A Framework for medical images diagnosing

Anca Loredana Udriștoiu (Ion)*, Ștefan Udriștoiu*

**University of Craiova, Faculty of Automation, Computers and Electronics, Romania (phone: 40-251-438-198; fax: 40-251-438-198; e-mail: aion@software.ucv.ro).*

Abstract: In this paper we propose a framework for medical image diagnosing. The proposed framework includes components for the extraction of low-level features along with methods for the incorporation of semantic knowledge into the diagnosing process. The semantic information is represented by semantic rules which are used for an interactive diagnosing of the image data. The experiments through the framework were realized on collections of medical images from the digestive apparatus.

Keywords: Content-based image retrieval, Image mining, Rough sets.

1. INTRODUCTION

The medical applications with automatic diagnosis capacity imply unique challenges, but at the same time new opportunities. In some way we understand an image from nature and in another way a medical image, if we are not physicians. 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.

Due to the great difficulty in recognizing and classifying the images, the methods that identify the semantic features of images recorded a great success. Many researches were realized to investigate automatic techniques for generating semantic description of the multimedia content. Also in the medical domain, a lot of researchers were realized to automate the image diagnosing process.

This study starts from the limitations regarding the researches in multimedia semantic modeling and exploits the opportunities of the medical domain, rich in formal representation of knowledge. The paper proposes new approaches for image diagnosing, like: methods for generation of rules which identify image diagnosis using rough set theory, a method for mapping low-level features to semantic indicators using the Prolog declarative language, the creation of a representation image vocabulary and syntax, and semantic image classification.

The remainder of this paper discusses the architecture of our system and describes the interactions between its components in Section 2. A comparative study on image visual representation is presented in Section 3. The capturing of semantic knowledge and steps required to generate semantic image representations are detailed in Section 4. Finally, in Section 5, we present a summary of the presented framework.

2. ARCHITECTURE

The principal objective of our diagnosing system is to provide physicians with an image retrieval system with the capabilities of image classification and assignment of the data to high level concepts. By analyzing the visual structure of already annotated images, the system provides a semi-automatic annotation which generates descriptions for a new and unlabeled image. Since the system is working semi-automatically, it depends on an expert at the mapping processing step.

The main components of the diagnosing system are:
Feature Extraction Component. This component mainly provides methods for extracting primitive (visual) characteristics of images. For example, the set of low-level features implemented in the system currently includes color, shape and texture features.
Image Segmentation Component. In order to find out the semantic relations between image categories and 'objects' contained in an image, it should be divided into objects. For that purpose, an automatic segmentation algorithm based on a color-set back projection approach (Smith et al., 1996) is provided by the segmentation component. The segmentation approach is based on color homogeneity criterion.
Content-Based Visual Retrieval Component. This component manages the visual retrieval process. Beginning with query formulation by image example, the component provides functions for similarity computation between the query object and the image data stored in the database.

Vocabulary Component. This component provides an intermediate level between visual and semantic representation of images. The semantic indicators are the visual elements of the vocabulary and their values are experimentally and manually established.

Semantic Image Annotation Component. This component has the role of deciding diagnosis for new unlabeled images. It includes the following subcomponents:

Image Mapping subComponent. This subcomponent has the role of mapping the low-level features of images to semantic indicators.

Rules Generation and Classification subComponent. This subcomponent provides methods to generate semantic rules used to recognize the diagnosis of unlabeled images.

Storage Component. This component manages the physically storage of images and their features.

3. CONTENT-BASED IMAGE RETRIEVAL

In our framework, we develop a study of comparing different representations of images visual features, to select the ones with the best retrieval performances that are not limited to any particular domain.

In this section, we present these features and the relationship between them and the results returned by the content-based visual retrieval component of the framework. The motivation of realizing this comparative study is the necessity of finding a good set of visual image descriptors, as a precondition for the accuracy of the retrieval and annotation processing.

3.1 Color feature

Color is a very important feature in many image domains and it is the most used feature in the content- based image retrieval systems, because the color characteristic is easy to be detected from images and objects.

The developed comparative study of methods for color representation includes the following descriptors, some of them from MPEG-7 standard (Ion, A.L. , 2010):

- The colour histogram represented in HSV color space quantized at 166 colors (Smith et al., 1996).
- The second descriptors is MPEG-7 – color structure descriptor (CSD) represented in HMMD color space quantized at 128 colors (Manjunath et al., 2001).
- The third descriptor is the dominant color descriptor represented in CIE-LUV color space (Manjunath et al., (2001)). The average normalized modified retrieval rate (ANMRR) (Manjunath et al., 2001) is computed for testing the efficiency and performance of the color descriptors.

To compare the developed methods, the following conditions were set:

1. The database of about 600 images was created, including images from medical collections (Gatrolab, (2010), Gastroenterology, (2010)): duodenal ulcer, gastric ulcer, gastric cancer, esophagitis, and rectocolitis.
2. The three color descriptors were computed for each image from the database.
3. The relevant images were established for each experiment.

For each image category, the average of the normalized modified retrieval rate was computed, as can be observed in Table 1.

Table 1. The average normalized modified rate for color feature

Category	HSV1 66	HMMD128	Dominant color
Esophagitis	0.22	0.21	0.3
Rectocolitis	0.2267	0.2267	0.29
Duodenal ulcer	0.2267	0.2267	0.29
Gastric ulcer	0.22	0.21	0.3
Gastric cancer	0.19	0.20	0.3

By analysing the results of the content-based retrieval using the color feature, the best results were obtained when using the color histogram in the HSV color space and the color histogram in HMMD color space.

3.2 Texture feature

The texture is another important characteristic taken into consideration for classifying and recognizing the material surfaces. In the current framework, two methods are developed: Gabor filter and co-occurrence matrices (Ion, A.L. , 2010).

The average normalized modified retrieval rate (ANMRR) (Manjunath et al., 2001) is computed for testing the efficiency and performance of the texture descriptors.

To compare the developed methods, the following conditions were set:

1. The database of about 600 images was created, including images from medical collections (Gatrolab, 2010, Gastroenterology, 2010): ulcer, gastric cancer etc.
2. The two texture descriptors were computed for each image from the database.
3. The relevant images were established for each experiment.
4. For each image category, the average of the normalized modified retrieval rate was computed, as can be observed in Table 2.

Table 2. The average normalized modified rate for texture feature

Category	Co-occurrence Matrix	Gabor Filter
Esophagitis	0.24	0.3
Rectocolitis	0.236	0.259
Duodenal ulcer	0.236	0.259
Gastric ulcer	0.24	0.24
Gastric cancer	0.231	0.235

By analysing the results of the content-based retrieval using the texture feature, we concluded that the best results were obtained by the co-occurrence matrices.

3.3 Shape feature

In this framework, three shape descriptors were developed: the geometric moment, the Zernike moment and the eccentricity (Ion, A.L., 2010).

The average normalized modified retrieval rate (ANMRR) (Manjunath et al., 2001) is computed for testing the efficiency and performance of the texture descriptors.

To compare the developed methods, the following conditions were set:
1. The database of about 600 images is creating, including images from medical collections (Gatrolab, 2010; Gastroenterology, 2010): ulcer, gastric cancer, etc.
2. The three shape descriptors were computed for each image from the database.
3. The relevant images were established for each experiment.
4. For each image category, the average of the normalized modified retrieval rate was computed, as can be observed in Table 3.

Table3. The average normalized modified rate for shape feature

Category	GM	Zernike	Ecc
Esophagitis	0.25	0.271	0.2
Rectocolitis	0.23	0.27	0.198
Duodenal ulcer	0.23	0.28	0.198
Gastric ulcer	0.23	0.27	0.198
Gastric cancer	0.21	0.27	0.2

By analysing the results of the content-based retrieval using the shape feature, we concluded that all three methods have proximate results.

4. MODELING IMAGE DIAGNOSIS USING ROUGH SETS

In this section the problem of discovery the images diagnosis using semantic rules are approached. The developed methods are based on:
-The automatic segmentation of images and the indexing of resulted colour regions.
-The mapping of visual features of images to semantic indicators.
-The definition of a knowledge database, using the declarative language, Prolog, which facilitates the mapping process.
-The automatic discovery of semantic rules for discovery the semantic concepts from images.
-The representation of the semantic rules, using the declarative language, Prolog, to easier infer them to any domain.

After the performing of a large set of experiments we inferred the importance of semantic concepts in establishing the similitude between images. Even if the semantic concepts are not directly related to the visual features (colour, texture, shape, position, dimension, etc.), these attributes capture the information about the semantic meaning.

Using the results of experiments realized in the third section, the HSV colour space quantized at 166 colours is used for representing the colour features. Before segmentation, the images are transformed from RGB to HSV colour space and quantized to 166 colours. The colour regions extraction is realized with the colour set back projection algorithm (Smith et al., 1996). This algorithm detects the regions of a single colour.

Each region is described by the following visual characteristics:

- The colour characteristics are represented in the HSV colour space quantized at 166 colours. A region is represented by a colour index, which is in fact an integer number between 0...165.
- The spatial coherency represents the region descriptor, which measures the spatial compactness of the pixels of same colour.
- A seven-dimension vector (maximum probability, energy, entropy, contrast, cluster shade, cluster prominence, correlation) represents the texture characteristic.
- The region dimension descriptor represents the number of pixels from region.
- The spatial information is represented by the centroid coordinates of the region and by minimum bounded rectangle.
- The eccentricity represents the shape feature.

The diagnosing method includes two phases, the learning/training phase in which the rules are generated for each image category and the diagnosing phase in which new images is diagnosed using the semantic rules. The algorithms for semantic rules generation are based on rough set theory.

4.1 Rough Sets Foundations

In this section, we recall some basic definitions from literature (Pawlak, 1986; Pawlak et al., 1994; Stepaniuk, 2008).

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: $U \rightarrow Va$, where Va 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,

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

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

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 4 represents the sick regions of images from different diagnoses represented in terms of semantic indicators values. For simplicity we consider only two semantic indicators as attributes, namely the colour and texture-entropy.

Table 4. 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

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 4 are given in Table 5.

Table 5. 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})	$\{R_1, R_2, R_3\}$, $\{R_4\}$, $\{R_5\}$, $\{R_6, R_7\}$, $\{R_8, R_9\}$, $\{R_{10}, R_{11}\}$

A rough set approximates traditional sets using a pair of sets named the lower and upper approximations of the set. Let $W = \{w_1, ..., w_n\}$ 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*.

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

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.

4.2. Dispensable features, Reducts and Core

An important notion used in rough set theory is the decision table. Pawlak (Pawlak, Z., 1986; Pawlak et al., 1994) 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 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*.

4.3 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:

$$m_{ij}^{a(R_i)} = \{(a \in C : a(R_i) \neq a(R_j)) \text{ and } (d(R_i) \neq d(R_j))\} \quad (6)$$

where *i, j*=1,2,…,11.

We construct the discernibility matrix, *DM(DT)*, 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 (Stepaniuk, J., 2008; Hassanien et al., 2008).

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

$$g_{R_i}(Colour, Texture) = \bigwedge_{R_j \in U} \left(\bigvee_{a \in m_{ij}} a \right) \quad (7)$$

The following conditions are equivalent (Stepaniuk, J., 2008):
1. $\{a_{i1}, \ldots, a_{in}\}$ is a reduct for the object R_i, $i = 1 \ldots n$.
2. $a_{i1} \wedge a_{i2} \wedge \ldots \wedge a_{in}$ is a prime implicant of the Boolean function g_{Ri}.

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

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.

4.4 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 (Stepaniuk, J., 2008).

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 (Stepaniuk, J., 2008). 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:

$$s(rule) = cardinality(featureSet \cap diagnosisSet) \quad (8)$$

$$a(rule) = \frac{cardinality(featureSet \cap diagnosisSet)}{cardinality(featureSet)} \quad (9)$$

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:

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

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 Rule 1, the support is 4, accuracy is 4/4 and coverage is 4/5. High accuracy and coverage are requirements of decision rules.

5. DECISION RULE EXTRACTION USING ROUGH SETS MODELS AND EXPERIMENTS

In this section 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.

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 (Gatrolab, 2010; Gastroenterology, 2010).

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 analyse 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. 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

6. CONCLUSION

The direct motivation for our work is the fact that physicians are highly interested in querying images at conceptual and semantic level, not only in terms of low-level features.

The need of enhancement of the retrieval performance and the importance of 'semantic meaning' makes a detailed image annotation indispensable. Presently, most of the image database systems utilize manual annotation, where users assign some descriptive keywords to images. Although this process takes away the uncertainty of fully automatic annotation, it requires a high effort in exchange. Another weak point is that indexers often use different descriptors and their perceptual subjectivity may differ.

In summary, since it is very difficult to automatically construct semantic knowledge from the extracted low-level features and map them on human perception, methods which combine both approaches are of great interest.

ACKNOWLEDGMENT

This work was supported by the strategic grant POSDRU/89/1.5/S/61968, Project ID61968 (2009), co-financed by the European Social Fund within the Sectorial Operational Program Human Resources Development 2007-2013.

REFERENCES

Smith, J.R., Chang, S-F. (1996). VisualSEEk: a fully automated content-based image query system. *Proceedings of the Fourth ACM International Multimedia Conference and Exhibition*. Boston, MA, USA, 1996. pp. 87-98.

Manjunath, B. S., Ohm, J.-R. , Vasudevan, V. V. and Yamada, A. (2001). Color and Texture Descriptors. *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11(6), pp. 703–715.

Gatrolab (2010). *http://www.gastrolab.net/.*

Gastroenterology (2010). *http://gicare.com/Endoscopy-Center/ Endoscopy-images.aspx*.

Pawlak. Z (1986). Rough Relations. *Bulletin of the Polish Academy of Sciences, Technical Sciences,* vol. 34(9-10), pp. 587–590.

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.

Stepaniuk, J. (2008). *Rough Granular Computing in Knowledge Discovery and Data Mining*. Springer-Verlag, Germany.

Ion, A.L. (2010). A Framework for semantic modeling of images. *Annals of the University of Craiova - Mathematics and Computer Science Series*, vol. 37(4), pp.37-49.

Ion, A.L., Udristoiu, S. (2010). Image Mining for discovering medical diagnosis. *Information Technology and Control*, vol. 39(1), pp. 123-129.

Hassanien, A.E., Abraham A., Peters, J.F. (2008). Rough Sets in Medical Imaging: Foundations and Trends. *Computational Intelligence in Medical Imaging: Techniques and Applications*, pp. 47-87. CRC Press, USA.