DIAGNOSIS OF BREAST LESIONS USING THE LOCAL CHAN-VESE MODEL, HIERARCHICAL FUZZY PARTITIONING AND FUZZY DECISION TREE INDUCTION

F. BOUTAOUCHE AND N. BENAMRANE

Abstract. Breast cancer is one of the leading causes of death among women. Mammography remains today the best technology to detect breast cancer, early and efficiently, to distinguish between benign and malignant diseases. Several techniques in image processing and analysis have been developed to address this problem. In this paper, we propose a new solution to the problem of computer aided detection and interpretation for breast cancer. In the proposed approach, a Local Chan-Vese (LCV) model is used for the mass lesion segmentation step to isolate a suspected abnormality in a mammogram. In the classification step, we propose a two-step process: firstly, we use the hierarchical fuzzy partitioning (HFP) to construct fuzzy partitions from data, instead of using the only human information, available from expert knowledge, which are not sufficiently accurate and confronted to errors or inconsistencies. Secondly,fuzzy decision tree induction are proposed to extract classification knowledge from a set of feature-based examples. Fuzzy decision trees are first used to determine the class of the abnormality detected (well-defined mass, ill-defined mass, architectural distortion, or speculated masses), then, to identify the Severity of the abnormality, which can be benign or malignant. The proposed system is tested by using the images from Mammographic Image Analysis Society[MIAS] database. Experimental results show the efficiency of the proposed approach, resulting in an accuracy rate of 87, a sensitivity of 82.14%, and good specificity of 91.42%.

1. Introduction

Breast cancer occurs when normal breast cells begin to grow out of control. Cells may remain in the breast or spread to other parts of the body through the lymphatic and blood vessels. Death rates for breast cancer have gradually decreased among women since 1990.This decrease in death rate represents significant progress in early detection, diagnosis, improved treatment, and decreased incidence rate (Incidence refers to the total number of new cases of cancer) [2]. Research has shown that X-ray mammography screening greatly improves a woman's chances for early breast cancer detection, and can provide highly accurate analyses of the sizes, shapes, and locations of the anomalies. In order to help the diagnosis, and to reduce the number of false positives, several systems have been developed over the past decade. The

Received: December 2014; Revised: June 2016; Accepted: April 2017

Key words and phrases: Breast cancer, Mass segmentation, Local Chan-Vese model fuzzy decision tree, Fuzzy partitioning, Computer-aided detection.

aim of this paper is to propose a novel technique, which contribute to the early detection of breast cancer, and improve accuracy in differentiating a benign breast tumor from breast cancer. In our study we are interested in four types of lesions: architectural distortions, stellate (spiculated) lesions, circumscribed and ill-defined masses. The proposed approach involves two main steps: the segmentation step in which we use the Local Chan-Vese segmentation algorithm [38], to separate the suspicious region from its background, and the classification step, in which fuzzy membership functions are automatically constructed; fuzzy decision trees are then, used to divide abnormalities into benign or malignant. In the mass segmentation step, we are interested in level set framework which is based on segmenting a single part from the whole image, this kind of method is called image selective segmentation. In our context segmenting a single region constitutes an interesting and reliable tool to isolate a suspected abnormality in a mammogram. Among the level set models, we will present and implement the local Chan-Vese model [38], which incorporates region-based information into the energy functional to stabilize the evolution of the curve. The functional energy is based on three terms: the global term, which includes global properties such as the intensity average, the local term which incorporates local statistical information to improve the segmentation process, and the regularization term, used to ensure curve evolution stability. This method has proved to be highly efficient and suitable to isolate and extract masses with weak boundaries, and intensity inhomogeneity, especially, when the abnormality presents physical characteristics similar to those of normal tissue, with blurred contours or contours hidden by superimposed or adjacent normal tissue. This is clearly visible when the breast is dense. Dense breasts can make mammogram harder to interpret because both tumors and dense breast tissues appear white. After segmentation, a set of photometric and geometric features is extracted to characterize the segmented suspicious regions, and will be used to train and test our logical decision system (classification step). In the classification step, we introduce the automatic construction of the fuzzy membership functions to represent each extracted feature. In this context, we propose to use a flexible approach based on hierarchical partitioning scheme called hierarchical fuzzy partitioning technique (HFP) [32],[33], for fuzzy discretization , and generation of fuzzy partitions from a multidimensional training dataset. This idea marks the originality of our approach, given that in most published studies based on fuzzy systems, partitions were expressed by the radiologist, and some of the reported data may not be accurate and valid, and may be subject to conflicting opinions. Fuzzy decision trees are then employed to solve the problem of abnormality detection and classification into benign or malignant, during diagnostic mammography. Decision trees have been introduced in various mammographic interpretation systems [22],[24],[25], [34]. However, the fuzzy character of some digital mammogram and the inaccuracy in data descriptions may disrupt a potential classification, so it was necessary to involve decision making under uncertainty. Another key motivation for building a fuzzy decision tree classifier is the interpretability of the obtained linguistic fuzzy rules regardless the considered problem (unlike some existing systems such as neural networks considered as black boxes), and the use of features according to their influence and their impact on

classification (decisive attributes). In our approach, we propose to use fuzzy decision trees in two main steps: first, an initial classification is performed to reveal if a potentially suspicious area appears to be a circumscribed mass, ill-defined mass.... Then, according to the type of lesion, a subsequent classification is performed to predict the severity of the observed abnormality (malignant or benign). This classification is carried out with high sensitivity and accuracy by taking into account the incertitude, and constraints on knowledge. The results of this classification will be presented, and conclusions will then be given. The outline of the paper is as follows: in section 2, we will review some related works about computer aided breast cancer detection and diagnosis. In section 3, we present the proposed approach, we briefly describe the general scheme of the method, and we then introduce the local Chan-Vese model, the hierarchical fuzzy partitioning, and fuzzy decision tree algorithm. Section 4 describes our experimental results. Finally, section 5 presents a summary and a conclusion of our work.

2. **Background**

Various works have been proposed to develop computer aided breast cancer detection and diagnosis. Many involve the use of data-mining and machine learning techniques (such as decision tree, artificial neural network, support vector machine (SVM) and pattern matching), as predictive models to detect and classify suspicious mammographic lesions.

Artificial neural networks present a traditional tool for classification. They have been used in several approaches for mass classification. Belloti [3] proposes to use an edge-based segmentation to extract suspicious regions, and applies an artificial neural network with supervision provided by the radiologist's diagnosis, to detect the nature of a suspicious region. The results indicate a good sensitivity. Campos et al. [5] propose a method based on an independent component analysis and multilayer neural networks to predict the severity (malignant or benign) of a mass. The work shows a good performance in terms of specificity and sensitivity. In [8], authors developed a computer aided breast cancer diagnosis system called EX-DBS. the system uses a hybrid approach based on neural network and fuzzy set theory to generate a set of fuzzy rules, that classifies mammogram images as benign, or malign. The results show good performance values. Martins et al. [15]presented a method in which textural features are extracted using co-occurrence matrix. then a Bayesian neural network is employed to classify mammogram images as benign or malignant. The results show a satisfactory accuracy rate. M. J.Islam et al. [18] proposed an approach based on the exploitation of statistical and textural features to characterize breast regions, then, neural networks are used to classify a mass as benign or malignant. Experiment results exhibit a good diagnostic performance resulting in significant values of accuracy, specificity and sensitivity. In their paper, Ramani et al. [27] propose an automated mammogram classification system in which, Symlet wavelets and weighted histogram are used for feature extraction. Coefficients from symlet are reduced with singular value decomposition, which reduces the variable data set to a lower feature set. The classification step is accomplished

using various techniques: naive bayes classifier, random forest and neural network. The classification accuracy for Neural Network achieves the best results.

Support vector machine (SVM) were employed in various medical diagnostic systems. Leonardo de Oliveira et al. [11] have proposed an approach in which they used a k-mean algorithm to segment the image, and a support vector machine to classify regions in two groups: masses and non-masses. The method achieves a good accuracy rate. In their paper [12], the authors used template matching with mutual information for mass detection in screening mammogram. Template matching is used to generate the likelihood map in which candidate pixels are assigned a likelihood value of belonging to a mass. Experiments yield a good performance and indicate that the developed system can be successfully applied to tomosynthesis. Another method of mammogram classification was presented by Salve and Chakkarwar [31], the proposed system is accomplished in three steps, first, Histogram equalization and median filter are applied for image enhancement, then, the Gabor and Discrete wavelet transforms DWT are employed to extract the most prominent features. Finally, the Support Vector Machine Classifier (SVM) is used to classify the images into benign or malignant images. Experimental results show that accuracy of DWT is more significant than Gabor Wavelets, but true positive recognition rate of Gabor Wavelet is higher.

Decision trees provide a good support to automatic detection of lesions in mammogram images. Khan et al [9], in their paper, studied the potential of fuzzy logic and decision tree combined in a single system to detect breast cancer. They have experimented different combinations of: the number of rules, types of fuzzy membership functions and inference techniques. The results indicate that the weighted fuzzy decision tree gives a good and accurate prognostic decision making system. Lei and Andrew Chan [10], through their paper, presented an approach based on the fractal dimension analysis to identify suspicious regions, and used decision trees to determine the abnormality class. Oliver et al. [21] developed an automatic breast tissue classification methodology in a few steps: first, they used gray-level information, and the fuzzy c-mean algorithm to segment fatty versus dense tissue types. Next, they extracted morphological and texture features from each region. To implement the classification step, they used a Bayesian combination of the KNN (K-nearest neighbours) and a decision tree. Pitchman et al. [25]developed a system to classify a tumor as either benign or malignant. He used the Gabor filter and equalization histogram for the preprocessing stage, the texture analysis and the watershed algorithm for the segmentation, then, the used decision trees for the classification task. This method yields very good accuracy in a minimum period of time. Shanthi and Bhaskaran [34] proposed a novel technique to detect and classify breast cancer using intuistionistic fuzzy c-mean to extract the suspicious regions. Decision trees are then employed to classify lesions. The results show a significant accuracy rate. In their paper, Miranda et all [19] membership functions are automatically generated using fuzzy omega algorithm, and classification is carried out by matching the input features and linguistic variables. The results yield good accuracy value.

After this quick overview of techniques dedicated to the classification, and interpretation of mammogram, we present our approach to solve the problems of detection and classification.

3. Methods

3.1. **General Scheme of the Proposed Approach.** The proposed system is divided into several steps: a pretreatment process aimed to improve image quality, by making use of contrast enhancement techniques. Next, the segmentation step, which aims to separate lesions from the background by extracting the location of their boundaries. In our system, we propose to use the Local Chan-Vese model [38], which is a region based level set algorithm, where the curve evolution is guided by local statistical information.

A set of discriminative features are extracted from the segmented regions, and submitted to our logical classification system. To improve the mammographic mass classification, we propose an approach based on fuzzy set theory. First, we introduce the HFP [32],[33] technique to automatically construct membership functions for each attribute extracted from segmented regions of mammogram. Usually, the attributes membership functions are determined by an expert radiologist, and expressed in a very natural way, using linguistic terms. But knowledge is often subject to interpretation, errors or inconsistencies (or even conflicts among experts). Secondly, we use fuzzy decision trees to solve the problem of extracting knowledge from data, and learn the association between attribute variables and classes from training examples. The potential benefit of decision trees is the homogeneous treatment of all the types of attributes. Decision trees filter attributes according to a relevant measure which estimate the ability to discriminate between different classes. In mammography, this will allow us to identify the crucial characteristics of regions that reveal the presence or absence of an abnormality. Furthermore, this method takes into account the presence of imprecise data and the semantic of vocabulary used to describe mammographic mass data. Figure1 shows the general scheme of our approach.

Mass classification is accomplished in two consecutive steps: first, an initial model is constructed to identify masses (circumscribed masses, spiculated masses, ill-defined masses), and architectural distortion. Then, a second classification is carried out to find out whether the lesion is benign or malignant. At this level, four decision tree models (corresponding to the four suspicious abnormalities) are built; each is designed to predict malignancy or benignity of a suspicious lesion. A more detailed description is given in the section on our experiment.

3.2. **Local Chan-Vese Model for Image Segmentation and Feature Extraction.** Before presenting the LCV model, let us recall the principle of the Chan-Vese method. The Chan-Vese algorithm is the most representative and popular among region-based level set methods [6],[7]. In this model, the curve evolution is guided by local statistics to stop the evolving curve on the boundary of the desired object, instead of the gradient which appears mostly inefficient, mainly, on objects with

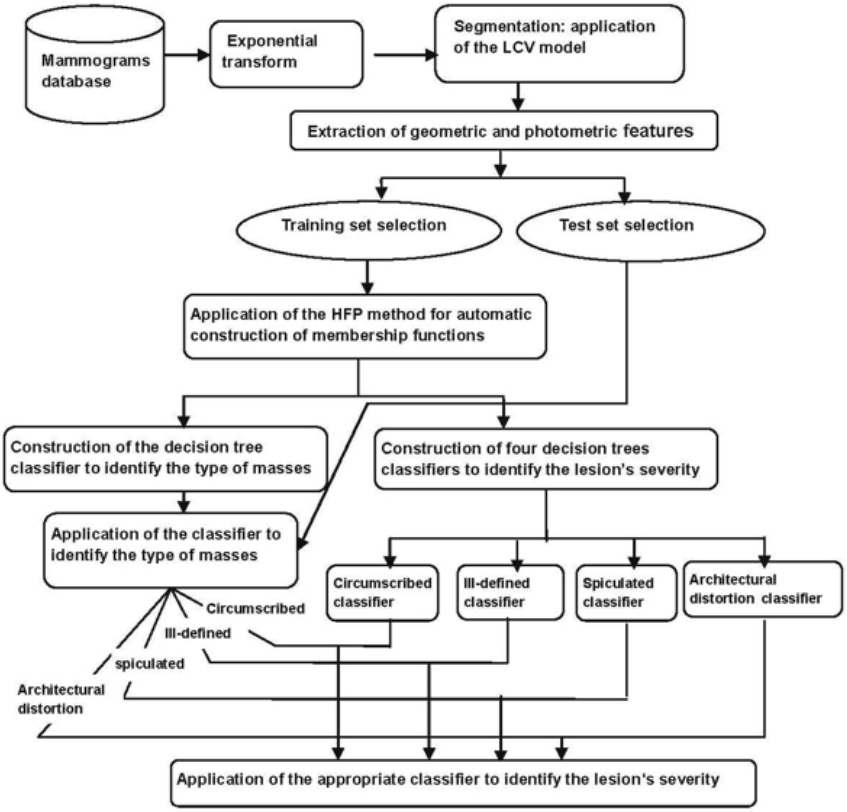


FIGURE 1. General Architecture of the Proposed Classification

low contrast. This model presents another advantage which occurs in its low sensitivity to the initialization of the curve. The basic idea of the CV model is the introduction of the fitting energy functional, which has to be minimized during the segmentation process. Several approaches have been reported for mammography segmentation using the Chan-Vese algorithm. In [1], authors used the model to isolate the region of interest and studied the sensitivity of different descriptors to the initial contour. In [26], authors used the Chan-Vese model to isolate spiculated masses. The curve evolution maximizes a likelihood function to divide images into foreground and background. The algorithm exhibits good performances. One of the major drawbacks of this model is its inadaptability for segmenting images composed of inhomogeneous regions. Intensity inhomogeneous problems can appear, particularly because of image acquisition conditions, and are often not visible to the human eye; this phenomenon can be observed, especially in screening mammography. To address this problem, a variant of this model was proposed by [38], and known as the Local Chan-Vese model (LCV). The LVC model is based on global and statistical information to overcome the inhomogeneous intensity distribution,

and thus, to drive the evolving curve towards the true boundaries. The evolution starts with a curve on the plane as the zero level set of a higher dimensional function : all pixels corresponding to the passage of the curve are initialized to zero. The whole surface can be then, divided into an internal region (pixels inside the curve), and an external region (pixels outside the curve). The evolution of the level set function is guided by a functional energy given by the formula:

$$Energy = \lambda_1.t_{global} + \lambda_2.t_{local} + t_{regul}. \tag{1}$$

with λ_1, λ_2, parameters specified by the user, they are chosen according to the illumination property of images in terms of inhomogeneity. For example, for images presenting intensity inhomogeneity,

λ_1, should be chosen less than λ_2. -t_{global} : is the global term, which includes global properties as the intensity average, it is given by the following equation:

$$t_{global} = \int_{\Omega} |I_0(x,y) - m_1|^2 \; H(\phi(x,y))dxdy + \int_{\Omega} |I_0(x,y) - m_2|^2 (1 - H(\phi(x,y)))dxdy \tag{2}$$

m_1 and m_2 are the intensity averages of the image I_0, inside and outside the curve C, respectively. $\phi$is the zero level set, it's given by:

$$\phi(x,y) = \begin{cases} > 0, & (x,y) \; inside \; \; curve \; C \\ = 0, & (x,y) \, on \; curve \; C \\ < 0, & (x,y) \, outside \; curve \; C \end{cases}$$

The Heaviside function $H(\phi(x,y))$indicates the set surrounded by curve C. It is used as an indicator of the set of pixels inside the curve.

- t_{local}: is the local term. It incorporates local statistical information to improve the segmentation process. Each pixel is analysed with respect to its neighbourhood, since smaller regions are more likely to have approximately homogeneous intensity. It is given by the following formula:

$$t_{local} = \int_{inside(c)} |g_k.I_0(x,y) - I_0(x,y) - d_1|^2 dxdy + \int_{outside(c)} |g_k.I_0(x,y) - \; I_0(x,y) - d_2|^2 dxdy \tag{3}$$

Where $g_k.I_0(x,y)$ is the convolution of the image I_0 with a filter g_k of size k * k. The contrast is calculated by subtracting the original image from the averaging convolution image$(g_k.I_0(x,y) - I_0(x,y))$. $d_1, d_2$are the intensity averages of $g_k.I_0(x,y) - I_0(x,y)$ inside and outside C, respectively.

-t_{regul}: is the regularization term, used to ensure curve evolution stability. The expression is formulated as follows:

$$t_{regul} = \alpha. \int_{\Omega} \delta(\phi(x,y)) |\nabla(\phi(x,y))| \, dxdy + \int_{\Omega} \frac{1}{2} (|\nabla\phi(x,y)| \; - 1)^2 dxdy. \tag{4}$$

α is used to control the size of objects to be detected: α takes small values if smaller objects should be detected, α takes large values otherwise. δ is the Dirac delta function (which corresponds to the derivative of Heaviside). The Heaviside and The Dirac delta functions are used to indicate pixels inside, outside and on the curve C.

The numerical implementation of the level set function evolution (from ϕ^n to ϕ^{n+1} at a time step ∇t), given previously, uses derivative approximations by a finite difference methods, the simplest scheme is of the following form:

$$\frac{\phi_{ij}^{n+1} - \phi_{ij}^{n}}{\nabla t} = \delta_\epsilon \left(\phi_{ij}^n\right)(T_1 + T_2) \tag{5}$$

$$T_1 = \left[-\lambda_1 (I_{ij} - m_1 (\phi^n))^2 + \lambda_2 + \left(\lambda_1 (I_{ij} - m_2 (\phi^n))^2 + \lambda_2 (g_k . I_{ij} - I_{ij} - v_2 (\phi^n))^2\right)\right]$$

$$T_2 = \left[\mu \delta_\epsilon \left(\phi_{ij}^n\right) k + \left(\phi_{i+1j}^n + \phi_{i-1j}^n + \phi_{ij+1}^n + \phi_{ij-1}^n - 4\phi_{ij}^n - k\right)\right]$$

Where K is the curvature, it is given as follows:

$$k = div\left(\frac{\nabla \phi}{|\nabla \phi|}\right)$$

3.3. **Feature Extraction.** After segmentation, a set of quantitative and descriptive features is extracted from segmented regions. Feature extraction is a critical task, because the particular features selected to characterize a specified region, directly affect the image classification accuracy. Feature extraction methods are based on 3 types of features: photometric features (grey level, variance, color histogram), geometrical features (surface, circularity, elongation) and texture features. In our work, each segmented region is characterized by a vector of six attributes, namely four shape attributes: surface, perimeter, compactness, and elongation, and two photometric attributes: average brightness and variance. This choice is justified by the fact that these features are simple, and especially well-suited to describe regions in mammogram. Furthermore, a study based on mutual information analysis [39], showed that mass margins and mass shape, are the most important and informative features in decision making for the diagnosis and classification of breast cancer. We limit the number of attributes to improve classification accuracy. In what follows we briefly describe these features [16]:

-Surface: the number of pixels within a specified region.

-Perimeter: the length of the path that surrounds a region.

-Mean Gray Value: corresponds to the intensity average.

-Compactness: this descriptor is used to define the regularity of an area. It is another measure of how circular an object is. Its given by the ration of perimeter to surface.

-Variance: is a measure of the dispersion of the gray levels around the mean gray value.

-Elongation (**eccentricity***):* this characteristic is very useful for describing the elongation degree of a lesion. It corresponds to the ratio of the length of major axis to the length of minor axis. The major and minor axes are calculated by the minimum bounding box method.

In our fuzzy systems for mammography interpretation, each input feature vector is characterized by uncertain data. The associated fuzzy partitions are defined using the hierarchical fuzzy partitioning (next section). Usually intervals of attribute values are represented with triangular and trapezoidal membership functions.

3.4. **Hierarchical Fuzzy Partitioning.** The HFP method was proposed by Serge Guillaume and Brigitte Charnomordic [32],[33]. The idea is to iteratively aggregate fuzzy sets from a multidimensional dataset. In our context, our dataset consists of six attributes, which are assigned a set of values calculated from the segmented regions. These values will constitute the basis upon which the fuzzy partitions will be built. The idea is to proceed iteratively by merging sets according to a merging criterion, based on minimizing the variation of the sum of distances between all pairs of points of the training set. This notion of distance implies an internal distance between points within the same group, a prototype distance and an external distance between points belonging to different fuzzy sets.

3.4.1. *Initial grouping.* The initial partitions are chosen according to the data distribution. Initial partitions can be built by sorting values in an ascending order, then one starts to form groups of values: a new partition is created, if the current value is superior to the sum of the current group value and a tolerance value (used to test the equality of two values).

For example: dataset={0.15, 0.2, **0.3**, 0.8, 0.85} and $\alpha = 0.1$

Current value $0.8 > (0.15+ 0.2+ 0.3) + 0.1$ so a partition composed of {0.15, 0.2, 0.3} increased.

Initial partitions can also be achieved by using a clustering algorithm, such an K-means [30].

Grouping makes it possible to define strong triangular fuzzy partitions, where:

- The fuzzy set center is defined as the average of all values belonging to a cluster.
- The left and right breakpoints of the defined fuzzy set correspond to the fuzzy set centers of the neighbouring fuzzy sets (see Figure 2 (a)).

Each fuzzy set f corresponding to an input variable in training dataset ϵ is assigned a weight G^f calculated by the formula:

$$G^f = \sum_{x \in \epsilon} \mu^f (x) \tag{6}$$

Where $\mu^f (x)$ denotes the membership degree of x in the fuzzy set

3.4.2. *Merging.* The "merging"process aims to reduce the fuzzy partitions number by considering adjacent sets. It proceeds recursively: at each stage, and starting with the initial partition, two sets are merged, according to a criterion based on distance metrics (given in the next section) between the data points of each set. Before introducing the notion of distance, we will first explain how to merge two fuzzy sets.

Merging two fuzzy sets $f_2$and f_3 involves two main operations:

1) Generate a new fuzzy set f_{23} such that:

-the resulting fuzzy set center is defined by the following formula:

$$C^{f_{23}} = \frac{G^{f_2} \; c^{f_2} + G^{f_3} \; c^{f_3}}{G^{f_2} + G^{f_3}} \tag{7}$$

Where G^f corresponds to the weight assigned to fuzzy set f (defined in(6))
-the left and right breakpoints are defined as follows:

$$left\,(f_{23}) = left\,(f_2)\,, right\,(f_{23}) = right\,(f_3)$$

This first operation is illustrated in Figure2(a).

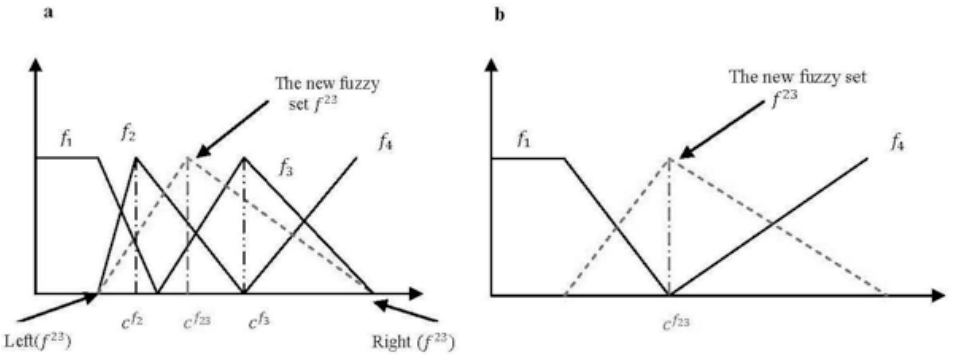


FIGURE 2. Merging Process: a) Merging Two Fuzzy Sets f_2 and f_3 to Form a New Fuzzy Set f_{23}, b) Updating Neighbouring Fuzzy Sets f_1 and f_4

2) Update the neighbouring fuzzy sets f1 and f4, by moving the right breakpoint of f1 and the left breakpoint of f4, to coincide with the new fuzzy set center (as shown in Figure2(b)). This definition makes it possible to preserve the strong fuzzy partition property (a strong fuzzy partition should satisfy the following condition: for each point, the sum of membership degrees in all the fuzzy sets is equal to one). The first and the last fuzzy sets in the partition are of semi trapezoidal shape.

3.4.3. *Distance metrics.* The merging criterion is based on the notion of distance, which is calculated from three partial distances: the internal distance, the prototype distance and the external distance.

Internal distance: It is the distance between two data points belonging to a fuzzy subset *f*. It is computed by differencing their membership degrees.

$$dist_{int}\,(a,b) = |\mu_a^f - \mu_b^f\;|, \quad 0 \leq \; dist_{int}\,(a,b) \tag{8}$$

μ_a^f, μ_b^f quantify the fuzzy memberships degrees of a and b in the considered fuzzy set *f*, respectively.

Prototype distance: It is defined as the distance between the prototype kernels of two fuzzy sets. Recall that a prototype is such that μ(x) =1. Usually we use

numerical prototype distances, which correspond to the distance between fuzzy set kernel locations.

$$dist_{prot}\left(prot_A, prot_B\right) = \sqrt{\left(prot_A - prot_B\right)^2} \tag{9}$$

Where $prot_A$, $prot_B$ represent the prototype kernels of fuzzy sets$A$ and B, respectively.

External distance: It is defined as the combination of the internal and the prototype distances. For two points a,b belonging to two fuzzy sets A_a, A_b respectively:

$$dist_{ext}\left(a, b\right) = dist_{int}\left(a, b\right) + dist_{prot}\left(A_a, A_b\right) + D_c \tag{10}$$

Where, D_c is a constant correction factor. It is used to ensure that the internal distance is always inferior to the external distance. Two types of corrections are considered: the multiplicative correction and the additive correction (more details can be found in[32]).

Now we can define the global formula of the pairwise data point distance given by the formula:

$$dist\left(a, b\right) = \frac{1}{\sum_{A=1}^{S} \mu_a^A} \sum_{A=1}^{S} \left[\mu_a^A \frac{1}{\sum_{B=1}^{S} \mu_b^B} \sum_{B=1}^{S} \left[\mu_b^B \; dist_{A,B}(a, b) \right] \right] \tag{11}$$

Whereμ_a^A, μ_b^B represents the membership degrees of a, b in the sets A, B respectively. S is the number of fuzzy sets composing the partition (partition size).

3.4.4. *item Merging criterion.* During the aggregation process, the sum of the distances is computed for each possible fusion of two adjacent fuzzy sets on the whole training set. This sum is defined as:

$$sum_dist = \frac{1}{n(n-1)} \sum_{a,b \in \{1..n\}} dist\left(a, b\right) \tag{12}$$

Where n is the number of examples of the training set. The retained fusion should satisfy the criterion of minimizing the variation of thesum_dist. Consequently, the size of the partition under consideration is reduced by 1 at each step.

3.5. **Fuzzy Decision Trees Algorithm.** Decision tree is one of the most powerful and practical algorithms in machine learning. This model classifies instances by sorting them down the tree from the root to leaf nodes, according to the discriminating power of attributes. A node is associated to an attribute describing the examples of the training set. The leaves correspond to a possible class of examples. An attribute that best partitions the data set is chosen as the splitting attribute and the training data are then partitioned into disjoint subsets following a given criterion.

The Top down Induction of Decision Tree (TDIT) consists of building a tree from its roots to its leaves. The process is heuristically guided by choosing the most informative attribute at each step. The general approach is summarized in the following steps [4],[20],[13],[14],[29]:

1- Select an attribute Att¸ of n modalities, according to the impurity measure. This attribute represents one internal node and has exactly as many arcs as its number of modalities.

2- Divide the training database according to a partitioning strategy, by creating a subset for each attribute modality.

3 Check the stopping criterion on each subset. A leaf is created if the stopping criterion is triggered.

4- Return to 1, to recursively apply the procedure on nodes that do not satisfy the stopping criterion.

In the context of fuzzy data

In the context of fuzzy data, the selected attribute must generate less impure data sets in child nodes, and must reduce uncertainty on the classes of examples, as well as possible. To determine the quality of attributes as branching attributes, we use the fuzzy events entropy measure (entropy star)[21]. It is an extension of the entropy of Shanon, where the traditional probabilities are replaced by the fuzzy event probabilities. In the context of fuzzy probabilities, we consider$V=\{V_1,V_2\ ,\ldots..,V_n\ \}$as a fuzzy set defined by the fuzzy membership function μ. P is considered as a value of a fuzzy probability of occurrence of, and μ quantifies the corresponding fuzzy membership degree. According to Zadeh [40], the probability of the defined fuzzy event V (since fuzzy events are considered as fuzzy sets) is given by the formula:

$$P\left(V\right)=\sum_{i=1}^{n}\mu(V_i)P(V_i) \tag{13}$$

In what follows, we introduce the fuzzy conditional entropy measure [37] to quantify the best attribute.

Let us consider:

ϵ: training data set.

Att: An attribute defined by the values $\{\ Att_1,\ Att_2,\ldots..Att_m\ \}$, where ATT_i is a fuzzy subset belonging to the domain of Att.

$P^*(\frac{C_i}{Att})$: Probability that an example having $ATTj$as value of Att, belongs to the class C_i. It is defined by the following formula:

$$P^*\left(C_i/Att_j\right)=\frac{|\epsilon_{C_i}\cap\epsilon_{Attj}|}{\epsilon_{Attj}} \tag{14}$$

Where $\epsilon_{C_i}$is the fuzzy subset of the examples belonging to the class C_i and is the fuzzy subset of the examples whose value for Att, belongs to Att_j. Thus, we define the fuzzy entropy of the training data set as follows:

$$ENT^*\left(\epsilon\right)=-\sum_{i=1}^{n}\ P^*\left(C_i/Att_j\right)\ log\ \ P^*\left(C_i/Att_j\right) \tag{15}$$

The entropy of the whole examples conditioned by the attribute Att is given by:

$$ENT^*\left(\epsilon/Att\right)=\sum_{j=1}^{m}\ \frac{P^*\left(Att_J\right)}{\sum_{k=1}^{m}P^*\left(Att_k\right)\log P^*\left(Att_k\right)}*I^*\left(\epsilon_{Attj}\right) \tag{16}$$

Now, we calculate the information gain as follows:

$$Gain^* (\epsilon/Att) = ENT^* (\epsilon) - ENT^* (\epsilon/Att) \tag{17}$$

The attribute that leads to the largest information gain, is selected as the branching attribute.

Recall that the information gain measures the quality of an attribute as the branching attribute. It quantifies the information content provided by the attribute modalities, after reduction of the examples set entropy when using it to split the training set.

To partition data set, several strategies were proposed, some suggest splitting up all the examples into sub classes according to a membership degree [28], others use an α-cut, the distribution is then, done according to the value of α [13].

3.6. **Inference.** A fuzzy inference system is a scheme which connects the output variables to the observed variables, from a set of fuzzy decision rules and a set of data represented by fuzzy membership functions. It is seen composed of three blocks: fuzzification, inference engine, and defuzzification. The fuzzification step is designed to convert numerical values to linguistic variables, by calculating membership degrees from the membership functions defined for each variable. the inference engine evaluates the membership degree of the premise for each applicable rule. Rules degrees are, then, combined to calculate the conclusion for each possible output. The defuzzification process transforms fuzzy data resulting from the aggregation rules into a value which corresponds to the final decision. In this paper, Takagi-Suggenos defuzzification method is used to classify a new object [36]. Recall that this model uses a single spike as the membership function of the consequent rule, instead of a fuzzy set (as Mamdani model). The process is as follows: after fuzzifying inputs, the activation degree (weight) of each rule is calculated by applying the AND fuzzy operation intersection of the membership degrees associated with the input variable values. The elected class can then be calculated by the weighted-average of all rule outputs as:

$$W = \frac{\sum_{i=1}^{n} W_i \;\; C_i}{\sum_{i=1}^{n} W_i} \tag{18}$$

Where W_i is the weight of the output class C_i, and n is the number of rules. Recall that The weight W_i is the degree of class C_i, It corresponds to the conclusion of the inference rule, resulting from a conjunction operation of the premise elements of the rule.

4. Results and Discussion

In this section, we present and discuss our experimental results

4.1. **MIAS Database of Digital Mammogram.** Experiments have been conducted on images of mini-MIAS database (Mammographic Image Analysis Society database (UK)) [35]. The database contains 322 digitized films and is available on 2.3GB 8mm (ExaByte) tape. It contains a description of four kinds of abnormalities (architectural distortions, stellate lesions, circumscribed and ill-defined masses,

and calcification). MIAS database is, actually, the most accessible database and the most used in experimental research.

4.2. **Segmentation and Features Extraction.** A series of pre-processing steps have been applied on the analysed images, thus, we applied the median filter, when necessary, to eliminate the impulse noise, followed by a contrast enhancement based on exponential distribution. This last processing is useful when the breast is dense, because both breast tissue and breast lesion appear white on a mammogram, and the contrast is low; thats why tumors are difficult to spot in women with dense breasts. Contrast enhancement based on the exponential distribution makes it possible to bring out regions with high intensity of the dynamic range.

Experiments show the ability of LCV models to segment more complex masses with weak boundaries, and with significant inhomogeneous**i**ntensity. We take as a baseline for the LCV model, the following set of parameters: λ_2=1, λ_1= {0.1, 0.01} for images with inhomogeneous intensity, and λ_1=1 for images without inhomogeneous intensity (in most cases), α is dynamically adjusted, the time-step Δt =0.1(this value must be chosen carefully, because a high value of Δt may lead to incorrect boundaries). Finally, we choose the value of 9 as the window size of the averaging convolution operator. It should be mentioned that in all our experiments, the initial form (circle) is placed on the mass to be segmented. Let us note that the selection of the segmentation parameters is particularly difficult, taking into account the great morphological variability of masses, and their contrast which can be particularly weak. For example, masses generally appear to be dense, bright regions, while, architectural distortions are defined as a derangement or disruption of the normal course, resulting in a radiating or haphazard pattern with no definite visible mass. So, a bad segmentation may invalidate all posterior process. Figure3 shows the segmentation results of a circumscribed mass. It can be seen that the LCV model has succeeded in the segmentation of the desired region, where the intensity decreases gradually from the left to the right (Figure 3(b)). The LCV curve successfully approaches the true boundaries after a few iterations. Figure 4(a) presents an architectural distortion difficult to distinguish from the surrounding parenchyma. The mass is difficult to be defined due to the great similarity between the mass region and neighbouring structures. Figure 4 (b), shows the contrast enhancement result. In Figure 4(c), we can see that the evolving curve of the LCV model quickly expands to surround the abnormality.

To assess the robustness of the LCV model, the results of the algorithm have been compared to a very popular set level algorithm: the Osher and Sethian model [3](an edge-based level set model), in which the curve evolution is guided by only the gradient information of the image to locate the boundaries of a region. Experiments show in some studied cases, that the Osher-Sethian model provides false boundaries detection results, namely when the breast is dense and the abnormality is difficult to distinguish from the surrounding parenchyma (presence of weak boundaries). This is due to the fact that this model is based on global information (gradient) to control the curve evolution. Since the LCV scheme incorporates local intensity

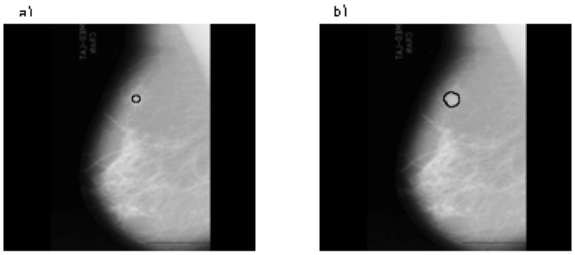


FIGURE 3. Circumscribed Mass Segmentation Without Contrast Enhancement: a) A Circumscribed Mass with Initial Contour, b) Final Segmentation Result Using the LCV Model (After 30 Iterations)



FIGURE 4. Segmentation Results of a Suspicious Region: a)An Architectural Distortion, b) Contrast Enhancement Based on Exponential Distribution, c) Final Segmentation Result Using the LCV Model (After 100 Iterations)

information, the evolving curve quickly expands to surround the suspicious region. In Figure5 and Figure6 we can see that the LCV model yields better results in segmenting lesions with weak boundaries, and the Osher-Sethian algorithm failed to reach the true contour. Mass analysis is mainly based on: form (oval, round, lobuled . . .), contour (circumscribed, spiculated, vll-defined . . .), and density (high, medium, weak) attributes. So, specific information on each region is quantified ba a vector of six attributes: surface, mean gray, compactness, variance, perimeter, and elongation. We noted that with the selected attributes, we obtain satisfactory results. Tab.1 presents the attribute vectors calculated for 3 suspicious regions from the MIAS database. If we take mdb 23 (which corresponds to Figure 3), it can be observed that the mass is of a light homogeneous color, so of low variance, elliptical and elongated in shape rather than circular.

4.3. **Fuzzy Partition Construction.** In section 3.4, we have presented the Hierarchical fuzzy partitioning method to construct fuzzy partitions. All membership functions being created are of triangular or semi-trapezoidal shapes. These forms are widely used, since they can represent fuzzy values well enough, and provide the best speed computing. The system accepts ar input data, a data file corresponding to the training set, and the desired number of fuzzy sets. By varying the tolerance value α, the fuzzy partitions vary more or less. For $\alpha = 0.25$, we obtained more interesting results. Figure7 shows the generated membership functions for the six attributes: compactness, elongation, variance, mean gray, surface, and perimeter.

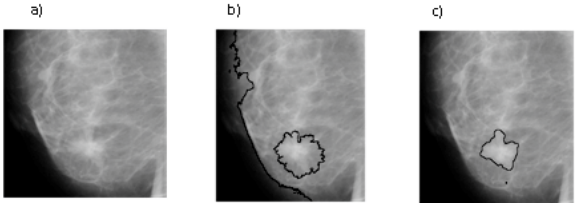


FIGURE 5. Segmentation Results of a Spiculated Mass. a) Original Image, b) The Osher Curve Evolution Pass Through the True Boundaries (After 120 Iterations), c) Segmentation Result Using the LCV Model (After 190 Iterations)

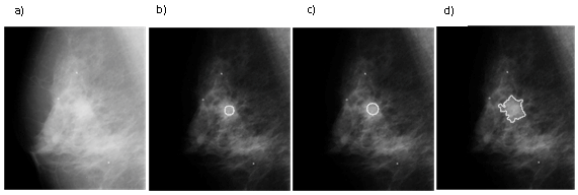


FIGURE 6. Segmentation Results of a Ill-defined Mass a) Original Image, b) Contrast Enhancement with Initial Contour, c) The Osher-Sethian Curve Evolution Stopped on High-frequency Area, d) Segmentation Result Using the LCV Model

MIAS mammogram	abnormality	intensity	surface	compactness	elongation	Variance	perimeter
Mdb 23 (figure 3)	circumscribed	189	1381	0.71	0.9	18.21	156
Mdb 184	spiculated	203	34475	0.04	0.8	79.47	3131
Mdb 130	Architectural distortion	185	1260	0.42	0.71	10.29	194

TABLE 1. Feature Vector Calculated for 3 Suspicious Regions From MIAS Satabase

4.4. **Model Construction and Classification.** To construct our classification model, we followed Takagi-Suggeno's defuzzification method (see section 3.6). Our interpretation system involves two main steps:

Step 1: mass type identification:

During this first step, an initial fuzzy decision tree is constructed, to identify for each candidate region the corresponding class: circumscribed mass, ill-defined mass, architectural distortions or spiculated mass.We have used 72 mammogram as training set describing four classes of abnormalities (16 circumscribed masses, 9 ill-defined masses, 15 spiculated masses, 13 architectural distortions and 19 normal cases). Test experiments were conducted with a set of 63 mammogram (18

circumscribed masses, 15 ill-defined masses, 17 speculated masses, 13 architectural distortions).

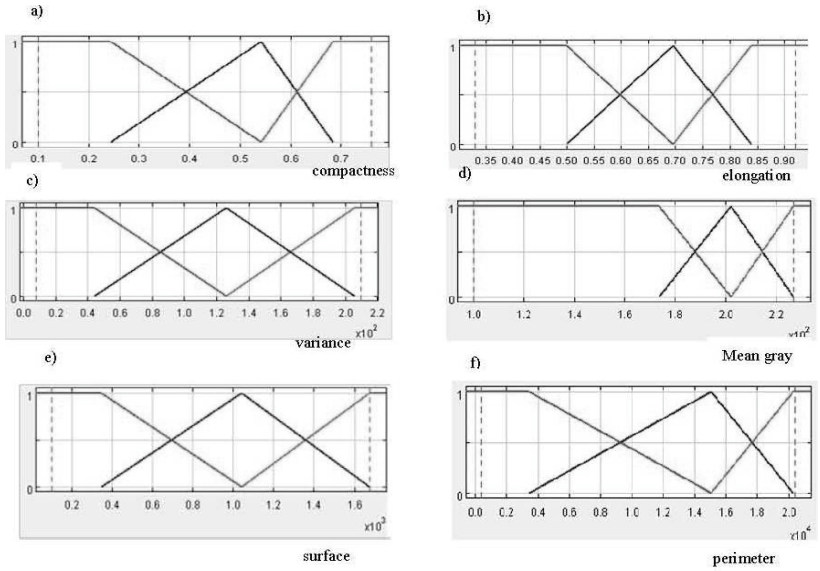


FIGURE 7. The Generated Membership Functions of the Input Variables: (a) Variance, (b) Compactness, (c) Elongation, (d) Mean Gray, (e) Surface, (f) Perimeter

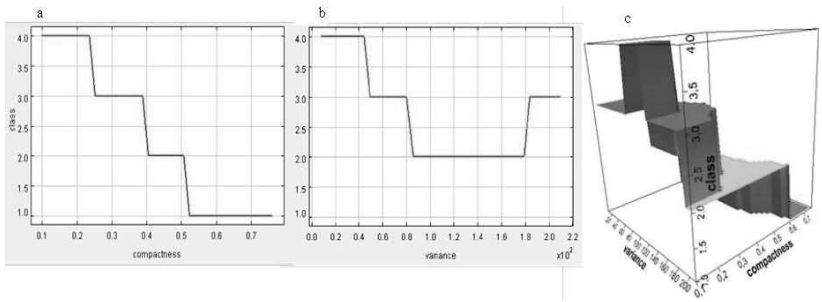


FIGURE 8. Mass Classification Features Impact on Classification (Classes are in Order: 1: Circumscribed, 2:Ill-defined, 3:Architectural Distortions, 4:Spiculated): (a)Compactness Impact, (b)Variance Impact, (c)3D Visualization of Compactness and Variance Impact on the Classification

The use of the entropy measure to select branching attribute allowed us to identify the most discriminating features, for our classification, they are given in order

Training set	Number of rules	Tested set	cases correctly classified	Error rate	coverage rate
72	16	63	53	0.15	100%

TABLE 2. Results of the First Classification

of priority : compactness, variance, and elongation, the surface is the less representative attribute. Figure8.(a) and Figure8.(b) show the impact of compactness and variance, respectively, on the lesion type identification. Circumscribed masses correspond to high values of compactness, unlike spiculated masses whose values of compactness and variance are relatively small. Figure8.(c) shows in 3D, the influence of these two features combined on the lesion classification result. The fuzzy inference system provides 16 linguistic rules. One of these rules is given as follows (see appendix):

-If (opacity=low) and (compactness=low) and (elongation=medium) and (variance=low) and (perimeter=medium) then class = 'architectural distortions'.

Classification results of the first step are summarized in Tab.2, (the coverage rate corresponds to the rate of using all generated rules), and some cases are shown in Figure9. We observed in a few cases, that types of masses are not easily discernible. This is due to some problems related to the location of the lesion, the nature of the breast, and even the tumor morphology (Figure 9.(e) and Figure9.(f)). Note that the classification fields better performance in fatty breasts, because in the glandular and dense beast, the difference between the mass and the normal tissue is less clear. In the mammogram of Figure9, we have applied the LCV model without contrast enhancement to prove the efficiency of this method of segmentation.

Step 2: severity Classification

Following the results obtained from the previous step, a second classification is performed to predict the severity of the abnormality (benign or malignant). We selected a training set for each type of mass. The idea is to restrict the training examples to those that faithfully characterize the mass. In this way, we could filter data, and enrich the information needed to build the classification model. From the training data, four fuzzy decision trees are constructed. After selecting the type of suspicious mass (resulting from the previous step), the corresponding decision tree is invoked, to recognize, more precisely the nature of the suspicion. The corresponding rule base system is defined by the same input variables, and generates a number of classification rules, to determine whether the lesion is benign or malignant, according to the input feature values. The following two rules are obtained from the spiculated tree classifier and the architectural distortions tree classifier, respectively:

-If (surface=small) and (opacity=high) and (compactness=medium) and (elongation=high) and (variance=low) and (perimeter=small) then class = 'malignant'.

-If (surface=small) and (compactness=high) and (elongation=low) and (variance=medium) and (perimeter=small) then class = 'benign'.

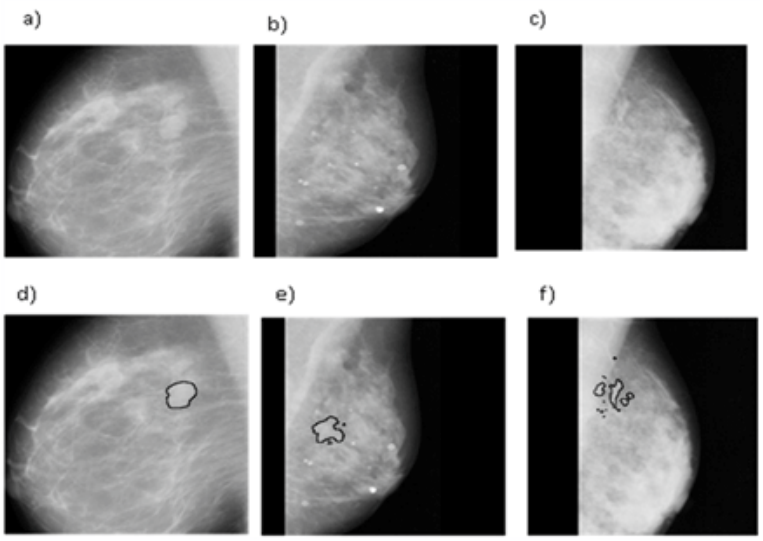


FIGURE 9. Identification of Some Masses by the Model: a), b), c) Original Mammogram Presenting Dense Breasts with Low Contrast and Blurred Contours (d) Mass Identified as Circumscribed Lesion, (e) Mass Identified as Spiculated Lesion, (f) Lesion Identified as Architectural Distortion

Performance evaluation is analysed with confusion matrix (table.3), and detailed results of the classification process are given in table.4. We can see that the proposed CAD system achieves good performances and very satisfactory results, With an overall accuracy rate of around *87. 30%*, a sensitivity (probability to identify disease in people who truly have the disease) of *82.14%*, and good specificity (probability that non-diseased subjects are identified as normal) which reaches a value of *91.43%*. Experiments show that the ill-defined classifier gives a higher accuracy value of *93.33%*, followed by the circumscribed classifier with *88.88%*. The best specificity results are provided by the spiculated classifier with a value of *100%*, followed by the circumscribed classifier with *93.75%*. The ill-defined classifier shows the best results of sensitivity with a rate of *100%*. The output results were low for the sensitivity prediction given by the circumscribed classifier showing a value of *50%*.

In some cases, masses are not correctly classified. This misclassification is due to the nature of some abnormalities difficult to be discerned and distinguished by the segmentation algorithm. Furthermore, the algorithm gives good results when dealing with small masses, because shape variation increases as the mass size increases, and the proposed algorithm is sensitive to some variations in feature values, such as compactness and elongation, which have a strong impact on the classification process, and which may be of close values when dealing with masses of large areas. Figure10 shows the results of the severity classification on two mammogram.

Circumscribed				Ill-defined		
	predicted				predicted	
actual	1 TP	1 FP		actual	5 TP	1 FP
	1 FN	15 TN			0 FN	9 TN
Spiculated				Architectural distortion		
	predicted				predicted	
actual	9 TP	0 FP		actual	8 TP	1 FP
	3 FN	5 TN			1 FN	3 TN

Table 3. Confusion Matrix of Lesion Severity Classification

	Specificity	Sensitivity	Accuracy	Error rate
Circumscribed	93.75%	50%	88.88%	0.11
Ill-defined	90%	100%	93.33%	0.06
Spiculated	100%	75%	83.33%	0.17
Architectural distortion	75%	88.88%	84.61%	0.15
	91.43	82.14	87.30	0.13

Table 4. Results of the Classification Process

author	database	segmentation	classifier	Mass identification	accuracy	specificity	sensitivity
Proposed approach	MIAS	LCVmodel	-Automatic construction of partitions -fuzzy decision trees	yes	87.30%	91.43%	82.14%
Radiologist [19]	Clinical cases	Human knowledge	Human knowledge	yes	Not available	64%	84%
de oliveira [11]	DDSM	K-mean	Support vector machine	no	86%	94.61%	92.63%
Olivier Arnau [22]		-Gray level -Fuzzy c-mean	-Bayes classifier -KNN -decision tree	no	80% no 86% (MIAS)	Not available	Not available
Mohamed J Islam [18]	MIAS	-	-ANN classifier -(MLP) classifier	no	86.97%	83.87%	90.91%
Salve M. S. [31]	MIAS	gray level	SVM classifier	no	86% 89%	85% 87%	89% 87%
Ramani R . [27]	MIAS	-Symlet -singular value decompo. -weighted histogram	-naïve bayes -random forest -neural network	no	81.21% 86.67% 89.7%	Not available	Not available

Table 5. Comparison Results with Other Approaches

A comparative analysis is a complex task for many different raisons, namely: the wide variety of schemes proposed to solve the mass classification problem, the database of mammogram used to check the system performances, and types of breast masses considered in the presented studies. For our evaluation, we have performed comparisons (see table.5) with previously published works. The performance evaluation indicates that our system gives significant and competitive results, and in

most cases are better, namely for the accuracy parameter (considered as the most important measure to evaluate a classifier), except for Slave[31] and Ramani[27], where values are slightly better, however, our system offers other major advantages: it can successfully recognize different types of masses, including architectural distortion which is not covered by the studied works.

Furthermore, the proposed CAD system offers radiologists a large amount of data to enrich the human expertise, with transparent diagnostic decisions based on understandable rules explaining how the classification is obtained, unlike neural network, SVM, ANN classifiers..., considered as black boxes, and showing direct results. More visual experimental results are shown in Figure 11, figure 12 andFigure 13. In Figure11 lesions are correctly classified as benign. In Figure 12 lesions are correctly classified as malignant. Figure13 shows two cases of misclassification.

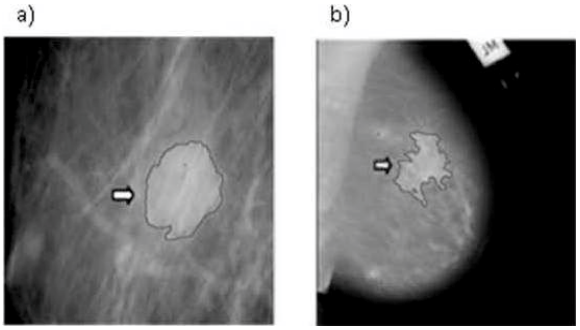


FIGURE 10. Mass Classification: (a) Ill-defined Mass Identified as Benign, (b) Spiculated Mass Identified as Malignant

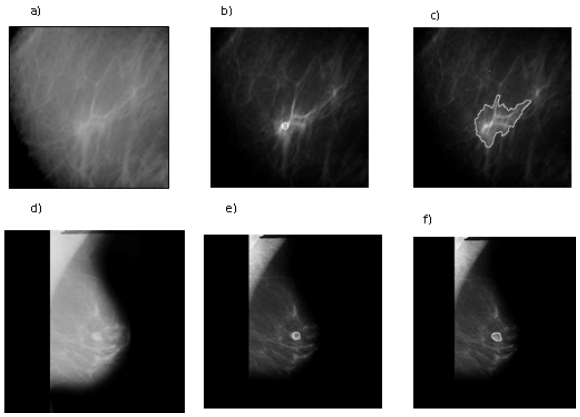


FIGURE 11. Lesions Correctly Classified as Benign. a) Image Presenting Spiculated Mass, d) Image Presenting Circumscribed Mass, b) and e) Image Enhancement with Initial Contour, c) Final Contour After 180 Iterations, f) Final Contour After 50 Iterations

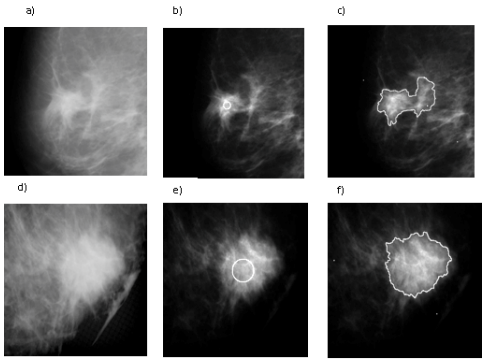


FIGURE 12. Lesions Correctly Classified as Malignant. a) Image Presenting Architectural Distortion, d) Image Presenting Spiculated Mass, b) and e)Image Enhancement with Initial Contour, c) Final Contour After 500 Iterations, f) Final Contour After 300 Iterations

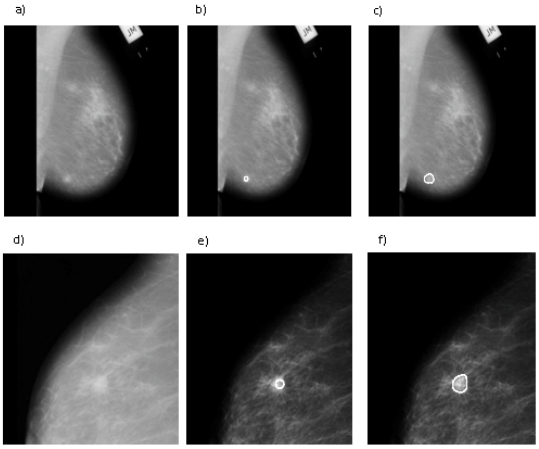


FIGURE 13. Two Misclassification Cases a) Image Presenting Spiculated Mass, d) Image Presenting Architectural Distortion, b) and e) Image Enhancement with Initial Contour, c) Final Contour After 60 Iterations, f) Final Contour After 100 Iterations

5. **Conclusions**

In this paper, we propose a level approach for automatic detection and classification of abnormalities in mammogram, in an attempt to improve severity prediction of a mammographic mass lesion. Experimental results have shown that the LCV segmentation model is an efficient and accurate method to isolate and extract masses in mammogram, and is better adapted to perform the segmentation

of regions with inhomogeneous intensity, weak boundaries and noise. To characterize a lesion we have selected shape features and photometric features. For each of these characteristics, fuzzy partitions are automatically generated using the hierarchical fuzzy partitioning technique (HFP), from a multidimensional training dataset. usually, these partitions are given by radiologists; however, they may not correspond exactly to reality, and may be subject to conflicts and contradictions between experts. Some generated partitions are given in the experimental section. Fuzzy decision trees were employed po solve tea problem of abnormality detection and classification. The proposed classification system is accomplished in two steps: First, we perform an automatic detection os the type of the suspicious regions (circumscribed mass, spiculated mass...), next, abnormalities are classified as benign or malignant. From the obtained decision tree classifiers, a set of interpretable fuzzy classification rules are generated, and presented in a form that can be easily interpreted and exploited by experts. They can even serve as a basis of rules for inference in other fuzzy systems. Experiment results show that the proposed approach gives satisfactory performance, since, it achieves an accuracy rate of *87.30%,* a sensitivity of *82.14%*, and good specificity of *91.42%*. The best classification results in terms of specificity are given by spiculated and circumscribed masses classifiers. Ill-defined and architectural distortion classifiers give better sensitivity rate. Higher accuracy values are obtained from ill-defined and circumscribed mass classifiers. Performance valuation indicates that our CAD system is often better than other performances reported in previous works, certainly it does not exhibit the best values in terms of accuracy if we take the average of the results, but it is better for the ill-defined classifier (with *93.33%*), and the circumscribed classifier (with *88.88%*). Additional benefits of our CAD system is its robustness to identify the mass type as well as the abnormality severity (benign or malignant), and its possibility to deal with multiple lesions including architectural distortion, which is often subject to specific works using specific schemes.

finally, we cannot ignore the major drawback of this approach which lies in choosing the different parameters, in particular for the LCV segmentation model and the hierarchical fuzzy partitioning algorithm. It was necessary to carry out a large number of tests to retain the best values. As future work, we can propose to experiment the system with further features, and to use feature selection algorithms (such as the greedy stepwior method, sequential forward selection algorithm. . .) to remove irrelevant and redundant features. Furthermore, we propose to examine different fuzzy partitions generated with different algorithms, to show the influence of input data structure on the classification accuracy.

6. **Appendix**

The set of inference rules generated to identify the type of masses:

(1) If (opacity=low or medium) and (compactness=low) and (elongation=low) then class = ‘architectural distortions’.
(2) If (opacity=high) and (compactness=low) and (elongation=low) then class = ‘spiculated ’.

(3) If (opacity=low or medium) and (compactness=low) and (elongation= medium) and (variance=low) and (perimeter=small) then class = 'spiculated '.
(4) If (opacity=low) and (compactness=low) and (elongation=medium) and (variance=low) and (perimeter= medium) then class = 'architectural distortions'.
(5) If (opacity=low or medium) and (compactness=low) ant (elongation= medium) and (variance=medium) then class = 'spiculated mass'.
(6) If (opacity=low or medium) and (compactness=low) and (elongation= medium) and (variance=high) then class = 'ill-defined mass'.
(7) If (opacity=medium) and (compactness=low) and (elongation=medium) and (variance=low) and (perimeter=medium or high) then class = 'architectural distortions'.
(8) lf(opacity=high)and(compactness=low)and(elongation=medium) then class = 'architectural distortion'.
(9) If (compactness=low) and (elongation=high) then class = 'spiculated mass'.
(10) If (opacity=low) and (compactness=medium) and (variance=low) and (perimeter=small) then class = 'circumscribed mass'.
(11) If (opacity=medium) and (surface=small) and (compactness=medium) and (variance=low) and (perimeter=small) then class = 'ill-defined mass'.
(12) If (opacity=medium) and (surface=medium) and (compactness=medium) and (variance=low) and (perimeter=small) then class = 'circumscribed mass'.
(13) If (opacity=high) and (compactness=medium) and (variance=low) and (perimeter=small) then class = 'ill-defined mass'.
(14) If (compactness=medium) and (variance=low or medium) and (perimeter=medium) then class = 'circumscribed mass'.
(15) If (compactness=medium) and (variance=high) then class = 'ill-defined mass'.
(16) If (compactness=high) then class = 'circumscribed mass'.

References

[1] S. N. Acho and W. I. D. Rae, *Dependence of shape-based descriptors and mass segmentation areas on initial contour placement using the chan-vese method on digital mammograms,* Computational and Mathematical Methods in Medicine, **2015** (2015), 1-16.
[2] American Cancer Society, *Cancer facts and figures,* Atlanta, Ga: American Cancer Society, (2013), 1-60.
[3] R. Bellotti, *A completely automated CAD system for mass detection in a large mammographic database*, Medical Physics, **33(8)** (2006), 3066–3075.
[4] L. Breslo and D. Aha, *Simplifying decision trees: a survey*, The Knowledge Engineering Review, **12(1)** (1997), 1-40.
[5] L. F. A. Campos, A. C Silva and A. K. Barros, *Diagnosis of breast cancer in digital mammograms using independent component analysis and neural networks*, X Iberoamerican, Conference on Pattern Recognition, Havana, Lecture notes in computer science, **3773** (2005), 460–469.

[6] T. F. Chan and L. A. Vese, *Active contours without edges*, IEEE Trans. Image Process, **10(2)** (2001), 266-277.

[7] R. Crandall, *Image segmentation using Chan Vese algorithm*, ECE532 Project fall, 2009.

[8] A. Keles and Y. Ugur,*Expert system based on neuro-fuzzy rules for diagnosis breast cancer*, Expert Syst Appl., **38(5)**(2011), 5719–5726.

[9] U. Khan, H. Shin, J. P. Choi and M. Kim, *Weighted fuzzy decision trees for prognosis of breast cancer survivability*, Proc of the Australasian Data Mining Conferenre, Glenelg, South Australia, **7(3)** (2008), 141-152.

[10] Z. Lei and K. Ardrew Chan, *An artificial intelligent algorithm for tumor detection in screening mammogram*, IEEE Trans. on Medical Imaging, **20(7)** (2009), 559-567.

[11] M. Leonardo de Oliveira, G. Braz Junior, C. S. .Aristofanes, A. Cardoso de Paiva and M. Gattass, *Detection of masses in digital mammograms using K-means and support vector machine*, Electronic Letters on Computer Vision and Image Analysis, **8(2)** (2009), 39-50.

[12] A. M.Maciej. Y.J.Lo, P.B.Harrawood and D. G Tourassc, *Mutual information-based template matching scheme for detection of breast masses: From mammography to digital breast tomosynthesis*, Journal of Biomedical Informatics **44(5)** (2011), 815-823.

[13] C. Marsala, *Apprentissage inductif en présence de données imprécises : construction et utilisation d'arbres de décision flous*, Thèse de doctorat, Université de Paris 6, (1988).

[14] C. Marsala, *Fuzzy decision trees to help flexible querying*, KYBERNETICA, **36** (2006), 689–705.

[15] L. Martins, A. Dos Santos, A. Silva and A. Paiva,*Classification of normal, benign and malignant tissues using co-occurrence matrix and bayesian neural network in mammographic images*, Proceedings of the Ninth Brazilian Symposium on Neural Networks, (2006), 479–486.

[16] A. Materka and M. Strzelecki, *Texture analysis methods, A review*, COST B11 Technical Report, Lodz-Brussels: Technical University of Lodz, (1998), 9-11.

[17] G. H. B. Miranda and J. C Felipe,*Computer-aided diagnosis system based on fuzzy logic for breast cancer categorization***,** Computers in Biology and Medicine, **64(1)** (2015), 334-34.

[18] J. I. Mohamed, M. Ahmadi and A. S. A. Maher, *An efficient automatic mass classification method in digitized mammograms using artificial neural network*, International Journal of Artificial Intelligence and Application (IJAIA), **1(3)** (2010), 1-13.

[19] E. Molins, F. Macia and F. Ferrer, *Association between radiologists' experience and accuracy in interpreting screening mammograms*, BMC Health Serv Res., **8(91)** (2008), 1-10.

[20] S. K. Murthy, *Automatic construction of decision trees from data: a multi-disciplinary survey*, Data Min Knowl Disc, **2(4)** (1998), 345-389.

[21] C. Olaru anf L. Wehenkel, *A complete fuzzy decision tree technique*, Fuzzy Sets and Systems, **138(2)** (2003), 221-254.

[22] A. Oliver, J. Freixenet, R Martí *et al.*, *A novel breast tissue density classification methodology*, IEEE Transactions on Information Technology in Biomedicine, **12(1)** (2008), 55–65.

[23] S. Osher and N. Paragios, *Geometric level set methods in imaging*, vision and Graphics, Springer-Verlag, 2003.

[24] G. Palma, G. Peters, S. Muller and I. Bloch, *Masses classification using fuzzy active contours and fuzzy decision tree*, SPIE Symposium on Medical Imaging, San Diego, CA, USA, **6915**(2008), 691509.1-691509.11.

[25] o. Pitchumani Angayarkanni and N. Banu Kamal, *Association rule mining based decision tree induction for efficient detection of cancerous masses in mammogram*, International Journal of Computer Applications, **31(6)** (2011), 1-5.

[26] P. Rahmati, A. Adler and G. Hamarneh, *Mammography segmentation with maximum likelihood active contours*, Medical Image Analysis, **16(6)** (2012), 1167–1186.

[27] R. Ramani and N. Suthanthira Vanitha, *Computer aided detection of tumours in mammograms*, international .Journal of Image, Graphics and Signal Processing, **6(4)** (2014), 54-59.

[28] M. Ramdani, *Système d'induction formelle à base de connaissances imprécises*, Thèse de doctorat, Paris 6, LIP6, 1994.

[29] R. S. Safavian and D. Landgrebe, *A survey of decision tree classifier methodology*, IEEE Transactions on Systems, Man, and Cybernetics, **3(21)** (1991), 660–674.

[30] G. Saborta, *Probabilités, Analyse des données et Statistique*, Ed. Technip, 1990.
[31] M. S. Salve and A. Chakkarwar, *Classification of mammographic images using Gabor Wavelet and discrete wavelet transform*, International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), **2(5)** (2013).
[32] G. Serge and B. Charnomordic, *Generating an interpretable family Of fuzzy partitions*, IEEE Transactions on Fuzzy Systems, **12(3)** (2004), 324– 335.
[33] G. Serge, *Induction de règles floues interprétables*, Thèse de Doctorat, INSA Toulouse, France, 2001.
[34] S. Shanthi and M. BhaskaraR, *Intuistionistic fuzzy C-means and decision tree approach for breast cancer detection and classification*, European Journal of Scientific research, **66(I3)** (2011), 345-351.
[35] J. Suckling, J. Parker, D. R. Dance et al., *The mammographic image analysis society digital mammogram database*, Excerpta Medica International Congress Series, **(1069)** (1994), 375-378.
[36] T. Takagi and M. Sugeno, *Fuzzy identification of systems and its applications to modeling and control*, IEEE Trans. Syst. Man Cybern, **15** (1985), 116–132.
[37] H. D. Thanh, *Mesures de discrimination et leurs applications en apprentissage inductif*, Thèse de doctorat, Université de Paris 6, 2007.
[38] X. F. Wang, D. S. Huang and H. Xu, *An efficient local Chan–vese model for image segmentation*, Pattern Recognition, **43(3)** (2010), 603-618.
[39] Y. Wu, O. Alagoz, M. U. S. Ayvaci, A. Munoz del Rio, A. D. J. Vanness, R. Woods and E. S. Burnsise, *A comprehensive methodology for determining the most informative mammographic features*, Journal of digital imaging, **26(5)** (2013), 941-947.
[40] L. A. Zadeh, *Fuzzy sets*, Information and Control, **8(3)** (1965), 338-353.

Fouzia Boutaouche*, laboraoire SIMPA, Departement d'informatique, Faculte des mathematiques et d'informatique, Universite des sciences et de la technologie d'Oran "Mohamed BOUDIAF", USTO-MB; BP 1505 El M'naouer 31000, Oran, Algerie
E-mail address: fouzia.boutaouche@univ-usto.dz

Nacéra Benamrane, laboratoire SIMPA, Departement d'informatique, Faculte des mathematiques et d'informatique, Universite des sciences et de la technologie d'Oran "Mohamed BOUDIAF", USTO-MB; BP 1505 El M'naouer 31000, Oran, Algerie
E-mail address: nacera.benamrane@univ-usto.dz

*Corresponding author