses; each sick region is described by a set of *k* semantic indicators, *SI*; the defined minimum support, *minsup*, and the defined minimum confidence, *minconf*.

Output: the set of semantic, *Rules*.

Method:

1 F_1 = {frequent 1-itemsets};
2. for(k=2;$F_{k-1} \neq$ null; k++) do
3. {$C_k = F_{k-1} \otimes \{SI_k\}$;
4. for each distinct diagnosis $d \in D$ do
5. {
6. for each $c \in C_k$ do
7. {c.SICount++;
8. if c.diagnosis = d then
9. c.ruleCount++;
10. }
11. }
12. F_k={$c \in C_k \mid c$.support $\geq minsup$}
13. DAR_k=genRules(F_k)
14. }
15. $Rules = \bigcup_k DAR_k$;

The function genRules(F_k) includes the following steps:

- for all the rules that have the same semantic indicator set (*SI*), the rule with the highest confidence is chosen.
- accurate rule: confidence ≥ minconf.

In the first step, the algorithm selects the frequent itemsets of length 1. The maximum length of itemsets is for $k = 14$. For each step (lines 2-15), the algorithm performs the following operations:

- the frequent itemsets Fk-1 found in the (k-1) step are joined to the values of the semantic indicator, SIk, to generate the itemsets Ck (line 3).
- it scans the transactional database and updates the support and confidence of itemsets from Ck (lines 4-11).
- new frequent itemsets Fk are found (line 12).
- the algorithm produces the set of rules DARk, using the function genRules (Fk) (line 13).
- the set, Rules is the output computed as the union of rules sets, DARk.

As an example, we apply Algorithm 2 on the transaction set from Table 1, by simplifying the number of semantic indicators of the itemsets, from 14 to 2. The results shown in Table 2 are obtained supposing that the minimum support is 20%, and the minimum confidence is 66.7% .

Table 1. Transactional sick regions having different diagnoses

Colour	Texture-entropy	Diagnosis
light-red	Small	gastric-ulcer
light-red	Small	gastric-ulcer
light-red	Small	gastric-ulcer
light-red	Big	gastric-ulcer
light-yellow	Big	gastric-ulcer
light-yellow	Medium	duodenal-ulcer
light-yellow	Medium	duodenal-ulcer
medium-yellow	Small	duodenal-ulcer
medium-yellow	Small	duodenal-ulcer
dark-yellow	Small	duodenal-ulcer
dark-yellow	Small	duodenal-ulcer

Table 2. Frequent itemsets obtained by applying Algorithm 2

1st step	F_1	<({(Colour, light-red)},40%),(gastric-ulcer,100%)>,<({(Colour,light-yellow)},30%), (duodenal-ulcer,66.7%)>,<({(Colour,medium-yellow)},20%),(duodenal-ulcer,100%)>,<({(Colour,dark-yellow)},20%),(duodenal-ulcer,100%)> <({(Texture, small)},70%), (gastric-ulcer,42.9%)>, <({(Texture, small)},70%), (duodenal-ulcer,57.1%)>,<({(Texture, big)},20%), (gastric-ulcer,100%)>, <({(Texture, medium)},20%), (duodenal-ulcer,100%)>
2nd step	C_2	<({(Colour, light-red),(Texture-entropy, small)},30%), (gastric-ulcer,100%)>, <({(Colour, light-red),(Texture-entropy, big)},10%), (gastric-ulcer,100%)>, <({(Colour, light-red),(Texture-entropy, medium)},0%), (duodenal-ulcer,0%)>, <({(Colour, light-red),(Texture-entropy, small)},0%), (duodenal-ulcer,0%)>, <({(Colour, light-yellow), (Texture-entropy, small)},0%), (gastric-ulcer,0%)>, <({(Colour, light-yellow), (Texture-entropy, big)},10%), (gastric-ulcer,100%)>, <({(Colour, light-yellow), (Texture-entropy, medium)},20%), (duodenal-ulcer,100%)>, <({(Colour, light-yellow), (Texture-entropy, small)},0%), (duodenal-ulcer,0%)>, <({(Colour, medium-yellow), (Texture-entropy, small)},0%), (gastric-ulcer,0%)>, <({(Colour, medium-yellow), (Texture-entropy, big)},0%), (gastric-ulcer,0%)>, <({(Colour, medium-yellow), (Texture-entropy, medium)},20%), (duodenal-ulcer,100%)>, <({(Colour, medium-yellow), (Texture-entropy, small)},0%), (duodenal-ulcer,0%)>, <({(Colour, dark-yellow), (Texture-entropy, small)},0%), (gastric-ulcer,0%)>, <({(Colour, dark-yellow), (Texture-entropy, big)},0%), (gastric-ulcer,0%)>, <({(Colour, dark-yellow), (Texture-entropy, medium)},0%), (duodenal-ulcer,0%)>, <({(Colour, dark-yellow), (Texture-entropy, small)},20%), (duodenal-ulcer,100%)>
	F_2	<({(Colour, light-red),(Texture-entropy, small)},30%), (gastric-ulcer,100%)>, <({(Colour, light-yellow), (Texture-entropy, medium)},20%), (duodenal-ulcer,100%)>, <({(Colour, medium-yellow), (Texture-entropy, medium)},20%), (duodenal-ulcer,100%)>, <({(Colour, dark-yellow), (Texture-entropy, small)},20%), (duodenal-ulcer,100%)>
DAR_1		(Colour, light-red)→ gastric-ulcer, (Colour, light-yellow)→ duodenal-ulcer, (Colour, medium-yellow)→duodenal-ulcer, (Colour, dark-yellow)→duodenal-ulcer, (Texture, big)→ gastric-ulcer, (Texture, medium)→ duodenal-ulcer
DAR_2		{(Colour, light-red), (Texture-entropy, small)}→ gastric-ulcer, {(Colour, light-yellow), (Texture-entropy, medium)}→ duodenal-ulcer, {(Colour, medium-yellow), (Texture-entropy, medium)}→ duodenal-ulcer, {(Colour, dark-yellow), (Texture-entropy, small)} → duodenal-ulcer
Rules		$DAR_1 \cup DAR_2$

4.2 Image annotation

The set of generated rules *Rules* represents the classifier. The classifier is used to predict which diagnoses the images from the test database belong to. Being given a new image, the classification process searches in the rules set for finding its most appropriate diagnosis.

The algorithm is described in pseudo-code:

```
Algoritm 3. Algorithm for determining an image diagnosis.
Input: new image, I, the set of generated rules, Rules; each rule has the confidence R.conf.
Output: the list of diagnoses attached to the image, I.
Method:
1.MaxConf = 0
2.DiagnosisSet = null
3.foreach rule R in Rules do{
4.   if  (match(R,I) = 1) then {
5.         if (maxConf <= R.conf) then {
6.          Add rule R to DiagnosisSet
7.          MaxConf = R.conf
8.          }
9.   }
10.}
11.Diagnose the image I with the diagnoses from DiagnosisSet.
```

The algorithm verifies if the image *I* is matched to any rule, *R* from the *Rules* set, and the semantic rules with maximum confidence are selected.

The function *match(R,I)* returns 1, if all the semantic indicators, which appear in the body of the rule are included in the characteristics of the image regions, otherwise it returns 0:

$$match(R,I)=\begin{cases}1, if\ SI(R)\subseteq SI(I)\\0, otherwise.\end{cases},$$

where *SI(R)* represents the set of semantic indicators of the rule, *R*, and *SI(I)* represents the set of semantic indicators of the image, *I*.

5. Experiments and conclusion

The image collections used in our experiments were taken from free repositories on the Internet [18], [19].

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.

We analyze the performance of the method proposed in this paper for duodenal ulcer diagnosis. Algorithm 2 generates 16 semantic rules that recognize this diagnosis.

The test database contains 450 images, from which 60 are relevant for duodenal ulcer diagnosis.

A part of images used for learning the duodenal ulcer diagnosis can be analyzed in Figure 4. Duodenal ulcers can come in different shapes, sizes, and textures, increasing the complexity to learn them and generate semantic rules. For example, the image from the first position shows a single, white-based ulcer; the image from the second position shows numerous small ulcers scattered across the duodenal cap; the image from the third position demonstrates a huge duodenal ulcer in yellow hues.

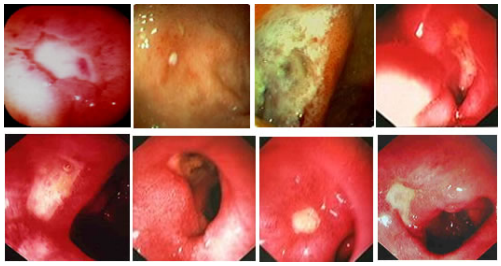

Figure 4. Images diagnosed with duodenal ulcer

After classification, we counted: TP the number of true positives (images correctly diagnosed with the duodenal ulcer diagnosis) and we found 51 images; FP the number of false positives (images incorrectly diagnosed with the duodenal ulcer diagnosis) and we found 6 images; TN the number of true negatives (images correctly diagnosed with a different diagnosis) and we found 385 images; FN 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 96.8%. The specificity, which measures the capability of duodenal ulcer rules not to miss the duodenal ulcer images, and not to diagnose images with a different diagnosis, is 98%. In our case, this set of rules is very specific.

An important improvement of this paper is in the generation of rules with very high specificity. The comparison between the method presented in this paper and the other ones presented in [2], [3] will be discussed in a future work. Methods proposed and developed in this study could assist physicians by doing automatic diagnosis based on visual content of medical images. For establishing correlations with diagnosis, we experimented and selected some low-level visual characteristics of images. So, each diagnosis is translated into visual computable characteristics and terms of sick regions. 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.

References

[1] **B. Liu, W. Hsu, Y. Ma.** Integrating classification and association rule mining. *Proceedings 4th International Conference on Knowledge Discovery and Data Mining*, 1998, 80–86.

[2] **S.C. Udristoiu, A.L. Ion**. Establishing Medical Diagnosis using Pattern Semantic Rules, *Electronics and Electrical Engineering*, *Vol*.2(98), 2010, 71–74.

[3] **A.L. Ion, S.C. Udristoiu, L. Stanescu, D. Burdescu.** Rule-Based Methods for the Computer Assisted Diagnosis of Medical Images. *Proceedings International Conference on Advancements of Medicine and Health Care through Technology*, *Series: IFMBE Proceedings, Vol*.26, *Springer-Verlag*, 2009, 247–250.

[4] **M.X. Ribeiro, A.J. Traina, C. Traina, N.A. Rosa, P.M. Marques.** How to Improve Medical Image Diagnosis through Association Rules: The IDEA Method. *Proceedings of the* 2008 21*st IEEE International Symposium on Computer-Based Medical Systems*, 2008, 266–271.

[5] **M.L. Antonie, O. R. Zaiane, and A. Coman.** Associative Classifiers for Medical Images. *LNAI* 2797, *MMCD*, *Springer-Verlag*, 2003, 68–83.

[6] **H. Pan, J.Li, Z. Wei.** Mining interesting association rules in medical images. *Advance Data Mining and Medical Applications*, 2005, 598–609.

[7] **J.R. Smith, S.-F. Chang.** VisualSEEk: a fully automated content-based image query system. *The Fourth ACM International Multimedia Conference and Exhibition*, *Boston, MA, USA*, 1996, 87–98.

[8] **A.L. Ion.** Methods for Knowledge Discovery in Images. *Information Technology and Control*, *Vol*.38, *No*.1, 2009, 43–49.

[9] **Y. Liu, D. Zhang, G. Lu, W-Y. Ma.** A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, *Vol*.40(1), 2007, 262–282.

[10] **T. Deselaers, D. Keysers, H. Ney**. FIRE-flexible image retrieval engine: ImageCLEF 2004 evaluation. *CLEF* 2004, *LNCS* 3491, 2004, 688–698.

[11] **T. Lehmann, M. Güld, C. Thies, B. Fischer, K. Spitzer, D. Keysers, H. Ney, M. Kohnen, H. Schubert, and B. Wein**. The IRMA project. A state of the art report on content-based image retrieval in medical applications. *Proceedings* 7*th Korea-Germany Joint Workshop on Advanced Medical Image Processing*, 2003, 161–171.

[12] **J. Vompras**. Towards adaptive ontology-based image retrieval. *Stefan Brass, C.G., editor,* 17*th GI-Workshop on the Foundations of Databases*, *Wörlitz, Germany*, 2005, 148–152.

[13] **L. Su, B. Sharp, C. Chibelushi.** Knowledge-based image understanding: A rule-based production system for X-ray segmentation. *Proceedings of Fourth International Conference on Enterprise Information System, Ciudad Real, Spain*, 2002, 530–533.

[14] **A. Mechouche, C. Golbreich, B**. **Gibaud.** Towards an hybrid system using an ontology enriched by rules for the semantic annotation of brain MRI images. *Lecture Notes in Computer Science*, *Vol*. 4524, 2007, 219–228.

[15] **R. Agrawal, T. Imielinski, A. Swami**. Mining Association Rules between Sets of Items in Large Databases. *Proceedings of the* 1993 *ACM SIGMOD International Conference on Management of Data*, *Washington, DC*, 1993, 207–216.

[16] **M. Houtsma, A. Swami**. Set-Oriented Mining of Association Rules. *Technical Report RJ* 9567, *IBM*, 1993.

[17] **R. Agrawal, R. Srikant.** Fast Algorithms for Mining Association Rules in Large Databases. *Proceedings of the* 20*th International Conference on Very Large Data Bases*, *Santiago, Chile*, 1994, 487–499.

[18] **Jackson Siegelbaum Gastroenterology**, 2008, *http://gicare.com/Endoscopy-Center/Endoscopy-images.aspx.*

[19] **The Gatrolab Image Library**. 2007: *http://www.gastrolab.net/***.**

Received September 2009.

DOI: 10.5755/j01.itc.39.2.12294