

## A new similarity measure based on Bayesian Network signature correspondence for brain tumors cases retrieval

Hedi Yazid\*, Karim Kalti, Najoua Essoukri Benamara  
SAGE Research Group, Sousse Engineers school, Sousse University,  
BP 264 Sousse Erriadh 4023, Sousse, Tunisia  
hedi.yazid@gmail.com, karim.kalti@gmail.com, najoua.benamara@eniso.rnu.tn

Received 21 July 2013

Accepted 1 February 2014

### Abstract

Case retrieval constitutes an interesting area of research which contributes to the evolution of several domains. The similarity measure module is a fundamental step in the retrieval process which affects remarkably on a retrieval system. In this context, we suggest in this paper a similarity measure applied to brain tumor cases retrieval. The rationale behind the proposed measure consists in quantifying the diagnosis correspondence followed by a clinician while comparing two medical cases. Our idea is characterized by the use of the Bayesian inference in the formulation of the proposed measure. The Bayesian network is applied in the classification task and it describes the decision-making process of a radiologist facing a tumor. The proposed similarity algorithm is based essentially on graph correspondence based on signature nodes comparison from the Bayesian classifiers. experiments were directed to compare the performance of the proposed similarity measure method with classical methods of similarity quantification. The performance indices of our proposition are promising.

*Keywords:* Medical case retrieval; Brain Tumors; Similarity Measure; Bayesian networks; Bayesian inference, graph signature.

### 1. Introduction

Information retrieval is an approach based on artificial intelligence (AI) techniques which is designed to facilitate the research of documents in complex databases<sup>1</sup>. Recently, Information retrieval has become a solicited research area. In the last years, several systems were implemented to retrieve data from large databases. Medical domain constitutes an ideal field for the retrieval research axis, which is marked by an increasing evolution of the medical images use. The archived medical cases can be re-used in computer aided-diagnosis process. The reuse of archived medical cases has become increasingly applied in the diagnosis process<sup>1</sup>. When doctors are confronted with a case of conflict, a retrieval system displays the cases classed as similar to the current case Refs. 2, 3 and 4. Treatment

process is stored on the system and it can be replicated on the current case. The relevance of a retrieval system depends closely on the chosen similarity measure which constitutes a fundamental module in the retrieval process. For this, the similarity-based retrieval has become an important area of computer vision research<sup>5</sup>.

In this study, we propose a similarity measure approach for medical cases retrieval application. The approach is applied on a medical problem of brain tumors retrieval application. The brain tumors retrieval problem requires a confirmed theoretical basis of representation and reasoning. Most of the current approaches retrieve a case by proposing a classical distance between the descriptors of the tested cases. These classic measures do not reflect the assimilation between two objects in a real context. The current approach relies on this observation to propose a similarity measure based on the comparison of the

---

\* Corresponding author: hedi.yazid@gmail.com

reasoning process leading to a medical diagnosis. The assessment of similarity between two medical cases is characterized by a probabilistic aspect. These points are embodied in this paper while referring to the principles of Bayesian inference. In this work, a Bayesian network is used to classify a medical case into a brain tumor class. The construction of a Bayesian classifier takes into account the radiologist reasoning in a real context of brain tumors diagnosis. The membership degree of a case to a tumor class is calculated after an inference exercise. Evidence propagation during a Bayesian inference plays a crucial role in the similarity measure process. Indeed, the proposed measure is based essentially on graph correspondence based on signature nodes comparison from the Bayesian classifiers.

The remainder of the paper is organized as follows. A background was briefly cited in Section 2. Section 3 presented the motivation and the overview of the proposed method. Section 4 described the brain tumor classification model based on Bayesian network. The description of the proposed similarity measure was elaborated in Section 5. Experiments were presented and discussed in section 6 followed by a conclusion and proposals for perspectives.

## 2. Related Work

### 2.1. Retrieval applied in the medical field

The studies, dealing with the retrievals contribution applied in medical field, are varied. This diversity is explained by the multitude of medical data modalities<sup>6</sup>. Information contained in a medical case can be visual or contextual. The visual information can be stated in general contexts in form of colors, textures, shapes, spatial relations, etc. The contextual information usually appears as an auxiliary data considered in the diagnostic process (such as age, sex, other diseases, etc.)<sup>7</sup>. Therefore, we mention retrieval contributions using the visual characteristics<sup>8</sup>, other contributions based on the contextual characteristics<sup>9</sup>, and works combining both types<sup>10</sup>.

The retrieval research axis is applied in various fields. In this paper, we will focus on works applied in the medical field. The computer aided-diagnosis is imposed as a fundamental tool for clinicians who refer to it in the diagnosis protocol. Thus, several are the

content based retrieval (CBR) systems that have been validated and installed in hospitals<sup>11</sup>. Most of the projects target a given pathology and cover a specific image type. For instance, we mention the work described in Ref. 12 which defines an automatic search and selection engine with retrieval tools (ASSERT). The ASSERT project is a content based image retrieval (CBIR) system for a high-resolution computed tomographic (HRCT) lung image database. In Ref. 13, authors proposed a general medical CBIR system named IRMA which encompasses an approach for content-based image retrieval in medical applications with a particular focus on its contextual layers of information modeling. Recently, Bayesian networks are increasingly used in the representation and interpretation of retrieval problems applied in the medical field. This is argued by its benefits in managing the uncertainty that appears in such types of problems<sup>14</sup>. Obviously, some researchers have relied on this reasoning since Bayesian networks appear in several medical image retrieval works. First, we mention The Bayesian CBIR system Pichunter<sup>15</sup> which represents a simple instance of a general Bayesian framework applied to relevance feedback. Second, authors proposed in Ref. 16 a framework for contextual image understanding. This work is a general-purpose knowledge integration framework that employs Bayesian Networks in integrating both low-level and contextual features. Finally, Bayesian networks are used for fusion information in multimodal medical case retrieval in Ref. 17.

### 2.2. Retrieval process schema

A retrieval process consists of a set of successive steps. First, the feature extraction and selection followed by information field representation which leads to the measure of similarity and decision. The last step of this process is the expert feedback. The similarity measure is considered as the most important module in a retrieval model. Fig.1 illustrates the position of a similarity measure module in a retrieval system.

The formulation of a logical and approved notion of similarity is fundamental for the pertinence of a retrieval system decisions. A similarity measure depends mainly on objects defining the treated field. A similarity measure technique is generally based on a comparison

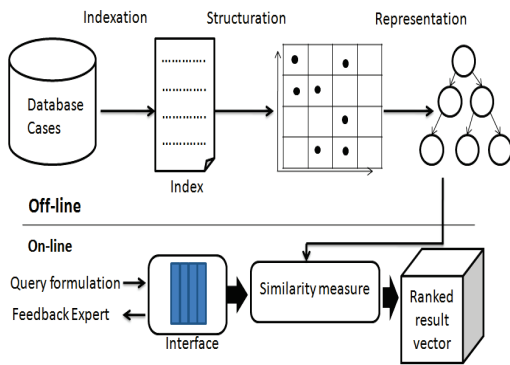


Fig. 1. Retrieval process schema

approach that embodies the specific nature of measurable objects. The similarity between two objects can be translated into a distance. A small distance between two objects is interpreted as a strong similarity between them.

**2.3. Cerebral Tumors diagnosis**

The present study addresses a very common pathology type characterized by a significant diversity which is brain tumors. A brain tumor is an intracranial solid neoplasm which is included inside the cranium or in the central spinal canal<sup>18</sup>. The classification of brain tumors into families is a difficult task due to the variety of tumor types; we talk about the histological variety. Usually, the cerebral tumors are classified into:

Intra-axial tumors developed starting from cerebral tissue;

Extra-axial tumors developed in under-arachnoidiens spaces (primarily starting from the meningeal envelope) or in the bony wall of the cranial cavity<sup>19</sup>.

The diagnosis of a tumor falls under a precise and complex clinical step leading thereafter to an adapted therapeutic decision. It is made up of several stages: suspicion, detection, observation and determination of its histological nature. The radiological examination is a fundamental step in the brain tumor diagnostic process. It aims at highlighting and characterizing the brain lesions. The radiological diagnosis is not definite, and the radiologic hypothesis is generally confirmed by biopsy of the lesion or paraclinical exams. The MRI exam is known for its supremacy in the radiological diagnosis of brain lesions<sup>19</sup>.

In this work, we focus on tumors belonging to the intra-axial SupraTentorial tumor family. Thus, seven tumors are concerned which are cited in Table 1.

When reading an MRI sequence, a radiologist prepares its interpretation by relying on the description of a set of attributes<sup>20</sup>. Thus, a final report is issued by the radiologist based on the observations, on the patient's age and on other clinical information provided by the referring doctor. This report is based on statistics. The decision process made by radiologists is statistical because it is referring to the occurrence frequency of the attributes according to the pathology type. In this context, Table 2 shows the attributes list in the brain tumors diagnosis with the possible states for each one of them.

Table1. Brain tumors constituting the study framework.

1	Metastasis
2	Pilocytic Astrocytoma
3	Glioblastoma
4	Oligodendroglioma
5	Lymphoma
6	Low grade glioma
7	Meningioma

**2.4. Bayesian network theory**

Bayesian Networks is a graphical representation of dependencies between a predefined set of variables (features) that aims at specifying the joint probability distribution for a domain. Variables are interconnected by arcs that encode conditional independencies between them. Bayesian networks are based on Bayes theorem<sup>21</sup> which expresses the posterior probability among variables of the network. Posterior probability measures the likelihood that an event will occur given that a related event has already occurred. Given a set of variables  $\{X_1, \dots, X_n\}$ , the joint distribution  $P(X_1, \dots, X_n)$  is defined as follows:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | parents(X_i)) \quad (1)$$

Where  $parents(X_i)$  represents the parents set of a node  $X_i$ .

To sum up, building a Bayesian network consists in delimiting arcs between variables that compose the graph model and defining a conditional probability table (CPT) for each variable while highlighting the cause-effect conditional dependence<sup>21</sup>.

2.4.1. Bayesian Network Construction

The Bayesian network structure definition is called the qualitative step. In this step, we intend to focus on the influence relationships that may exist between variables taken two by two. This leads naturally to a graphical representation of relationships between variables<sup>22</sup>.

The parameters initialization in a Bayesian network can be ensured by two methods, either by probability tables' elicitation or by an application of a learning algorithm. First, field experts are interviewed to elicit probability parameters. The probability distributions must correspond to the reality of the treated domain. Second, Bayesian network parameters learning consist in calculating the probability of each hypothesis, giving the data, and making predictions on that basis<sup>23</sup>. Thus, and after an exercise of parameter learning, the conditional probability tables (CPT) could be constructed from empirical evidence<sup>24</sup>.

Table 2. Attributes referring List of MRI case Interpretation.

	Attributes	Possible Values
Visual Attributes	Siege	SupraTentorial, SubTentorial
	Localization	Superficial, Deep, brain stem, Vermis, Cerebral
	Limit	Good Limitation, Bad Limitation
	Size	Small, medium, Big, Variant
	Number	Single, Multiple
	Oedema Presence	0, +, ++, +++
	Mass Effect	0, +, ++
	Contrast Taking Importance	Intense, Small, Absent
	Contrast Taking Type	Nodular, Annular, Annular/ Nodular
	Composition	Fleshy, Mixed, Cystic
	T1 Weighting	Hyposignal, Isosignal, Hypersignal, Hyposignal and
	T2 Weighting	Hyposignal, isosignal, Hypersignal, isosignal and
	Cystic constituent	Yes, Rare, No
	Calcification	Yes, Rare, No
	Bleeding	-, +
Corpus Callusum Invasion	-, +	
Semantic Attributes	Age	Child, Young Adult, Adult, +60
	Sex	Man, Woman
	Other Diseases	Yes, No
	First Infection	Yes, No

2.4.2. Inference

Bayesian inference is a useful technique used in the problems of induction because it is based on individual

cases and it is validated in probabilistic terms. A Bayesian network can be represented by a set of random variables for which we know a number of dependency relationships. From this initial state, additional information on one or more variables can be injected

into the network. An update of the parameters distribution is performed following this injection; it is called an inference based on Bayes rule<sup>24</sup>.

### 3. Motivation and overview of the method

The case representation level on retrieval research has been fully discussed. Therefore, various approaches have been proposed. Any approach depends on the specificity of the field knowledge<sup>17</sup>.

In this work, the problem fits in a probabilistic context in order to model a pertinent similarity measure approach that imitates the real process of medical diagnosis decision making. For this, we chose to work with Bayesian networks which are approved in several areas of retrieval applications<sup>14</sup>. In addition, Bayesian networks demonstrate a pertinent treatment of uncertainty situations. Starting from this observation, the main idea is to develop a similarity measure incorporated in the retrieval approach which is based on the Bayesian networks classifiers. This similarity measure is calculated using the data generated from the Bayesian network classifiers. Thus, the interest is amplified in the signature formulation of the nodes that compose the compared graphs. Several studies in the literature follow this reasoning to quantify the similarity between two graphs such as Refs. 25 and 26.

Indeed, the inference data is used to calculate the degree of similarity between a query case and existing cases in the knowledge base. The main idea of this approach is to imitate the conducted reasoning adopted by a clinician in a real diagnosis process. Thus, a radiologist builds its conclusion by observing all diagnostic data to guide him to the most probable decision.

The two steps of defining the proposed retrieval approach (see fig. 2) are the following:

- The first step defines a module of brain tumors classification based on Bayesian classifiers. We associate a classifier for each tumor class of the study framework. The construction of such models requires two phases: the definition of the structure and the estimation of network nodes probabilities. For each classifier, a membership degree and propagation evidence trace is recovered while running an inference exercise. This collected information will be used in the formulation of the similarity measure in the next step.

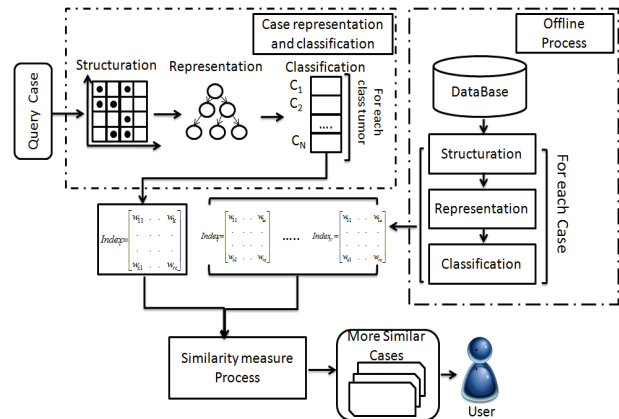


Fig. 2. Overview of the retrieval system based on Bayesian Classifiers

- The second step concerns a similarity measure based on the comparison of signatures deduced from the Bayesian classifiers nodes. This process is followed by the final decision of the retrieval task. The similarity measure aims to determine the degree of correspondence between two medical cases. The main idea of this approach consists of an association of a local signature for each node of Bayesian classifier associated to a tumor class. The process is repeated in all Bayesian classifiers (each one represents a tumor class) in the classification task. In order to achieve this idea, we chose to use the Pearl algorithm as the inference approach in the. This algorithm uses the principle of evidence propagation as received and sent messages between the network nodes. Thus, a node signature is formulated from the messages propagation process in the time of inference. After this procedure, each case will be represented by a set of parameters that can be considered as its global signature. The comparison step between two cases is performed based on their signatures.

## 4. Brain tumors classification based on a Bayesian Network

### 4.1. Construction of a Bayesian Network for tumors classification

This actual work applies the structured representations category and more precisely the network-based representation Refs. 27 and 28. Case representation based on networks approach is proposed in several research works. It has shown its performance in analogical reasoning research area<sup>29</sup>. We promote this

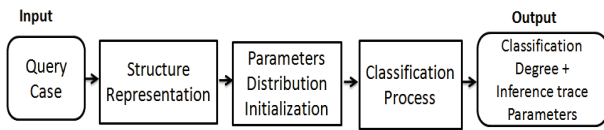


Fig. 3. Classification based Bayesian networks process

choice in our work by designing structure that represents the problem domain based on Bayesian networks.

The idea here is to develop Bayesian networks of medical cases representation that aim to test the appurtenance degree of a cerebral to a given tumor. The construction of a Bayesian network requires a transition through two stages; structure modeling and network nodes probabilities distribution.

Figure 3 illustrates the process steps. As an input, a query case is injected and, as an output, the classification degree and the inference tracking parameters are recovered for each tumor class from the study framework.

First, the Bayesian network structure is built in cooperation with medical experts. For this, Experts have emulated the diagnosis process while interpreting a cerebral tumor case in the desired network. The built network describes the causality distribution of the attributes included in a diagnostic process. Each attribute is schematized by a node. Each node may take one value of the refereed attribute possible states (Table 1). The causal relationship between two attributes is represented by an arc between two nodes. Thus, we propose to decompose the network structure into three levels; the first one is called “observed layer” which represents variables corresponding to attributes used in the interpretation process. The second level is called “the intermediate layer” and it plays the role of state nodes in a subset of the attributes. In a real context, a set of attributes give an idea of precise information of diagnosis. For instance, ‘Contrast Taking Importance’ and ‘Contrast Taking Type’ gives a decision about ‘Contrast Taking state’. Besides, this state is described into a node in the intermediate layer. The last and third level is a result node. This node deals with the task of decision “What is the membership probability of the query case to the tumor?” As an illustration, we present in Fig.4 a sub graph which contains nodes extracted from a Bayesian network.

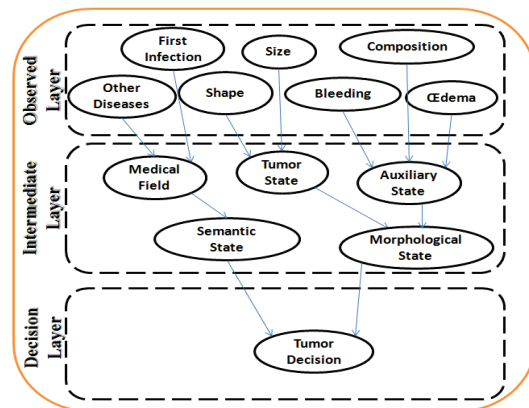


Fig. 4. Bayesian network architecture for Brain tumors classification

Second, the initialization of Bayesian networks nodes is provided according two methods. A first method is based on nodes elicitation referred to the expertise of radiologists collaborating in this work. Medical experts propose an elicitation of probabilities for each node of the Bayesian network. Indeed, we perform an elicitation to each treated tumor. The second method ensures a parameter initialization by applying a learning algorithm. A variety of parameter learning algorithms in Bayesian network are defined in literature. The learning process uses cases attributes from the knowledge base of our work. Indeed, in both methods, nodes initialization varies from tumor to another. Moreover, we count a CPT distribution for each concerned tumor.

**4.2. Case inference description**

An inference in a network of causalities consists of a propagation of one or more uncertain information within this network to deduce how the beliefs concerning the other nodes are modified<sup>30</sup>. In our current situation, the objective of an inference is to classify a query case into a tumor class. This procedure is done with all tumors included in the study framework. A query case is defined by a set of attributes (morphological and contextual). A cerebral MRI case interpretation is ensured by the radiologists via an interface. At the end of an interpretation procedure, we have a complete description of visual and contextual attributes for each new case. Those attributes constitute the evidence set to be injected into an inference exercise. Thus, we have defined a glossary of symbolic values representing the attributes interpretation in a

brain tumor diagnostic report. From a brain tumor case, a clinician translates the attributes description mentioned in the case report interpretation to symbolic values. At the end of this procedure, we recover an appurtenance probability for each tumor class. The final decision of the tumor class of a tested case is referred to the maximum probability returned.

**4.3. Contribution of the Brain tumors classifiers in the Similarity measure**

So far, we develop our similarity measure that is based mainly on data recovered from the Bayesian classifiers. The main idea of this measure is to develop a score for each medical case from nodes values after an inference exercise. The recovered data is used to calculate the similarity degree between a query case and a case from database. In reality, when a radiologist interprets a medical case, he builds his conclusion following a reasoning way. This reasoning explores all diagnostic data to guide interpretation to the most probable decision. Thus, we have translated this reasoning way in our approach. Besides, measuring a distance between two reasoning is defined by measuring the distance between two inferences. The similarity measure procedure is described in the fore coming section.

**5. Similarity measure for Retrieval based on Bayesian Network nodes signature comparison**

**5.1. Idea principals**

The objective of this paper is to present a similarity measure based on Bayesian Networks nodes signature comparison. This similarity measure aims to determine the correspondence degree between cases. Our main idea is based on the principle of association of local signature for each node of the graph. This will allow us to have a set of local descriptions of the different graph components.

The formulation of the signature refers mainly to the propagation of information between the graph's nodes during an inference. The Pearl inference algorithm is the most appropriated algorithm that refers essentially to the evidence propagation between the network nodes. This is called the belief propagation. The pearl algorithm is considered as an exact inference algorithm. It is similar

to the classic ‘forward – backward’ algorithm of information sharing<sup>22</sup>.

Given a graph, the Pearl belief propagation underlines the information through the links of it<sup>31</sup>. Flows are

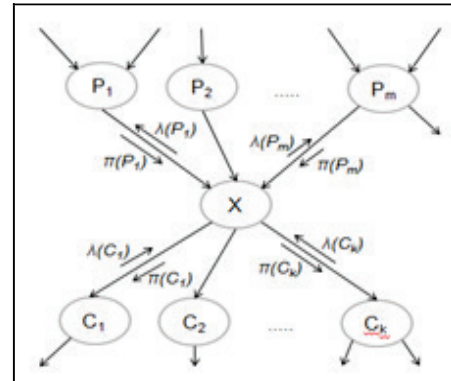


Fig.5. Messages' propagation in Pearl's inference algorithm

viewed as messages between nodes. We propose the graph  $G$  which consists of a set of nodes  $\{X_1, \dots, X_i\}$ . Consider a node  $X$  having a set of parents  $\{P_1, \dots, P_m\}$  and a set of children  $\{C_1, \dots, C_n\}$  as shown in the following figure (Fig.5):

We note  $E$  as the evidence of the local computation of  $X$ .  $E$  is composed of two parts; the first part is accessible from the parents of  $x$  and known as evidence at node's ancestors (noted as  $E^+$ ). This causal evidence is passed downward in  $\pi$  messages. The second part is called diagnostic evidence and is passed upwards in  $\lambda$  messages and noted as  $E^-$ <sup>31</sup>. Thus, knowledge can flow in both directions into a network, from parent to child (down) or from child to parent (up), and the belief of  $x$  where its parents and children propagation is defined as:

$$Bel(X) = P(X|e) = \alpha \lambda(X) \pi(X) \quad (2)$$

Where  $\alpha$  is the normalization constant and  $\lambda(x)$  denotes the sent messages, given as:

$$\lambda(x) = P(e^- | x) \quad (3)$$

$\lambda(x)$  can be described as follows:

$$\lambda(x) = \prod_{j=1}^k \lambda_{C_j X_i}(x) \quad (4)$$

$\pi(X)$  denotes the messages received from parents and defined by:

$$\pi(X) = \sum_{\vec{p}} (x_i | \vec{p}) \pi(\vec{p}) \quad (5)$$

$\vec{p}$  is the possible values vector. Then,  $\pi(X)$  can be formulated as follows:

$$\pi(X) = \sum_{p_1, \dots, p_m} P(X_i | p_1, \dots, p_m) \prod_{j=1}^m \pi_{p_i, X_i}(p_j) \quad (6)$$

Once  $Bel(x)$  has been updated,  $X$  must send  $\pi$  and  $\lambda$  messages to its children and parents, respectively.

### 5.2. Similarity measure prerequisites

Up to now, we are aimed to transform our similarity problem to a comparison task of two ‘Weighed graphs’. The formalism of this idea refers to the principals of the ‘Belief propagation method’ described above.

For each tumor type, we extract a prior sub graph which varies from one tumor to another. Actually, this choice is recommended by the medical experts. In a real diagnostic context, a radiologist orients its approximations by observing primarily the values of a small set of attributes. This attributes’ prior set varies from one tumor to another tumor. For example, the prior attributes of the Glioblastoma tumor are size, siege, Corpus Callusum Invasion, Oedema Presence and calcification. This idea is taken up in our similarity measure procedure by proposing the concept of prior sub graph. Moreover, we assign a weighting for all nodes belonging to the prior sub graph.

In the algorithm process, we intend to explore the following points:

- We have a set of graph  $G_r = \{G^1, \dots, G^T\}$  that each one of them represents a tumor  $T = \{\tau_1, \dots, \tau_n\}$  from the list mentioned above,
- In order to obtain a set of local descriptions that define a weighted graph, we associate a signature at each node (a vector of probability). The signatures of the nodes are designed to determine if two graphs may be similar. The node signature construction is a crucial step in the process of measuring the correspondence between the two graphs,
- For a node to node comparison, the score computation process consists in a distance measure between the probabilities of these two nodes in a

first stage and a distance measure on the set of parent nodes and children nodes (Comparison Matrix - Matrix) in a second stage.

For further details, we propose an algorithm that reveals these points. This algorithm is described in the next section.

### 5.3. Node signature definition

Therefore, we propose  $x$  as the current treated node selected from the set of prior graph nodes noted as  $X\{x_1, x_2, \dots, x_i\}$ . In occurrence,  $P\{p_1, p_2, \dots, p_m\}$  is the set of this node’s parents and  $C\{c_1, c_2, \dots, c_n\}$  is the set of its children.

For the weighted graphs, the signature is defined as the degree of the node and th\*e weights of all the incident edges. Given such a graph, the node signature is formulated as follows:

$$w_{sign}(x) = \{d(x), \pi(x), \lambda(x)\} \quad (7)$$

Where  $x \in X$ ,  $d(x)$  gives the degree of the node  $x$ ,  $\pi(x)$  is the sum of inflows from parents and  $\lambda(x)$  is the sum of inflows from children.

Considering these data, we aim at schematizing the received messages of arcs for a node  $x$ . First, the received messages from parents which are given as follows:

$$\pi(x) = \sum_{p_1, \dots, p_m} P(x | P_1 \dots P_m) \prod_{j=1}^m \pi_{p_j, x_i}(P_j) \quad (8)$$

Thus, we need to calculate:

$$\pi_{XY_j} = \alpha \pi_X(x) \lambda_{Y_k X}(X) \quad (9)$$

Second, the messages sent by the child nodes can be formulated as follows:

$$\lambda(N) = \prod_{j=1}^n \lambda_{p_j N_i}(x) \quad (10)$$

### 5.4. Node to node correspondence

The first step in this process is a comparison of a node that belongs to  $G_i$  with its counterpart in the  $G_r$  graph. This correspondence equation is given as:

$$M(x_i, x_r) = M(w_{sign}(x_i), w_{sign}(x_r)) \quad (11)$$

More precisely, the equation is written as follows:



$$M(x_i, x_r) = + \prod_{i=1}^n (\lambda_{y, x_i}(x_r) - \lambda_{y, x_i}(x_i)) + \prod_{j=1}^m ((\pi_{p_{x_i}}(x_r) - \pi_{p_{x_i}}(x_i)) + (\Pr(x_i) - \Pr(x_r))) \quad (12)$$

We denote by  $\Pr$ , the probability of the compared node after inference.

**5.5. Graph to Graph correspondence**

The Graph to Graph correspondence is viewed as a node-to-node assignment for a pair of graphs. We define  $\hat{M}(G_i, G_r)$  as the final cost of correspondence and it is the sum of the correspondence operations for the case of two graphs. The distance formula between two graphs  $G_i$  and  $G_r$  is given as follows:

$$\hat{M}(G_i, G_r) = (\sum_{x=1}^X M(w_{sign}(x_i), w_{sign}(x_r))) \quad (13)$$

$\hat{M}$  is the size of the correspondence function that represents the number of matching operations (number of nodes that constitute the prior graph).

**5.6. Case signature definition**

This process is repeated while dealing with all brain tumors' classifiers (tumors list is described in section 2.3). For each case's iteration in a Bayesian classifier, we recover a vector containing the nodes signatures. Those signatures represent a collection of local descriptions for each one of them. After browsing all brain tumors classifiers, the global signature of a case is presented as a matrix of local signatures. Each local signature concerns a local node. We denote this matrix as a cost-matrix and it is schematized as follows:

$$Index_c = \begin{bmatrix} w_{11}^c & \cdot & \cdot & w_{1T}^c \\ \cdot & \cdot & w_{ij}^c & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ w_{x1}^c & \cdot & \cdot & w_{xT}^c \end{bmatrix} \quad (14)$$

Where,  $w_{ij}$  denotes the signature of the  $i^{th}$  node from the network representing the tumor  $j$ . The current structure of the matrix is considered the index of a query case. The comparison step in our algorithm between two cases is translated into a computation of distance between the indexes of concerned cases. Besides, the system must be able to find the most similar cases to the query case on the basis of this index.

In this final step, this score is considered as the correspondence degree between  $C_r$  and  $C_i$  and it is formulated as the following:

$$Score_{C_r, C_i} = \sum_{t=1}^T \hat{M}(G_i^t, G_r^t) \quad (15)$$

The problem can be translated into a correspondence of a cost matrix that defines a vertex-to-vertex assignment for a pair of cases (query case and test case). This correspondence is defined as the following:

$$\hat{M}(Index_i, Index_r) = \begin{bmatrix} w_{11}^i & \cdot & \cdot & w_{1T}^i \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ w_{x1}^i & \cdot & \cdot & w_{xT}^i \end{bmatrix} * \begin{bmatrix} w_{11}^r & \cdot & \cdot & w_{1T}^r \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ w_{x1}^r & \cdot & \cdot & w_{xT}^r \end{bmatrix} \quad (16)$$

The conduct of this measure process is applied to all database cases. At the end of the experiment, we will have a set of scores. Each score describes the degrees of similarity between  $C_r$  and  $C_i$ . Then, this vector will be ranked and the cases that have the smallest degrees are considered as the most similar cases to the query case. Thus, the best measure of correspondence in this problem is considered as a problem of index correspondence based on node signatures.

**6. Experiments and results**

**6.1. Experimental data**

To ensure experiments, we have collected 124 cases from the MRI acquisition center in Sahloul Hospital, Tunisia. Each case folder contains between 8 and 16 images selected by experts. Those cases were previously confirmed and interpreted by radiologists before adding them to the database. We intended to build a heterogeneous database that regroups all tumors indicated above in the table 1. These tumors exist in the database with different rates. In this study, we intend to test the performance of our similarity measure in conflict situations. Indeed, we propose to conduct two experimental sets in order to compare our method performance with other similarity measures.

The first experiments set is a comparison between our present measure (denoted SIM-GC algorithm in what follows) with a classical method based on Bayesian inference and the Euclidean distance for similarity

quantification which is described in Ref. 32 (denoted SIM-BN algorithm in what follows) and a similarity measure based on the k-nearest neighbor method (denoted SIM-KNN in what follows)<sup>33</sup>.

The second one is a comparative study between our similarity measure and a set of distance measures (five distances) selected from the literature. Thus, a brief presentation of those distances is proposed on what follows. The evaluation of these measures is provided by a diverse set of experiments.

In the first experiments set, we opt to diversify initializations in order to observe the influence of nodes initial values on the retrieval results precision. Each test set consists of three sub-classes:

A first experiment based on a parameters distribution referring to an intervention of medical experts. This experiment class is noted as ‘Sim-Expert’ in what follows.

A second experiment is defined by applying complete data learning. We have chosen to apply the maximum Likelihood algorithm<sup>24</sup>. This class is noted as ‘Sim-Train1’. The maximum likelihood training algorithm considers the probability of an event as being its frequency of appearance:

$$\hat{P}(X_i = x_k | Pa(X_i) = x_j) = \frac{N_{i,j,k}}{\sum_k N_{i,j,k}} \quad (17)$$

Where  $N_{i,j,k}$  is the occurrence number of  $X_i, \{X_i = x_k \text{ and } Pa(X_i) = x_j\}$ .

A third experiment with an exercise of incomplete data learning method is proposed. In this context, we propose to apply the EM algorithm<sup>31</sup>. EM is an iterative optimization method to estimate some unknown parameters  $\Theta$ , given measurement data  $U$ . This experiment class is noted as ‘Sim-Train2’. EM estimation consists of maximization of the posterior probability of the parameters  $\Theta$  given the data  $U$ , marginalizing over  $J$  as follows:

$$\Theta^* = \arg \max_{\Theta} \sum_{J \in \mathcal{J}^n} P(\Theta, J | U) \quad (18)$$

In the second set of experiments, we orient experiments to compare our current similarity measure with similarity measures based on distance. The distance measure, often called a dissimilarity measure which measures the rate of difference from one object to

another. Most of the works on similarity measures, in the framework of propositional representations, have as basis the mathematical notion of distance (inverse of the similarity concept) which has been extensively studied in data analysis. In the state of the art, many distances have been proposed. Thus, we selected five distances, which are the followings:

- Bhattacharyya Distance:

The Bhattacharyya distance (denoted BHAT-D in what follows) is related to the Bhattacharyya coefficient which is a measure of the amount of overlap between two statistical samples or populations. For two vectors  $A$  and  $B$ , the Bhattacharyya distance is given as follows<sup>34</sup>:

$$D_B(A, B) = -\ln(BC(A, B)) \quad (19)$$

Where  $BC$  is the Bhattacharyya coefficient.

- Hausdrauff distance:

Apparently, the most natural distance measure for such objects  $A$  and  $B$  is the Hausdorff distance (denoted HAUS-D in what follows) where for each point on one object we consider the closest point on the other one and then maximize over all these values<sup>34</sup>. The Hausdorff distance between  $A$  and  $B$  is formulated as:

$$\delta(A, B) = \max(\bar{\delta}(A, B), \bar{\delta}(B, A)) \quad (20)$$

Where  $\delta(A, B) = \max_{x \in A} \min_{y \in B} \|x - y\|$

- Chi-2 Distance:

The chi-square distance (denoted CHI2 in what follows), always denoted by  $X$ , between two vectors  $s$  is a weighted Euclidean distance. Given two profiles  $x = [x_1, x_2, \dots, x_j]$  and  $y = [y_1, y_2, \dots, y_j]$ , the chi-square distance is defined as:

$$X = \sqrt{\sum_{j=1}^J \frac{1}{c_j} (x_j - y_j)^2} \quad (21)$$

$C_j$  denotes the  $j^{\text{th}}$  element of the average profile, that is the abundance proportion of the  $j$ -th species in the whole data set<sup>34</sup>.

- Vector Cosine Angle distance:

Cosine similarity (denoted COS-D in what follows) is a measure of similarity between two vectors by measuring the cosine of the angle between them<sup>34</sup>. The cosine of two vectors A and B is given as the following:

$$\cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} \quad (22)$$

- Kullback leibler distance:

The Kullback Leibler distance (denoted Kull-D in what follows) is a natural distance function from a ‘true’ probability distribution to a ‘target’ probability distribution<sup>34</sup>. Given a probability distributions  $A = \{a_1, \dots, a_n\}$  and  $B = \{b_1, \dots, b_n\}$ , the KL-distance is defined as:

$$KL(A, B) = \sum_i p_i * \log_2(P_i/q_i) \quad (23)$$

In all experiments described in what follows, we apply Cross validation to assess those methods efficiency. Cross validation is called to compare the performance of two or more different algorithms and find out the best algorithm for the available data, or alternatively to compare the two or more variants of a parameterized model<sup>35</sup>. In this context, we have opted for using the k-fold cross validation. The advantage of this method is the way that the data gets divided matters less. Every data point gets to be in a test set once exactly, and gets to be in a learning set k-1 times<sup>36</sup>. In this work, 5 fold

cross-validations are established in the evaluation process and the process is repeated to provide average results. In each iteration, 4/5 th of each category have been used as a learning set for the computation of experiment results provided by the test set. In each round, the average performance indices are calculated over three randomly selected queries. In total, we have tested 15 queries that cover all fold cross validations.

In the two experiment levels, we used two indices to evaluate the retrieval effectiveness of the two models which are the precision rate and the recall rate. Precision rate P is the number of true retrieved cases in a result window fixed as 10, in accordance with an expert’s needs. It is defined as:

$$P = \frac{\text{relevant cases in top } N}{N} \quad (24)$$

Where N is the preset result window and it is fixed at 10.

The recall rate (R) defines the number of relevant retrieved cases per the number of relevant cases in the database. It is described as follows:

$$R = \frac{\text{relevant retrieved cases in top } N}{\text{Number of total relevant cases}} \quad (25)$$

Therefore, the recall rate and precision can measure the accuracy of retrieval to show the strength of a system for diagnosis.

On what follows, we schematize the performance indices derived from the first experiments set in Table 3.

Table 3. Performance rates comparison in terms of precision (P) and recall (R) indices (%)

Initialization Index	SIM-GC						SIM-BN						SIM-KNN	
	Sim-Expert		Sim-Train1		Sim-Train2		Sim-Expert		Sim-Train1		Sim-Train2		-	
	P	R	P	R	P	R	P	R	P	R	P	R	P	R
1 Fold tests	90	93	60	62	53	62	67	69	30	31	17	17	27	28
2 Fold tests	70	58	60	50	47	50	53	44	53	44	30	25	37	33
3 Fold tests	67	61	70	64	50	45	57	52	27	24	33	30	40	36
4 Fold tests	67	61	50	45	63	58	57	52	23	21	33	30	47	42
5 Fold tests	83	64	60	46	80	62	67	51	37	28	33	26	47	39
average rate %	78	67,4	60	53,5	58,7	51,8	60	53,6	34	29,9	29,4	25,7	39,6	35,6

In a second level, we conducted a comparative study with a varied set of metric distances. This set of comparisons is used to situate our proposed similarity measure in relation to the more used metric distances used in more indexing problems. The conduct

of experiments is similar to those described above. Table 4 illustrates the average precision and recall rates.

Table 4. Performance rates comparison with distance metrics (%)

Experiment Set	SIM-GC		Metric Distances									
			HAUS-D		CHI2-D		COS-D		BHAT-D		KULL-D	
Method	P	R	P	R	P	R	P	R	P	R	P	R
Index	P	R	P	R	P	R	P	R	P	R	P	R
1st Fold tests	90	93	50	52	53	55	53	55	0,33	34	53	55
2 Fold tests	70	58	30	27	30	27	33	30	0,30	27	37	33
3 Fold tests	67	61	40	36	37	33	37	33	0,43	39	37	33
4 Fold tests	80	73	47	42	43	39	43	39	0,33	30	43	39
5 Fold tests	83	64	47	39	50	42	43	36	0,43	36	53	44
average rate %	78	69,77	42,67	39,33	42,67	39,37	42,0	38,86	36,67	33,51	44,67	41,14

6.2. Discussion

This work described a similarity measure incorporated into a contribution of medical cases retrieving. In the experimental phase, we aimed to validate this method by testing it on cases classified as conflict cases. The following figure (fig. 6) shows the performance variation of the applied algorithms.

As shown in this figure, the SIM-GC algorithm represents a clear supremacy compared to the SIM-BN algorithm and the SIM-KNN algorithm. As an example, the average precision of Sim-Expert experiments from the SIM-GC algorithm is at 78% while it is about 57.5% from SIM-BN. Moreover, the experiment results based on initialization that refers to expert knowledge (Sim-expert) exceed those coming from the experiments

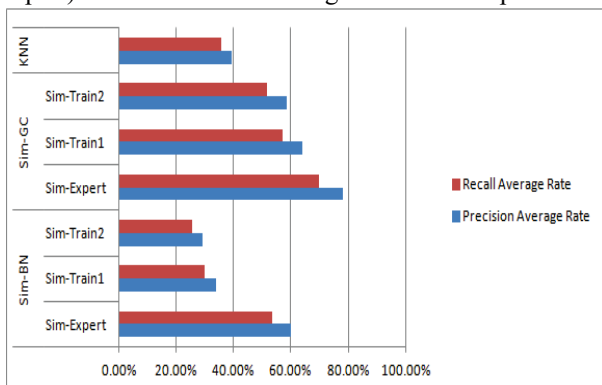


Fig.6. Performance indices comparison between Sim-GC, Sim-BN and Sim-KNN

referred to parameters learning algorithm. Equally, our algorithm clearly exceeds the (SIM-KNN) algorithm in terms of the same performance indices.

In the second set of experiments, the supremacy of the current method continues to be confirmed. This observation is deduced in regard to the performance rates (see fig. 7). By measuring the correct answers rate of the current measure, we can deduce that it clearly exceeds the same rate of other distances included in desk sets. So, the best overall results were obtained using our proposed similarity distance (SIM-GC). Indeed, it was as high as 78.00% of correct answers. In addition, the Bhattacharya distance was the worst distance with a rate of 37%.



Fig.7. Performance indices comparison between the various tested distances

As a summary, those results proved that the actual similarity measure gives the best performance in terms of precision of our retrieval task. Thus, these results are acceptable and it is advisable to further refine them. Developing the performance is closely dependent on the improvement of the cases database.

## 7. Conclusion

In this paper, we have proposed a similarity measure algorithm applied to a MRI brain tumors cases retrieval contribution. This algorithm is based essentially on a Bayesian Network graph nodes signature comparison. In real context, the decision concerning a brain tumor interpretation process is complicated and it is deduced from several information sources. Our efforts have been oriented to define an appropriate representation of this problem using Bayesian network and to formulate a similarity measure deduced from all rules of radiologist reasoning facing a cerebral case diagnosis. Thus, we proposed an experiment sets that are based on clinical data collected from a MRI acquisition center. Experiments were directed to test the algorithm performance while comparing them to conflict cases chosen by the experts. The retrieval results are satisfactory and they mark a supremacy compared to several methods of similarity quantification.

The topic area discussed in this paper offers several research perspectives. We plan to investigate more about Bayesian inference algorithms for the formulation of a relevant similarity measure. In this context, we aim to develop an approach that highlights the impact of the most influential attributes in determining the medical interpretation result. Hence, the objective is to imitate the reasoning of a radiologist. For this, we need to provide the tracing propagation path information leading to the final diagnosis.

## References

1. M. Nilsson et al. (eds.), *Advancements and Trends in Medical Case-Based Reasoning: An Overview of Systems and System Development*, (2004).
2. B. V. Ginneken et al., Computer-aided diagnosis in chest radiography: A survey, in *IEEE Transactions on Medical Imaging* (2001), pp. 1228–1241.
3. B. J. Erickson and B. Bartholmai, Computer-aided detection and diagnosis at the start of the third millennium, *J Digit Imaging*, 15(2),(2002), 59–68.
4. M. O. Güld et al., A generic concept for the implementation of medical image retrieval systems, *Int J Med Info* (2006).
5. L. G. Shapiro et al., Similarity-Based Retrieval for Biomedical Applications. *Case-Based Reasoning on Images and Signals Studies in Computational Intelligence*, 73 (2008) 355-387.
6. I. Bichindaritz and C. Marling, Case-based reasoning in the health sciences: What's next?, *Artif Intell Med.* 36(2)(2006) 127–135.
7. S. Montani, How to use contextual knowledge in medical case-based reasoning systems: a survey on very recent trends, *Artif Intell Med.* 51(2)(2011) 125–131
8. H. Greenspan and A. Pinhas, Medical Image Categorization and Retrieval for PACS Using the GMM-KL Framework, *IEEE Transactions on Information Technology in Biomedicine*, 11(2)(2007) 190–202.
9. Y. Liu et al., Semantic-based biomedical image indexing and retrieval, in *Proc. International Conference on Diagnostic Imaging and Analysis*, (2002), pp. 18–20.
10. S. Clinchant et al., Semantic combination of textual and visual information in multimedia retrieval, in *Proc. 1<sup>st</sup> ACM International Conference on Multimedia Retrieval*, (New York, USA, 2011), pp. 1-8.
11. C. B. Akgül et al., Content-Based Image Retrieval in Radiology: Current Status and Future Directions, *J Digit Imaging*, 24 (2) (2011) 208–222.
12. C.-R. Shyu et al., ASSERT: A Physician-in-the-Loop Content-Based Retrieval System for HRCT Image Databases, *Computer Vision and Image Understanding*, 75(1) (1999) 111–132.
13. T. M. Lehmann et al., The IRMA Project: A State of the Art Report, in *Proc. Korea-Germany Workshop on Advanced Medical Image*, (2003), pp. 161–171.
14. D. Nikovski, Constructing Bayesian Networks for Medical Diagnosis from Incomplete and Partially Correct Statistics, *IEEE Trans. on Knowl. and Data Eng.*, 12(4) (2000) 509–516.
15. I. J. Cox et al., The Bayesian image retrieval system, PicHunter: theory, implementation, and psychophysical experiments, *IEEE Transactions on Image Processing*, 9(1) (2000) 20–37.
16. J. Luo et al., A Bayesian network-based framework for semantic image understanding, *Pattern Recognition Journal*, 38(6) (2005) 919–934.
17. G. Quellec et al., Multimodal medical case retrieval using Bayesian networks and the Dezert-Smarandache theory, in *Proc. 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, (2008), pp. 245–248.
18. D. N. Louis et al., The 2007 WHO Classification of Tumours of the Central Nervous System, *Acta Neuropathol Journal*, 114(2) (2007) 97–109.
19. Y.-E. L. Koo et al., Brain cancer diagnosis and therapy with nanoplateforms, *Adv. Drug Deliv. Rev.*, 58(14) (2006) 1556–1577.

20. R. N. Al-Okaili et al., Advanced MR Imaging Techniques in the Diagnosis of Intraaxial Brain Tumors in Adults, *Radiographics Journal*, 26(1) (2006) 173–189.
21. O. Pourret et al. (eds.), *Bayesian Networks: A Practical Guide to Applications*, (1st ed. Wiley, 2008).
22. J. Pearl (eds.), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, 1st edn. (Morgan Kaufmann, 1988).
23. N. Harris et al., Criticizing Conditional Probabilities in Belief Networks, in Proc. *Annu Symp Comput Appl Med Care*, (1990), pp. 805–809.
24. J. Cheng et al., Learning Bayesian Networks from Data: An Information-Theory Based Approach, *Artificial Intelligence*, 137(1-2) (2002) 43–90.
25. S. Jouili and S. Tabbone, Graph Matching Based on Node Signatures, in Proc. *the 7th IAPR-TC-15 International Workshop on Graph-Based Representations in Pattern Recognition*, (Berlin, Heidelberg, 2009), pp. 154–163.
26. L. Torresani et al. “Feature Correspondence Via Graph Matching: Models and Global Optimization,” in Proc. the 10th European Conference on Computer Vision, (Berlin, Heidelberg, 2008), pp. 596–609.
27. R. Bergmann and A. Stahl, Similarity measures for object-oriented case representations, *Advances in Case-Based Reasoning Lecture Notes in Computer Science*, 1488 (1998) 25–36.
28. H. Bunke and B. T. Messmer, Similarity measures for structured representations, in *Topics in Case-Based Reasoning*, (Eds. Springer Berlin Heidelberg, 1994), pp. 106–118.
29. P. Cunningham, A Taxonomy of Similarity Mechanisms for Case-Based Reasoning, *IEEE Transactions on Knowledge and Data Engineering*, 21(11) (2009) 1532–1543.
30. F. Campillo et al., Computational probability modeling and Bayesian inference, *ARIMA*, 9 (2008) 123–143.
31. K. P. Murphy et al., Loopy Belief Propagation for Approximate Inference: An Empirical Study, in Proc. of *Uncertainty in AI*, (1999), pp. 467–475.
32. H. Yazid et al., MRI cases containing cerebral tumors retrieval using Bayesian networks, in Proc. *the 10th IEEE International Symposium on Signal Processing and Information Technology*, (Washington, DC, USA, 2010), pp. 7–12.
33. P. Zezula et al. (eds.), *Similarity Search - The Metric Space Approach*, (1st Springer Publishing, 2010).
34. L. Leydesdorff, *Similarity Measures, Analysis, and Information Theory*, CoRR, abs/0911.4292 (2009).
35. R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, in Proc. *the 14th international joint conference on Artificial intelligence*, (San Francisco, CA, USA, 1995), pp. 1137–1143.
36. B. Efron, Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation, *Journal of the American Statistical Association*, 78(382) (1983) 316–331.