



HAL
open science

High Level Mammographic Information Fusion For Real World Ontology Population

Yosra Ben Salem, Rihab Idoudi, Karim Saheb Ettabaa, Kamel Hamrouni,
Basel Solaiman

► **To cite this version:**

Yosra Ben Salem, Rihab Idoudi, Karim Saheb Ettabaa, Kamel Hamrouni, Basel Solaiman. High Level Mammographic Information Fusion For Real World Ontology Population. *Journal of Digital Information Management*, 2017, 15 (5), pp.259 - 271. hal-01836406

HAL Id: hal-01836406

<https://hal.science/hal-01836406>

Submitted on 9 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

High Level Mammographic Information Fusion For Real World Ontology Population

Yosra Ben Salem^{1,2}, Rihab Idodi^{1,2}, Karim Saheb Etabaa², Kamel Hamrouni¹, Basel Solaiman²

¹ Sound, Image and Information Processing laboratory (LR-SITI)
Université Tunis El Manar, Ecole Nationale d'Ingénieurs de Tunis
Campus Universitaire farhat HACHED el Manar; BP 37
Le Belvedere 1002 Tunis, Tunisia



*Journal of Digital
Information Management*

²Image and Information Processing Laboratory (ITI)
Campus de Brest; Technopôle Brest-Iroise
CS 83818, 29238 Brest cedex 03, France

ABSTRACT: *In this paper, we propose a novel approach for ontology instantiating from real data related to the mammographic domain. In our study, we are interested in handling two modalities of mammographic images: mammography and Breast MRI. Firstly, we propose to model both images content in ontological representations since ontologies allow the description of the objects from a common perspective. In order, to overcome the ambiguity problem of representation of image's entities, we propose to take advantage of the possibility theory applied to the ontological representation. Second, both local generated ontologies are merged in a unique formal representation with the use of two similarity measures: syntactic measure and possibilistic measure. The candidate instances are, finally, used for the global domain ontology populating in order to empower the mammographic knowledge base. The approach was validated on real world domain and the results were evaluated in terms of precision and recall by an expert.*

Subject Categories and Descriptors

J.3 [Medical information systems]; I.4.10 [Image Representation] H.3.1 [Content Analysis and Indexing]

General Terms:

Mammographic Image Processing, Ontology

Keywords: Ontology, Possibility Theory, Local Ontologies Merging, Similarity Measure, Mammographic Domain.

Received: 19 April 2017, Revised 28 May 2017, Accepted 4 June 2017

1. Introduction

Due to the complexity and susceptibility of the mammographic domain, the necessity of having a reliable domain knowledge base is crucial (Idoudi and al., 2014). Since, ontologies provide a formal, structured knowledge representation, and have the advantage of being reusable and shareable; numerous ontological representations have been proposed in literature MammOnto (Hu and al., 2003) GIMI Mammography Ontology (Taylor and Toujilov, 2012), BCGO (Tutac, 2010). Although these proposals are rich in terms of concepts relations and axioms, they do not consider instances or individuals. In fact, the latter are relevant in the scope of the mechanisms related with Information Retrieval (Park and al., 2011) (Chen and al., 2010), Service Discovery (Zhang and al., 2009), Question Answering (Yang and al., 2009), and Searching for contents and information (Jiang and al., 2009). To overcome this limitation, we propose an ontology learning (i.e, ontology population) approach related to the mammographic domain. In fact, the implementation of mammography acquisition systems makes available a large volume of information (mammography, IRM, etc.) providing a decisive competitive knowledge. In order to be reusable, these information needs to be processed and archived within a structured knowledge base. To this end,

we propose, initially, to transform the medical image into a local ontological representation in order to semantically formalize the image content in terms of concepts and relations. However, this transformation is pervaded with uncertainty and ambiguity. To this end, we propose to handle uncertain information with the use of the possibility theory. After generating, a possibilistic local ontology for each imaging modality (RMI, mammography), we proceed to merge, for each case, its relative ontologies in order to obtain a unique representation which refers to an instance or individual. The case's entities are then mapped to concepts, attributes and semantic relations of a given global ontology (the knowledge base target).

The remainder of the paper is organized as follows: Section 2 presents a survey of recent works dealing with image based knowledge modelling using ontologies and ontology population; the following section gives the basic notions of the ontology paradigm and the possibility theory; the next section describes the proposed approach as well as its main steps; and finally, the last section provides experimentations tests and validation results.

2. Related Works

Ontologies are widely used in different domains to conceptualize the domain knowledge. To construct domain ontologies multiple sources are used: text, images, signals, etc. Most studies used text to construct domain ontologies in order to formalize the corresponding knowledge. Several approaches used images to define a domain knowledge. This requires a transforming step of the image to an ontology. In this context, Gowsikhaa and al. (Gowsikhaa and al., 2012) proposed an approach to build an ontology from an image. First, three types of features are extracted from the image: Shape features, color features, and texture features. Then, a classification step is performed to group similar entities according to the domain specification in classes. These entities used to manually construct the ontology by defining concepts and relations in accordance with the domain terminology. Finally, the ontology is instantiated by the features values extracted from the image. Wang and al. (Wang and al., 2008) constructed a multi-modal ontology. It is constructed from an image and the Wikipedia text. The ontology is defined by words from Wikipedia and features extracted from the image. From the visual characteristics, 16 concepts are defined to describe the color and the texture of the image. Iskandar and al. (Iskandar and al. 2007) use the "Equal Weight Linear Combination" (EWLC) technique for annotating an image with the corresponding concepts. The latter, are used to specify the ontology. The ontology is enriched by new concepts obtained from other images. In (Liu and al., 2010), authors proposed to transform an image to an ontology using low-level features (color features, shape features and texture features) and SIFT (Scale Invariant Feature Transform) characterizing entities contained in the image. Wang and Yang (Wang and Yang, 2011) proposed an approach to model the content of an image with a specific ontology. At first, different features

are extracted from an image. These features are grouped into classes characterized by relations. This step provides a knowledge base used to annotate the image. Finally, the ontology is constructed from this annotation.

Ontology population is applied in multiple fields. It uses, mainly, texts to extract instances and then matches concepts and relations of instances to the domain ontology. In (Reyes-Ortiz and al., 2016) an approach is proposed for ontology population from web services to facilitate their discovery. Ontology population is performed from classified web services based on features extracted from the text. The algorithm creates an instance into the global ontology with the corresponding category assigned to a web service. In (Lima and al. 2014) an approach based on Inductive Logic Programming is proposed for ontology population from the web. To perform the ontology population, a semantic similarity is measured between ontology classes and instance classes. In (Imsombut and Sirikayon, 2016), an ontology population approach is proposed for the tourism domain. The methodology is based on Instance Extraction by CRF method, then Relation Extraction applying Lexical and syntactic Patterns. For medical studies, Amato and al. proposed an approach (Amato and al., 2015), in order to structure unstructured knowledge from medical records by an ontological approach. The main object of this study is the enrichment of a domain ontology. However, it reserves a population layer for the instantiation of the global ontology automatically. The population process is performed with a statistical and semantic approach by matching entities with RDF triples extracted using semantic tags.

3. Basic Notions

3.1 Ontology

Ontology is a new paradigm used mainly in the semantic web application. It is defined as a hierarchical specification (Manzoor and al., 2015) of a particular reality conceptualization. Ontologies are used in order to formalize knowledge, in a generic form, by providing a commonly agreed upon understanding of the domain of interest. They have been widely put in service of the medical domains to catch knowledge and to formalize the medical lexicons. An ontology is constituted basically of concepts (are the pertinent entities for specification), relations between concepts, individuals, and axioms. Each concept has three basic components: labels, attributes, and relations. Labels designate the names used to refer to a particular concept and can include a set of synonyms that refer to the same entity. Attributes represent the set of characteristics (or features) of a concept that describe the concept in details. Finally, relations are used to represent the different associations among the concepts and to represent consequently the general structure to the ontology (Guber, 1993). For example, a mammogram is a concept within the mammographic domain, whereas relations designate the interactions between revealed concepts, for example, mammogram classified_as Bi-Rads3 (Breast Imaging Reporting and Data System), that's

to say, the concepts mammogram and Bi-Rads3 are related via the relationship classified_as.

An ontology O is a four-tuple (C, R, I, A) , where:

- C is a set of concepts,
- R is a set of relations,
- I is a set of instances,
- A is a set of axioms

3.2 Possibility Theory

Extracted information from images often suffers from imperfections due to some conditions. The most considered imperfections are:

- **Imprecision:** which means that the information is insufficient to understand the case. Ex. "The mass diameter is between 2cm and 3cm."
- **Uncertainty:** which means that the information is unknown if it is true or false. Ex. "The mass can be malignant". There is two cause for this imperfection: the variability of phenomena and the incomplete nature of the information (Dubois and Prade, 1988)
- **Ambiguity:** which means that the information admits different interpretations. Ambiguity is the miss of the specificity of a situation. Ex. "The left breast contains a pathological zone".

Many theories have been introduced to handle these imperfections. Possibility theory is a relatively new theory devoted to handle uncertainty in the context where the available knowledge is only expressed in an ambiguous form. This theory was first introduced by Zadeh (Zadeh, 1978) and developed by Dubois and Prade (Dubois and Prade, 1988). It is, principally, based on the possibility measure, called π which is a membership function with values in $[0,1]$. That is to say a coefficient is assigned to every case that express the degree of the possibility occurrence of the case. $\pi(C) = "0"$ refers to impossible case, while $\pi(C) = "1"$ signifies that the case is totally possible. Nevertheless, some cases and their complements could exist at the same time (the possibility of a case C is evaluated to 1 and the possibility of its complement C^c is also evaluated to 1). For this situation, the possibility measure is not enough to describe the uncertainty of the case. Hence, a necessity measure function (denoted N) is developed by Dubois (Dubois, 1980). This measure, express the degree of the certainty of a case characterized by a specific possibility distribution. $N(C) = "0"$ implies that the case is very uncertain while $N(C) = "1"$ indicates that the certainty of the case is absolute.

Relations between Possibility and Necessity Measures are expressed by these equations:

$$\forall A \in \rho(\Omega) N(A) = 1 - \Pi(A^c) \quad (1)$$

$$\forall A \in \rho(\Omega) N(A) = \text{INF}_{x \in A} (1 - \pi(x)) \quad (2)$$

$$\Pi(A) \geq N(A) \quad (3)$$

$$\text{Max}(\Pi(A), 1 - N(A)) = 1 \quad (4)$$

$$\text{If } N(A) \neq 0 \text{ Then } \Pi(A) = 1 \quad (5)$$

$$\text{If } \Pi(A) \neq 1 \text{ Then } N(A) = 0 \quad (6)$$

$$N(A) \leq Pr(A) \leq \Pi(A) \quad (7)$$

with $Pr(A)$: the probability measure associated to A .

3. Proposed Approach

The proposed approach process two images as input to populate a domain ontology. It is composed of four stages as illustrated in in *Figure 1*: the first stage aims to model the knowledge extracted from an image in a local ontology based on an existing domain ontology; the next stage attempts to fuse the two local ontologies into one global ontology which represents an instance of the domain knowledge; the third stage performs the ontology instantiating; and finally, the last stage aims to validate the ontology making by checking its consistency.

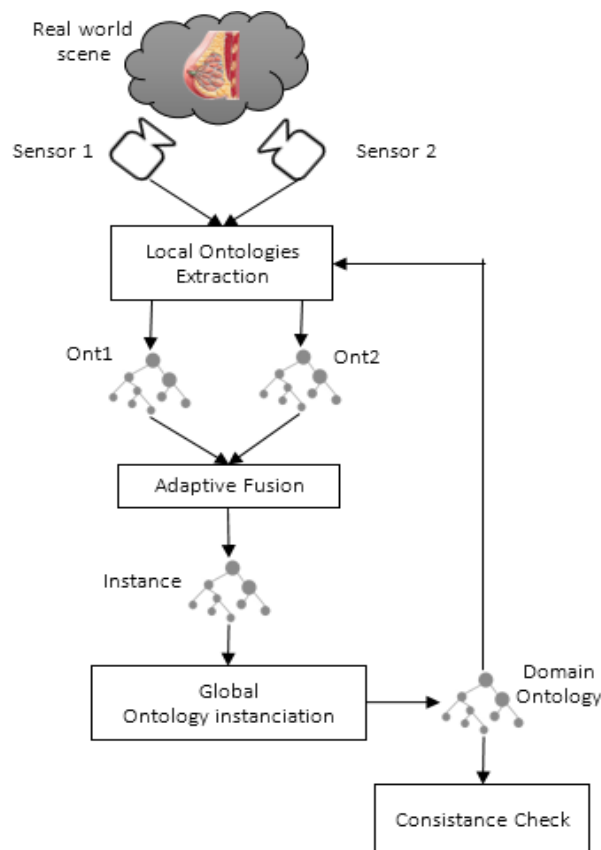


Figure 1. The proposed approach

3.1 Image based Ontology Extraction

The proposed method for the local ontologies construction using an image is based on an anterior work (Ben Salem

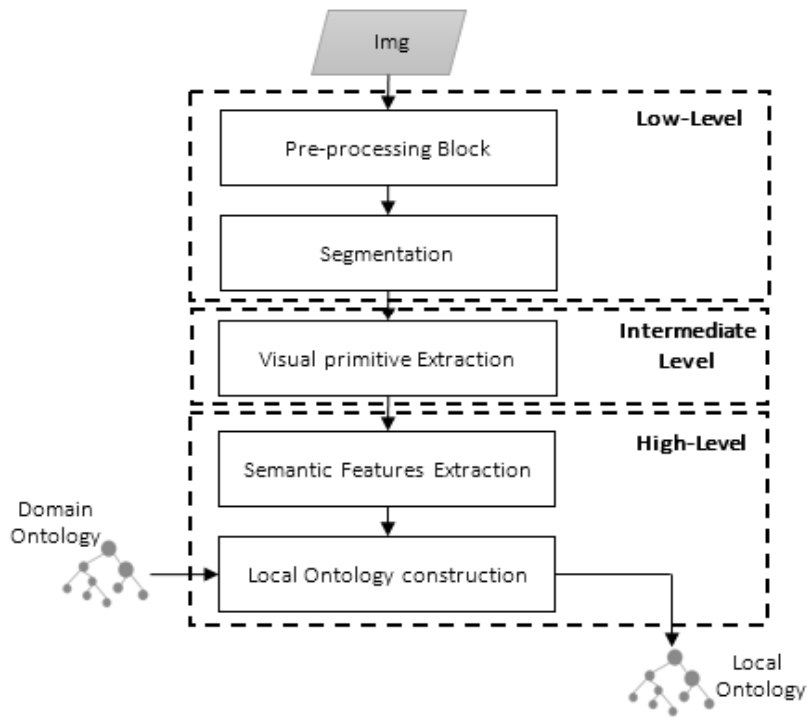


Figure 2. Multi-level procedure to extract ontology from an image

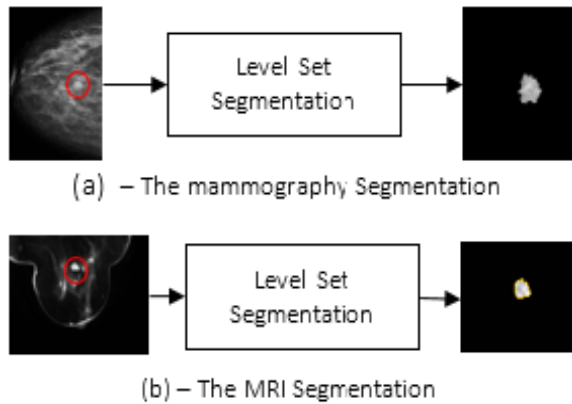


Figure 3. The segmentation step

and al., 2016). As explained in this works, the method is defined according three levels: the low level processing in order to extract the region of interest, the intermediate level for the visual features extraction and the high level to extract semantic features. The latter are mapped to a domain ontology then used to specify the local ontology. *Figure 2* shows the different steps followed to obtain this local ontology.

The low level processing aims to extract the salient regions following two steps. The first step is a preprocessing step aiming to enhance the quality of the image, reduce the noise and eliminate the non significant forms (labels, artifacts, etc). This step is performed by applying the RID (Regularized Isotropic Diffusion) Filter. In the segmentation step, the relevant regions of the image are extracted using the Level Set method. *Figure 4* depicts examples of the segmentation step of a pathological form from a

mammography image (*Figure 3 (a)*) and another one from a breast MRI image (*Figure 3 (b)*).

The intermediate level aims to specify visual features, which can be extracted from segmented forms. Using these forms we compute several characteristics such as the shape (area, diameter, circumscribed ellipse, etc), the contour (perimeter, convexity, etc) and the texture (Gray Level Co-occurrence Matrix GLCM). The different features are collected into a numeric signature file (cf. *Figure 4*). This level represents the bridge allowing passing from the low level representation of data to the semantic meaning of entities by computing some parameters (Circularity, compactness, etc).

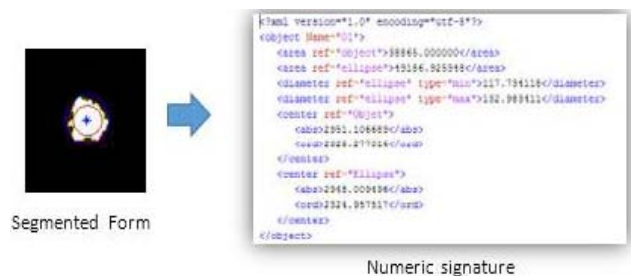


Figure 4. The feature extraction step

The last level (High level) of processing, attempts to, firstly, deduce the semantic characteristics from the visual primitives according to certain rules predefined in a rule base. As example of rules, we propose:

- “If Circularity = 1 Then The form is circle”;
- “If the compactness < s Then the contour is irregular”

(where s: a threshold value).

These rules are defined according directions from experts. To deal with imperfections that can affect the deduced information, we specify some features (shape, contour, localization, distribution, size) as possibilistic features. Then, a possibility measure is assigned to each one of them. This measure indicates the degree of the correctness of a semantic characteristic to a corresponding descriptor. Different possibility distributions are defined by some parameters according to expert's recommendation. The possibility measure assigned to a certain feature is, then, computed as the minimum values of measures selected from distributions defined for parameters specifying this feature. These semantic characteristics are specified in a semantic signature illustrated by an extract in the Figure 5.

In a second step, these semantic features are mapped to a domain ontology. In the context of the breast knowledge definition, we use the "Mammo" Ontology specified in (Idoudi and al., 2016). Figure 6 represents an extract from the "Mammo" Ontology. The different features are mapped

to the concepts defined in the ontology specification. This process results in a local ontology characterized by weights that represent a possibilistic measures. The local ontology is described by the Figure 7 and the Figure 8.

```

<descriptors>
  <form>
    <round>0.3</round>
    <oval>0.6</oval>
    <irregular>0.2</irregular>
    <lobular>0.05</lobular>
  </form>
  <contour>
    <regulier>0.8</regulier>
    <indistinct>0.2</indistinct>
    <irregulier>0.05</irregulier>
  </contour>
  <distribution>
    <singular>100</singular>
    <dispersed>0</dispersed>
    <concentrated>0</concentrated>
  </distribution>
</descriptors>

```

Figure 5. Extract from the semantic signature of an image

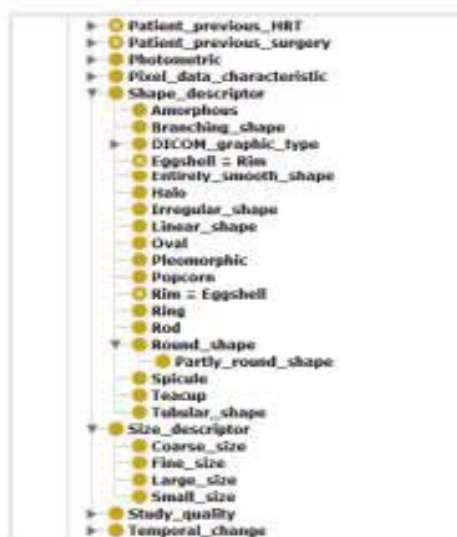


Figure 6. Extract from The Mammo Domain Ontology

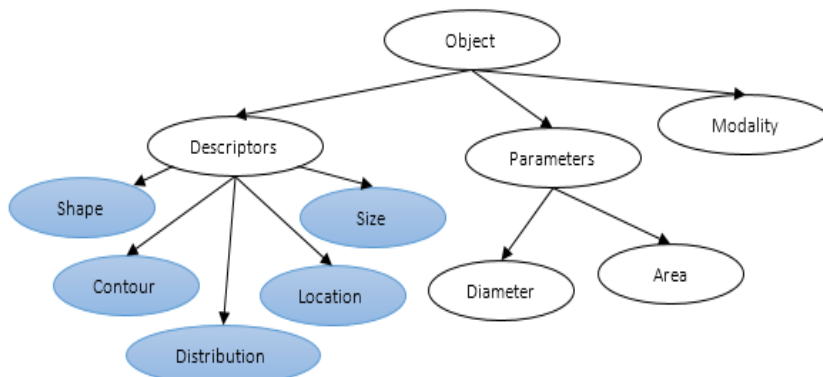


Figure 7. The Local Ontology Definition

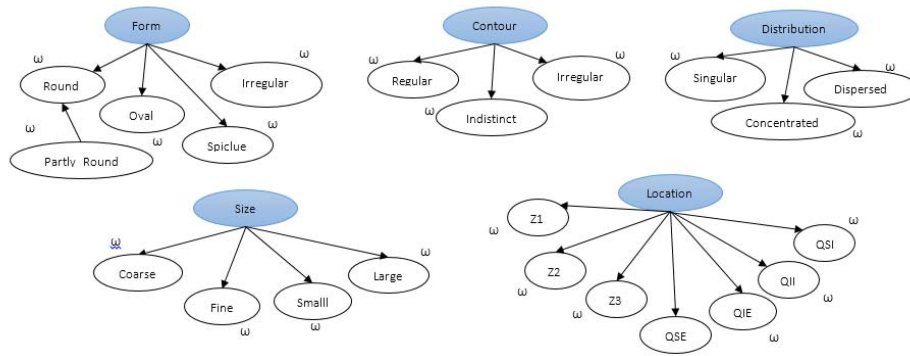


Figure 8. Detailed Concepts of the local ontology

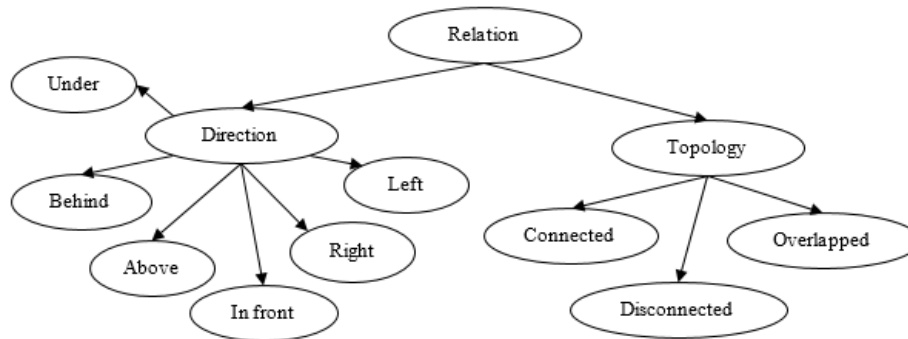


Figure 9. The relation specification of the local ontology

Relations between entities are defined according to the position of the different objects in the image. As shown in the *Figure 9* we propose two types of relations:

- Direction relations which specifies the spatial position of an object in relation with another object;
- Topology relation which defines the link between two objects (connected, disconnected or overlapped). The resulting local ontology describe formally the set of the pathological objects contained in an image.

3.2 Ontologies Fusion

In a fusion process, the most important step is the similarity measure. To perform the fusion between the local ontologies constructed from each modality, two main steps are followed: the similarity measure step applied to the extracted local ontologies and the fusion step (cf. *Figure 10*). As shown in *Figure 10*, the similarity measure and fusion are performed according to the type of the corresponding concept.

3.2.1 Similarity Measure

To compare local ontologies it is necessary to measure the similarity between their correspondent objects. Two similarity measures are used to compare the ontologies: a possibility measure applied to concepts associated with possibilistic weights and a syntactic measure applied to the rest.

Syntactic Similarity Measure

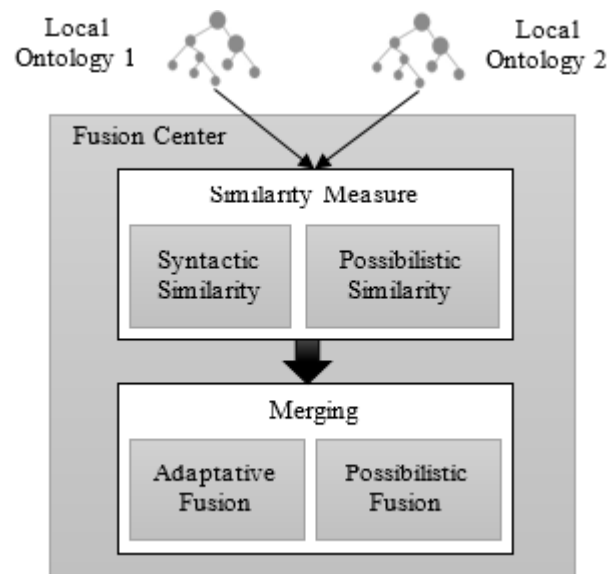


Figure 10. The fusion process

The syntactic similarity measure is applied to attributes defined syntactically according to the Equation Eq 08.

$$S_{sa} = \begin{cases} 1 & \text{if } val(a_{c11}) = val(a_{c12}) \\ 0 & \text{if } val(a_{c11}) \neq val(a_{c12}) \end{cases} \quad (8)$$

where :

- S_{sa} is the similarity measure of the ontologies according to the attribute a .

• $Val(a_{ci})$: The value assigned to the attribute 'a' of a concept c in the local ontology i.

Possibilistic Similarity Measure

As some concepts in the local ontology are defined with attributes to which we assigned a possibility values, it is necessary to apply a possibilistic similarity measure with these attributes. Many possibilistic similarity measures are defined in the literature. We use the Information Affinity (InfAff) proposed by Jenhani and al. in (Jenhani and al., 2007). This measure is defined by the Equation Eq 09.

$$S_{IA}(\pi_i, \pi_j) = 1 - \frac{D_M(\pi_i, \pi_j) + Inc(\pi_i, \pi_j)}{2} \quad (9)$$

Where:

• D_M : One of the variants of the Distance of Minkowski (Hertz, 2006). For example, the Distance of Manhattan defined by Equation 10:

$$L_1(\pi_i, \pi_j) = \frac{1}{N} \sum_{k=1}^N |\pi_i(x_k) - \pi_j(x_k)| \quad (10)$$

• $Inc(\pi_i, \pi_j)$: The Inconsistency measure between π_i and π_j . This measure considers the inconsistency of the intersection of the two possibility distributions.

Overall Similarity

To measure the overall similarity between the concepts of two local ontologies, we use the equation below:

$$Sim(C_{01}, C_{02}) = \frac{1}{L} \sum_{i=1}^L S_i(A_{01}, A_{02}) \quad (11)$$

Where L represents the number of attributes of a concept.

Example:

Let's take the concept 'form' as an example. It disposes of these attributes: {Round, Oval, Partially_round}

$$S_{IA}(\pi_{Round1}, \pi_{Round2}) = 0, 3; S_{IA}(\pi_{Oval1}, \pi_{Oval2}) = 0, 4; S_{IA}(\pi_{Partially_Round1}, \pi_{Partially_Round2}) = 0, 8$$

The overall similarity for the concept 'form' is: $Sim(C_{01}, C_{02}) = 1/3 (0.3 + 0.4 + 0.8) = 0.5$

Combination

The two kinds of proximity computed above are combined to form the overall proximity. To this end, we propose the equation below:

$$Sim(O_1, O_2) = \sum_{k=1}^N w_k S_k(A_{01}, A_{02}) \quad (12)$$

where :

• N : is the number of concepts

• $S_k(A_{01}, A_{02})$ is the similarity measure between the two set of attributes A_{01} and A_{02} of the concept k; where:

• $S_k(A_{01}, A_{02}) = S_{IA}(A_{01}, A_{02})$ if A_{01} and A_{02} are attributes associated with possibilistic weights.

• $S_k(A_{01}, A_{02}) = S_{sa}(A_{01}, A_{02})$ if A_{01} and A_{02} are defined syntactically.

• w_k : A ponderation coefficient defined by the expert as shown in Table 1.

Concepts	Ponderation Coefficient w_k
Form	0.03
Contour	0.04
Distance	0.1
Center	0.12
Size	0.18
Location	0.3
distribution	0.2
Weight	0.01
Diameters	0.02

Table 1. Ponderation coefficients for attributes

At the end of the comparison process, we obtain a possibilistic similarity matrix as shown in Figure 11.

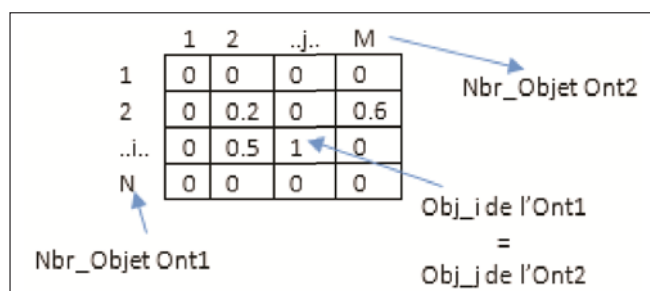


Figure 11. The similarity matrix

3.2.2 Fusion Process

The fusion is performed over the local ontologies using the similarity degree between different objects. As we use two types of attributes (possibilistic and deterministic), two fusion operators are, also, used: a possibilistic operator for possibilistic attributes and another operator for deterministic attributes.

Deterministic Fusion

For the deterministic fusion, we combine values of attributes in an interval. In fact, the sources can provide not exactly the same value for an attribute. Then to model this imperfection, we propose to combine the two values in an interval. The expert, will have an approximate range of the value.

Example:

For the distance concept, one source provide 2.5 for the distance from the nipple and the other source provide the value 2.1 for the same distance. These two values are combined in the interval [2.1, 2.5] which means that the distance of the mass from the nipple is between 2.1 and 2.5.

Possibilistic Fusion

For the possibilistic fusion, we distinguish three fusion operators:

- The convergent operators that are used if we assume that sources are reliable and not contradictory. The most used operators are T-Norm Operators (Martin, 2005)
- The divergent operators that are used when we have to collect the maximum of information. The most used operators are the T-CoNorm Operators (Martin, 2005)
- The adaptive operators that are used if we want to get a maximum of reliable information when we can not decide about the reliability of sources. An adaptive operator is proposed by Dubois and Prad in (Dubois and Prade, 1994). It consists of a convergent operator if the source is reliable and a divergent operator if the source is not reliable. It is defined by the equation below:

$$\pi_{DP}(x) = \max\left[\frac{\pi_A(x)}{h(\pi_1, \pi_2)}, \min(\pi_V(x), 1 - h(\pi_1, \pi_2))\right] \quad (13)$$

For our study, as we can not judge the reliability of sources, we apply the adaptive fusion. We use then the Dubois and Prad's equation.

The Figure 12 presents the fusion process of the two objects ontologies.

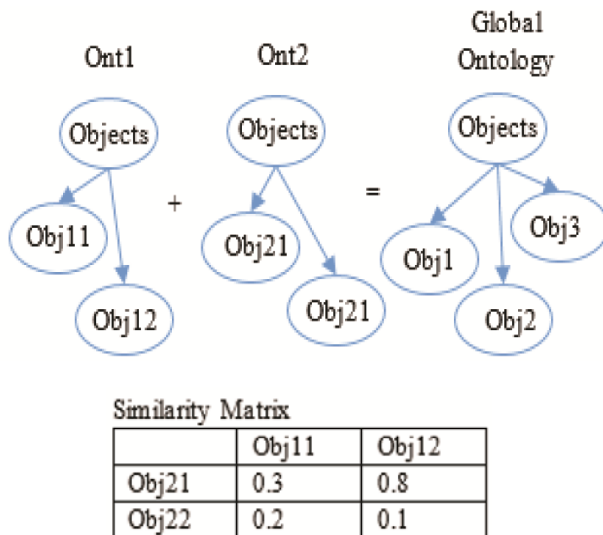


Figure 12. The Fusion process of two local ontologies

3.2.3 Ontology Population

The resulting ontology represents an instance of the domain ontology. The latter is then instantiated by the resulting ontology. The instantiating process can be performed only with deterministic values. Nevertheless, our ontology is defined by possibilistic weights. For this reason, we propose a specific ontology population for possibilistic concepts using a xml file associate with the corresponding instance in the domain ontology. The ontology population and enrichment process iterates through the following steps:

Algorithm -1

For each case:

- (1) **Ontology population:** The instance of the domain ontology annotating the medical images is mapped to the existing concepts as well as relations.
- (2) Create a xml file (cf. Figure 13) comprising the possibilistic degrees related to ambiguous concepts.
- (3) **Validation:** The ontology consistency is checked by the inference engine.

End For

```

▼<Individuals>
  ▼<individual>
    <NamedIndividual IRI="Cas1"/>
    ▼<Concept IRI="Round_shape">
      <degree value="0"/>
    </Concept>
    ▼<Concept IRI="Oval">
      <degree value="0.8"/>
    </Concept>
    ▼<Concept IRI="Irregular_shape">
      <degree value="0.3"/>
    </Concept>
    ▼<Concept IRI="Large_size">
      <degree value="0.1"/>
    </Concept>
    ▼<Concept IRI="Small_size">
      <degree value="0.8"/>
    </Concept>
  </individual>
  ▼<individual>
    <NamedIndividual IRI="Cas2"/>
    ▼<Concept IRI="Partly_round_shape">
      <degree value="0.4"/>
    </Concept>
    ▼<Concept IRI="Irregular_shape">
      <degree value="0.8"/>
    </Concept>
    ▼<Concept IRI="Large_size">
      <degree value="0.1"/>
    </Concept>
    ▼<Concept IRI="Small_size">
      <degree value="0.75"/>
    </Concept>
  </individual>
</Individuals>

```

Figure 13. Extract from an example of an instantiation file

4. Results and Experimentation

The evaluation stage is performed to prove the performance of our approach and the effectiveness of results. For this purpose, we use three data sets: the Mini-Mias DataBase which contains 98 pathological cases for evaluating the mammography processing; the TCIA (The Cancer Imaging Archive) Breast MRI DataBase which contains 429 volumes for 89 subjects for evaluating the MRI processing; and for the evaluation of the fusion step we dispose of 35 mammography images correspondent to 5 patients, and 134 MRI volume correspondent to the same 5 patients.

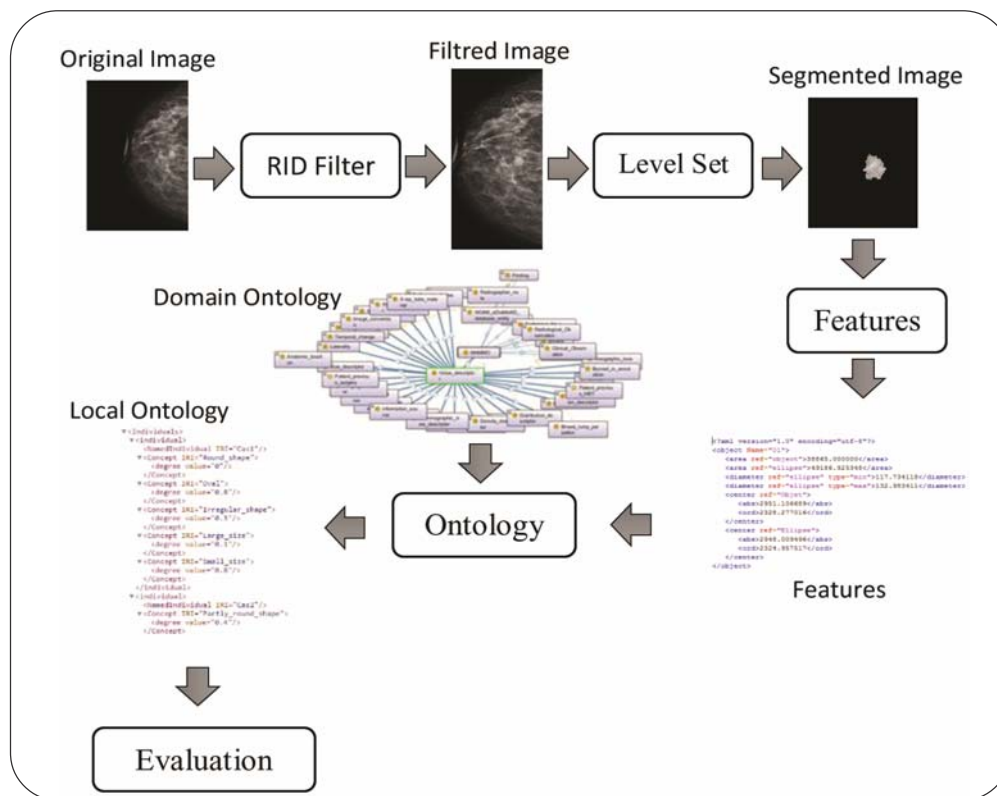


Figure 14. The local ontology construction experimentation process

Data Sets	Number of images	Number of pathological objects	Error Rate
Mammography	98	115	11.504%
MRI	88	38	9.988%

Table 2. Evaluation of the pathological objects specification

These images are provided by the regional hospital Ben Arous – Tunisia. The evaluation process is performed according two steps:

- The local ontology Evaluation
- The fused ontology Evaluation

The instantiating phase is evaluated by the consistency check as mentioned above.

4.1 Local Ontology Evaluation

The local ontology is constructed from the different measures calculated from features extracted from the images processed and it is based on the domain ontology definition. The Figure 14 shows the experimental procedure to construct the local ontology from an image.

The local ontology is composed of pathological objects contained in the used image. In order to evaluate the accuracy of the calculated values for features extraction, we use a Form Base built from manual extraction by the expert. We calculate the Mean Square Error (MSE) for each calculated value and then the average for a patho-

logical object. Table 2 depicts the evaluation performance of the pathological objects specification from images. The error rate represents the rate of the invalid objects detected. These objects have very different values compared with the manual extraction made by the expert. This rate is due, principally, to the high density of the breast imaged, and yet, to the imaging conditions.

Then in order to check the performances of local ontologies, we adopt a two-step procedure:

- **First Step:** Verification of the ontology contradiction and its syntactic and semantic consistence.
- **Second Step:** Validation of the treatment performance.

For the semantic contradiction and semantic consistency verification, we use a Rule Base build previously and containing all rules allowing the best verification for possibilistic concepts. For example:

- “An object can not be a singular distribution and dispersed distribution simultaneously” is a contradiction rule;

- “If an object has a round or oval form, it can not have an irregular contour.” is a consistency rule.

A binary validation matrix is generated verifying the semantic contradiction and consistency of the ontology according to the following assumptions:

- If an object contains a contradiction, the corresponding cell is evaluated to 1, otherwise to 0;
- If the object contains an inconsistency, the corresponding cell is evaluated to 0, otherwise to 1.

To evaluate the performance of the local ontology construction step we compute the effectiveness measured by the equation below:

$$effectiveness = \frac{\text{number of correct constructed ontologies}}{\text{total number of ontologies constructed}} \quad (14)$$

To evaluate the performance of the local ontology construction step we compute the effectiveness measured by the equation below:

As shown in the Table 3, the small number of the contradictory local ontologies and the inconsistent ones proves the effectiveness of the mammography processing at the semantic level. In fact, the possibilistic weights assigned to some attributes, land information that is more adequate to describe correctly the pathological objects.

Data Set	Number of ontologies	Number of contradictory ontologies	Number of inconsistent ontologies	effectiveness
Mammography	98	4	2	93,877 %
MRI	89	6	4	88.764 %

Table 3. The evaluation of the local constructed ontologies

Objects in MGP*	Objects in MRI	FP	TP	TN	FN	P	S	A	R
3	7	1	2	2	0	66.666%	66.666%	80%	100%

Table 4. The evaluation of the ontologies fusion phase

The resulting values evince the performance of the possibilistic local ontologies for medical knowledge modeling from mammographic images.

4.2 Fused Ontology Evaluation

The combination phase is evaluated by the results of the similarity measure step using *FP* (False Positive), *TP* (True Positive), *FN* (False Negative) and *TN* (True Negative) measures as follows:

- *FP* : Two objects are considered as not similar whereas really they are similar.

- *TP* : Two objects are considered as similar and really they are so.

- *TN* : Two objects are considered as not similar and really they are so.

- *FN* : Two objects are considered as not similar and really they are similar.

Two objects are considered as similar if the correspondent possibility measure is greater than 0.8.

Table 4 depicts results obtained for the proposed approach. The different *TP*, *FP*, *TN*, and *FN* are used to compute the precision (*P*), specificity (*S*), the accuracy (*A*), and the recall (*R*) rates according to the equations below:

$$P = \frac{TP}{TP + FP} \quad (15)$$

$$S = \frac{TN}{TP + FN} \quad (16)$$

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

$$R = \frac{TP}{TP + FN} \quad (18)$$

* *MGP*: Mammography

The small number of the data set is due, principally, to the fact that the MRI investigation tool is too expensive and it is not intended in all cases. In a breast diagnosis case, generally, there is few pathological objects presented in one image, so the comparison procedure is not, yet, complex. Despite that, results obtained proves the effectiveness of our solution to fuse information from two different breast modalities. Comparing results with the approach proposed in (Hopp and al.,2012) shows the closeness of the results. In fact, to combine data from two modalities, authors proposed to make a projection

from the MRI image and registered it with the mammographic image. As a result, authors indicate that the specificity rate is 70% and the accuracy rate is 84%.

The main contribution of our study is to overcome the problem of the numeric registration of the breast images. In fact the numeric registration can affect the precision of results, because of the non-rigidity of the breast organ and the big difference between the imaging conditions. Our main focus is to perform the registration between knowledge extracted from images and then make a high level combination. The use of ontologies allow us to pass from the low level representation of the data imaging to the high level of knowledge representation.

5. Conclusion

In this paper, a new approach was presented for knowledge fusion extracted from breast medical images taking advantages of both ontology paradigm and the possibility theory. On one hand the use of the ontology paradigm to formalize the image knowledge can reduce significantly the semantic gap between low level representation of the features in the image and the high level interpretation of the studied situation. On the other hand, the possibility theory can handle the different imperfections which generally affect medical images. The fusion result is a local ontology, which describes all the pathological objects present in the breast organ. This ontology is used to instantiate a domain mammographic ontology. In this population phase, we proposed to associate an xml file to the domain ontology instances. This file contains the different weights affected to the possibilistic concepts. Results show the effectiveness of the approach to build a consistent description of a pathological case in order to enrich a global domain ontology. The utility of the ontology population is, mainly, the enrichment of the domain ontology. In the medical case, this enriched ontology can be used to help doctors to make a diagnosis. It consists of a support that can be used to evaluate a situation according to similar cases that represent instances in the domain ontology.

References

- [1] Idoudi, R., Hamrouni, K., Solaiman, B. (2014). 'Ontological approach to mammographic knowledge representation'. Paper Presented at the *1st International Conference on Advanced Technologies for Signal and Image Processing*, 31 - 34.
- [2] Hu, B., Dasmahapatra, S., Shadbolt, N. (2003). 'From Lexicon To Mammographic Ontology: Experiences and Lessons'. *At International Workshop of Description Logics*, 229-233.
- [3] Taylor, P., Toujilov, I. (2012). 'Mammographic knowledge representation in description logic.' Paper Presented at the *International Workshop on Knowledge Representation for Health Care*, 158 – 169.
- [4] Tutac, A. (2010). 'Representation et raisonnement formels pour le pronostic basé sur l'imagerie médicale microscopique. Application à la graduation du cancer du sein' Phd Thesis: Université de Franche-Comté, p. 98-136.
- [5] Park, H. M., Lee, Y. L., Noh, B. N., Lee, H. H. (2011). 'Ontology-based generic event model for ubiquitous environment. *International Journal of Innovative Computing, Information and Control*, p. 4317 – 4326.
- [6] Chen, R. C., Bau, C. T., Tsai, M. Y., Huang, C. Y. (2010). 'Web pages cluster based on the relations of mapping keywords to ontology concept hierarchy'. *International Journal of Innovative Computing, Information and Control*, p. 2749 – 2760.
- [7] Zhang, Y., Huang, H., Yang, D., Zhang, H. (2009). 'A hierarchical and chord-based semantic service discovery system in the universal network'. *International Journal of Innovative Computing Information and Control*, 3745 – 3753.
- [8] Yang, Y., Jiang, P., Tsuchiya, S., Ren, F. (2009). 'Effect of using pragmatics information on question answering system of analects of confucius'. *International Journal of Innovative Computing Information and Control*, 5 (5), 1201-1212
- [9] Jiang, X., Tan, A. H. (2009). 'Learning and inferencing in user ontology for personalized Semantic Web search'. *Information sciences - Elsevier*, 179 (16), 2794 – 2808.
- [10] Gowsikhaa, D., Abirami, S., Baskaran, R. (2012). 'Construction of Image Ontology using low-level features for Image Retrieval'. Paper Presented at the *International Conference on Computer Communication and Informatics (ICCCI) - IEEE*, 1 – 7.
- [11] Wang, H., Jiang, X., Chia, L. T., Tan, A. H. (2008). 'Ontology Enhanced Web Image Retrieval: Aided by Wikipedia & Spreading Activation Theory'. Paper Presented at the *1st ACM International Conference on Multimedia Information Retrieval*, 195 – 201.
- [12] Iskandar, D. N. F. A., Thom, J. A., Tahaghoghi, S. M. M. (2007). 'Content-based Image Retrieval Using Image Regions As Query Examples'. Paper Presented at the *9th Conference on Australasian Database*, 38 – 46.
- [13] Liu, X., Shao, Z., Liu, J. (2010). 'Ontology-Based Image Retrieval with SIFT Features'. Paper Presented at the *First International Conference on Pervasive Computing, Signal Processing and Applications*, 464 – 467.
- [14] Wang, Y., Yang, X. (2011). 'Image Domain Ontology Fusion Approach Using Multi-level Inference Mechanism'. Paper Presented at the *International Conference on Computer Science and Network Technology (ICCSNT) IEEE.*, p.2662—2664
- [15] Lima, R., Oliveira, H., Freitas, F., Espinasse, B. (2014). 'Ontology Population from the Web: An Inductive Logic Programming-Based Approach'. Paper Presented at the *International Conference on Information Technology: New Generations (ITNG) - IEEE*, 473 – 478.
- [16] Imsombut, A., Sirikayon, C. (2016). 'An alternative

technique for populating Thai tourism ontology from texts based on machine learning, *In: The International Conference on Computer and Information Science (ICIS) - IEEE*, 1 – 4.

[17] Amato, F., De-Santo, A., Moscato, V., Picariello, A., Serpico, D., Sperli, G. (2015). 'A Lexicon-Grammar Based Methodology for Ontology Population for e-Health Applications. *In: International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS) - IEEE*, 521 – 526.

[18] Manzoor, U., Balubaid, M. A., Zafar, B., Umar, H., Khan, M. S. (2015). Semantic image retrieval: An ontology based approach, *International Journal of Advanced Research in Artificial Intelligence (IJARAI)*, 1 (4) 1 – 8.

[19] Gruber, T. R. (1993). A translation approach to portable ontologies, *Knowledge Acquisition Journal*, 4 (5) 199-229.

[20] Reyes-Ortiz, J. A. Bravo, M., Pablo, H. (2016). Web services ontology population through text classification'. *In: Federated Conference on Computer Science and Information Systems (FedCSIS) - IEEE*, 491- 495.

[21] Dubois, D., Prade, H. (1988). Theory of possibility an approach to computerized processing of uncertainty. Plenum Press, 1-31.

[22] Zadeh, L. (1978). Fuzzy sets as a basis for a theory of possibility'. *Fuzzy Sets and Systems*, 3-8.

[23] Dubois, D. (1980). Fuzzy Sets and Systems: Theory

and Applications'. San Diego, CA, USA: Academic, 144.

[24] Ben Salem, Y., Idoudi, R., Hamrouni, K., Soleiman, B., Bousetta, S. (2016). Image Based Ontology Learning'. Paper Presented at the *11th International Conference on Intelligent Systems: Theories and Applications*, 1 - 5

[25] Idoudi, R., Saheb-Ettabaa, K., Solaiman, B., Mnif, N. (2016). 'Association rules based ontology enrichment'. *International Journal of Web Applications*, 9 (1) 16 – 25.

[26] Jenhani, I., Ben Amor, N., Elouedi, Z., Benferhat, S., Mellouli, K. (2007). Information Affinity: A New Similarity Measure for Possibilistic Uncertain Information'. *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, 840 – 852.

[27] Hertz, T. (2006). Learning Distance Functions: Algorithms and Applications'. Phd Thesis, 9-14.

[28] Martin, A. (2005). 'La fusion d'informations', Course Manual, ENSIETA, 77-82.

[29] Dubois, D., Prade, H. (1994). 'La fusion d'informations imprécises'. *Traitement du signal*, 447-458.

[30] Hopp, T., Baltzer, P., Dietzel, M., Kaiser, W. A., Ruitter N. V. (2012). 2D/3D image fusion of X-ray mammograms with breast MRI: visualizing dynamic contrast enhancement in mammograms, *International Journal of Computer Assisted Radiology and Surgery (CARS)*, (7) p 339–348.

Author Biographies



Yosra BEN SALEM was born on the 23th of June 1987 in Bizerte from Tunisia. She received her engineer degree in computer science in June 2012 from the national engineering school of Tunis. Actually, she is a phd student in the national engineering school of Tunis and IMS Atlantique France. Her main research interest is the medical image processing and analysis, Ontology, knowledge management and reasoning for computer aided diagnosis for medical application.



Rihab IDOUDI was born on the 28th of May 1988 in Mahdia from Tunisia; she graduated from the national engineering school of Tunis, with a state Engineering degree in computer science, in June 2012. She is a doctoral student since 2012. Her main research interest is the ontology approach for the mammography domain knowledge, reasoning about acquired knowledge and knowledge mining using ontology.



Karim SAHEB ETTABAA received a first master degree from ENSICA (Toulouse - France), a second master degree from ENSI (Tunis - Tunisia), the phd degree from ENSI and ENST Bretagne (France) and the HDR diploma from the ENST Bretagne. He is, actually a professor at the ISTIC Institute (Sousse - Tunisia) . His main research interest include satellite image processing and analysis, remote sensing, spatial analysis, etc... He published multiple research studies in scientific journals and international conferences.



Pr. Kamel HAMROUNI received his Master and PhD degrees from “Pierre and Marie Curie” University, Paris, France. He received his “HDR” diploma from “University of Tunis El Manar, Tunisia”. He is, actually a professor at the National Engineering school of Tunis, teaching graduate and undergraduate courses in computer science and image processing. His main research interests include image segmentation, texture analysis, mathematical morphology, biometry and medical image application. He supervises a research team composed of around thirty researchers preparing master thesis, Phd thesis and HDR diploma. He published more than eighty papers in scientific journals and international conferences.



Pr. Bassel SOLAIMAN, received the telecommunication engineering degree from ‘Ecole Nationale Supérieure des télécommunications de Bretagne’, France, and the PhD degree from the ‘University de Rennes, Brest, France. From 1984 to 1985, he was a Research Assistant in the Communication Group at the ‘Centre d’Etudes et de recherche Scientifique’, Damascus, Syria. He joined the Image and Information Processing Department at Télécom Bretagne, Brest, France, in 1992, and now he is the responsible of Image and Information Processing department of Telecom Bretagne. His current research interests include the fields of remote sensing, medical image processing, pattern recognition, and artificial intelligence.