IMCAD: Computer Aided System for Breast Masses Detection based on Immune Recognition

Leila Belkhodja*, Djamila Hamdadou

Computer Science Laboratory of Oran (LIO) - University Oranl, Ahmed Benbella, Oran (Algeria)

uniR
LA UNIVERSIDAD
EN INTERNET

Abstract

Computer Aided Detection (CAD) systems are very important tools which help radiologists as a second reader in detecting early breast cancer in an efficient way, specially on screening mammograms. One of the challenging problems is the detection of masses, which are powerful signs of cancer, because of their poor apperance on mammograms. This paper investigates an automatic CAD for detection of breast masses in screening mammograms based on fuzzy segmentation and a bio-inspired method for pattern recognition: Artificial Immune Recognition System. The proposed approach is applied to real clinical images from the full field digital mammographic database: Inbreast. In order to validate our proposition, we propose the Receiver Operating Characteristic Curve as an analyzer of our IMCAD classifier system, which achieves a good area under curve, with a sensitivity of 100% and a specificity of 95%. The recognition system based on artificial immunity has shown its efficiency on recognizing masses from a very restricted set of training regions.

Keywords

I. Introduction

WHEN biological immune system fails to distinguish between what belongs and what does not belongs to the body, the defense against the invaders weakens. Such invaders include microorganisms, parasites and cancer cells.

We call a cancer all diseases in which abnormal cells divide without control and can invade other tissues. The National Cancer Institute cites more than 100 kinds of cancer in the world [1].

Lung and bronchus, colorectum and breast cancer are the three most commonly diagnosed women cancers. Breast cancer alone is expected to account for 30% of all new cancer diagnosis in women [2]. It is the most frequently observed cancer among women in France, European Union and the United States and remains one of the leading cause of women cancer death.

Breast cancer starts in the breast tissue by an uncontrolled growth of cells in the mammary gland. These cells may remain in the breast or migrate into the body via the blood and lymph vessels. The majority of cancers start in the milk channels. If they remain in the channels the cancer is called in situ or non-invasive. However if the cells leave the wall of the channels, the term "invasive cancer" is used [3].

If detected at an early stage, cancer can be cured in 9 cases out of 10. Presently, there is no effective way to prevent this disease. However, to improve survival rates recently researches showed that screening mammograms helps finding precancerous lesions before they become cancerous by early detection. Thereby, the number of new cases can be reduced and deaths caused by this kind of cancer can be prevented.

Mammogram can be used for breast screening or diagnosing

abnormalities. The screening mammogram is an x-ray exam of the breasts and the most effective tool for early detection. It is used when women have no breast signs, in order to find breast cancer when it is too small to be felt by a woman or her doctor. This will greatly improve a woman's chance for successful treatment.

Masses and microcalcifications are two powerful cancer indicators that are commonly used in evaluating mammogram. Radiologists consider mass detection a more challenging problem than micro-calcifications detection because of the poor image contrast of masses, not only for the large variation in size and shape in which masses can appear in a mammogram, but also because masses often exhibit poor image contrast due to breast density [4]. So, it is often difficult to separate normal and abnormal breast tissues. This engenders false positive cases that look like cancer.

Because of fatigued or inexperienced physicians during screening campaigns and the complex structure of the breast, radiology interpretation by visual perception can often miss true positive readings. To fix this problem, strategies, such as a second reading of screening mammograms, have been selectively used, which yield an increase in the cancer detection rate. This is a heavy major challenge for governments, medical organizations, and a difficult task to interpret screening mammograms in large numbers [5].

Computer aided systems for detection and diagnosis on mammograms are one of the automatic solutions that help the radiologist in detecting abnormalities in an efficient way as a second reader of digital mammograms.

To this end, we propose, in this paper, a methodology for computer-aided detection of breast masses on screening mammograms, which joins multidisciplinary axes such as medical domain, image processing and biological pattern recognition. For this, we focus on minimizing false positive findings and increasing true positive cases using all benefits of fuzzy processing and artificial immune recognition system.

* Corresponding author.

E-mail address: belkhodja_leila@yahoo.fr

In the following paragraphs, we first outline, in section II, some related work on computer detection for breast cancer and particularly masses symptoms, followed in section III by a state of the art on the artificial immune systems. Our contribution is described in section IV. Then we present the description of the proposed CAD system in section V. The adopted approach is given in detail in section VI. Finally, we present our experimental results and discuss them, respectively, in section VII and VIII.

II. Related Work

Generally, computer aided systems on breast medical images take two forms:

1. **Computer aided detection system (CAD)** which is able to identify the regions of Suspicion (ROS),
2. **Computer aided diagnosis system (CADx)** which can make a decision whether a ROS is benign or malignant.

We are going to focus on the first form. For more details of the second form the reader can refer to [6]

In the context of the CAD detection systems, the goal of the detection stage is to assist radiologists in locating abnormalities on asymptomatic women mammogram images especially during screening campaigns where a large numbers of mammograms must be analyzed.

In radiological routine the practice consists of applying visual perception by looking at a mammogram and then using cognition for interpreting what is seen. Prospective clinical studies have demonstrated an increase in breast cancer detection with CAD assistance [5].

CAD algorithms, which refers to pattern recognition software, must explore digital or digitized mammograms and search particular signs, which may be the first alarm of cancer. In this way many researchers have focused on particularly two markers: masses and calcifications. They take into the account single image or multiple images.

It is reported in [7] that the American College of Radiology (ACR) and the Breast Imaging Reporting and Data System (BI-RADS) define a mass as a three dimensional space occupying lesion which can be seen at least in two different projections (Cranio Caudal/Medio Lateral). It is characterized by its margins and shape.

Most CAD algorithms operate on a single image and are performed on oblique medio-lateral mammograms (MLO views) or cranio-caudal mammograms (CC views), on the right breast or the left breast, either to detect suspicious regions on a mammogram or to classify them as normal tissue or abnormal one. This engenders a large number of false positives which may be removed when classifying them. Reference [6] notes two types of algorithms: pixel based and region based detection methods.

The pixel based techniques work on features extracted from the local neighborhood of the pixel. In [8] for example, the authors proposed a method for lesion site selection using a morphological filtering enhancement combined with the stochastic model-based segmentation. Their results showed that with the proposed algorithm, the subtle masses could be segmented more accurately than those when the original image is used for extraction without enhancement. Another way to enhance masses is the adaptive thresholding. Reference [9] presented a dual stage adaptive thresholding method to identify the suspicious mass region. They used global histogram, to perform coarse level segmentation in order to locate abnormal regions, and local window thresholding method for each pixel to provide precise and fine segmentation results. Matsubara et al. [10] developed an adaptive threshold technique that uses histogram analysis to divide mammograms into three categories based on the density of the tissue ranging from fatty to dense. Masses were detected using multiple threshold values based on the category of the mammogram.

Contrary to pixel methods, region based detection ones use filtering techniques or segmentation to extract regions of interest and their features, which are later classified as suspicious or normal. [6] notes that a number of these methods are based on the idea of matched filtering where the image is filtered with a filter that is used as a model for a mass. For example, [11] [12] focused on circumscribed masses. The method in [11] uses modified median filtering to enhance mammogram images and template matching to detect the tumors. Herredsvela et al. [12] presented a method based on morphological hierarchical watersheds in the segmentation process. For circular and stellar masses [13] proposed a fuzzy pyramid linking method to detect tumors in mammogram image and classified detected regions to benign and malignant.

It is reported in [14] that because of many complex and changing characteristics of mass in mammogram images, with great difficulty in mass segmentation, region growing becomes a reliable method to accomplish it. Other segmentation techniques are cited in [15].

On the other hand, some methods use multiple images from right and left breasts or CC and MLO projections to search for asymmetries, which can be potential abnormalities. An example of such methods is developed in [16]. Multiple modalities, like mammography, ultrasonography and magnetic resonance imaging, can also be used [17]. More details for different imaging modalities are given in a review established by [18].

Note also that calcifications are the second very important marker of benign or malignant process after masses. They are tiny or big deposits of calcium. A number of different approaches have been applied for detection of calcifications. We cite as an example the work of [19].

Once regions of interest are extracted, some researchers, using CAD or CADx, focus on classifying them to normal and abnormal tissues or benign or malignant region respectively. The goal of this stage is to reduce false positive number for CAD and specify the nature of the mass for CADx. To deal with that, shape and texture features are used by a lot of researchers, for example [20].

To improve classification accuracy and stability of the system, various classification techniques have been used for classifying Regions of Interest (ROI) as normal or suspicious or as benign or malignant. Most of them use supervised methods. Up to present, popular approaches mainly include: artificial neural networks [21], swarm intelligence for neural network optimization [22], support vector machine [23], linear discriminant analysis [24], bayesian network [25], etc. In addition, deep learning, which is based on learning data representations and convolutional neural network, has shown its higher performance and has been effectively applied to breast mass detection and classification [26][27]. However, [28] mentioned that these methods work well on large data sets but exhibit certain limitations on small data sets. They propose a new exploratory method for the automatic detection of lesion based on gestalt psychology, which combines human cognitive characteristics and radiologist's knowledge. Most researches are applied on digitized mammograms from the MIAS [29] and the DDSM [30] databases, or on real clinical images from screening centers.

Recently, natural computing techniques have emerged in the artificial medical processing image domain and have proved their robustness to improve interpretation quality for radiologists. In this work, a particular attention is given to artificial immune recognition system (AIRS) whose details are outlined in the following section.

These kinds of systems (CAD/CADx) are generally evaluated using ROC plot (Receiver Operating Characteristic) and FROC plot (Free-response Receiver Operating Characteristic) [6]. They are standard methodologies for measurement of performance of detection and diagnosis algorithms in CAD systems. Raman et al. present more details about ROC in [31].

III. ARTIFICIAL IMMUNE RECOGNITION SYSTEM: STATE OF THE ART

Also called immunological computation, AIS is a field of study dedicated to develop computational models based on biological immune mechanisms, which are used to solve hard computational problems.

The human immune system is a robust, complex, highly distributive learning system that is able, through adaptation, to distinguish between dangerous foreign antigens and the body own cells. It learns how to identify patterns and then uses memory cells to remember previously identified patterns. There are two types of defense mechanism: innate and acquired. Innate defense acts without taking into account the type of disease and is achieved by some specialized cells. The acquired response involves specialized cells called lymphocytes [32]. The immune system contains B lymphocytes originated from the bone marrow and T- lymphocytes originated from thymus. When a pathogen is identified, stimulated B cells, helped by T cells, use the mechanism defense to lock on it. Undergoing somatic and hypermutation cloning, these B cells produce antibodies and distribute them all over the body to prepare the next attack from the antigen which is destroyed by the T cells. Detailed information about the immune system can be found in [33].

Many researchers have been motivated by these self-defense biological concepts. Artificial Immune Recognition Systems (AIRS) have not emerged since a collection of other supervised and unsupervised artificial immune systems algorithms have been developed [34] [35] [36] [37] [38].

AIRS is a supervised learning approach inspired from the biologic immune system for pattern recognition proposed in the Masters work by Watkins [39] who published the first version AIRS1 in [40]. It was replaced after that by a new efficient version of the algorithm called AIRS2 proposed by Watkins and Timmis in [41]. The specified version of this algorithm is detailed in section VI.

Meng et al. demonstrate in [42] the reliability and accuracy of AIRS on benchmarking experiments. They find that AIRS consistently outperforms other algorithms and can be used for real world classification tasks.

In diagnosing disease, AIRS have been largely applied as a decision making tool in medicine for example for heart disease [43] and diabetes [44]. For breast cancer diagnosis, many developed AIRS based researches show very important accuracies on Wisconsin breast cancer dataset [45] [46]. AIRS was also applied by Katsis et al. [17] to detect early breast cancer using different examinations (i.e. mammography, ultrasonography and magnetic resonance imaging) with promising results.

The main factors handled by the artificial immune system are antigens, antibodies or B- memory cells and Artificial Recognition Ball (ARB). Antigens (AG) are a set of n training data AG with labelled instances C. Antibodies are feature vectors of potential solutions matching more to antigens. An ARB represents a number of identical B-Cells which are employed within a mechanism to reduce duplication and dictate survival within the population [41].

The idea of the technique is to prepare a set of real-valued vectors to classify patterns. The system generates a set of memory cells from training data. If these cells are insufficiently stimulated for a given input pattern, candidate memory cells are then generated to replace them by a process of cloning and mutation of cells for the most stimulated memory cell. To join the memory pool, clones compete based on stimulation and on the amount of resources used by each cell [47].

To fine tune the training process, the AIRS algorithm uses a set of configurable parameters outlined in the following:

- **Initialization Instances**: are randomly selected training antigens to initialize the memory cells pool.

- **Affinity Threshold (AT)**: is the mean affinity value between all the antigens in the training set (1).

$$AT = \frac{\sum_{i=1}^{n} \sum_{j=i+1}^{n} affinity(ag_i, ag_j)}{\frac{n(n-1)}{2}} \tag{1}$$

Where:

- n is the number of training antigens,
- affinity (ag_i, ag_j)=Euclidean distance (ag_i, ag_j),
- ag_i and ag_j are the ith and jth training antigen.

- **Affinity Threshold Scalar (ATS)**: It is a parameter used with AT for memory cell replacement in the training process. Its value is between 0 and 1.

- **Stimulation Threshold**: In the range of [0,1], and generally around 0,9 this parameter regulates the ARB refinement process.

- **Clonal Rate**: It is an integer value, which determines the number of clones of the best memory cell or the number of clones and resources of each ARB in the refinement stage.

- **Mutation Rate**: It is used to determine the number of mutated clones that the best matching memory cell can create. The common value is 2.

- **Maximum Resources**: It limits the number of ARBs in the system. Usual values are between 150 and 300.

IV. CONTRIBUTION

We present, in detail, our contribution named IMCAD, which is a computer aided masses detection for screening mammography that acts as a second reader, and have the goal of improving the detection performance.

The proposed system IMCAD offers to radiologist's community an important tool, during screening campaigns of breast cancer, to detect abnormalities and missed masses that can be fatal for women life.

The main purpose of IMCAD is not to provide a perfect decision because of the lack of information about asymptomatic patients, but much more attract the attention of radiologist on regions that can be the beginning of cancer.

In this study, our contribution consists in imitating exactly biological immune self-defense of human body by developing a full automatic system based on a powerful recognition classifier AIRS. IMCAD acts, on reduced data, as well as an adaptive natural immune system, which can learn via experience.

For this aim, we propose a methodology, over different research areas (medical image processing, pattern recognition, computer vision…) which processes, mammograms as input data and produces results decision as output.

V. IMCAD SYSTEM DESCRIPTION

IMCAD is the proposed computer aided detection for breast masses based on a pattern recognition immune system which automatically identifies abnormal regions on screening mammograms. It is designed to provide a second opinion to aid rather than substituting the radiologist. Our CAD scheme is applied to a mammogram database and it is based on four sequential modules:

- **Subsystem1**: Preprocessing.

  The goal of this module, which uses real mammogram as input data, is to minimize time and memory allocation and reduce noise on mammograms.

- **Subsystem2**: Fuzzy Segmentation.

The second module uses a fuzzy classifier to extract homogenous classes from mammogram. The classified images are then labelled by a recursive labelling method.

- **Subsystem3**: Characterization.

The third module of the CAD scheme converts results of the second module to quantitative information. In this stage, IMCAD computes a set of features for each region in the segmented mammogram.

- **Subsystem4**: Immune Recognition.

Extracted features are then injected in the last module based on immune learning and recognition to detect suspecting regions as being positive masses. Details are given in the following section.

Note that IMCAD contains an offline immune treatment which is detailed later. The overall methodology schema of the proposed method is illustrated in Fig. 1.
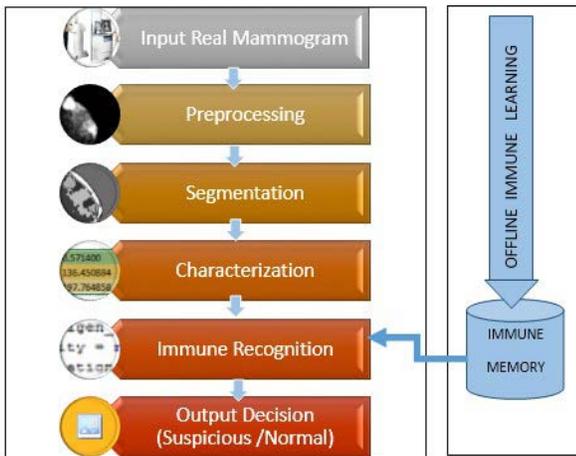
Fig. 1. The proposed system IMCAD: An overview.

VI. IMCAD Adopted Approach

The decisional process, adopted by IMCAD, presented previously, is detailed on a flowchart in Fig. 2.
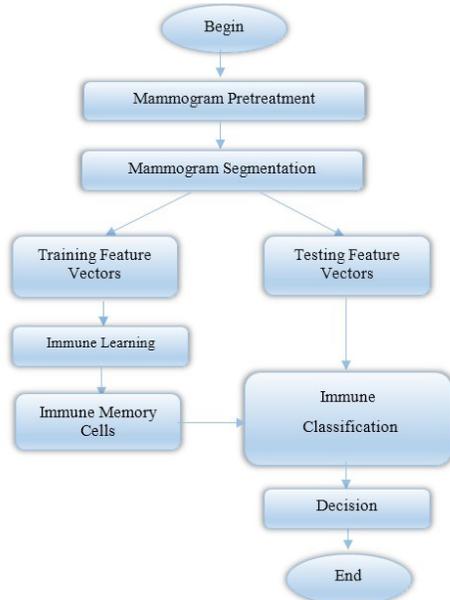
Fig. 2. Flowchart of the general approach adopted by IMCAD.

A. About Data

IMCAD was conducted on the *Inbreast Database* which is created by the Breast Research Group from INESC Porto and acquired at the Breast Center in Centro Hospitalar of São João at Porto [48]. In opposition to usual digitized mammograms, *Inbreast* is built with full-field digital mammograms, with a wide variability of cases. The acquisition equipment was the MammoNovation Siemens FFDM (Full Field Digital Mammogram), with 14-bit in contrast resolution.

InBreast has a total of 115 cases (screening, diagnostic and follow up). Images are in DICOM format (Digital Imaging and Communications in Medicine), with matrix size equal to 3328×4084 or 2560×3328 pixels, depending on the compression plate used in the acquisition (according to the breast size of the patient). This format gathers not only the image but also some related metadata.

The database contains examples of normal mammograms, mammograms with masses, mammograms with calcifications, architectural distortions, asymmetries, and images with multiple findings. An example of mammograms is given in Fig. 3.
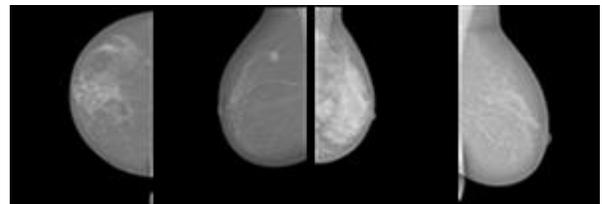
Fig. 3. Examples of mammograms from *InBreast* database.

B. Mammogram Preprocessing

Because of mammogram's hardness interpretation, any CAD system needs a preprocessing and a preparation stage to improve image quality, remove noise and make more correct the image segmentation outcome.

Our IMCAD preprocessing algorithm, shown in Fig. 4, proceeds first by performing a Gauss pyramid reduction to mammograms because of high image sizes and slowness of running time. Then we apply iteratively a mean filter to reduced images.
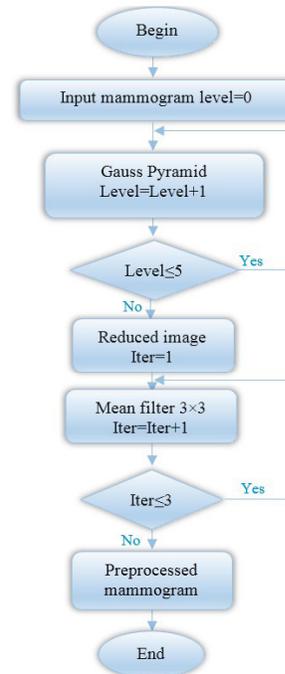
Fig. 4. Flowchart of the IMCAD preprocessing.

The Gaussian Pyramid is a multiple scale representation of the image. It allows the processing algorithm to work from the details up to the rough. To generate a pyramid, we iterate between two steps: smoothing and down-sampling. The smoothing operation removes high frequency components, which engenders fast changes that down-sampling would miss. The down-sampling reduces the image size by ½ at each level [49].

Note that images in level 0 are 3328x4084 or 2560x3328 pixels. The kernel used is cited below (2), with α=0,375.

$$\left[\frac{1}{4} - \frac{\alpha}{2} \quad \frac{1}{4} \quad \alpha \quad \frac{1}{4} \quad \frac{1}{4} - \frac{\alpha}{2}\right] \tag{2}$$

After reduction to level 5 mammograms are then filtered using a median filter, three times in succession, to avoid later over-segmentation and reduce the number of very small classes and regions. The median filter is a nonlinear digital filtering technique, which is used to remove noise from images and to improve the results for later processing. The main idea is to run a square window 3×3 through the image pixel by pixel replacing each entry with the median of neighboring pixels.

C. Segmentation

Segmentation is the mid-level image processing which consists in partitioning an image into regions or objects and reducing them to a form suitable for high computer processing level (recognition). One of the most difficult tasks in digital image processing is automatic segmentation. The approach to apply usually depends on the context and type of image to be segmented.

Mammograms are delicate images to analyze even when masses are blur and hidden in dense tissue. The concept of fuzziness corresponds exactly to this problem, that is why we choose a partitioning method based on the fuzzy algorithms. Note that our goal is to locate only the anomaly in a computer aided detection. Therefore, a successful segmentation must preserve the whole real mass and avoid creating negative ones.
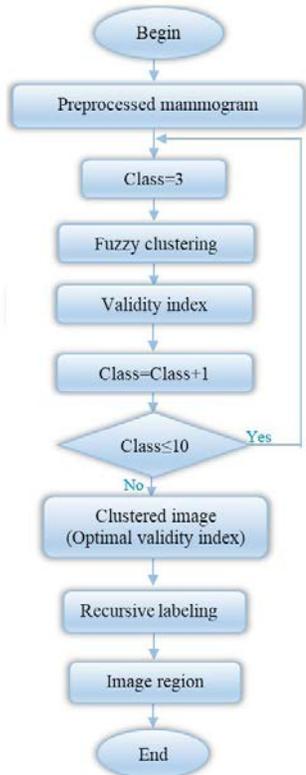
Fig. 5. Flowchart of the IMCAD fuzzy segmentation.

The proposed segmentation algorithm is carried out as follows:

- *Phase1*: Loading reduced and filtered images from the preprocessed step.
- *Phase 2*: Automatic grayscale fuzzy image clustering.
- *Phase 3*: Iterative evaluation of the validity index for the clustering process.
- Phase 4: Recursive labeling of the clustered image.

A detailed description of the segmentation algorithm is illustrated in Fig. 5.

1) Fuzzy Clustering

The fuzzy classification of the preprocessed mammogram is conducted by the Fuzzy C-means algorithm. It requires beforehand, the knowledge of the number of classes c and produces them in such a way that the objective function J_{fcm} (3) minimizes the total weighted mean-square error so:

$$J_{fcm}(\mu, x, c^{(k)}) = \sum_{i=1}^{c} \sum_{k=1}^{n} (\mu_{ik})^m \left\| x^{(k)} - c^{(i)} \right\|^2 \tag{3}$$

Where :

$C^{(i)}$: is the center of the class i,

m: is a real number greater than 1 to control the fuzziness of cluster,

$X^{(k)}$: is the kth pixel in the preprocessed mammogram.

The optimization of this function is done iteratively. At each iteration the membership degrees μ_{ik} of each pixel to the classes C and prototypes $C^{(i)}$ of classes are updated respectively according to the following relations (4) and (5):

$$\mu_{ik} = \frac{\left\| x_k - c_i \right\|^{\frac{2}{1-m}}}{\sum_{j=1}^{c} \left\| x_k - c_j \right\|^{\frac{2}{1-m}}} \tag{4}$$

$$c_i = \frac{\sum_{k=1}^{n} \mu_{ik}^{m} x_k}{\sum_{k=1}^{n} \mu_{ik}^{m}} \tag{5}$$

We use a variant of the fuzzy algorithm proposed in [50], which is based on the original Fuzzy C-Means (FCM) [51]. The main steps of the algorithm are:

Step 1: Set the parameters

- m: the fuzziness index, a real number greater than 1 to control the fuzziness of cluster,
- c: the number of classes,
- ε: convergence error = 0,001,
- X: vector which contains all pixels of the preprocessed mammogram.

Step 2: Initialize the membership degree matrix μ with random values.

Step 3: Standardize the initial weight over K (6).

$$\mu_{qk}^{normK} = \frac{\mu_{qk}}{\sum_{i=1}^{K} \mu_{qi}} \tag{6}$$

Step 4: Standardize cluster weights over Q (7).

$$\mu_{qk}^{normQ} = \frac{\mu_{qk}^{normK}}{\sum_{i=1}^{Q} (\mu_{ik}^{normK} - min)/(max - min)} \tag{7}$$

Step 5: Compute new prototype centers according to relation (5).

Step 6: Compute new weights according to relation (6).

Step 7: Repeat steps 4, 5 and 6 until reaching the maximum number of iterations or satisfy the criterion (8).

$$\left\| J^{old} - J^{new} \right\| < \varepsilon \qquad (8)$$

Step 8: Evaluate the K clustering by calculating the Xie-Beni validity index.

2) The Xie-Beni Validity Measure

Cluster validity measures are methods that evaluate clustering either by comparing the results of two different sets of cluster analysis to choose the best one or by determining the correct number of clusters in the data set.

Various indexes have been proposed in the literature. In this work we use the Xie-Beni measure XB [52] which is an index of fuzzy clustering and also applicable to crisp clustering. The XB index (9) focuses on two properties: compactness of the fuzzy partition and separation of clusters. A well-defined partition produces a small value of compactness and well separated centroids will give a high value of separation. Consequently, minimizing XB for c=2,3,…cmax will determine the optimal partition of data.

$$XB = \frac{\sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^{m} \left\| x_j - c_i \right\|^2}{n \cdot D_{\min}} \qquad (9)$$

Where

C_i: Centroid of the cluster i,

X_j: Pixel j,

μ_{ij}: The fuzzy membership of the pixel j belonging to class i,

$D_{\min}$: indicates the minimum Euclidean distance between centroids.

3) Labeling of the Fuzzy Clustered Image

The purpose of this phase is to obtain the image region with connected components (Fig. 6.b). The main problem is that the fuzzy clustered image is class labeled (Fig. 6.a). A class may contain one or more regions semantically separated. To obtain these regions, we separate classes in different binary images, where we label each class 1 and the other pixels background 0. A region is then a connected set of 1 pixels.

3	3	3	3	3	3	3	3	3	3		7	7	7	7	7	7	7	7	7	7
3	3	3	3	3	3	3	3	3	3		7	7	7	7	7	7	7	7	7	7
3	3	3	2	1	1	1	3	3	3		7	7	7	3	1	1	1	7	7	7
3	3	2	2	4	4	4	3	3	3		7	7	3	3	5	5	5	7	7	7
3	3	2	2	2	4	4	3	3	3		7	7	3	3	3	5	5	7	7	7
3	3	4	2	2	2	4	3	3	3		7	7	6	3	3	3	5	7	7	7
3	3	4	4	4	1	2	3	3	3		7	7	6	6	6	2	4	7	7	7
3	3	3	1	1	1	2	3	3	3		7	7	7	2	2	2	4	7	7	7
3	3	3	1	1	2	2	3	3	3		7	7	7	2	2	4	4	7	7	7
3	3	3	3	3	3	3	3	3	3		7	7	7	7	7	7	7	7	7	7
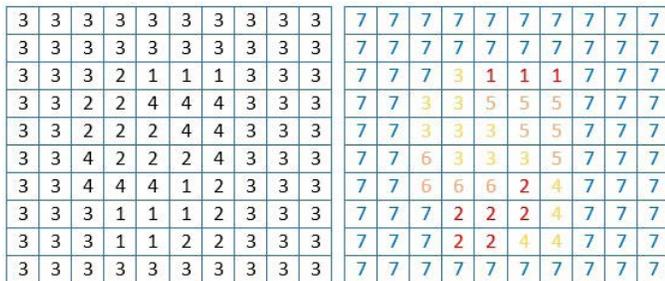
(a) (b)

Fig. 6. (a) Example of fuzzy clustered image (b) and image region with connected components.

This means that from any pixel labeled 1 there is a path of 1's to any other pixel in its region. This path can be found by searching recursively all roots from the starting 1 pixel and its 8 neighbors until the destination 1 is reached. These steps are detailed in the recursive labeling algorithm 1.

Algorithm 1: Fuzzy Clustered Recursive Labeling

Input data: Clustered image *Clus*

Output data: Labeled image *Lab*

Procedure Browse(Lab,label,L,P)

BEGIN$_{Browse}$
    Lab[L][P] ←label;
    if (Lab[L][P-1]=-1) Browse(Lab,L,P-1);
    if (Lab[L][P+1]=-1) Browse(Lab,L,P+1);
    if (Lab[L-1][P]=-1) Browse(Lab,L-1,P);
    if (Lab[L+1][P]=-1) Browse(Lab,L+1,P);
END$_{Browse}$

BEGIN
Copy each class in *Clus* in a single image *Im*
foreach image *Im* **do**
Bin← Binarize (*Im*);
Lab← negate (bin); //Put all pixels with value 1 to -1
Label←0;
for i←0 to N
    for j←0 to M
    if Lab[i,j]=-1 **then**
        label←label+1;
        Browse(Lab,label,i,j);
    Endif
    Endfor
Endfor
END.
return *Lab*;

D. Feature Selection

This process follows the output of a segmentation stage, which are generally pixel data. The purpose of this step is to extract attributes which give quantitative information able to distinguish between normal and abnormal regions.

It comes out, after a discussion with radiologists and a preliminary study, that the description of a mass is based on its intensity, size, density, shape, position, and edge characteristics. Radiologists define a critical size from which metastatic spread occurs. When the patient is treated before this size, she will not have metastases.

Given the wide variety of masses, it is extremely difficult to define a common set of attributes. As a result, we limited our research on circumscribed, spicules and poorly defined masses. The set of features we have calculated, for each region in the segmented image, includes only intensity and shape features to localize abnormal regions. IMCAD features are:

Area (S): Is the number of pixels inside the boundary of the region. We believe that this is an important feature because it corresponds to the size feature defined by the radiologists.

Average gray level (AGL): This feature (10) defines the average intensity of the region. It corresponds exactly to radiologists density attribute. Density is a measure used to describe mammogram's masses. A hyperdense region tends to be clearer than an hypodense region.

$$AGL = \frac{\sum_{k=1}^{k=S} Pixel_k}{S} \qquad (10)$$

Compactness (C): This parameter can be used to detect compact and circumscribed regions (11).

$$C = \frac{4\pi S}{P^2} \tag{11}$$

Where P represents the region perimeter (number of border pixels).

E. Artificial Immune Masses Recognition

Recognition, based on Artificial Immune Recognition System, is the higher-level processing step in the proposed IMCAD system. It involves making sense to previous extracted features by performing cognitive functions normally associated with radiologist vision. Our challenge and purpose in using artificial immunity in detection of abnormalities in IMCAD is to simulate globally immune human defense and imitate the extraordinary powers of brain against danger.

AIRS proceeds in two steps: offline training and online classification. It tries to make memory cells, which are representative of the extracted training regions the model is exposed to, and are suitable for classifying unseen mammograms.

1) Airs Training Phase:

The immune training process consists in building supervised classes from specific extracted features of normal and abnormal regions. AIRS learning [41][47], outlined in the Algorithm 2, turns on four stages: initialization, memory cell identification, competition of resources and refinement of memory cells.

1. **Initialization**

The initialization step consists in:

- Normalizing all antigens (Input regions).
- Calculating the *affinity threshold* AT (1).
- Initializing $Cells_{Memory}$ by choosing regions similar to antigens.
- Initializing ARB population to ∅.

2. **Memory cell identification ($Cells_{Memory}$) and ARBs Generation ($Cells_{clones}$)**

This process and the others described later are run for each antigenic input region one at a time, which makes AIRS a one-shot learning algorithm.

For each input region, this process consists of:

- Stimulating each initial memory cell B to the current antigen A (12).

$$Stimulation(B,A) = 1 - Affinity(B,A) \tag{12}$$

- Selecting MC_{best}, the memory cell that more stimulates the antigen and adding it to ARB pool.
- Making ARB pool by cloning MC_{best} *Nclones* (13) times and introducing diversification by mutating randomly each cloned cell.

$$Nclones = Stimulation(MC_{best}, ag) \times clonal_{rate} \times mutation\_rate \tag{13}$$

3. **Competition of resources and development of a candidate memory cell**

AIRS must maintain a population of memory cells for each class of antigen at the end of the algorithm. For this purpose the strongest ARBs ($Cells_{clones}$), with important resources (14) must survive in this stage. This is performed via a resource allocation and competition mechanism which is used to control the size of the ARB pool.

$$Ressources = Normalised\ stimulation * clonal\_rate \tag{14}$$

Before allocation of resources, diversification is also introduced. Each ARB is cloned *num_clones* times and then mutated. After that, resources are allocated to each ARB in the pool. The total of resources is computed and compared against *max_ressources*. ARBs with low resources are then removed from the pool. The stop condition for this

process occurs when the mean normalised stimulation exceeds the stimulation threshold

$$num\_clones = Stimulation * clonal\_rate \tag{15}$$

Note that mutation and competition of resources allocation subroutines are referred to [47].

4. **Memory Cell Introduction**

Once the competition stop condition is reached, the optimal ARB pool is selected. The ARB with maximum normalized stimulation value is designated to become the memory cell candidate: $MC_{candidate}$. This cell joins the memory cells ($Cells_{Memory}$) if its stimulation value is better than MC_{best} stimulation, which is removed if their affinity is less than the product of the affinity threshold and the affinity threshold scalar.

Algorithm 2: Artificial immune Learning Of Abnormal Regions

Input data: Input regions, stimulation_threshold, affinity_threshold, mutation_rate, clonal_rate, max_ressources,

Output data: $Cells_{Memory}$

Initialization

foreach $Input\_region_i \in$ Input regions **do**

    Stimulate the $Input\_region_i$ with each memory cell;

    $MC_{best} \leftarrow$ most stimulated memory cell;

    if (class(MC_{best})≠ class($Input\_region_i$) **then**

        $Cells_{Memory} \leftarrow Input\_region_i$;

        Class($Cells_{Memory}$) $\leftarrow$ Class($Input\_region_i$);

    else

        Nclones $\leftarrow$ Stim(MC_{best}) $\times$ clonal_rate $\times$ mutation_rate;

        $Cells_{clones} \leftarrow MC_{best}$;

        for j to Nclones do

            $Cells_{clones} \leftarrow$ Clone and Mutate MC^{best};

        end

    Stimulate ($Cells_{clones}$, $Input\_region_i$);

    while AverageNormalisedStimulation(Cellsclones) $\leq$ stim_threshold **do**

        foreach $Cell_j \in$ Cellsclones **do**

            . $Cells_{clones} \leftarrow$ Clone and Mutate $Cell_j$;

        end

        Stimulate ($Cells_{clones}$, $Input\_region_i$);

        Ressources allocation

        ReducePoolToMaximumRessources($Cells_{clones}$, $max\_ressources$);

    end

    $MC_{candidate} \leftarrow$ **GetMostStimulated** ($Input\_region_i$, $Cells_{clones}$);

    if Stim ($MC_{candidate}$) > Stim (MC_{best}) **then**

        $Cells_{Memory} \leftarrow MC_{candidate}$;

        if Affinity($MC_{candidate}$, MC_{best}) $\leq$ *affinity_threshold* **then**

        DeleteCell(MC_{best}, $Cells_{Memory}$);

        end

    end

    end

end

return $Cells_{Memory}$;

---

2) AIRS Testing Phase:

At the end of the training phase, all antigens are represented by a set of antibodies in the memory cell pool. A segmented mammogram is stimulated by all antibodies in order to classify the new antigenic regions. The criterion of classification is to attribute the class of the most stimulated antibody to each region in the image.

VII. Experimental Results

The objective of this work is to develop the automatic computer aided detection system IMCAD for screening mammograms, using artificial immunity, to help radiologists in preventing breast cancer early. IMCAD acts as a natural defense of the human immune system facing cancer. All experiment results were conducted on selected masses cases from Inbreast database.

All methods in IMCAD are performed using Embarcadero C++ Builder 2010 software, running on a laptop PC with a 2.50GHZ CPU(i5) and 4G RAM.

1) Nature of Data

Recall that the InBreast database includes a set of high quality mammogram images with different pathologies. IMCAD focus is on cases annotated as masses and normal, which are sorted and extracted manually. Original dicom mammograms are converted to BMP format. Corresponding annotated images are PNG files.

There are a total of 116 masses among 107 images (=1.1 masses per image). The average mass size is 479 mm2 (with a standard deviation of 619 mm2), the smallest mass has 15 mm2 and the biggest has an area of 3689 mm2.

The annotations were made by a specialist in the field, and validated by a second specialist, between April 2010 and December 2010. When there was a disagreement between the experts, the case was discussed until a consensus was obtained.

2) IMCAD Running

The proposed system proceeds in several steps.

1. Preprocessing

This step prepares mammograms for high levels treatments. The aim objective is to improve segmentation in areas of interest. We applied a 5 level multiresolution with a gauss pyramid reduction (Fig. 7.b), followed by 3 times median filtering operation for the last level of the pyramid (Fig. 7.c). A case study with abnormal mass region is shown in Fig. 7.a.

(a)   (b)   (c)
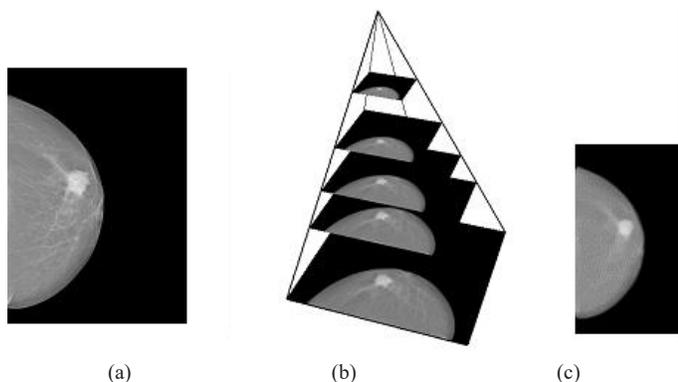
Fig. 7. Original mammogram with abnormal region (mass) (a) Five level Gauss Pyramid (b) Reduced filtered image (c).

2. Fuzzy Segmentation

Automated FCM clustering, intensity based, gives a very interesting result for our IMCAD segmentation step. It is applied on the fifth level image in the gauss pyramid. Through our implementation, we set the parameter fuzzy index $\mu = 2$ and we have varied the number of classes from 2 to 10. For each number the process stops when the optimal objective function is reached with $\varepsilon = 0.001$
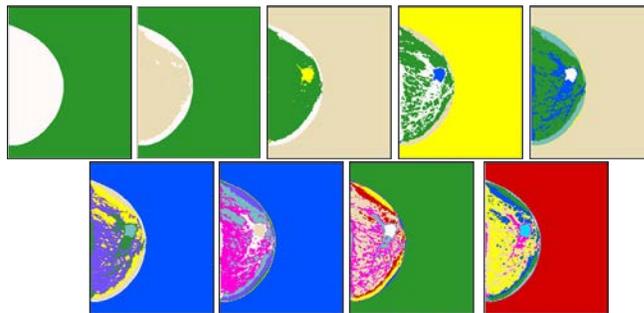
Fig. 8. FCM clustered images on original reduced mammograms (c=2..10).

Several tests were done with different settings of iteration number. First FCM works manually on original data considered as clean data (0% noise). Clustered images are shown in Fig. 8.

In order to reduce noise, avoid over-segmentation and minimize the number of regions to be processed, a median filter is applied to original reduced images. We can observe in Fig. 9 that results are much better and masses are more valued.
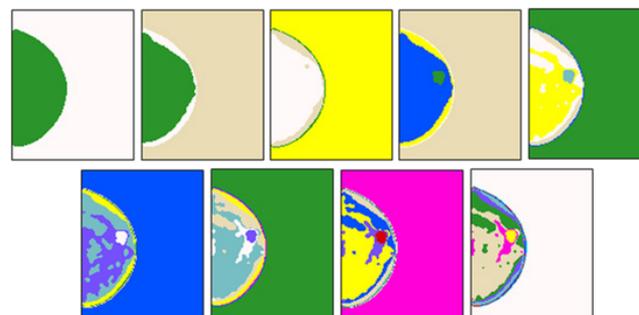
Fig. 9. FCM clustered images on reduced filtered mammograms (c=2..10).

To obtain connected regions, we submit clustered images to recursive labeling. An example is outlined in Fig. 10.
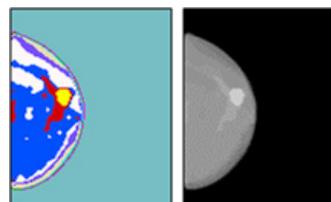
Fig. 10. Connected regions with recursive labeling on clustered mammogram(c=10).

With the objective to complete the IMCAD independence to parameters, we automated the FCM segmentation task by accurately identifying the optimal number of clusters from the Xie-Beni validity index. Results, shown on the graph in the Fig. 11, give 5 classes as optimal number with Xie-Beni=0.012375 over 50 iterations.
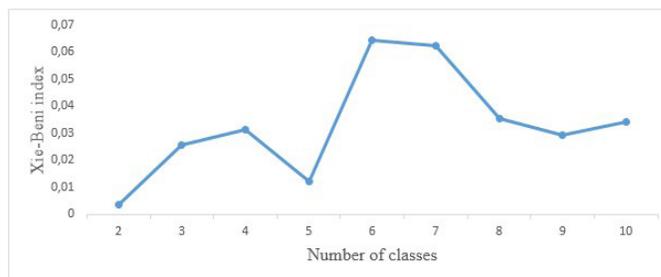
Fig. 11. Evolution of the Xie-Beni validity index over number of classes.

The number of iterations corresponds to minimum error between the objective function in the epoch t-1 and the final epoch t. The best visual results (Fig. 9, c=5) which conserve the whole abnormal region and avoid over-segmentation were obtained with the following parameters: $\mu = 2$, 70 iterations and $\varepsilon \sim 0$ (Fig. 12).
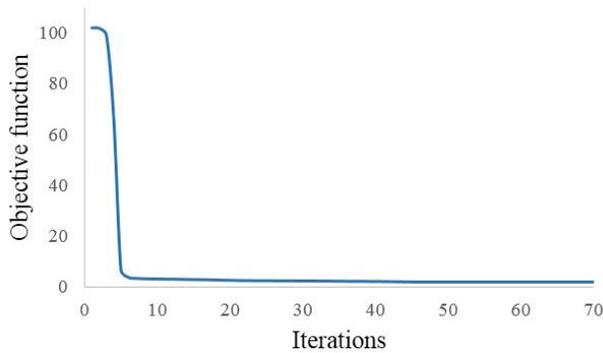
Fig. 12. Evolution of the objective function.

3. Immune learning

To train the artificial immune system, four abnormal mammograms are chosen from Inbreast Database (Fig. 13). Only masses and normal regions are featured and area, compactness and AGL are extracted to be presented to the system.
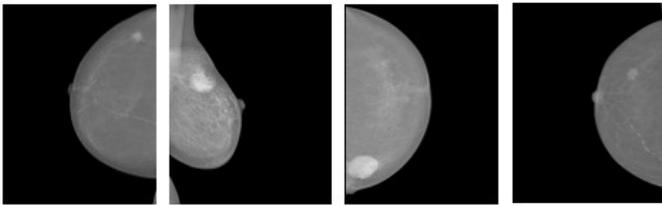
Fig. 13. Learning mammograms.

The set of training data is composed of 6 antigens which represents 4 masses regions and only 2 normal regions. Important masses and average masses form the class 1 and class 2 (Abnormal class) and the normal region forms the class 3 (Normal class). The affinity threshold obtained is 0,563. Values of training parameters are given in the table in Fig. 14.

During the ARB refinement process, the artificial recognition balls enter in competition of their resources. The average mean normalised stimulation value obtained is 0,97 (>stimulation threshold) for all iterations. The average value of departure ~0.45

Note that the purpose of this step is to compute a set of memory cells which are used later to recognize regions of segmented mammograms. To obtain more Bcells, we choose to run the training process iteratively seven times.

The total number of the memory cells after the training process is 20. Antigens of the normal class (C3) are recognized by a set of 10 Bcells. We obtain 3 Bcells in the abnormal class (C1), and 7 Bcells in the abnormal class (C2).

Parameters	Values
Max Ressources	200
Mutation Rate	2
Clone Rate	10
Hypermutation Rate	2
Affinity Threshold Scalar	0,1
Stimulation Threshold	0,9

Fig. 14. Table of the immune training parameters.

4. Immune classification

After 7 iterations, the B memory cells generated from the learning process are used in the classification step (test) on a total of 342 regions of 32 mammograms (16 with masses and 12 normal). Very small regions were ignored.

We stimulate each region of the segmented mammogram test by all the memory cells. Then we affect each region in the class of the most stimulated Bcell if the stimulation exceeds a certain threshold *T*.

VIII. IMCAD EVALUATION AND DISCUSSION

The proposed system IMCAD is a computer aided detection system for breast masses computerized to support radiologists in achieving their interpretation task in detecting abnormalities on screening mammograms.

In order to evaluate the effectiveness of our IMCAD system and to extract the best threshold *T* for classification of regions, we choose the Receiver Operating Characteristic curve and compare our results to radiologist's annotations in the Inbreast database.

The ROC curve is a graphical representation of the false positive rate (1-specificity) on the X axis and the true positive rate (sensitivity) on the Y axis, calculated for all possible thresholds *T*.

For all training and testing regions of all studied mammograms, we compute true positive and false positive rates (TPR, FPR). Results with different cut-off used are outlined on the ROC plot in Fig. 15.
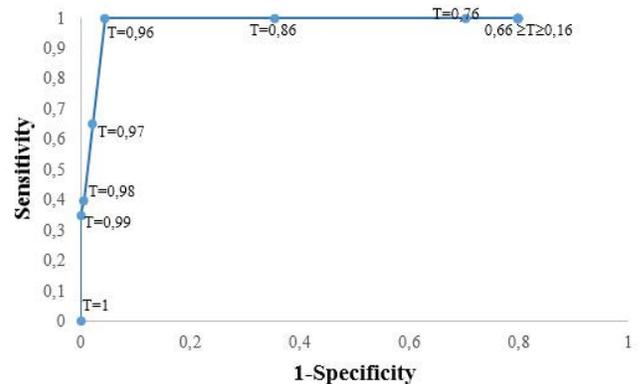
Fig. 15. IMCAD Receiver Operating Characteristic Curve.

The accuracy of the IMCAD test is measured by the Area Under Curve (AUC) which is equal to 0,78. The best cut-off, where sensitivity and specificity are close to 1, corresponds to T=0,96 (Sensitivity=1, Specificity=0,956). Decisional results on some abnormal and normal mammograms are exposed, respectively, in Fig. 16 and Fig. 17.

Normal images are processed exactly in the same way as abnormal mammograms. Some results are displayed in Fig. 17.
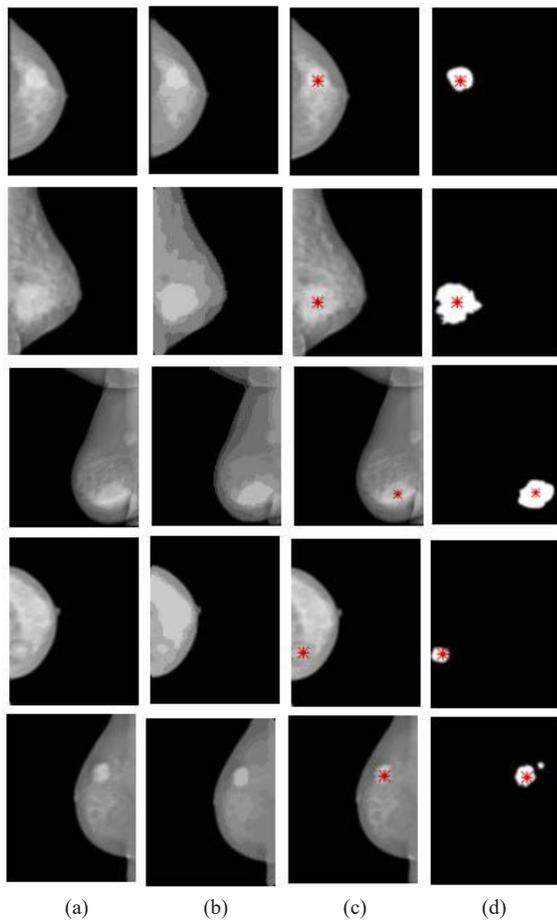
Fig. 16. Examples of IMCAD segmentation images (b) and detection results (c) on original images (a). Comparison by superposition on annotated masks of Inbreast Database (d).
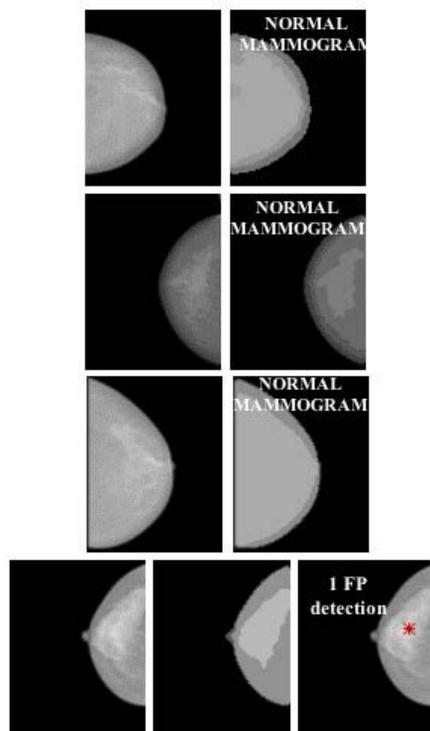
Fig. 17. IMCAD results on normal mammograms.

From these results, we can note that:

- Statistically, an area under curve of 0.5 represents a worthless test, while an area under curve of 1 represents a perfect test. IMCAD classifier has reached an AUC~0,8 which makes it a good system capable of identifying more true positives while minimizing the number of false positives.

- We obtain a very good sensitivity and a good specificity, which does not affect the objective of our IMCAD system.

- The detection of a mass is affected by the automatic computation of the number of classes *C*. In the data sample used, the only case, with 2 masses per mammogram (Example 5 in Fig. 16), one detected and one missed, is due to FCM segmentation. Indeed, the second mass was merged during the segmentation process (Example 5 in Fig. 16.b) and thus lost for the immune system (Fig. 16.d)

- Large original images require a high computation time and a powerful hardware. We solve this problem by applying a five level multiresolution Gauss pyramid that had no effect on defining ROIs. However, a very high level of reduction can make disappear some masses.

- We believe that the false positive rate reported, on some normal mammograms and normal regions in abnormal mammograms, by the IMCAD system, is not alarming since these suspicions could be removed by adding other complementary exams or performing other features to eliminate them.

- The artificial immune recognition system shows its efficiency on recognizing masses and normal regions from a very restraint set of training data. However, the choice of immune parameters is a very delicate task and it is done in an empirical way.

- Because abnormalities are often hidden in dense breast tissue, some learning mammograms were contrast adjusted.

- Unlike the deep learning which needs a large amount of data, our automatic IMCAD manages to classify a large number of regions from only 4 learned masses and 3 features (intensity, compactness and area).

- The proposed computer aided detection system based on artificial immunity works automatically from input mammogram to final decision. It takes into account the set of memory cells computed during the learning step. However, it has a significant time consuming (~2mn), when computing the Xie-Beni validity index and the optimal number of classes.

- The direct comparison of systems for detecting mammographic abnormalities is difficult because few studies have been reported on a common database and have not the same working conditions (for example [45][46] on Wisconsin database and [17] on multimodality images). For this reason, we relied on the decision making model of H. Simon [53] which refers to the expert in the evaluation step and we compared our results with the annotations of Inbreast Database radiologists (Fig 16.d).

IX. Conclusion and Future Work

We presented in this work an automatic Computer Aided Detection system IMCAD which combines medical image processing, bio-inspired pattern recognition areas and others methods in computer vision.

The aim of this work is to support radiologists as a second reader of screening mammograms to search subtle lesions that might otherwise be missed visually, and thus a contribution to reduce the mortality rate caused by breast cancer.

The methodology presented in this paper takes advantages of several robust approaches. Reduced data by a multiresolution Gauss

pyramid allows the system to work easily by reducing the processing time and resource allocation. Automatic segmentation was performed by the fuzzy c-means approach with a recursive labeling of regions. One of the important features of FCM algorithm is the membership function and the belonging of an object to several classes, with different degrees. This is an important supportive tool for medical CAD systems. Artificial immune recognition system in a morphological feature space lets our CAD act like a natural human system facing danger. It was designed to differentiate masses from normal regions from only three features.

It can be concluded that IMCAD succeeded widely in automatic detection of abnormal regions. Indeed, AIRS achieved a good ROC curve with AUC of 0,78, sensitivity of 100% and specificity of 95%, on the studied images.

Studies are in progress to increase specificity, treat ill-defined masses and minimize time processing. In our future works, we plan to extend our system to a computer aided diagnosis by analyzing the ROI extracted from the actual system and specifying its degree of malignancy and benignity.

References

[1] National Cancer Institute (2015, Feb). What Is Cancer? Available: http://www.cancer.gov

[2] Rebecca L. Siegel, Kimberly D. Miller, Ahmedin Jemal (2017). Cancer statistics. CA: A Cancer Journal for Clinicians, 65(1), pp. 7–30, Wiley Publishing.

[3] Ahmad Taher Azar, Sundarapandian Vaidyanathan (2014). Computational Intelligence Applications in Modeling and Control. Springer.

[4] Cheng HD, Cai X, Chen X, Hu L, Lou X (2003). Computer-aided detection and classification of microcalcifications in mammograms: a survey. Pattern Recognition, 36(12), pp. 2967–91.

[5] Ronald A Castellino (2005). Computer aided detection (CAD): an overview. Cancer Imaging, 5(1), pp. 17–19.

[6] Mehul P. Sampat, Mia K. Markey, and Alan C. Bovik (2005). Computer-Aided Detection and Diagnosis in Mammography, Handbook of image and video processing, 2(1), pp. 1195-1217.

[7] Debra M. lkeda (2010). Breast imaging: the requisites. Mosby Elsevier, 2nd edition.

[8] H. Li, Y. Wang, K. J. Liu, Lo SC, Freedman MT. (2001). Computerized radiographic mass detection—part I: Lesion site selection by morphological enhancement and contextual segmentation. IEEE Transactions on Medical Imaging, 20(4), pp. 289–301.

[9] J. Anitha, J. Dinesh Peter, S. Immanuel Alex Pandian (2017). A dual stage adaptive thresholding (DuSAT) for automatic mass detection in mammograms. Computer methods and programs in biomedicine. 138©, pp. 93–104.

[10] T. Matsubara, H. Fujita, T. Endo, et al. (1996). Development of mass detection algorithm based on adaptive thresholding technique in digital mammograms. K. Doi, M. L. Giger et al. eds. pp. 391–396.

[11] S. M. Lai, X. Li, and W. F. Bischof (1989). On techniques for detecting circumscribed masses in mammograms. IEEE Transaction on Medical Imaging. 8(4), pp. 377–386.

[12] Jostein Herredsvela, Thor Ole Gulsrud, Kjersti Engan (2005). Detection of circumscribed masses in mammograms using morphological segmentation. SPIE Proceedings Volume 5747, Medical Imaging 2005: Image Processing.

[13] Brzakovic, D., Luo, X.M., Brzakovic, P. (1990). An Approach to Automated Detection of Tumors in Mammograms. IEEE Transaction on Medical Imaging. 9(3), pp. 232–241.

[14] Lixin Song, Yanan Lv, Bin Yang, Yuhong Wang (2013). Segmentation of breast masses using adaptive region growing. 8th International Forum on Strategic Technology (IFOST), Ulaanbaatar, Mongolia, 28 June-1 July 2013.

[15] Marcomini K.D., Schiabel H. (2015). Investigating automatic techniques in segmentation accuracy of masses in digital mammography images. In: Jaffray D. (eds) World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada. IFMBE Proceedings, vol. 51. Springer, Cham.

[16] Frédéric J. P. Richard, Laurent D. Cohen (2003). A new Image registration technique with free boundary constraints: application to mammography. Computer Vision and Image Understanding, Vol. 89, pp. 166-196.

[17] C.D. Katsis, I. Gkogkou, C.A. Papadopoulos, Y. Goletsis, P.V. Boufounou (2013), Using Artificial Immune Recognition Systems in Order to Detect Early Breast Cancer, I.J. Intelligent Systems and Applications,vol.2, pp. 34-40.

[18] Bushra Mughal, Muhammad Sharif, Automated Detection of Breast Tumor in Different Imaging Modalities: A Review (2017). Current Medical Imaging Reviews, 13(2), pp. 121-139.

[19] Zobia Suhail, Mansoor Sarwar, and Kashif Murtaza (2015), Automatic detection of abnormalities in mammograms, BMC Medical Imaging, Elsevier, 2015.

[20] G. M. te Brake, N. Karssemeijer, and J. H. Hendriks (2000). An automatic method to discriminate malignant masses from normal tissue in digital mammograms. Physics Med. Biol. 45(10), pp. 2843–2857.

[21] M. M. Mehdy, P. Y. Ng, E. F. Shair, N. I. Md Saleh, and C. Gomes (2017). Artificial Neural Networks in Image Processing for Early Detection of Breast Cancer. Computational and Mathematical Methods in Medicine, vol. 2017, Article ID 2610628, 15 pages, 2017. https://doi.org/10.1155/2017/2610628.

[22] Dheeba J., Albert Singh N., Tamil Selvi S. (2014). Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach. Journal of Biomedical Informatics 49, pp. 45–52.

[23] Mohamed Meselhy Eltoukhy, Ibrahima Faye (2013). An Adaptive Threshold Method for Mass Detection in Mammographic Images. IEEE international conference on Signal and image processing applications (ICSIPA), pp.374-378.

[24] Petrick N, Chan HP, Wei D, Sahiner B, Helvie MA, Adler DD (1996). Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification, Med Phys., 23(10), pp. 1685-1696.

[25] B. Zheng, Y.H. Chang, X.H. Wang, W.F. Good, D. Gur (1999). Application of a Bayesian belief network in a computer-assisted diagnosis scheme for mass detection. SPIE Conference on Image Processing, 3661 (2), pp. 1553–1561.

[26] D Ribli, A Horváth, Z Unger, P Pollner, I Csabai (2018). Detecting and classifying lesions in mammograms with Deep Learning. Scientific reports 8 (1), 4165.

[27] R. Agarwal, O. Diaz, X. Lladó, R. Martí (2018). Mass detection in mammograms using pre-trained deep learning models. Proc. SPIE 10718, 14th International Workshop on Breast Imaging (IWBI 2018).

[28] Hongyu Wang, Jun Feng, Qirong Bu, Feihong Liu, Min Zhang, Yu Ren, and Yi Lv (2018). Breast Mass Detection in Digital Mammogram Based on Gestalt Psychology. Journal of Healthcare Engineering. Volume 2018, Article ID 4015613, 13 pages.

[29] Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, I., Boggis, C., Ricketts, I., et al. (2015). Mammographic Image Analysis Society (MIAS) database v1.21 [Dataset]. https://www.repository.cam.ac.uk/handle/1810/250394.

[30] Michael Heath, Kevin Bowyer, Daniel Kopans, Richard Moore and W. Philip Kegelmeyer (2001). The Digital Database for Screening Mammography. Proceedings of the Fifth International Workshop on Digital Mammography, M.J. Yaffe, ed., 212-218, Medical Physics Publishing.

[31] Valliappan Raman, Putra Sumari, H.H.Then, and Saleh Ali K. Al-Omari (2011). Review on Mammogram Mass Detection by Machine Learning Techniques. International Journal of Computer and Electrical Engineering, Vol. 3, pp. 873-879.

[32] Dipankar Dasgupta, Luis Fernando Niño (2009), Immunological computation: Theory and applications, CRC Press Taylor & Francis Group.

[33] Anthony Brabazon, Michael O'Neill, Seán McGarraghy (2015), Natural Computing Algorithms, Springer.

[34] Carter, J. H. (2000). The immune systems as a model for pattern recognition and classification. Journal of the American Medical Informatics Association, 7(1), pp. 28-4.

[35] Jon Timmis, Mark Neal, and John Hunt (2000). An Artificial Immune System for Data Analysis. Biosystems, 55(1-3), pp. 143-150.

[36] Jon Timmis and Mark J. Neal (2000). A Resource Limited Artificial Immune System for Data Analysis Research and Development in Intelligent Systems XVII, pp. 19-32.

[37] Jon Timmis and Mark Neal (2000). Investigating the evolution and stability of a resource limited artificial immune system. Special Workshop on Artificial Immune Systems, Genetic and Evolutionary Computation, Conference (GECCO) 2000, Las Vegas, Nevada, U.S.A., pp. 40-41.

[38] Leandro N. de Castro and Fernando J. Von Zuben (2000). The Clonal Selection Algorithm with Engineering Applications. GECCO 2000, Workshop on Artificial Immune Systems and Their Applications, Las Vegas, USA, pp. 36-37.

[39] A. B. Watkins (2001). AIRS: A resource limited artificial immune classifiers. PhD Thesis, Mississippi State University.

[40] A. B. Watkins and L. C. Boggess (2002). A Resource Limited Artificial Immune Classifier. In Part of the 2002 IEEE World Congress on Computational Intelligence held in Honolulu, 2002.

[41] A. Watkins and J. Timmis (2002), Artificial Immune Recognition System (AIRS): Revisions and Refinements. 1st International Conference on Artificial Immune Systems (ICARIS2002).

[42] Meng L, van der Putten P,Wang H (2005). A comprehensive benchmark of the artificial immune recognition system (AIRS). Advanced Data Mining and Applications, Springer Berlin/Heidelberg, vol.3584, pp. 575-582.

[43] Polat K, Güneˌs S, Tosun S (2006) Diagnosis of heart disease using artificial immune recognition system and fuzzy weighted preprocessing. Pattern Recogn 39, pp. 2186–2193.

[44] Chikh MA, Saidi M, Settouti N (2012). Diagnosis of diabetes diseases using an artificial immune recognition system2 (AIRS2) with fuzzy k-nearest neighbor. Journal of Medical Systems, 36(5), pp. 2721–2729.

[45] Saybani, M.R., Wah, T.Y., Aghabozorgi, S.R. et al. (2016). Diagnosing breast cancer with an improved artificial immune recognition system. Soft computing, 20(10), pp. 4069-4084.

[46] B.V.Kavethaa, Dr. P. Mohankumarb (2016). An artificial immune system based investigation of the breast cancer classification. International Journal of Pharmacy & Technology, 8(4), pp. 23097-23107.

[47] Jason Brownlee (2012). Clever algorithms, Nature-Inspired Programming Recipes, ISBN:1446785068 9781446785065.

[48] Moreira IC, Amaral I, Domingues I, Cardoso A, Cardoso MJ, Cardoso JS (2012). INbreast: Toward a Full-field Digital Mammographic Database. Acad. Radiol., 19(2), pp. 236-248.

[49] Burt and Adelson (1983). The Laplacian Pyramid as a Compact Image Code. IEEE Transactions on Communications, COM-31(4), pp. 532-540.

[50] Liyan Zhang (2001). Comparison of fuzzy c-means algorithm and new fuzzy clustering and fuzzy merging algorithm, Professional Paper, University of Nevada, May 2001.

[51] Liyan Zhang (May 2001). Comparison of Fuzzy c-means algorithm and New Fuzzy Clustering and Fuzzy Merging Algorithm. Professional Paper, Computer Science Department, University of Nevada, Reno.

[52] Xuanli Lisa Xie and Gerardo Beni (1991). A validity measure for fuzzy clustering. IEEE Trans. Pattern Anal. Mach. Intell., 13(8), pp. 841–847.

[53] Herbert Siomon (1977). The New Science of Management Decision. Prentice Hall, 1977.

Djamila Hamdadou

D. Hamdadou received her Engineering degree in Computer Science and her Master of Science degree from the Computer Science Institute in 1993 and 2000, respectively. She also obtained her Doctorate in 2008. She received her PHD in 2012 from the Computer Science Department. Currently, responsible of the Research Team "Spatio- Temporal Modeling and Artificial Vision: from the Sensor to the Decision" of the Computer Science Laboratory of Oran LIO. Her research interests include Business Intelligence, Artificial Intelligence, Spatio Temporal Modeling and Image Processing.

Leila Belkhodja

Leila Belkhodja is an Assistant Professor at the national institute of industrial security and maintenance in the University of Oran2. She received her engineering degree in computer science (2002) and her Master in electronic filed (2006) from the University of Sciences and Technology Mohammed Boudiaf, Algeria. She is preparing her PhD thesis within the Computer Science Department in the University of Oran1. In research field, she works on medical image processing, Computer Aided systems, bio-inspired algorithms, and artificial recognition.