Biomedical Signal Processing and Control 79 (2023) 104043

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Breast cancer anomaly detection based on the possibility theory with a clustering paradigm☆

Jihen Frikha Elleuch [a,b,*], Mouna Zouari Mehdi [a,b], Majd Belaaj [a,b], Norhène Gargouri Benayed [a,c], Dorra Sellami [a,b], Alima Damak [a,b]

[a] CEM research Lab, Electrical Department at the National Engineering School of Sfax, University of Sfax, 3038, Tunisia
[b] National Engineering School of Sfax, 3038, Tunisia
[c] Digital research center of Sfax, Technopole of Sfax, PO Box 275, Sakiet Ezzit, 3021, Sfax, Tunisia

ARTICLE INFO

Keywords:
Possibility theory
Clustering
Gaussian mixture model
Mass
Microcalcification

ABSTRACT

Breast cancer early diagnosis is a major concern for reducing deadly cases. Related breast tissue anomalies are masses and micro-calcifications, which characterization is essential to obtain a correct diagnosis. At this step, high false negative rates are observed, revealing a high ambiguity, which are crucial to handle with conventional approaches. Possibility theory offers a powerful paradigm enabling to handle a high uncertainty level. For enhancing the classification's accuracy in case of scattered classes, we have redefined our classes based on clustering. Under this paradigm, we adopt in the clustering optimization a criteria based on the Dubois and Prade possibility transformation. In this research, we propose a new possibilistic modeling where a clustering based aggregation of features is applied followed by a fusion of the possibility distributions, for decision making. The above strategy has been applied on two case studies in breast cancer screening: mass and micro-calcification detection, where a set of challenging features are considered. Validation on the CBIS-DDSM public dataset has been undertaken. The proposed system, yielding a mass and micro-calcification detection accuracy respectively of 95.4% and 99.4% (at a specificity respectively of 96% and 100% and with an AUC respectively of 0.96 and 0.996), can be used to assist medical practitioners.

1. Introduction

The anomaly detection is a crucial problem in medical image processing field, seeking at a successful identification at the earliest stage of the disease. Indeed, only experienced practitioner can be confident in the decision making process and most of the time, they organize consulting stuffs for making more reliable decisions. In the case of breast cancer detection, classification is a crucial step, for ensuring a correct diagnosis in further steps. Automation of breast tissue anomaly detection is becoming increasingly important given the tiny scale of either masses or micro-calcifications at an early stage on the one hand, and the risk of confusing masses to normal parenchyma on the other hand.

Recently, medical imaging researchers have applied innovative methods to overcome challenges in digital pathology. Deep learning techniques have shown surprisingly strong results, challenging the accuracy of human experts in image based tasks [1–5]. These studies provide various intelligent deep learning (DL)-based methods for automating detection and diagnosis using 1-D (Electro-encephalography (EEG) signals), 2-D signals (spectrograms of EEG signals) and 3-D signals (Magnetic Resonance Imaging (MRI) related electroencephalography signals and its variants and cardiac MRI signals). Although deep learning theory is becoming the trend to resolve the detection problem, it is not well exploited in the medical field due to the lack of data on specific pathologies. Besides, it is a difficult and a time-consuming task to train a Convolutional Neural Network (CNN) model completely even with the computational power available today [6]. Also, the complexity of their architecture and the different involved parameters to adjust, can induce low performances, that may not be acceptable for clinical use [7,8].

As an alternative, many recent researches are proposed to tackle the problem of detection within imperfect information representation, especially those dealing with uncertainty. The best known are probability, fuzzy logic, belief functions and possibility theory.

☆ This paper describes the results of the research project Hatem Bettaher, under the reference PEJC-2018, funded by the Ministry of Higher Education in Tunisia and the digital research center of Sfax.

* Corresponding author at: CEM research Lab, Electrical Department at the National Engineering School of Sfax, University of Sfax, 3038, Tunisia.
E-mail address: Jihen.frikha@fsegs.usf.tn (J.F. Elleuch).

https://doi.org/10.1016/j.bspc.2022.104043
Received 24 April 2022; Received in revised form 22 July 2022; Accepted 26 July 2022
Available online 11 August 2022

Statistical and probabilistic classification algorithms can be efficient in representing relevant information ensuring a good representation of the different classes as well as discrimination between classes [9, 10]. According to a wide and recent bibliographic research conducted by [11] on 30 years of research in the treatment of uncertain information in the medical field, only 7% of the related work is based on the imprecise probability, including basic probability theories and other advanced probability theories. Although some above researches demonstrate challenging results, they are still limited in modeling ambiguity.

Plausibility of such approaches can be significantly improved, in a possibility theory based framework. Indeed, both possibility theory and probability theory are intended for modeling data uncertainty and coping with ambiguity. Nevertheless, unlike the probability concept, the possibility theory is taking into account the distribution of the probability on the universe, allowing to integrate the necessity concept, which is nearest to human reasoning, yielding more realistic predictions and decisions.

In medical field and particularly in breast screening, some uncertainty is affecting the data due to acquisition sensors, and to the breast tissue compression while screening, causing also some ambiguity in the information representation at medium and then high level, during the interpretation, especially in the case of dense breast tissue.

Although possibility distributions are useful for representing imperfect knowledge, only a few works have used possibilistic classifiers. It was applied either at low level image segmentation in case of breast cancer [12] or for decision making in information theory [13,14]. Nevertheless, no previous work have addressed the problem of feature-based breast anomaly detection within a possibilistic framework.

Moreover, clustering can offer a refinement of the possibilistic model, resulting in more efficient classification performances [10,13, 14]. Indeed, for enhancing classification's accuracy, two aspects may be taken into consideration: scattering within classes and the class discrimination power. A high inter-class variance may be regarded as an evidence of good discrimination. Nevertheless, in case of highly scattered classes, miss-classifications of samples at the class margins are still encountered, even at high inter-class variance. Clustering can recover from this problem by offering a redefinition of the classes, based on more compact and less scattered sets (the obtained clusters). The clustering model, either applying a Gaussian Mixture model or k-Means, can build more compact sets, by fitting the data probability distribution [4,15]. In this framework, optimization of the clustering parameters (number of adequate clusters for each feature), fitting the data, can adopt as performance criterion many metrics such as accuracy, sensitivity, specificity and the AUC (Area Under the ROC Curve). In the AUC, a confidence degree level is varied and we plot accordingly the obtained sensitivity rate versus (1-specificity) rate. Generally, the confidence degree level is taken as threshold the distance separating two samples in a same class. No previous work has considered as a confidence degree level a possibility related metric.

In this paper, we propose a possibility theory-based classification algorithm making use of the clustering paradigm. To address ambiguity and uncertainty affecting medical imaging breast screening data, our major contributions involved in this work are listed below:

- Development of general formalism for possibility transformation of medium level features and their fusion, using the asymmetric Dubois and Prade transformation.
- Optimization of the clustering parameters (number of adequate clusters for each feature) based on a new related metric, which is the AUC taking as a confidence degree level the possibility consistency measure.
- High rates of breast tissues anomaly detection accuracy:for the mass detection, an accuracy rate of 95.4% is obtained, while an accuracy rate of 99.4% is obtained for the micro-calcification detection.

Table 1
A list of glossary terms and abbreviations.

Abbreviation	Term
Acc	Accuracy
ACR	American College of Radiology
ANLS	Automatic Non Linear Stretching
AUC	Area Under the Curve
CADe	Computer Aided Detection
CBIS-DDSM	Curated Breast Imaging Subset of Digital Database for Screening Mammography
DL	Deep Learning
EM	Expectation Maximization
FN	False Negative rate
FP	False Positive rate
FPN	Feature Pyramid Network
GLCM	Gray Level Co-occurrence Matrix
GMM	Gaussian Mixture Model
kNN	k Nearest Neighbors
LBP	Local Binary Pattern
MLBP	Magnitude Gray Level Pattern
MSER	Maximally Stable Extremal Regions
N	Necessity measure
nga	number of clusters in the mass/microcalcification class
ngb	number of clusters in the mass/microcalcification free class
Π	Possibility distribution
π	Possibility measure
ROC curve	Receiving Operator Characteristics curve
SFTA	Segmentation with Fractal Texture Analysis
SLBP	Sign Local Binary Pattern
TN	True Negative rate
TP	True Positive rate
TWQ	Textural Wavelet Quantization

The remainder of this paper is organized as follows: In Section 2, we will describe the related work and draw our main research issues. In Section 3, we will describe the proposed system. In this section, we will start by presenting the possibility based new modeling formalism and the adopted possibility based performance metric. Then, a review of the Gaussian Mixture model and k-means based clustering will be also done in this section. In Section 4, we will describe our experimental set-up, including the validation dataset and the set of the features extracted for our two case studies, as well as the performance metrics. Then, experimental results will be illustrated. Finally, in Section 5, we draw our conclusions and discussion and we give our future work.

A list of the used common glossary terms and abbreviations is summarized in Table 1.

2. Related work

In this section, we start with some previous works on anomaly detection in breast cancer diagnosis, that will be considered later in our further comparative study. Then, we will describe some related work within the possibility theory framework.

A set of conventional approaches have been proposed in the literature for mass and microcalcification detection. In [16], authors propose a mass detection algorithm, which consists of a set of image enhancement steps followed by textural and shape feature extraction and a Support Vector Machine (SVM) based classification. The detection accuracy rate is about 90.85% for the Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM). In [17], authors propose a mass detection system based on the gray difference weight and Maximally Stable External Regions (MSER) descriptor. Experimentation on the CBIS-DDSM shows that the proposed method performs breast mass detection with an accuracy of 94.66%. In [18], authors address the problem of mass detection by proposing an architecture where they make use of anchor free and feature pyramid, yielding a detection accuracy of 93.4%. In [15], authors propose a methodology to detect masses from mammograms, based on k-means clustering algorithm for spatially splitting the mammograms into regions. Each region is then classified through a Support Vector Machine

Table 2
Related work on mass and microcalcification detection.

Works (year)	Dataset (s)	Anomaly	Preprocessing	Main methods	Performances
[22] (2020)	DDSM	Mcs	ANLS	TWQ+SVM	**AUC =0.97**
[19] (2015)	DDSM	Mcs	–	ufilter	**AUC=0.80**
	BCDR	Mcs	–	+ SVM	**AUC=0.83**
[20] (2019)	MIAS	Mcs	median filter	FTA	**Acc=75%**
			+ CLAHE	+SVM	
[15] (2009)	DDSM	Mass	–	GLCM+ k-means	**Acc=85%**
[16] (2019)	MIAS	Mass	CLAHE	shape	**Acc =94.2%**
	MIAS			and	**AUC=0.95**
	CBIS-DDSM			texture features	**Acc=90.44%**
	CBIS-DDSM			+SVM	**AUC=0.90**
[17] (2021)	DDSM	Mass	CLAHE	gray difference	**Acc=94.66%**
			+Decorrelation	weight	
	MIAS		stretch	+ MSER	**Acc=97.64%**
[18] (2021)	DDSM	Mass	truncation normalization	Feature	**Acc=94.3%**
			method	Pyramid	
	INbreast		+ AHE	Network (FPN)	**Acc=93%**

(SVM) as mass or non-mass region. Authors use a set of textural and shape measures to detect suspicious regions. Each textural measure (contrast, homogeneity, inverse difference moment, entropy and energy) is computed through the co-occurrence matrix, for discriminating mass from non-mass elements. Moreover, for microcalcification's detection, in [19], authors have used an u-filter in order to extract a feature vector which has been considered thereafter as input to an SVM classifier. Other related works applied texture features such as in [20] where authors have used the Fractal Texture Analysis called (FTA) followed by an SVM classifier. In addition, authors in [21] have classified breast tissues using the binary Random Forest, for breast microcalcification detection. Accordingly, a detection rate of 95% has been obtained. In [22], authors have used a new Textural descriptor based on Wavelet Quantization approach (TWQ). The obtained feature vector has been considered as input in k-NN (k Nearest Neighbors) and SVM classifiers, resulting in a classification accuracy of 97%. Although the above researches demonstrate challenging results, they do not consider the uncertainty aspect in their analysis. At this step, a high false negative rate can thus be observed, leading to high ambiguity. In this context, conventional approaches are unable to handle such ambiguity.

To cope with such ambiguity, we have considered possibility based advanced theories. In the following, we deal with some related work. Although they do not address the problem of anomaly detection, these works can be inspiring in our methodology. Different theories have been applied to address the classification within uncertainty framework. Table 2 illustrates the above different detailed approaches.

In this optic, the probabilistic framework can be exploited for medical diagnosis of breast cancer for addressing the uncertainty problem. The study of possibility based classifiers is motivated by the ability of possibility theory to handle poor data [23]. Although possibility distributions are useful for representing imperfect knowledge, only a few works have used possibilistic classifiers [24,25]. Decision making within a possibility based framework exploits the possibility distributions of the different classes. Accordingly, for a given request sample, we compute its consistency with each class. At this step, two main strategies which differ basically in the way they handle the consistency values can be adopted:

- In the first category, the sample consistency values π resulting from the possibility distributions of the different classes are handled as features and a classification paradigm is applied for decision making. The related work here is concerned with the problem of feature selection and classification, following a conventional architecture [26,27]. In [26], authors proposed a system devoted to iris recognition, based on possibility theory. Accordingly, authors have handled the consistency levels of the extracted descriptors as simple features to be classified. In an analog possibility based system architecture, authors in [27] have focused on a feature selection using the Shapley importance index. Hence, two aspects are proposed to evaluate the feature importance: the scattering of possibility distribution within one class and its class discrimination power.
- In a second category, the decision making has exploited an essential built-in basic duality in the possibility theory relying on both possibility–necessity based reasoning, yielding a more plausible decision [28,29]. In [28], authors have tackled the problem of foreground segmentation in an outdoor environment, which is highly affected by uncertainty at the source level. Accordingly, a decision based on the possibility–necessity measures allow to cope with the ambiguity. Moreover, in [30], authors proposed a segmentation algorithm using a map of the possibility distribution of the gray level images. A successful segmentation rate has been obtained using a possibility–necessity based concept.

Although both above categories yielded acceptable performances in solving hard classification problems, the first category seems to be under-using some features of the possibility theory, while the second category is naturally applying its built-in principle, including the possibility–necessity measures.

Moreover, we illustrate in the following some related work on clustering, since it is an important phase in our system. Many researches have demonstrated the benefits of data clustering in the classification problem [10,13,14]. Indeed, clustering can create a more relevant partitioning of data, yielding a new redefinition of classes, based on the different clusters, which can be more efficient in case of sparse classes. In general, clustering ends with defining new subsets with less sparsity, that can reduce the miss-classification problem. In [10], authors have adopted in clustering a well known algorithm, which is the Gaussian mixture model. The latter is a probabilistic model for describing normally distributed samples within an overall set. Mixture model's parameters are fitted from data, using the expectation maximization algorithm. The clustering step enhances the classification accuracy rate [10]. Recently, in [13], authors have proposed a new hybrid decision-making scheme for classification, exploiting the duality of probability–possibility in data modeling. In order to enhance classification rates, a new Possibility Maximum Likelihood (PML) criterion was developed. Depending on whether the data are probabilistic or possibilistic, i.e. they are matching a Gaussian Mixture model respectively at the probability distribution or at the possibility distribution, a clustering based on the Expectation maximization is more or less efficient. Accordingly, the possibility distributions are constructed with respect to the clusters. This criterion allows exploiting jointly both probabilistic and possibilistic sources within the same decision-making framework. In [14], authors have proposed a new paradigm for the aggregation of possibility distributions based on k-means clustering. In contrast with the previous approaches of computing with words, the authors proposed an algorithm driven by k-means clustering, to

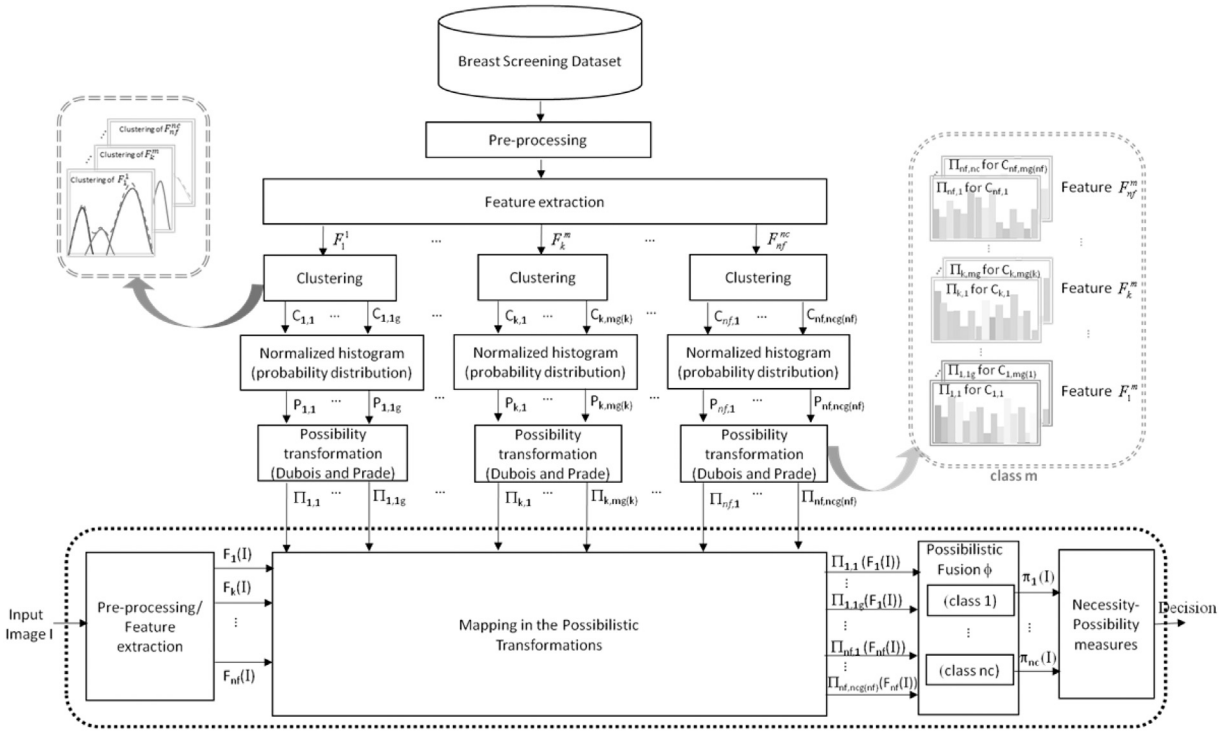

Fig. 1. The general framework of our proposed system where F_k^m denotes the kth feature for the mth class, $\Pi_{k,p}(p = 1, \ldots, mg(k))$ denotes the possibility distribution of the kth feature in the pth cluster, nf denotes the number of features, while nc denotes the number of classes and $mg(k)$ denotes the optimum numbers of clusters for the mth class and the kth feature.

aggregate possibility distributions under the framework of statistical data analysis, intended for reducing information loss and distortion.

In this paper, we will develop the general formalism for possibility modeling allowing data fusion. A clustering step will be integrated in our system for improving the classification accuracy. We will adopt as performance criterion the AUC, taking as a confidence degree level the possibility consistency value.

3. The proposed system

The general framework of our proposed system is illustrated in Fig. 1. In this study, a prior clustering is achieved, for tackling the potential sparsity in the different classes. According to the data distribution, either a Gaussian Mixture model based or k-means based clustering can be fitted. The clustering algorithm is applied for each feature, and each class. A best choice of the cluster number results in an optimal AUC, taking as degree of confidence level the intra-class and inter-class possibility consistency measure of the different samples.

Then, a possibilistic modeling of the obtained clusters is done based on Dubois and Prade transformation. A fusion process is finally applied for generating the whole possibility distribution map. Finally, given a request sample, the different consistency levels to the different clusters are computed, based on their possibility distributions and a decision is generated based on the obtained possibility–necessity measures. In the following, we conduct a detailed description of the different blocs of our system.

3.1. Pre-processing and feature extraction

3.1.1. Mass feature extraction

This work is based on very common features of texture analysis used in this application, which are local binary patterns related features. Local binary Pattern (LBP) operator and its variants proposed by Ojala are a powerful tool for texture classification [31]. In the following, we present a brief formulation of the LBP operator. Let us consider a central pixel and its neighboring pixels. We accordingly start with thresholding the different neighboring pixels by the central gray value. This operator takes "1" if the neighboring pixel is greater or equal to that of the center, and "0" otherwise. Accordingly, LBP code may be formulated as follows:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \tag{1}$$

$$s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{2}$$

where g_c is the central pixel value, g_p is the gray level value of the pth pixel, $(p = 0, 1, \ldots, P-1)$ and P, R are respectively the neighbors' number and the neighboring radius.

LBP has the ability to map all the texture details in the image. However, it cannot extract information relative to the tissue density in a mammographic image. Accordingly, we consider an operator where we extract both gray level and local difference. Let g_{cmean} be the mean value of the central pixels and neighborhood. A new local difference is obtained as follows:

$$D_p = (g_p - g_{cmean}), p = 0, 1, \ldots, P-1 \tag{3}$$

This vector can be decomposed into sign and magnitude, as described in the following equations:

$$D_p = MLBP_p.SLBP_p \tag{4}$$

knowing that:

$$SLBP_p = \begin{cases} 1, D_p >= 0 \\ -1, D_p < 0 \end{cases} \tag{5}$$

and

$$MLBP_p = |D_p| \tag{6}$$

We apply here a local pattern model based on the generation of the difference of the central pixel and its neighbors. A first order

statistics with respect to the obtained image, yields the histogram of the Magnitude of the gray level difference (MLBP) and a histogram of the local difference of gray levels signs(SLBP). In our work, we consider a multi-scale version of these features as a first and second feature, denoted respectively by $MLBP_{MS}$ and $SLBP_{MS}$.

3.1.2. Micro-calcification feature extraction

Here, the input image is pre-processed for enhancing the micro-calcification appearance. Then, we proceed by the extraction of a set of texture features that are conventionally used in such pattern recognition problem. They can be described as follows:

- **Statistical features**
  On the one hand, first order statistical characteristics are generally informative of the breast density and provide a useful mapping of the image background gray level and its first order distribution. Accordingly, we adopt here the following features:
  - Mean of gray level values: The mean value is mapping here the breast density.
  - Standard deviation: is a measure of the variation or dispersion of a set of values. A low standard deviation means that the values are close to the set's mean (also known as the anticipated value), whereas a high standard deviation means that the values are spread out over a wider range.

  On the other hand, second order statistical features are describing the internal relative distribution of gray levels at different resolutions, via the co-occurrence matrix. The gray level co-occurrence matrix (GLCM) is a highly effective texture description technique based on second statistical description. It maps spatial interactions between pixels in multiple orientations and sizes, that translates gray level co-occurrence probabilities. GLCM is used to extract our differential structure features. One GLCM is computed for a specific distance and angle between the pixels with the co-occurring levels. We consider two distances (1 and 2) and four angles (0°, 45°, 90°, and 135° degrees) for each distance in our work, yielding 8 GLCMs. The normalized GLCMs denote the probability of co-occurrence of pixels. In a normalized GLCM, let p(i,j) be the $(i,j)^{th}$ element. Let ng represent the total number of gray levels in the image. It is possible to compute on this matrix classic statistical attributes such as the mean and variance, as well as other attributes such as energy, contrast, entropy, homogeneity, and correlation. They are given by:

$$Mean = \frac{1}{ng^2}\sum_{i=1}^{ng}\sum_{j=1}^{ng}p(i,j) \tag{7}$$

$$Energy = \sum_{i=1}^{ng}\sum_{j=1}^{ng}[p(i,j)]^2 \tag{8}$$

$$Variance = \sum_{i=1}^{ng}\sum_{j=1}^{ng}(i-Mean)^2 p(i,j) \tag{9}$$

$$Contrast = \sum_{i=1}^{ng}\sum_{j=1}^{ng}(i-j)^2 p(i,j) \tag{10}$$

$$Entropy = -\sum_{i=1}^{ng}\sum_{j=1}^{ng}p(i,j)log[p(i,j)] \tag{11}$$

$$Homogeneity = \sum_{i=1}^{ng}\sum_{j=1}^{ng}\frac{1}{1+(i-j)^2}p(i,j) \tag{12}$$

$$Correlation = \sum_{i=1}^{ng}\sum_{j=1}^{ng}(i-\mu_x)(j-\mu_y)p(i,j)/(\sigma_x\sigma_y) \tag{13}$$

where μ_x, μ_y and σ_x, σ_y are the respective means and variances of the two marginal distributions $P_x(i)$ and $P_y(j)$:

$$\mu_x = \sum_{i=1}^{ng} ip_x(i), \qquad \sigma_x^2 = \sum_{i=1}^{ng}(i-\mu_x)^2 P_x(i), \tag{14}$$

$$\mu_y = \sum_{j=1}^{ng} jp_y(j), \qquad \sigma_y^2 = \sum_{j=1}^{ng}(j-\mu_y)^2 P_y(j) \tag{15}$$

Knowing that $P_x(i)$ and $P_y(j)$ are calculated as follows:

$$P_x(i) = \sum_{j=1}^{ng} p(i,j), \qquad P_y(j) = \sum_{i=1}^{ng} p(i,j) \tag{16}$$

The different above statistical features are concatenated in a feature vector that we denote by F_{stat}.

- **Pyramidal Local Binary Pattern features**
  The LBP features are adopted in most texture recognition problems for their high ability to characterize the different local extrema and their orientation in the image. Here, we adopt a pyramidal version, intended for characterizing three image resolutions.
- **Multi-resolution textural wavelet**
  The basic principle here is to quantify the textural information of microcalcifications transformed in the wavelet domain [22]. Let us consider a mammography image to which we have already applied the Automatic Non Linear Stretching algorithm (ANLS) proposed by [32]. After that, a decomposition of the wavelet coefficients has been applied to both enhanced image and Gaussian Derivative filter. In order to obtain micro-calcification's texture information, a product has been applied to both resulting wavelet coefficients. Thereafter, an optimization of the obtained coefficients has been done here to reduce the number of texture feature's vector, based on Principal Component Analysis algorithm (PCA), yielding a reduced textural wavelet quantization feature (RTWQ). Accordingly, the RTWQ is defined by:

$$RTWQ_i = \begin{cases} -1 & if \quad V_i \le -r.s \\ 0 & if \quad -r.std \le V_i \le r.s \\ 1 & if \quad V_i \ge r.std \end{cases} \tag{17}$$

  where s is the standard deviation of V_i. The Final feature vector is obtained by multiplying V_i by 2^{n-1}, n is the row number of V.

3.2. Data clustering

3.2.1. Gaussian Mixture Models (GMM) in high dimensional spaces

Definitions of the specific notations used in this paper are detailed as follows: $T = \{x_i, \dots, y_i\}_{i=1}^{n}$ denotes the training set where: x_i is the features' vector of the ith sample; $x_i \in R^d$; d is the number of features, y_i is the associated label, where $i = 1, \dots, k$ and k is the total number of classes, n is the number of samples and nk the number of samples of class k.

A set of GMMs are used to recover the joint probability distribution map as follows:

$$p(x) = \Sigma_{k=1}^{N}\pi_k f_k(x \mid \theta) \tag{18}$$

$$f_k(x \mid \mu_k, \Sigma_k) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma_k|^{\frac{1}{2}}} exp(-\frac{1}{2}(x-\mu_k)^t \Sigma_k^{-1}(x-\mu_k)) \tag{19}$$

where π_k, μ_k, Σ_k are the related prior parameters such as the mean and the covariance matrix. Maximum Likelihood Estimation (MLE) is an optimization algorithm used for iteratively determining the GMM parameters.

$$\hat{\Theta} = arg_\theta max log(P(X \mid \theta)) \tag{20}$$

However, in practical use, the GMM can be trapped into local minima. To avoid such a case, an initialization step based on k-means allows to recover from such situation.

3.2.2. The k-means clustering algorithm

In some cases, where the Gaussian Mixture model does not fit to the data probability distribution, the EM encounters convergence problems, and the k-means becomes the best clustering alternative. The k-means algorithm is one of the non-supervised learning algorithms that can achieve a data partitioning into a prior number of groups k. We start by computing the k centroïds of each group. After that, a redistribution of the different samples in the data is achieved to assign each sample to the nearest centroïd. This step is followed by an update of the centroïd of each class taking into account the new obtained sample assignment. We repeat the above steps, until convergence. The solution corresponds to a least square minimum of the following criterion:

$$J = \sum_{i=1}^{m} \sum_{j=1}^{n} \|y_j^i - c_i\|^2 \tag{21}$$

where $\|y_j^i - c_i\|^2$ represents the distance that separate an arbitrary point y_j^i to the center of c_i. In this way, J is the measure of similarity between samples and their groups.

3.3. Description of the possibilistic formalism

Possibility theory is one of the well-known theories for handling uncertain information. It is defined as a compromise issue between fuzzy sets and probability. It was introduced by [33] as an extension of fuzzy set theory, suggesting that possibility is able to capture imprecision numerically. In addition, in [34], De Cooman presents a self-contained measure-theoretic framework and suggests more subjective view by focusing on ranking the probabilities of individual events. Equipped with the concepts of necessity and possibility, possibility theory can be employed for quantitative reasoning with imprecise probabilities. Firstly, the possibility formalism, coined by Dubois and Prade in [23], belongs to imprecise probability theory. It is also considered in connection with the belief function. Moreover, quantitative possibility refers to the case when possibility degrees range in the interval [0,1] [35]. The use of the possibility measure, denoted by Π, and the necessity measure, denoted by N, make the main difference compared to the theory of probabilities. For each subset A, the response to the query can be obtained by computing degrees of possibility and necessity, respectively (if the possibility scale $L = [01]$) [36].

$$\Pi(A) = \sup_{s \in A} \pi(s) \tag{22}$$

$$N(A) = \inf_{s \notin A} (1 - \pi(s)) \tag{23}$$

where π is the consistency measure of each element in A and $\Pi(A)$ evaluates to what extent A is consistent with Π, while N(A) evaluates to what extent A is certainly implied by π.

- $\Pi(A) = 1$ if $A \cap E \neq \emptyset$, and $\Pi(A) = 0$ otherwise: function Π checks whether proposition A is logically consistent with the available information E or not.
- N(A) = 1 if $E \subseteq A$, and 0 otherwise: function N checks whether proposition A is logically entailed by the available information E or not.

To apply this formalism in our case, we consider the different features extracted for the whole images within the different classes. For each feature vector F_k^m in class C_m, we derive for each ith component the histogram vector of a number of bins (nbin=10) denoted by H_{F_k}. Thus, we get a matrix of the histogram values for the whole class, given by:

$$H_{F_k^m} = [H_{F_k^m(1)} \ldots H_{F_k^m(i)} \ldots H_{F_k^m(p_k)}] \tag{24}$$

where p_k is the size of the feature F_k^m. The size of this matrix is given by $bin \times p_k$.

Then, $H_{k^{(i)}}^m$ is normalized to get the probability distribution for the same number of bins, as follows:

$$P_{F_k^m(i)} = \frac{1}{N_{C_m}} H_{F_k^m(i)} \tag{25}$$

where N_{C_m} denotes the number of samples in class C_m.

Then, we compute the corresponding probability distribution $P_{F_k^m}$ and then the possibility distribution $\Pi_{F_k^m}$ using Dubois and Prade transformation, which produces the following representation of the feature F_k^m:

$$\Pi_{F_k^m} = [\Pi_{F_k^m(1)} \ldots \Pi_{F_k^m(i)} \ldots \Pi_{F_k^m(p_k)}] \tag{26}$$

where $\Pi_{F_k^m(i)}$ is the possibility distribution of the i^{th} component of the kth feature in the mth class F_k^m. The transformation $p \rightarrow \pi$ of Dubois and Prade is guided by the principle of maximum specificity, which aims at finding the most informative possibility distribution [29]. While the transformation $\pi \rightarrow p$ is guided by the principle of insufficient reason which aims at finding the possibility distribution that contains as much as uncertainty as possible but that retains the features of possibility distribution [29]. This leads to write the consistency principle of Dubois and Prade, given by:

$$\forall A \subset X : \pi(A) \geq p(A) \tag{27}$$

The transformation $p \rightarrow \pi$ is defined by:

$$\pi_i = \Sigma_{j=i}^{n} p_j \tag{28}$$

This transformation, defined by Eq. (28), can be named as asymmetric.

We accordingly get the different possibility distributions of the different features for each class, as follows:

$$\Pi_{C_m} = [\Pi_{F_1^m} \ldots \Pi_{F_k^m} \ldots \Pi_{F_{n_f}^m}] \tag{29}$$

where n_f is the number of features.

The possibility map, taking into account the different class possibility distributions is thus given by:

$$\Pi = [\Pi_{C_1} \ldots \Pi_{C_i} \ldots \Pi_{C_{nc}}] \tag{30}$$

where nc is the total number of classes (here $nc = 2$).

In case of a redefinition of each class by its corresponding clusters, as follows:

$$C_m = C_{m,1} \cup C_{m,2} \cup \cdots \cup C_{m,mg}, \quad m = 1, \ldots, nc$$

where mg is the number of the involved clusters in class C_m and class $C_{m,j}$ is the jth cluster in class C_m.

We thus get for each feature possibility distribution and for each class C_m, the following concatenation of the distributions of the corresponding clusters:

$$\Pi_{F_k^m} = [\Pi_{F_k^{C_{m,1}}} \Pi_{F_k^{C_{m,2}}} \ldots \Pi_{F_k^{C_{m,mg}}}] \tag{31}$$

We accordingly get for each feature a set of possibility distributions with respect to the clusters within the different classes. Such formalism requires a fusion at a source level (denoting here the different features) and at a class level for fusing the different possibility distributions of the different clusters in a single class.

3.4. Possibilistic fusion

Possibility theory could be an excellent framework for information fusion that can take into account both incompleteness and conflict. The conjunctive and disjunctive modes are the basic fusion modes [37, 38]. The conjunctive mode corresponds to the use of the minimum operation which avoids sources dependence. If they are, the product rule can be applied, whereby low plausibility degrees reinforce toward impossibility. Very often, the results of a conjunctive aggregation are sub-normalized, which indicates a conflict. Then, when applying a re-normalization step, this mode of combination becomes fragile especially in case of strong conflict, and in any case the conflict increases with the increase of the sources. The disjunctive mode is more interesting because it avoids conflict, even if inducing a loss of information. The result becomes totally uninformative when many

Table 3
The different hyper-parameters adopted in the proposed approach.

Parameter	Description
Number of bins in the possibility distributions	**10**
Maximum number of clusters within each class	**3**
Applied distance in possibility fusion	**Manhattan distance**

Table 4
Main characteristics of the CBIS-DDSM validation dataset.

Characteristic	description
Number of patients	2620
Number of volumes	43
Number of views	two views (CC and MLO)
Available meta-knowledge	Date of study, age of the patient, type of anomaly, etc
ROI localization	Identification of the outlines of the abnormalities
Mass anomalies	1160 breast tissues including 160 mass anomaly and 1000 mass free.
Mcs anomalies	510 breast tissues including 375 Mcs free tissues and 135 tissues with Mcs

sources are considered. In that case, for merging the distributions of the set of possibilities, we can use a weighted average of the degrees of possibility, or a distance, ensuring a compromise for information preserval [39,40]. Any kind of distance can be used in this fusion. In our work, we considered the Manhattan distance for possibilistic fusion.

3.5. Possibility–necessity based decision making

Given the obtained possibility map, we generate for each sample, a set of consistency measures, with respect to the different defined clusters. One can adopt a threshold for the possibility value, for prediction of a given sample class. Nevertheless, a more efficient decision making can adopt a possibility–necessity based reasoning that will compare the possibility that the sample be in a given class, versus the possibility of the same sample to belong to the other classes. If the former is lower than the latter, the sample will be affected to the first class even though the possibility value is absolutely small.

Table 3 summarizes the different hyper-parameters that we have adopted in our approach.

4. Experimental results and discussion

During the last decade, several works have been carried on Computer Aided Detection (CADe) of breast cancer. Generally, the main difficulty remains the detection of breast tissue lesion screened in mammography. Medical practitioners are commonly based on the Breast imaging-Reporting And Data Systems (Bi-RADS) for breast cancer diagnosis. Such system describes six categories with disease severity levels, where the potential existing anomalies such as microcalcification and masses in the breast are at the heart of the discussion. In this work, we propose an efficient methodology for mass and microcalcification lesion detection, a first and crucial step in breast cancer diagnosis.

4.1. Database overview

We validate our method on the Digital Database for Screening Mammography (CBIS-DDSM) [41]. Table 4 summarizes the main characteristics of this dataset.

CBIS-DDSM is a public dataset collected from 2620 patients, including 43 volumes. Each volume corresponds to different cases, while a case is a set of a meta-knowledge describing the mammography screening relative to one patient. In each case, the CBIS-DDSM database incorporates two images: a cranio-caudal and medio-lateral oblique view (CC and MLO respectively). It includes also the information on the exam such as: the date of study, the breast density, the assessment

Table 5
The adopted performance metrics (TP, FP, TN and FN denote respectively True Positive, False Positive, True Negative, and False Negative).

Performance metric	Expression
Accuracy	$Acc = \frac{TP+TN}{TP+FP+TN+FN}$
Sensitivity	$Sens = \frac{TP}{TP+FN}$
Specificity	$Spec = \frac{TN}{TN+FP}$
AUC	Area Under the ROC curve

categories, etc. of each pathology by means of the lexicon incorporated in the Bi-RADS. It provides, also, the localization of the suspicious regions. With such localization, the outlines of the abnormalities may be identified. The CBIS-DDSM provides delineations of mass regions as well as potential calcifications.

In our work, we have been based in the validation of the proposed mass detection system on a first dataset, which is composed of a set of 1160 breast tissues including 160 breast tissues with mass anomalies and 1000 mass-free tissues; while we validated our proposed microcalcification detection system on a second dataset composed of 510 breast tissues in which 375 tissues are microcalcification-free and 135 tissues are with microcalcifications.

4.2. Performance metrics

In the following, we will describe the different metrics used in the performance evaluation of the proposed system.

4.2.1. Classification accuracy

The accuracy rate is reported by the precision of the system to correctly identify True Positives (TPs) and True Negatives (TNs) respectively. TP rates and TN rates define two common performance metrics in classification which are respectively sensitivity and specificity.

In medical diagnosis, a system with low sensitivity rate in anomaly detection leads to a risk of unnecessary biopsies. While low specificity rates of detection lead to a risk of missing the anomaly, which is much more risky on the patient health. In our experimental validation, we will adopt these performance metrics for the classification.

4.2.2. ROC curve within a possibilistic framework

The most commonly used approaches for evaluating the quality of a given classification paradigm are formally based on the quality of inter-class separation on the one hand and intra-class compacity on the other hand. Both criteria provide an evaluation of the quality of the descriptions in a given classification problem. However, considering two metrics in a classification problem can raise some conflicts between descriptors [27]. This is the case where a selected descriptor has a high inter-class variance, but also it has a relatively high intra-class variance, failing thus at representing classes. Therefore, an evaluation metric, that provides a simultaneous inter-class/intra-class assessment can prevent from such common conflict. Accordingly, the Receiver Operator Characteristic (ROC) curve can be plotted by considering as confidence level, the threshold of intra-class/inter-class distance. Thus, it can perfectly evaluate the quality of class separation as well as intra-class compacity. Thus, the ROC curve and the area under the curve (AUC) can yield a more efficient performance evaluation, by assessing the inter class and intra-class properties in one single metric.

Classification efficiency is rather a measure of the system ability to detect TPs and contrast it with its ability to detect False Positives (FPs). The ROC curves meet this need and were originally developed in the field of signal detection [42]. It is created by plotting the rate of TPs versus the rate of FPs, while varying a degree of a confidence level. In a possibilistic framework, varying the degree of a confidence level for plotting the ROC curve can be naturally done using a variable threshold on possibility consistency measure.

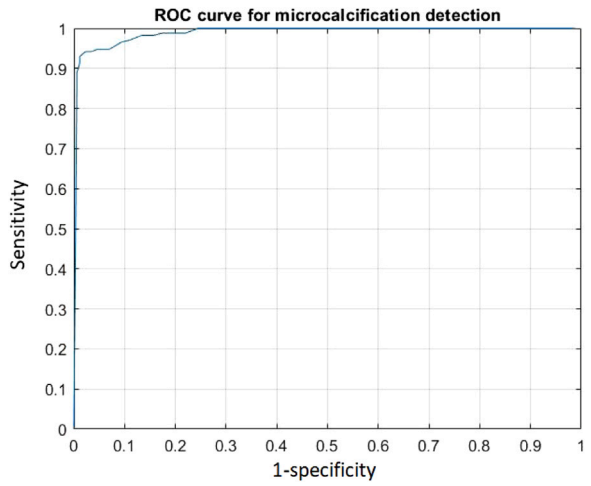

Fig. 2. The obtained ROC curve for micro-calcification detection, with the optimum number of clusters which is respectively for the F_{stat} feature, the pyramidal LBP feature and the $RTWQ$ feature $nga = 2, nga = 3$ and $nga = 2$ respectively for the class micro-calcification and for the same features respectively $ngb = 3$, $ngb = 1$ and $ngb = 1$ for the micro-calcification-free breast tissue class.

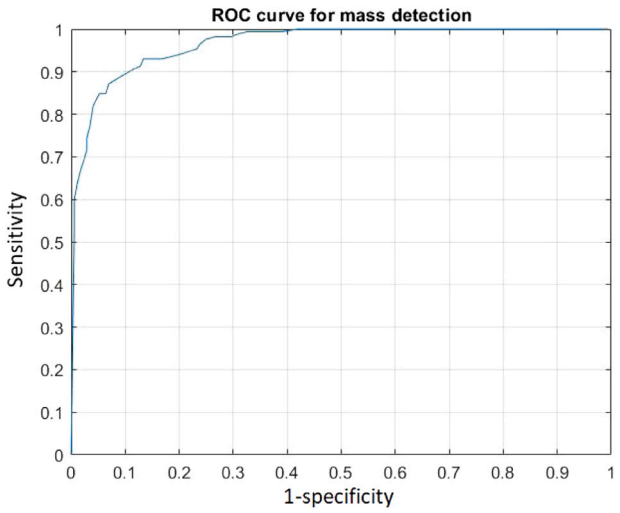

Fig. 3. The obtained ROC curve for mass detection, with the optimum number of clusters which is respectively for the $MLBP$ feature, the $SLBP$ feature $nga = 3$ and $nga = 2$ respectively for the class micro-calcification and respectively $ngb = 2$ and $ngb = 3$ for the micro-calcification-free breast tissue class.

Accordingly, for a given possibility confidence level, samples with higher possibility consistency level, with respect to a particular class, will be associated to that class. While samples with lower possibility consistency levels will be discarded. In this way, we can define the TPs as well as the FPs for a given possibility confidence level. By varying this level, we can plot the ROC curve.

Table 5 summarizes the different performance metrics adopted in our work.

4.3. Experimental results

4.3.1. Clustering optimization

In our both case studies, the clustering models fitting our data are a k-means based clustering for the micro-calcification detection problem and a Gaussian Mixture model for the mass detection case. We accordingly compute the AUCs, while tuning the number of clusters for each class, respectively for the different features. An illustration, the best numbers of clusters, yielding the best AUC, for the micro-calcification detection problem for the F_{stat} feature is $nga = 2$ for the mammographic image with micro-calcification and $ngb = 3$ for the micro-calcification-free breast tissue image. For the pyramidal LBP feature, the best clusters yielding the best AUC is obtained for a cluster number $nga = 3$ and $ngb = 1$ respectively for the mammographic images with and without microcalcifications. Regarding the wavelet quantization feature $RTWQ$, the best AUC is obtained with a number of clusters $nga = 2$ and $ngb = 1$ respectively for the mammographic images with and without microcalcifications.

Figs. 2 and 3 illustrate the ROC possibilistic curves, obtained after optimization of the cluster number, respectively for the micro-calcification detection problem and the mass detection problem.

4.3.2. Mass detection

In this case study, we are based in data clustering on the Gaussian Mixture Model. The adequate choice of the number of clusters in this model should provide the best classification accuracy. The problem of cluster number best choice has been addressed in different researches [10]. It can be tuned and optimized with respect to data distribution, using a performance metric such as Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC). Accordingly, we have optimized the number of clusters for the different features, in the different classes, based on the AUC (using the consistency measures as a confidence degree level).

Table 6
Mass detection Accuracy, Sensitivity, Specificity and AUC for our proposed approach and some basic approaches.

Method	Accuracy	Sensitivity	Specificity	AUC
Used features + k-NN	**93%**	**90%**	**93.5%**	**0.93**
Used features + Decision Tree	**92%**	**89%**	**92.5%**	**0.92**
Used features + Poss model without clustering	**94%**	**93%**	**94%**	**0.94**
The proposed approach	**95.4%**	**91.6%**	**96%**	**0.96**

The Gaussian Mixture Model based clustering improves the AUC for the different considered features. Moreover, the whole system performances are tested with the different Gaussian models, developed for the different features and different classes. Based on these clusters, we have built the different possibility distributions. Then, a possibility–necessity based reasoning yields a classification accuracy of 95.4%. In the basic approaches, we have applied to our features a k-NN classifier and a Decision Tree classifier for highlighting the importance of the possibilistic formalism. In Table 6, we illustrate the results of the proposed mass detection system as well as the basic approaches. In light of these results, the proposed system outperforms approaches based on a classical feature representation, followed by either SVM or k-NN classifiers. Therefore, k-means based clustering, combined with possibilistic modeling, improves classification rates.

For illustrating the effectiveness of our approach, we show in Fig. 4, a set of true positives and true negatives detected masses.

4.3.3. Micro-calcification detection

In this case study, we apply a k-means based clustering. It yields an improvement of the $AUCs$, for the different features. Depending on the data distributions at the different classes, the improvement can be less more important. Moreover, the whole system performances are tested with different parameters, developed at feature level and at class level. Once the different clusters are built, we compute the different new related possibility of the different classes. Then, based on the possibility–necessity measure, we get a classification accuracy rate of 99.4%.

In Table 7, we illustrate the results of the proposed micro-calcification detection system as well as some basic approaches, where we have applied to our features a k-NN classifier and a Decision Tree, to highlight the importance of the possibilistic formalism. According to these

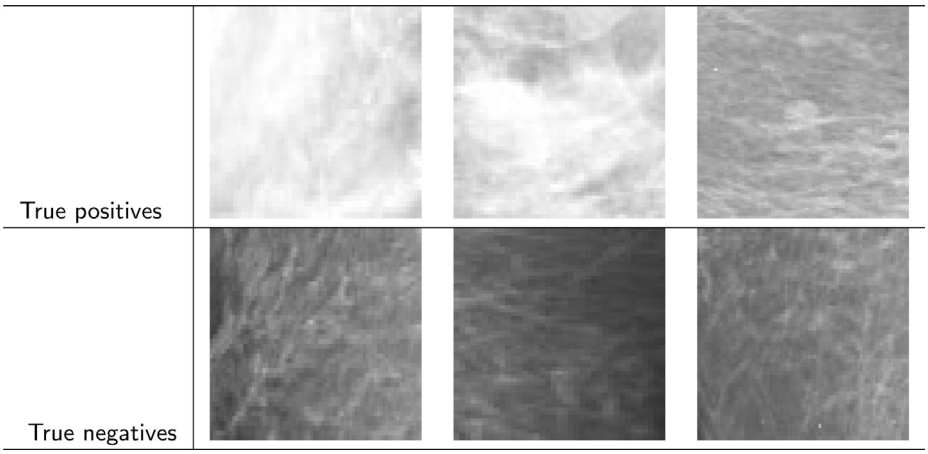

Fig. 4. Illustration of the effectiveness of our approach with some detected True Positive masses and some detected True Negative masses from the CBIS-DDSM images.

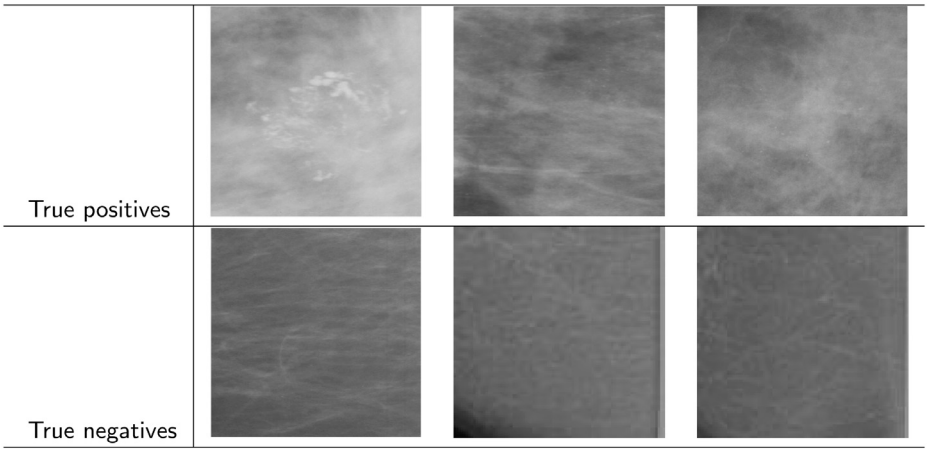

Fig. 5. Illustration of the effectiveness of our approach with some detected True Positives and some True Negatives micro-calcifications from the CBIS-DDSM images.

Table 7
Micro-calcification detection Accuracy, Sensitivity, Specificity and AUC for our proposed approach and some basic approaches.

Method	Accuracy	Sensitivity	Specificity	AUC
Used features + K-NN	**96%**	**93.2%**	**97%**	**0.93**
Used features + Decision Tree	**96.3%**	**94.3%**	**97%**	**0.96**
Used features + Poss model without clustering	**97%**	**96%**	**97.3%**	**0.97**
Our proposed approach	**99.4%**	**97.7%**	**100%**	**0.99**

results, the proposed system outperforms, in term of accuracy of detection, the approaches based on a classical representation of features, followed either by SVM or k-NN classifiers. Therefore, the k-means based clustering, associated to the possibilistic modeling framework, enhances the classification rates.

For illustrating the effectiveness of our approach, we show in Fig. 5 a set of true positive and true negative detected microcalcifications.

4.3.4. Limitations of the study

This section is devoted to discussing the limitations of the proposed method. In our study, we have applied the Dubois and Prade Transformation for measuring the different consistencies and determining the different possibility distributions. Some other possibility transformations may bring more genericity in the probability model such as the Probability Intervals based modeling. Indeed, the latter can partition the universe into probability intervals and propose a more tunable study with respect to the confidence level and the reliability of the decisions.

Moreover, in the case of mass anomaly detection, lower performance rates are observed, compared with the micro-calcification detection case. This can be attributed to the effect of breast density level on the efficiency of the detection process. Accordingly, breast tissues with higher densities have lower contrast, making mass detection more crucial, while there is less challenge in mass detection at lower breast tissue density and thus higher image contrast. The ACR (American College of Radiology) Bi-RADS classification reported four categories of breast densities, from lowest breast tissues to highest ones [43]. Taking into account such meta-parameter can help optimizing the mass detection by an adequate tuning of our approach parameters with respect to the tissue density level.

5. Conclusion, discussion and future work

In this paper, we propose a new system devoted to anomaly detection for both breast tissue anomalies: masses and microcalcifications. The proposed system is based on a possibilistic framework within a clustering paradigm for the definition of the different related distributions. The clustering yields a redefinition of classes, coping with data sparsity. Accordingly, Dubois and Prade possibilistic transformation are then built, for mapping the possibility distributions of the different classes. Dubois and Prade possibilistic transformation are adopted for their higher consistency level. As a performance criterion, we considered the AUC, based on the possibility as a confidence degree

Table 8
Comparison of the proposed methods with the related work.

Method	Anomaly	Accuracy	Sensitivity	Specificity	AUC
Texture & shape features + k-means + SVM [15]	Mass	**85%**	**86%**	**94.6%**	–
Texture & shape features + SVM [16]	Mass	**90.4%**	**90.8%**	–	**0.90**
Gray difference weight + MSER [17]	Mass	**94.66%**	–	–	–
Feature Pyramid Network + Anchor-free [18]	Mass	**94.3%**	–	–	–
The proposed approach	Mass	**95.4%**	**91.6%**	**96%**	**0.96**
TWQ + SVM [22]	Mcs	**97%**	–	–	–
ufilter+ SVM [19]	Mcs	**96.8%**	–	–	–
FTA + SVM [20]	Mcs	**75%**	–	–	–
Haar filter + SURF + Random Forest [21]	Mcs	**97%**	**94.6%**	**100%**	**0.98**
Our proposed approach	Mcs	**99.4%**	**97.7%**	**100%**	**0.99**

Table 9
Comparison of the proposed method with popular deep learning models related work.

Method	Anomaly	Accuracy	Sensitivity	Specificity	AUC
ResNet-50 [44]	Mass	**89%**	**86.1%**	**80.1%**	**0.9**
YOLOv4-based histogram [45]	Mass	**95.74%**	–	–	–
The proposed approach	Mass	**95.4%**	**91.6%**	**96%**	**0.96**
CNN [46]	Mcs	**88.59%**	**88.43%**	**86.89%**	**0.93**
ResNet-50 [47]	Mcs	**96.7%**	**96.7%**	**96.7%**	**0.98**
Our proposed approach	Mcs	**99.4%**	**97.7%**	**100%**	**0.99**

level. In the validation of the proposed approach, we were based on the CBIS-DDSM public dataset. The proposed approach yields an improvement of the breast tissue anomaly detection. A classification accuracy rate of 95.4% has been obtained in mass detection, while the yielded classification detection accuracy rate for the micro-calcification detection is 99.4%. The proposed strategy can be applied for improving any classification problem. A significant improvement is obtained in case of high data sparsity. In Table 8, we illustrate the results of some existing methods for both mass and micro-calcification detection on the same database (CBIS-DDSM). For mass detection, in [15,16], authors proposed to extract features a textural descriptors such as the co-occurrence matrix in addition to a shape descriptor (eccentricity, circularity, perimeter, circularity, area...). For classification, in [15], authors have adopted a mixture of k-means and SVM classifiers and achieved an accuracy rate of 85%, and in [16], they have used only the SVM classifier and they have reached 90.44% as accuracy rate. In addition, authors have used the FPN network in order to extract feature and the Anchor-free classifier to detect breast mass. As a result of this work, they have achieved 94.3% of accuracy.

For micro-calcifications detection, most of authors have used the SVM classifier with the different feature extractor techniques and they have reached high accuracy values: in [22], authors have developed a new one based on the quantization of textural information (TWQ), whereas in [19], authors have proposed the U-filter algorithm. While in [20], authors have used the FTA descriptor. They have adopted the random forest classifier with the combination of the Haar filter and SURF algorithm, yielding an accuracy rate of 97%.

In this table, we note that mainly two classifiers have been used: the SVM and the Random Forest. Nevertheless, a classifier based decision does not integrate the necessity–plausibility measure; which are likely to generate more realistic decisions. This is the reason for which we have adopted a possibilistic based decision given the different consistency measures, yielding to higher sensitivity, specificity, and accuracy rates. Moreover, the higher specificity that we get is highly appreciated by experts since an anomaly free assertion in the case of cancer disease, should be done at a high confidence level.

The proposed formalism yielded better results than existing approaches. However, its strength is that the obtained results are more reliable and more reproducible than those obtained with conventional approaches, since they take into account the data uncertainty and intra-class and inter-class dispersion which are likely to be met in practical situations by doctors.

Considering the run time, our proposed approach is challenging (it includes only an evaluation of the possibility values of a request sample, given the different elaborated distributions, and a classification of that sample based on these values. Nevertheless, the run-time seems to be not so significant, because the existing related work does not evoke this parameter. Indeed, in medical decision systems, the most important is the ability of a system to generate an accurate decision, especially when the system is not real-time. Moreover, in Table 9, a comparison of our model is carried on with recent popular deep learning models. For Mass detection we referred to [44] in which an accurate classification of screening mammograms has been achieved considering a deep learning model trained in an end-to-end structure that relies on clinical ROI annotations. Two deep learning methods for breast cancer classification namely ResNet-50 and VGG16 were proposed. It was found that ResNet-50 outperforms the randomly initialized one by a large margin, achieving an accuracy of 89%, however the accuracy of the pre-trained VGG16 is only 84%. In [45] authors proposed an anomaly detection model of mammography using a YOLOv4-based histogram. Based on image edges enhancement a new YOLO learning model is developed. Performance of this model was evaluated by comparison with ResNet18, ResNet50, GoogleNet, and VGG16 popular deep learning models. According to mass detection experiment, the proposed method had the highest accuracy of 95.74%, followed by GoogleNet (89.9%), VGG16 (88.93%), ResNet50 (87.77%), and ResNet18 (87.67%).

In [46] suspicious region of interested calcification area was extracted by an automatic image pre-processing method, followed by learning radiomics feature based on CNN model. The achieved classification accuracy is 88.59%. Recently, authors in [47] has proposed Patch Learning Approach (PLA) based on ResNet-50 Convolutional Neural Network (CNN) for micro-calcification detection. This approach performed a classification accuracy of 96.7%. The PLA achieved sensitivity and specificity of 96.7% and 96.7%, respectively, with an AUC 0.98.

It is worth noting that our model is as good as YOLOv4 model for mass detection. Indeed, our model applies also for micro-calcifications detection where it transpires an excellent metric performance especially for a specificity of 100%, an accuracy rate of 99.4% and an AUC of 0.99. These findings are higher than all results available in the literature corresponding to micro-calcification detection based on deep learning models. In our future research, an investigation of other possibility transformation than Dubois and Prade may bring a more effective modeling of the uncertainty affecting data, such as Klir transformation, the symmetric transformation, and the probability interval.

Moreover, we plan to take into account the annotated breast density in mass classification for further enhancing its detection rate.

Regarding our system, it allows a reliable detection of the anomalies. Next step, we will be concerned by classification of the severity of the detected anomaly (benign or malign). The possibilistic framework, taking into account data incertitude and imperfections, will be attractive as well in such diagnosis.

Moreover, we will focus on applying the proposed system on covid-19 related lung anomaly detection, which can be very crucial in some cases.

CRediT authorship contribution statement

Jihen Frikha Elleuch: Conceptualization, Methodology, Writing – review & editing, Software. **Mouna Zouari Mehdi:** Methodology, Writing – review & editing, Software. **Majd Belaaj:** Methodology, Software. **Norhène Gargouri Benayed:** Methodology, Software. **Dorra Sellami:** Methodology, Writing – review & editing, Software. **Alima Damak:** Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] Andrew Janowczyk, Anant Madabhushi, Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases, J. Pathol. Inform. 7 (2016).

[2] Afshin Shoeibi, Navid Ghassemi, Marjane Khodatars, Parisa Moridian, Roohallah Alizadehsani, Assef Zare, Abbas Khosravi, Abdulhamit Subasi, U Rajendra Acharya, Juan M Gorriz, Detection of epileptic seizures on EEG signals using ANFIS classifier, autoencoders and fuzzy entropies, Biomed. Signal Process. Control 73 (2022) 103417.

[3] Afshin Shoeibi, Navid Ghassemi, Marjane Khodatars, Mahboobeh Jafari, Parisa Moridian, Roohallah Alizadehsani, Ali Khadem, Yinan Kong, Assef Zare, Juan Manuel Gorriz, et al., Applications of epileptic seizures detection in neuroimaging modalities using deep learning techniques: methods, challenges, and future works, 2021, arXiv preprint arXiv:2105.14278.

[4] Danial Sharifrazi, Roohallah Alizadehsani, Javad Hassannataj Joloudari, Shahab Shamshirband, Sadiq Hussain, Zahra Alizadeh Sani, Fereshteh Hasanzadeh, Afshin Shoaibi, Abdollah Dehzangi, Hamid Alinejad-Rokny, CNN-KCL: Automatic myocarditis diagnosis using convolutional neural network combined with k-means clustering, 2020, Preprints.

[5] A Shoeibi, D Sadeghi, P Moridian, N Ghassemi, J Heras, R Alizadehsani, JM Gorriz, Automatic diagnosis of schizophrenia using EEG signals and CNN-LSTM models. arXiv, 2021, arXiv preprint arXiv:2109.01120.

[6] Salman Ahmed, Maria Tariq, Hammad Naveed, PMNet: A probability map based scaled network for breast cancer diagnosis, Comput. Med. Imaging Graph. 89 (2021) 101863.

[7] Ana C. Perre, Luís A. Alexandre, Luís C. Freire, Lesion classification in mammograms using convolutional neural networks and transfer learning, Comput. Methods Biomech. Biomed. Eng. Imag. Vis. (2018).

[8] Hasan Nasir Khan, Ahmad Raza Shahid, Basit Raza, Amir Hanif Dar, Hani Alquhayz, Multi-view feature fusion based four views model for mammogram classification using convolutional neural network, IEEE Access 7 (2019) 165724–165733.

[9] Vikas Chaurasia, Saurabh Pal, B.B. Tiwari, Prediction of benign and malignant breast cancer using data mining techniques, J. Algorithms Comput. Technol. 12 (2) (2018) 119–126.

[10] Mehrbakhsh Nilashi, Othman Ibrahim, Hossein Ahmadi, Leila Shahmoradi, A knowledge-based system for breast cancer classification using fuzzy logic method, Telemat. Inform. 34 (4) (2017) 133–144.

[11] Roohallah Alizadehsani, Mohamad Roshanzamir, Sadiq Hussain, Abbas Khosravi, Afsaneh Koohestani, Mohammad Hossein Zangooei, Moloud Abdar, Adham Beykikhoshk, Afshin Shoeibi, Assef Zare, et al., Handling of uncertainty in medical data using machine learning and probability theory techniques: A review of 30 years (1991–2020), Ann. Oper. Res. (2021) 1–42.

[12] Marwa Hmida, Kamel Hamrouni, Basel Solaiman, Sana Boussetta, Mammographic mass classification based on possibility theory, in: Ninth International Conference on Machine Vision, Vol. 10341, ICMV 2016, SPIE, 2017, pp. 384–388.

[13] Basel Solaiman, Didier Guériot, Shaban Almouahed, Bassem Alsahwa, Éloi Bossé, A new hybrid possibilistic-probabilistic decision-making scheme for classification, Entropy 23 (1) (2021).

[14] Zhen-Song Chen, Xuan Zhang, Witold Pedrycz, Xian-Jia Wang, Kwai-Sang Chin, Luis Martínez, K-means clustering for the aggregation of HFLTS possibility distributions: N-two-stage algorithmic paradigm, Knowl.-Based Syst. 227 (2021) 107230.

[15] Leonardo de Oliveira Martins, Geraldo Braz Junior, Aristófanes Corrêa Silva, Anselmo Cardoso de Paiva, Marcelo Gattass, Detection of masses in digital mammograms using K-means and support vector machine, ELCVIA ISSN 8 (2) (2009) 39–50.

[16] Ilhame Ait Lbachir, Imane Daoudi, Saadia Tallal, Automatic computer-aided diagnosis system for mass detection and classification in mammography, Multimedia Tools Appl. 80 (6) (2021) 9493–9525.

[17] B.V. Divyashree, G. Hemantha Kumar, Breast cancer mass detection in mammograms using gray difference weight and MSER detector, SN Comput. Sci. 2 (2) (2021) 1–13.

[18] Haichao Cao, Shiliang Pu, Wenming Tan, Junyan Tong, Breast mass detection in digital mammography based on anchor-free architecture, Comput. Methods Programs Biomed. 205 (2021) 106033.

[19] Noel Pérez Pérez, Miguel A Guevara López, Augusto Silva, Isabel Ramos, Improving the Mann–Whitney statistical test for feature selection: An approach in breast cancer diagnosis on mammography, Artif. Intell. Med. 63 (1) (2015) 19–31.

[20] Ashgan Mohamed Omer, Detection of Breast Cancer in Mammogram Images Using Texture Analysis Methods (Ph.D. thesis), Sudan University of Science and Technology, 2019.

[21] Annarita Fanizzi, Teresa MA Basile, Liliana Losurdo, Bellotti, et al., A machine learning approach on multiscale texture analysis for breast microcalcification diagnosis, BMC Bioinformatics 21 (2) (2020) 1–11.

[22] Mouna Zouari Mehdi, Norhene Gargouri Ben Ayed, Alima Damak Masmoudi, Dorra Sellami, A textural wavelet quantization approach for an efficient breast microcalcifcation's detection, Multimedia Tools Appl. 79 (33) (2020) 24911–24927.

[23] Didier Dubois, Henri Prade, Possibility theory: qualitative and quantitative aspects, in: Quantified Representation of Uncertainty and Imprecision, Springer, 1998, pp. 169–226.

[24] Myriam Bounhas, Khaled Mellouli, Henri Prade, Mathieu Serrurier, Possibilistic classifiers for numerical data, Soft Comput. 17 (5) (2013) 733–751.

[25] Didier Dubois, Henri Prade, Towards a logic-based view of some approaches to classification tasks, in: International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Springer, 2020, pp. 697–711.

[26] Bellaaj Majd, Imen Khanfir Kallel, Dorra Sellami, Probability-possibility theories based iris biometric recognition system, ELCVIA Electron. Lett. Comput. Vis. Image Anal. 18 (1) (2019) 21–36.

[27] S Ammar Bouhamed, I Khanfir Kallel, D Sellami Masmoudi, Basel Solaiman, Feature selection in possibilistic modeling, Pattern Recognit. 48 (11) (2015) 3627–3640.

[28] Jihen Frikha, Dorra Sellami, Imen Khanfir Kallel, Indoor/outdoor navigation system based on possibilistic traversable area segmentation for visually impaired people, ELCVIA Electron. Lett. Comput. Vis. Image Anal. 15 (1) (2016) 60–76.

[29] Didier Dubois, Henri Prade, Sandra Sandri, On possibility/probability transformations, in: Fuzzy Logic, Springer, 1993, pp. 103–112.

[30] Waël Ezziddin, Julien Montagner, Basel Solaiman, An iterative possibilistic image segmentation system: application to breast cancer detection, in: 2010 13th International Conference on Information Fusion, IEEE, 2010, pp. 1–8.

[31] Timo Ojala, Matti Pietikainen, Topi Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell. 24 (7) (2002) 971–987.

[32] Mouna Zouari Mehdi, Norhene Gargouri Ben Ayed, Alima Damak Masmoudi, Dorra Sellami, Riadh Abid, An efficient microcalcifications detection based on dual spatial/spectral processing, Multimedia Tools Appl. 76 (11) (2017) 13047–13065.

[33] Lotfi A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 100 (1999) 9–34.

[34] Gert De Cooman, Possibility theory I: the measure-and integral-theoretic groundwork, Int. J. Gen. Syst. 25 (4) (1997) 291–323.

[35] Dominik Hose, Michael Hanss, A universal approach to imprecise probabilities in possibility theory, Internat. J. Approx. Reason. 133 (2021) 133–158.

[36] Didier Dubois, Henry Prade, Possibility theory and its applications: Where do we stand? in: Springer Handbook of Computational Intelligence, Springer, 2015, pp. 31–60.

[37] Didier Dubois, Henri Prade, Possibility theory and data fusion in poorly informed environments, Control Eng. Pract. 2 (5) (1994) 811–823.

[38] Didier Dubois, Henri Prade, Ronald Yager, Merging fuzzy information, in: Fuzzy Sets in Approximate Reasoning and Information Systems, Springer, 1999, pp. 335–401.

[39] Salem Benferhat, Didier Dubois, Henri Prade, From semantic to syntactic approaches to information combination in possibilistic logic, in: Aggregation and Fusion of Imperfect Information, Springer, 1998, pp. 141–161.
[40] Salem Benferhat, Didier Dubois, Souhila Kaci, Henri Prade, Possibilistic merging and distance-based fusion of propositional information, Ann. Math. Artif. Intell. 34 (1) (2002) 217–252.
[41] Rebecca Sawyer Lee, Francisco Gimenez, Assaf Hoogi, Kanae Kawai Miyake, Mia Gorovoy, Daniel L Rubin, A curated mammography data set for use in computer-aided detection and diagnosis research, Sci. Data 4 (1) (2017) 1–9.
[42] David Marvin Green, John A. Swets, et al., Signal Detection Theory and Psychophysics, Vol. 1, Wiley New York, 1966.
[43] Alima Damak Masmoudi, Norhen Gargouri Ben Ayed, Dorra Sellami Masmoudi, Riad Abid, Lbpv descriptors-based automatic ACR/BIRADS classification approach, EURASIP J. Image Video Process. 2013 (1) (2013) 1–9.
[44] Li Shen, Laurie R Margolies, Joseph H Rothstein, Eugene Fluder, Russell McBride, Weiva Sieh, Deep learning to improve breast cancer detection on screening mammography, Sci. Rep. 9 (1) (2019) 1–12.
[45] Chang-Min Kim, Kyungyong Chung, Roy C. Park, Anomaly detection model of mammography using YOLOv4-based histogram, Pers. Ubiquitous Comput. (2021) 1–12.
[46] Hongmin Cai, Qinjian Huang, Wentao Rong, Yan Song, Jiao Li, Jinhua Wang, Jiazhou Chen, Li Li, Breast microcalcification diagnosis using deep convolutional neural network from digital mammograms, Comput. Math. Methods Med. 2019 (2019).
[47] Mahmoud Shiri Kahnouei, Masoumeh Giti, Mohammad Ali Akhaee, Ali Ameri, Microcalcification detection in mammograms using deep learning, Iranian J. Radiol. 19 (1) (2022).