



Article

Selection of the Discriminating Feature Using the BEMD's BIMF for Classification of Breast Cancer Mammography Image [†]

Fatima Ghazi ^{1,*}, Aziza Benkuider ¹, Fouad Ayoub ² and Khalil Ibrahim ¹

¹ Computer Science Research Laboratory (LaRI), Ibn Tofail University, Kenitra 14000, Morocco; aziza.benkuider@uit.ac.ma (A.B.); ibrahimi.khalil@uit.ac.ma (K.I.)

² LAREAMI, CRMEF, Ibn Tofail University, Kenitra 14000, Morocco; ayoub@crmekf.ma

* Correspondence: ghazifatima@gmail.com

[†] This paper is an extended version of our paper Published Fatima, G; Aziza, B; Mohamed, Z; Fouad, A; Khalil, I. Neighborhood Feature Extraction and Haralick Attributes for Medical Image Analysis: Application to Breast Cancer Mammography Image. In Proceedings of the 2023 10th International Conference on Wireless Networks and Mobile Communications (WINCOM), IEEE, Istanbul, Turkey, 26–28 October 2023, <https://doi.org/10.1109/WINCOM59760.2023.10323028>.

Abstract: Mammogram exam images are useful in identifying diseases, such as breast cancer, which is one of the deadliest cancers, affecting adult women around the world. Computational image analysis and machine learning techniques can help experts identify abnormalities in these images. In this work we present a new system to help diagnose and analyze breast mammogram images. To do this, the system a method the Selection of the Most Discriminant Attributes of the images preprocessed by BEMD "SMDA-BEMD", this entails picking the most pertinent traits from the collection of variables that characterize the state under study. A reduction of attribute based on a transformation of the data also called an extraction of characteristics by extracting the Haralick attributes from the Co-occurrence Matrices Methods "GLCM" this reduction which consists of replacing the initial set of data by a new reduced set, constructed at from the initial set of features extracted by images decomposed using Bidimensional Empirical Multimodal Decomposition "BEMD", for discrimination of breast mammogram images (healthy and pathology) using BEMD. This decomposition makes it possible to decompose an image into several Bidimensional Intrinsic Mode Functions "BIMFs" modes and a residue. The results obtained show that mammographic images can be represented in a relatively short space by selecting the most discriminating features based on a supervised method where they can be differentiated with high reliability between healthy mammographic images and pathologies, However, certain aspects and findings demonstrate how successful the suggested strategy is to detect the tumor. A BEMD technique is used as preprocessing on mammographic images. This suggested methodology makes it possible to obtain consistent results and establishes the discrimination threshold for mammography images (healthy and pathological), the classification rate is improved (98.6%) compared to existing cutting-edge techniques in the field. This approach is tested and validated on mammographic medical images from the Kenitra-Morocco reproductive health reference center (CRSRKM) which contains breast mammographic images of normal and pathological cases.

Keywords: bidimensional empirical multimodal decomposition "BEMD"; breast cancer; CT-Scan; the gray level co-occurrence matrix "GLCM"; Haralick characteristics; medical mammography image classification; supervised iterative selection; computer-aided diagnosis; artificial intelligence



Citation: Ghazi, F.; Benkuider, A.; Ayoub, F.; Ibrahim, K. Selection of the Discriminating Feature Using the BEMD's BIMF for Classification of Breast Cancer Mammography Image. *BioMedInformatics* **2024**, *4*, 1202–1224. <https://doi.org/10.3390/biomedinformatics4020066>

Academic Editor: Arcangelo Merla

Received: 2 March 2024

Revised: 10 April 2024

Accepted: 4 May 2024

Published: 9 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The most common cancer among women worldwide is breast cancer. Different organizations publish a number of reports on the epidemiology and degree of breast cancer [1]. As of December 2020, data from the International Agency for Research on Cancer, or "IARC",

indicated that breast cancer had eclipsed lung cancer as the most prevalent disease diagnosed worldwide. The total number of cancers diagnoses over the last 20 years has nearly doubled, from about 10 million in 2000 to 19.3 million in 2020. In the modern world, one in five people will experience cancer at some point in their lives. According to projections, there will be a notable rise in the number of cancers diagnoses in the upcoming years, with estimates indicating a nearly 50% increase from 2020 to 2040 [2]. Mammography is the most popular modality among them all for screening for various breast abnormalities. Mammography, however, has a number of flaws that make it a less accurate or conclusive test. This method sometimes produces unclear images, which causes radiologists to make diagnostic mistakes that fall into two categories: false positives and false negatives.

Artificial intelligence “AI”, whose first concepts for many years, many scientists have tried to help radiologists detect and diagnose these anomalies in the medical field. The use of natural language data entry. This enables the development of algorithms for the automatic processing of medical pictures from enormous volumes of radiological data, Computer-aided diagnosis, (CAD), systems and mammography image processing techniques are created for this purpose.

According to the American Cancer Society, radiologists miss about 20% of breast cancers during mammograms, and half of women who undergo screening tests over a ten-year period receive a false-positive result. The integration of medical imaging diagnostics with clinical, biological, and genetic data creates a true “integrated medicine,” and artificial intelligence, machine learning algorithms in the areas of breast cancer diagnosis of the breast opening the way to personalized, predictive medicine. The great difficulty of medical data is the multiplicity of their sources and their mode of expression (coding).

The future of AI in breast cancer screening is bright, as it is increasingly being used to analyze mammograms and tomosyntheses. Research is being conducted to expand its use to MRI, particularly for high-risk women. In the long run, AI could also help determine which patients need chemotherapy; the rapid development and implementation of AI has the potential to revolutionize healthcare.

Several studies are conducted in the medical field:

Co-learning and the integration of features from clinical texts and medical imaging are made possible by medical vision language models. Nevertheless, this technique, called Medical Vision Language Pretraining with Frozen Language Models and Latent Space Geometry Optimization (M-FLAG), enhances the pre-trained model’s potential for medical image classification, segmentation, and object detection by utilizing a frozen language model for stability and training efficiency in addition to a new orthogonality loss to harmonize the latent space geometry [3]. English and Spanish are the two most often used languages. Medical Vision-Language Pre-Training (Med-UniC) aims to combine multimodal medical data from both of these languages. Specifically, multilingual semantic representations of medical reports from different language communities are explicitly unified using multilingual text alignment regularization (CTR) [4]

Textural analysis is part of these computer operations, such as: regularity, homogeneity, contrast, etc. In the context of medical images, this type of analysis can be used for various reasons for distinguishing between healthy and pathological areas [5,6]. Many feature extraction techniques based on texture, shape, and gray level features have been presented recently for mass classification of mammography images. In order to differentiate between benign and malignant tumors in mammography images, grayscale features—first-order statistics including mean, standard deviation, and variance—are used to evaluate intensity variation [7–9]. Using pixel layouts in space, shape parameters including circumference, area, compactness, and circularity are retrieved to differentiate between breast tissue that is mass and that which is normal [10–13]. Texture is a key feature for mass classification in mammography images since it explains how pixel brightness vary spatially [13]. This determines whether the image’s features are soft or coarse. Well-defined primitives are used in structural approaches to represent the texture and yield an accurate symbolic representation of the image [14,15]. The distribution and connections of gray levels in

an image are determined by features that are the basis of statistical techniques. GLCM and GLRLM calculations are used for the most popular statistical techniques [16–18]. Transformation methods rely on image processing in the transformation domain. The most common methods for extracting texture features in various orientations are Gabor wavelets and the Contourlet transform [15,19,20]. Once the features from the mammography mass ROI have been extracted, To categorize the mass as benign or malignant, classifiers are employed. Researchers most frequently utilize LDA, K-closer neighbors, decision trees, and SVM as classifiers for mass mammography categorization. Using Gabor filters, S. Khan et al. suggested texture features [15]. SVM was utilized for classification, and the accuracy on average was 93.95%. Based on neighborhood structural similarity, R. Rabidas et al. [14] proposed a technique for extracting texture information. Using LDA, 94.57% accuracy in classification was attained.

X. Liu et al. [16] present a method that incorporates geometric and textural features into the SVM classifier for mammography mass classification. This method yielded an accuracy of 94%. Ioan B et al. [19] characteristics that were extracted with direction from Gabor wavelets and classified the data using proximal SVM. The classification of masses into benign or malignant mammograms BEMD and MBEMD, two feature extraction techniques, are used [21], with accuracy respectively 90%, 92.59%.

We present a novel approach in this work to automate the assessment of breast cancer using mammographic medical images called the selection of the most discriminative attributes (SMDA-BEMD). Existing feature extraction methods have resulted in a large set of features and hence dimensionality reduction of the feature set or feature selection needs to be done.

We present a novel approach in this work to automate breast cancer assessment using mammographic medical images. The configuration of the proposed method is as follows: The image is decomposed into its constituent elements using the BEMD algorithm “BIMF + residual levels” [22,23], Then, the extraction of descriptors using the fourteen Haralick attributes of the co-occurrence matrix [24] of the BIMF, Residue and image components constructed after decomposition and also for the original images. the total number of features is very high, the determination of a subspace of the most discriminating features of these descriptors, where the one-dimensional observations associated with each medical mammography image—both abnormal and healthy—are decomposed into discrete components. The subspace is selected by an iterative procedure based on a supervised learning program [25,26]. Our project will be validated by a review of the healthy and pathological medical photos that were chosen from the Kenitra Morocco Reference (CRSRKM) database and the Reproductive Health Center.

This article is structured as follows: The theoretical framework is presented in the Section 2 where we present the two-dimensional empirical modal decomposition (BEMD), Firstly, and an overview of texture analysis methods, particularly analysis methods based on co-occurrence matrices is presented in a second step. Section 3 presents the overall functional architecture on the Selection of the most discriminating attributes. Section 4 presents the general architecture of the model. The results, along with their explanations and a discussion, are shown in Section 5. Concluded in the final Section 6.

2. The Theoretical Farmwork

2.1. Bidimensional Empirique Multimodal Decomposition Algorithm

Empirical Mode Decomposition or EMD is a flexible signal decomposition technique. Its principle is based on an adapted decomposition by locally describing the signal as a succession of contributions from fast oscillations to slower oscillations. Because of its self-adaptive nature, it can recognize variations in the signal’s amplitude or frequency. It has become one of the most well-known decomposition filters due to its ability to analyze non-stationary and non-linear signals on multiple scales with great power, EMD was introduced in image analysis and processing (BEMD). It consists of decomposing the image into different modes called BIMFs (Bidimensional Intrinsic Multimodal Functions), which

have a good physical meaning. Knowledge of these modes makes it possible to intuitively understand the frequency content of the image and therefore a better description of the image. This technique has been successfully applied to real data such as oceanography and the study of climatic phenomena [27], seismology [28], non-destructive testing [29], underwater acoustics [30], biology [31], image denoising [32], image compression [33], image feature extraction [34,35], texture synthesis [36], to image texture classification [37] and watermarked images [38]. The BEMD is defined by a process called “sifting” to decompose the image into basic contributions (BIMF) which are of the AM-FM type, each with zero a mean [39].

2.1.1. BEMD Algorithm

The EBMD algorithm is described as follows:

- Step 1.** Initialization: $r_0(t) = x(t), k = 1$ ($x(t)$ is the original signal and $r_0(t)$ is the residual).
- Step 2.** Extraction of the k th BIMF noted $d_k(t)$ (the sifting process)
 - (a) Initialization: $h_0(t) = r_{k-1}(t), j = 1$
 - (b) Extract the local maxima and minima of $h_{j-1}(t)$
 - (c) Let us calculate the upper and lower envelope by interpolating the maxima and minima of $h_{j-1}(t)$ ($EnvMax_{j-1}(t), EnvMin_{j-1}(t)$)
 - (d) Determine the average envelope of the two envelopes. $m_j(t) = \frac{1}{2}(EnvMax_{j-1}(t) + EnvMin_{j-1}(t)) = (EnvMoy_j(t))$
 - (e) $h_j(t) = h_{j-1}(t) - m_j(t)$
 - (f) If the stopping criterion is satisfied then $d_k(t) = h_j(t)$ otherwise return to (b) with $j = j + 1$.
- Step 3.** $r_k(t) = r_{k-1}(t) - d_k(t)$
- Step 4.** If $r_k(t)$ has at least 2 extrema, return to (2) with $k = k + 1$ otherwise the decomposition is finished; $r_k(t)$ consists of the residue $r(t)$ of this decomposition.

The flowchart below illustrates the BEMD decomposition algorithm (Figure 1):

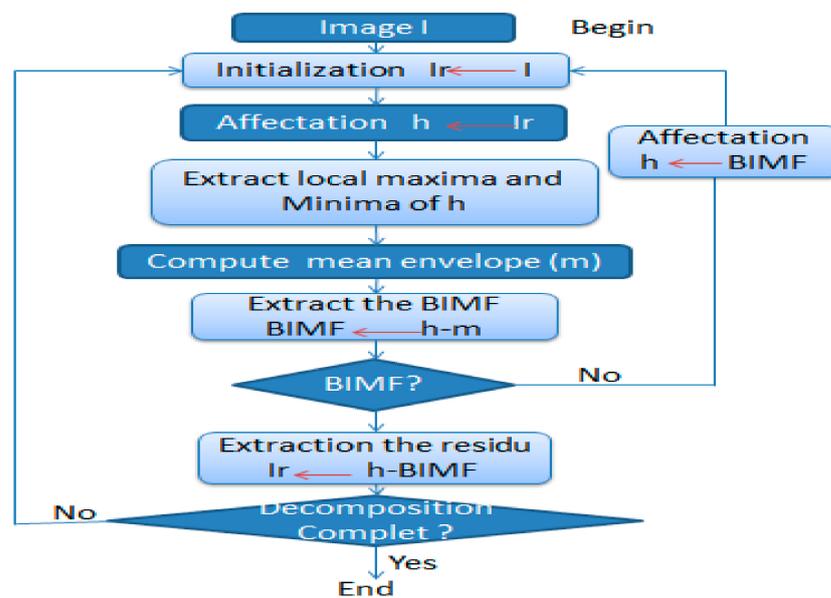


Figure 1. The flowchart of the Bidimensional Empirique Multimodal Decomposition BEMD algorithm.

The flowchart described below has two nested loops. The main loop corresponds to the same concept as the sifting loop which It stops when it is no longer possible to extract an oscillating component. This makes it possible to define the level of depth of the decomposition The sifting loop which corresponds to the sifting process described below to obtain a BIMF.

Thus, the reconstruction of image I is carried out by summing all the components $(BIMF_k)_{k=1\dots n-1}$ and the residual I_r (1)

$$I = \sum_{k=1}^{n-1} BIMF_k + I_r, n \in \mathbb{N} \tag{1}$$

where $BIMF_k$ is the k th oscillation, I_r is the residue of the decomposition and $n - 1$ is the number of BIMFs (Bidimensional Intrinsic Multimodal Functions).

2.1.2. Sifting (SD)

This is the most important step in the algorithm. BIMF is extracted at this stage. This process, known in the literature as “sifting” or “sifting process”, is iterative. Formally, we define this process as an iteration of the elementary sieve operator S . The oscillatory components are then defined by iterations of this operator (2):

$$x_k(t) = d_k(t) = S^j(r_{k-1}(t)) \tag{2}$$

where j is the number of iterations determined according to certain criteria to obtain the k th BIMF. It is worth noting that a stop condition is required to stop the screening process. This condition is associated with the criteria defined in the next paragraph.

2.1.3. Criteria for Stopping The Decomposition

The criteria for stopping the screening process are based on the properties of the BIMF. Furthermore, the quality of BIMF extraction depends on the quality of the previous BIMF. Therefore, the choice of criteria for stopping the screening process is very important. Against this background, several recommendations were made:

- Convergence of the screening process is assumed. Second bullet;
- Perform a certain number of iterations without verification testing of the extracted BIMF (not recommended).
- Define a stopping criterion during sifting.

Cauchy Criterion in L2 Standard

In the work [40], the authors propose a stopping criterion SD (j) based on the standard deviation (Standard Deviation SD) defined by (3)

$$SD(j) = \sum_{t=0}^T \left[\frac{(d_{j-1}(m, n) - d_j(m, n))^2}{d_{j-1}^2(m, n) + \zeta} \right] \tag{3}$$

where j is the index of the j th BIMF, $d_{j-1}(m, n)$ and $d_j(m, n)$ are the results of two consecutive siftings and ζ is a (weak) term intended to eliminate possible divisions by zero.

The Double Stopping Step

As an improvement on the termination criterion proposed in [40], article [41] proposes a more subtle criterion. To do this, two steps must be defined θ_1 and θ_2 : the first to limit small amplitudes and the second to limit large amplitudes. We use the upper and lower envelopes to define the mean and amplitude of a given signal (4), (5):

$$m(t) = \frac{UpperEnv(t) + LowerEnv(t)}{2} \tag{4}$$

$$a(t) = \frac{UpperEnv(t) - LowerEnv(t)}{2} \tag{5}$$

where $UpperEnv$ and $LowerEnv$ are the upper and lower envelopes respectively

The evaluation function is (6):

$$\sigma(t) = \left| \frac{m(t)}{a(t)} \right| \tag{6}$$

We then iterate the sifting process until

$\sigma(t) < \theta_1$ for $(1 - \alpha)$ the duration of the signal.

$\sigma(t) < \theta_2$ for α the duration of the remaining signal.

Usually, we choose $\alpha = 0.05$, $\theta_1 = 0.05$ and $\theta_2 = 10\theta_1$.

When the number of extrema is less than two [23], which is typically when the decomposition ends, there are no more oscillations to extract. The number of BIMFs selected is influenced by the application details. For example, when denoising images, we only need the first BIMF [42].

2.1.4. Determination of Extrema in 2D Signal

Morphological reconstruction is one of the alternative methods used in the search for extrema points. This method is based on geodesic operators, namely erosion and dilation [43,44]. The following definition applies to an image's morphological dilation with the structuring element (7):

$$\delta_B(I) = I \oplus B = \sup\{I(x - q) | x \in I, q \in B\} \tag{7}$$

The same goes for the morphological erosion of the image with the structuring element. It is defined as follows (8):

$$\delta_B(I) = I \ominus B = \inf\{I(x - q) | x \in I, q \in B\} \tag{8}$$

2.1.5. Interpolation

One of the most important processes in the estimation and extraction of BIMFs from extrema is interpolation. The two upper and lower envelopes can only be obtained by interpolating the extrema after they have been extracted. The points that need to be interpolated in the two-dimensional case are scattered, or "scattered data", rather than arranged on a regular grid. As such, a more complex method tailored to this kind of problem is needed. Radial basis functions, also known as RBFs or Radial Basis Functions, are typically used [45]. A radial basis function has the form (9):

$$\forall x \in \mathfrak{R}^d; S(x) = p_m(x) + \sum_{k=1}^n \lambda_k \phi(\|x - x_k\|) \tag{9}$$

with $x_k \in \mathfrak{R}^d$ et $\lambda_k \in \mathfrak{R}$

where p_m is a low degree polynomial, x_k with $0 \leq k \leq n$ are the interpolation centers; $\|\cdot\|$ is the Euclidean norm in \mathfrak{R}^2 , ϕ is the basis function for functions with an assumed fixed radial basis and λ_k are the RBF coefficients to be determined.

The equation $S(x)$ resulting from an RBF interpolation is defined by the coefficients of the polynomial P and the weights λ_k . Given $f = (f_1, f_2, \dots, f_N)$, we look λ_k for the weights for the RBF to check (10):

$$s(X_i) = f_i, i = 1, \dots, N \tag{10}$$

with $X_i = (x_i, y_i)$

By doing this, we are able to reduce our problem from one in an infinite dimensional space to one in a finite dimensional space (the set of coefficients that determine s) and to a linear system that can be solved using standard linear algebra techniques.

2.1.6. Application Example

In this part, we will present an example of the application of BEMD that illustrates the principle of BEMD decomposition, that is, a multi-scale decomposition from high frequencies to low frequencies.

We illustrate in the figures below the result of a BEMD decomposition of a synthetic 2D signal (11).

$$S(x, y) = \exp(- (x^2 + y^2)) + \cos(2y) + \sin(3x) + \text{bruit_blanc}(x, y) \quad (11)$$

The decomposition is carried out without any prior knowledge of the characteristics of the signal S . It makes it possible to extract, initially, the high frequency components like BIMFs 1 and 2 in Figure 2. The low frequency components begin to appear in the 3rd BIMF. We can say that the BEMD behaves like a self-adapting filter bank. The BEMD makes it possible to select, on a local scale the oscillations corresponding to high frequencies, to and gradually tend towards those corresponding to the lowest frequencies Figure 2.

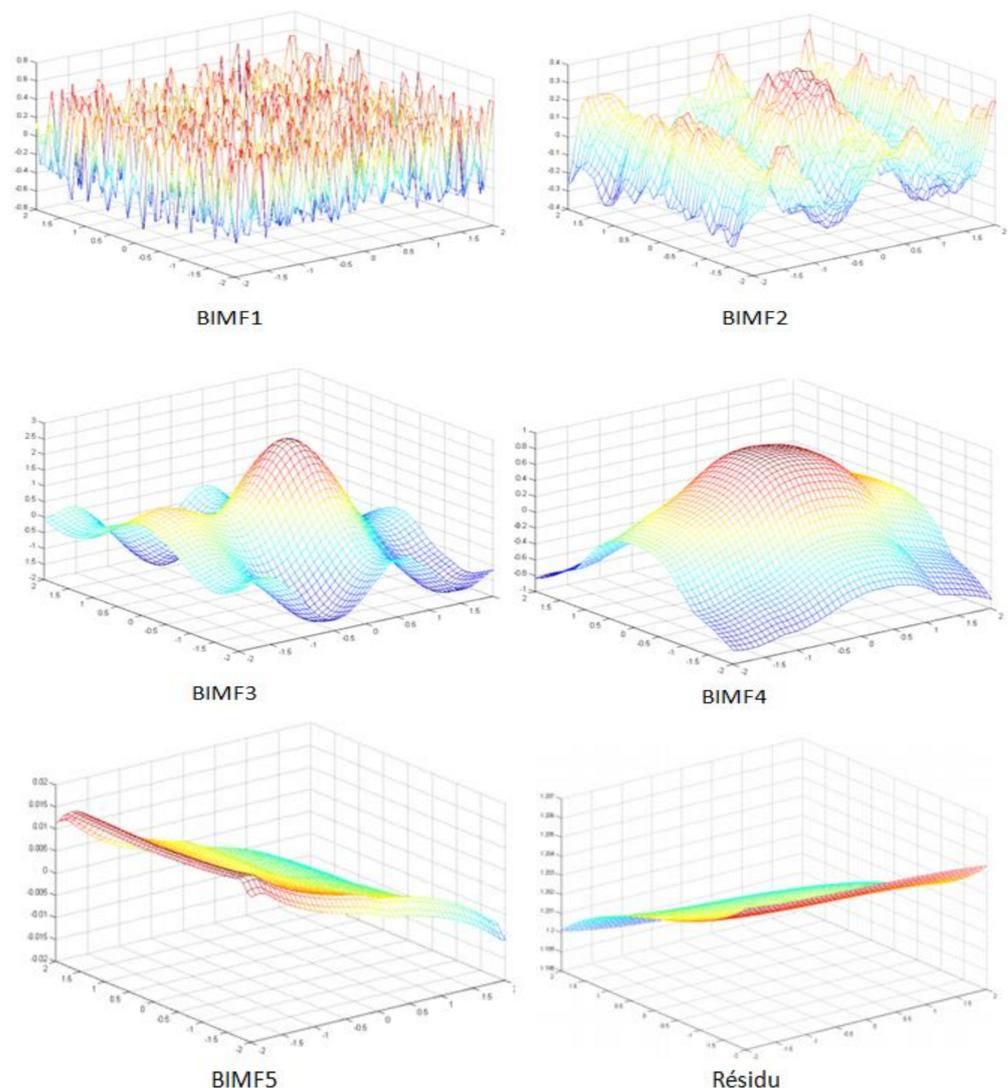


Figure 2. Bidimensional Empirique Multimodal Decomposition BEMD of the signal $S(x,y)$.

2.2. Texture Analyze Methods

2.2.1. GLMC Co-Occurrence Matrices Methods

To the extent that we seek certain forms of regularity within images made up of a very large number of pixels, the statistical analysis seems a priori interesting. Indeed, the use

of gray level co-occurrence matrices of pixels of an image are the most popular statistical technique [46–50] for extracting texture descriptors for different types images, as well as for the analysis, segmentation and classification of images of types varied [51].

A co-occurrence matrix is used to characterize the local distribution of gray levels in a spatial neighborhood of each image pixel [52].

In the classical framework of image processing, an element $M(d, i, j)$ of a co-occurrence matrix is determined by counting the number of times a pixel $P = [x1', x2']^T$ with a gray level i is positioned with respect to a pixel $P = [x1, x2]$ with a level of gray j knowing that

$$P_{r'} = P_r \begin{pmatrix} d \cos \theta \\ d \sin \theta \end{pmatrix} \tag{12}$$

where d is the distance in the θ direction between the two pixels. Consider an image to be analyzed I rectangular. Its size in x is N_x and that in y is N_y . Suppose the gray level is quantized over a set of N_g discrete values. Let $L_x = \{0, 1, 2, \dots, N_x - 1\}$, $L_y = \{0, 1, 2, \dots, N_y - 1\}$ the horizontal and vertical spatial domains, and $G = \{0, 1, 2, \dots, N_g - 1\}$ all the quantized N_g gray levels. The set $L_x \times L_y$ is the set of pixels in the image. The image I can be represented as a function which assigns a value from the set G to each pixel from the set $L_x \times L_y$.

$$I: L_x \times L_y \rightarrow G$$

Therefore, the formula defining the non-normalized co-occurrence matrix element, $M(d, \theta, i, j)$ of I is as follows

$$M(d, \theta, i, j) = \text{card} \left\{ \begin{matrix} (n, m) \in (L_x \times L_y) / \\ I(n, m) = i, I(n \pm d \cos \theta, m \pm d \sin \theta) = j \end{matrix} \right\} \tag{13}$$

This matrix is then normalized:

$$m(d, \theta, i, j) = \frac{M(d, \theta, i, j)}{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} M(d, \theta, i, j)} \tag{14}$$

θ is the particular direction in which the relationships between pixels go to be analyzed ($\theta = 0^\circ, 45^\circ, 90^\circ$ or 135°). Figure 3 illustrates the different directions considered is the distance at which the neighbors of the pixel to be analyzed are located (d) (see Figure 3) or an image having N_g distinct gray levels, the co-occurrence matrix admits $N_g \times N_g$ elements for each direction θ . Even if we see that this matrix is symmetrical (the number of transitions from gray level i to level j is equal to the number transitions from level j to level i), the number of elements is generally far too important for the matrix to be directly used to characterize the texture in question. A first reduction of this number consists, for example, of reducing the number of gray levels of the image by histogram equalization. In practice, this general formulation is particularized for the main directions of θ .

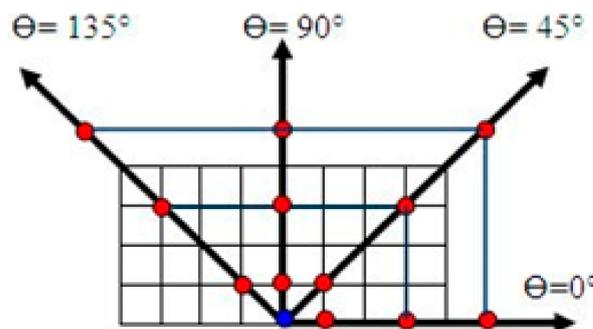


Figure 3. Co-occurrence matrix directions.

2.2.2. Statistical Attributes

The co-occurrence matrices are full of information, but because they are difficult to alter and include a lot of data, they are nevertheless difficult to use in their totality. One method of using the co-occurrence matrix elements as texture properties is to decrease the number of gray levels; this is why there are fourteen in the table. The elements that I have chosen, as defined by Haralick [53,54], can be computed from these matrices in various directions to represent the descriptive characters of the textures and provide an overview of all the information that the co-occurrence matrix offers, such as homogeneity, contrast, energy, entropy, correlation, and so forth Table 1.

Table 1. Texture attributers proposed bay haralick.

Haralick Attributes			
f1	Second moment angular	f8	Entropy of sums
f2	Contrast	f9	Entropy
f3	Correlation	f10	Variance of differences
f4	Variance	f11	Entropy of differences
f5	Moment differential inverse (or homogeneity)	f12	Correlation Information –1
f6	Average sums	f13	Correlation Information –2
f7	Variance of sums	f14	Maximum correlation coefficient

3. Selection of the Most Discrimling Attributes

3.1. Slection of the Most Discrimling Attributes of Decomposed Images

Amount Considering every pixel in the 14-dimensional attribute space displayed in Table 1. The vector I has 14 attributes linked to it. The total number of qualities (Nf = 14) is usually too high in dimension. It is interesting to determine which one-dimensional data analysis techniques are the most discriminant. Based on the calculation of their discriminating ability, we assume that the clusters associated with the classes of mammography images are the best capable of differentiating between the two classes (healthy and pathological). Finding an attribute space where the one-dimensional observations (healthy and diseased) associated with each preprocessed mammography medical image are sufficiently separated to form readily distinguishable categories is the goal. This will assist in locating the area with the greatest discriminating qualities. By considering an increasing number of distinctive qualities, we consider subspaces $R_1, \dots, R_n \dots, R_{14}$ connected with 14 attributes of increasing dimensions. The 14 two-dimensional subspaces R1 are connected to the 14 table. At each stage of this process, an informational criterion J is calculated to assess the discriminative power of each candidate feature space and to gauge the compactness and separability of the groups of observations linked to the various breast mammography images under consideration. The 14 two-dimensional subspaces R1 connected to the 14 table are the attributes where the algorithm begins [55]. The attribute space that maximizes the informational criterion J , like in (15), is the most discriminating:

$$J^f = trace \frac{\sum_B^f}{\sum_W^f + \sum_B^f} \tag{15}$$

With (\sum_w^f) The intra-cluster dispersion value defines how compact the texture classes are (16):

$$\sum_W^f = \frac{1}{n} \sum_{k=1}^{Nc} \sum_{m^2 \in c_k} (X_m - \bar{X}_k)^2 \tag{16}$$

(\sum_B^f) The inter-cluster dispersion value defines the separability of classes (17).

$$\sum_B^f = \frac{1}{n} \sum_{i=1}^{Nc} n_i (X_i - X)^2 \tag{17}$$

With $(\text{Card } Q = f)$ represents the set of characteristics, $(\text{card } C = Nc)$ represents the set of classes, and E represents the set of individuals.

$X_i = (X_i^j)_{1 \leq j \leq q}$, vector of its f attributes. The values of the j th attribute are stored in the vector $X^j = (X_i^j)_{1 \leq i \leq n}$, and each individual belongs to a class $C_k, k = 1, \dots, Nc$.

We notice: $\bar{X}^j = \sum_{i=1}^n X_i^j$ the mean of attribute j , \bar{X} the vector of means of each attribute, \bar{X}_k the vector of means of each attribute of the individuals of C_k , and n_k the cardinality of C_k .

The feature that maximizes J is the best option for distinguishing between texture classes. In the first phase, C is selected, and in the second step ($s = 2$), it is linked to each of the remaining $(Nf - 1)$, After analyzing the trace criterion in the 14 spaces R_2 , the pair of attributes with the highest discriminating power is retained. After that, each of the other attributes in the sub-spaces R_3 is linked to this pair of attributes in order to determine which of the mammography images' qualities is the best. The pair of attributes that is the most discriminating is retained by evaluating the trace criterion in the 14 spaces R_2 , in order to choose the mammography images' best feature. Next, every other attribute in the sub-spaces R_3 is linked to this pair of attributes.

For images decomposed using BEMD decomposition which decompose an image into several BIMFs ($s = 1, \dots, 5$) and residual modes, we have 6 co-occurrence matrices and therefore 6×14 Haralick characteristics aken out of these matrices We examine that $Nf = 6 \times 14 = 84$ texture characteristics of images broken down (see the Figure 4).

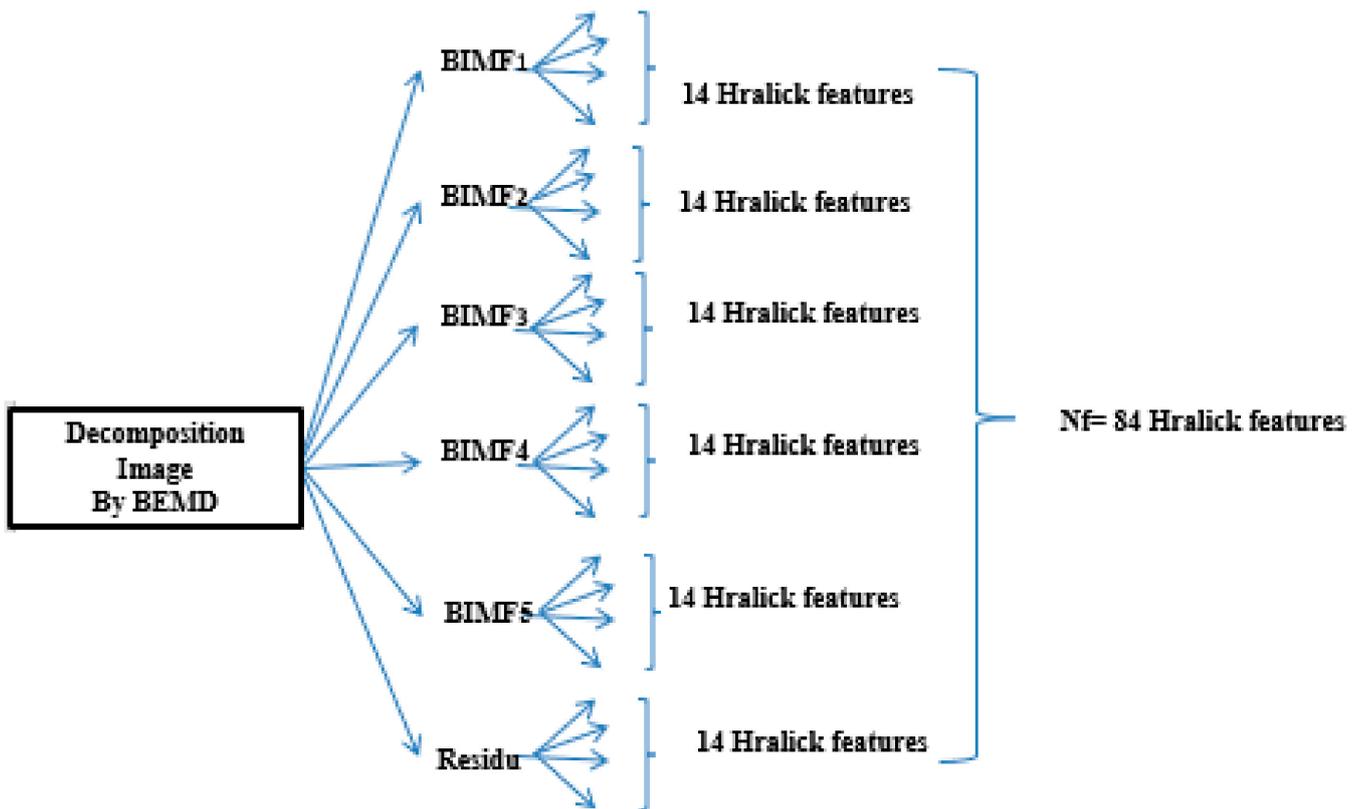


Figure 4. Haralick features for decomposing images by Bidimensional Empirique Multimodal Decomposition BMED.

Given that the total number of candidates, Nf , Nf decomposed texture features is particularly high, it can deteriorate the quality of the discrimination because the risk of considering correlated attributes is then significant. Finally, the time required for classification depends on the number of attributes considered.

Processing each training image's N_f characteristics is what this stage entails. The following iterative selection technique then enables the process by which choose the best features, i.e., those that are the most discriminative for the texture of NC classes: The discriminative strength of every potential feature space is determined at each stage of this process by computing an informative criteria J .

At the start of this technique ($n = 1$), The candidate feature in one-dimension N_f , the spaces formed by each of the N_f features of the available BIMFs, are taken into consideration. The best candidate characteristic for texture class classification is the one that maximizes J . To produce $(N_f - 1)$ candidate feature spaces in two dimensions, it is chosen as the first step and linked to each of the $(N_f - 1)$ remaining mode characteristics in the second stage of the technique ($n = 2$). We believe that the ideal two-dimensional space for discriminating is the one that maximizes J between the classes of healthy and pathological mammographic images.

Certain criteria must be defined to stop the search process on the most discriminating subsets of characteristics. For our method, the commonly used stopping criterion is based on the order of features, ranked according to certain relevance scores (generally statistical measures). Once the features are ordered, those with the highest scores will be chosen and used by a classifier.

The search process can stop when there is no longer any improvement in precision when the selection of strongly discriminating attributes, In other words, when there is no longer the possibility of finding a subset better than the current subset. See (Section 3.2 Stopping Criteria).

3.2. Stopping Criteria

Most writers favor using criteria that are based on assessments [56]:

In our study, the halting criteria determine a threshold by comparing the evaluation value at stage m with the evaluation value at stage $m + 1$, which, when altering an attribute's value doesn't sufficiently boost discrimination, terminates the search. The method looks at the subspaces R_{m+1} of dimension $(m + 1)$ when n attributes are kept out of the 14 available. Next, it chooses the $(m + 1)$ th feature that maximizes the trac' criterion when paired with the n qualities that were previously chosen his procedure is then repeated with the 14 qualities in descending order of relevance. The final features selected correspond to the evaluation rate growth of the criterion's initial stage in the sequence J_m [57], where J_m is the trace J 's criterion value at that particular iteration. This assessment A rate is provided by the following relationship, (18).

$$T_V = \left| \frac{J_{m+1} - J_m}{J_{m+1}} \right| \quad (18)$$

Any qualities in the sequence J_m that are lower than the greatest value of this evaluation rate are not appropriate. This iterative selection procedure can effectively and simply retain subsets of strongly discriminating attributes, but it does not ensure that the optimal subset of attributes is obtained in terms of the separability and compactness of the groupings of observations associated with the different types of mammogram images.

4. General Architecture of the Model (SMDA-BEMD)

This methodology consists of the global analysis of breast cancer mammography images using gray level co-occurrence matrix methods. The proposed methodology consists of the estimation of the properties of images relating to second-order statistics for a more precise analysis in the calculation of the parameters, the statistical analysis gives the relationships between a pixel and its neighbors and defines discriminant parameters of the texture, using co-occurrence matrices which consists of calculating the repetitions of gray levels in the image, it gives the differential characteristics of the local variation of gray levels, to assign to each pixel a set of texture properties known as Haralick attributes (14 attributes), these attributes are called local attributes, making it possible to summarize

the textural information contained in the GLCM, the proposed Methodology is divided into three stages:

The first stage: This methodology consists of calculating its haralick attributes on the original images:

The second stage of this process involves calculating these attributes on images after decomposing the mammographic images using BEMD decomposition which decomposes an image into several BIMF modes “BIMF + residual levels”.

The third stage: This methodology consists of calculating the haralick attributes on the reconstructed images.

The goal is to extract the database containing the attributes that are required to classify images as either healthy or pathological, as well as to identify the range of the most distinctive features for BIMFs and both the original and reconstructed images.

The proposed approach algorithm is described as follows:

Phase 1:

- (i) Determine the 14 haralick attributes of all original Images healthy s in the studied database.
- (ii) Determining the most reduced fat using the iterative selection method
- (iii) Mammography image classification based on the best discriminating space for both healthy and diseased classes.
- (iv) Determination of the classification rate (%)

Phase 2:

- (i) For each $k = 1$ to 5
 - Extract from k th BIMF of all healthy people and pathology people with cancer images in the database studied by the BEMD application.
 - Determine the 14 haralick attributes of all each BIMF and the Residue.
- (ii) Determining the most reduced fat using the iterative selection method with $NT = 84$ attributes.
- (iii) Classification of classes of healthy and pathological mammographic images according to the most discriminating space.
- (iv) Determination of the classification rate (%).

Phase 3:

- (i) Determine the 14 haralick attributes of all healthy and cancerous Reconstructed Images in the database studied.

$$\text{Images Reconstructed} = \sum_{k=1}^5 \text{BIMF}_k + \text{The RESIDU} \quad (19)$$

- (ii) Determining the most reduced fat using the iterative selection method
- (iii) Using the most discriminating space, classes of normal and abnormal mammography images are classified.
- (iv) Determination of the classification rate (%)

Phase 4:

Comparison of the classification rates of the proposed methods and other methods.

The choice of these attributes is delicate and achieving efficient discrimination of the classes involved leads us to consider an important set, an iterative selection procedure based on a supervised learning program is used to determine this space [58] to form easily identifiable groupings that are well separated based on which level of BIMFs. The method of iterative selection is described in Section 3. Figure 5 presents the diagram of the Selection the Most Discriminant Attributes of the images preprocessed by BEMD “SMDA-BEMD”.

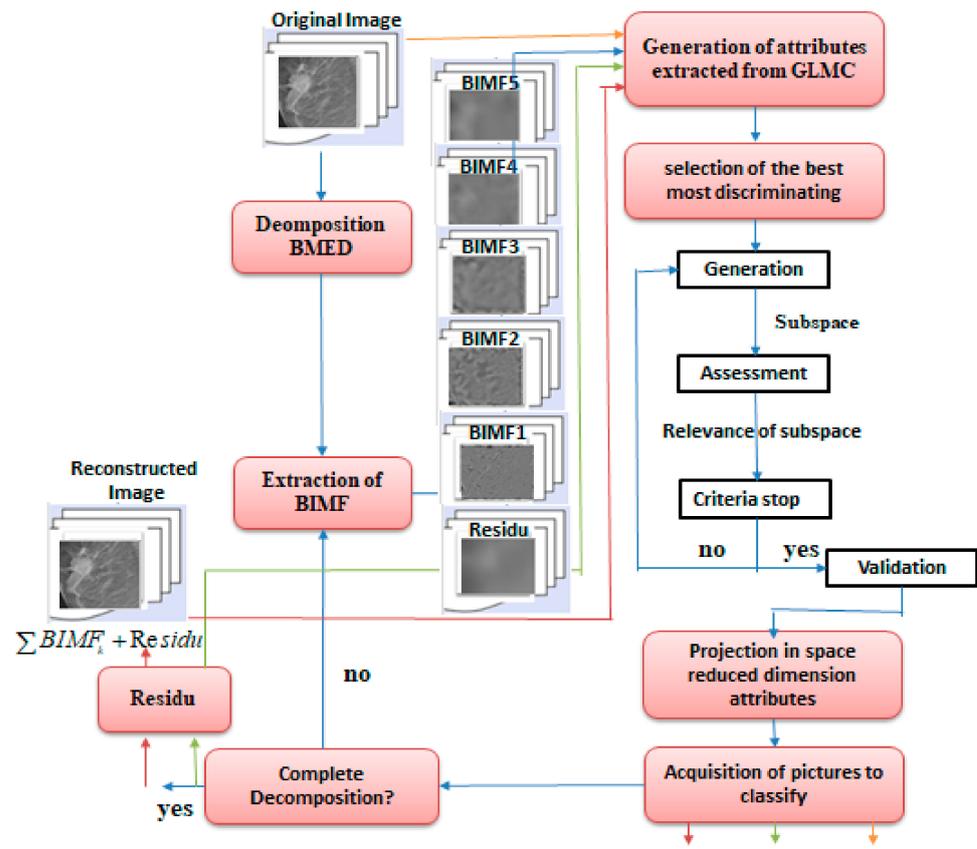


Figure 5. The architecture of the methods suggested for the diagnosis of breast cancer through the identification of the most distinctive features of the images broken down using Bidimensional Empirical Multimodal Decomposition (BMED).

5. Experimental Results and Discussion

5.1. Data Set

Mammography is the main diagnostic and screening method for breast cancer. It is a radiographic (X-ray) examination of both breasts, and studies have demonstrated that early detection of breast abnormalities considerably increases the chance of survival. For the application of our method, we have a data set of medical images of the that was collected from the reference center for reproductive health in Kenitra-Morocco “CRSRKM” This data set contains 150 mammography images in their normal and cancerous cases, Figure 6 shows some examples.

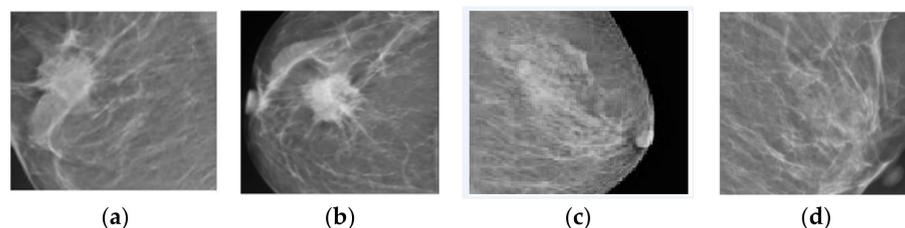


Figure 6. Examples breast mammogram images: healthy (c,d) pathological (a,b). from the reference center for reproductive health in Kenitra-Morocco (CRSRKM).

5.2. Experimental Results and Discussion

Figure 7 illustrates pathological and healthy images decomposed by the bidimensional empirical multimodal decomposition “BEMD” method, they also to present the BIMFs level, residue, and reconstructed images.

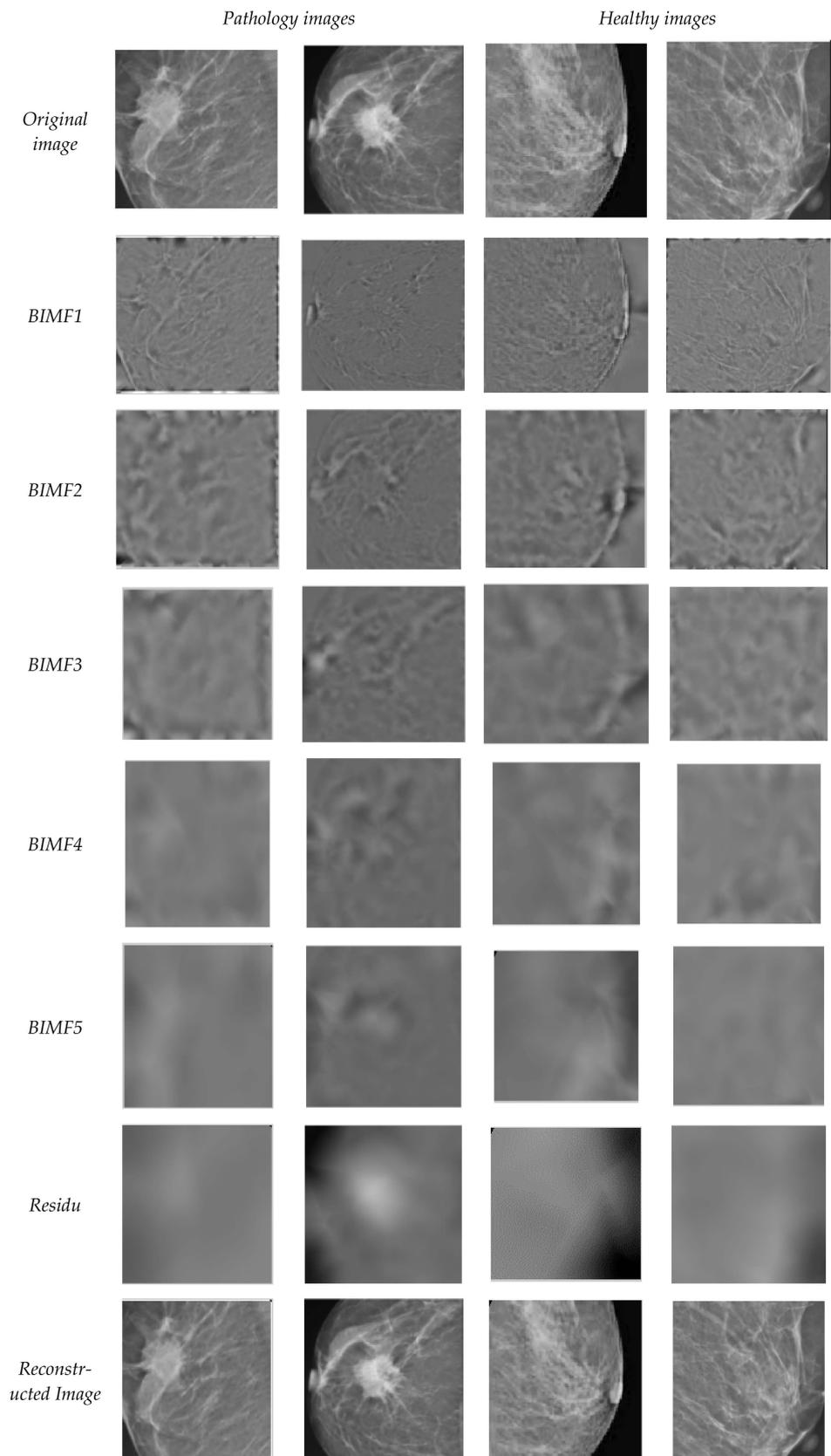


Figure 7. The decomposition of pathology and healthy images decomposed by the bidimensional empirical multimodal decomposition “BEMD” method.

We carried out the 14 characteristics’ computation of Haralick from the table taken from GLCM on mammography images of the original images (Figure 6) and of each BIMFs and also of the reconstructed images of his (healthy and cancerous) Figure 7. It is therefore necessary to choose a subset of these 14 texture attributes at each pixel of these images in the attribute space of dimension 14, as this dimension is too high. This will enable the two healthy and pathological observation classes to be easily separated by the reduction of the representation space’s dimension, Table 2 compiles the most effective and discriminating selections for traits of both healthy and pathological breast mammography images of each BIMFs, residue and the reconstructed and original image, we employed the Section 3 iterative selection process.

Table 2. The Most Discrimling Attributes the for each BIMFs level, residue, reconstructed images and originals images.

Type of Image	Haralick Attributes	The Most Discriminating Power Jf for Healthy Images	The Most Discriminating Power Jf for Pathological Images
BIMF1	f5	0.668	0.672
BIMF2	f4	0.777	0.403
BIMF3	f8	0.982	0.645
BIMF4	f10	0.582	0.423
BIMF5	f1	0.442	0.552
Residue	f3	0.222	0.422
Reconstructed Image	f11	0.868	0.724
Original image	f6	0.862	0.771

The most discriminating features listed in the Table 2 are represented in the related space by the observations taken from the two series of healthy images and pathologies from our data set that were taken into consideration, are shown as point clouds, with one representing a disease and the others representing a healthy one. Figure 7 illustrates the results obtained from the most discriminating BIMFs level, the reconstructed images and the original images.

Table 2 illustrates the most discriminative feature characteristics which are maximized the discriminative power of the decomposed healthy and pathological breast mammogram images for each BIMFs, the residuals and the reconstructed and original images. The most discriminative The following is the Haralick feature that was taken out of the co-occurrence matrix:

For the original images: the most discriminating attribute selected for the normal and pathological images is f11, their maximum discriminating power Jf is and 0.868 and 0.771 respectively.

For the reconstructive images: the most discriminating attribute selected is for the normal and pathological images is, f6, their maximum discriminating power Jf respectively is 0.862 and 0.724.

For the decomposed images: the most discriminating attribute selected for the healthy and pathological decomposed images at the BIMF3 level is f8, their most discriminating power Jf is respectively 0.982 and 0.645.

The results in Table 2 show that f8, extracted from the BIMF3 image co-occurrence matrix, is the most discriminating in the cases mentioned above. In addition, we observe that there is a difference in its ability to distinguish healthy populations from pathological populations. The results obtained from the duration of the interval [Jfp max, Jfsmin] are given in Table 3 and calculated using the following equation:

$$L = |Jfs \text{ min} - Jfp \text{ max}| \text{ such as:}$$

Jfs min is the minimum value of Jf healthy images

Jfp max is the maximum value of Jf des pathology images.

Table 3. Values obtained from the interval between the Jfs min and the Jfp max.

Type of Images	The Interval $L = Jfs\ min - Jfp\ max $
BIMF1	$L = 0.752 - 0.672 = 0.08$
BIMF2	$L = 0.318 - 0.403 = 0.085$
BIMF3	$L = 0.875 - 0.645 = 0.23$
BIMF4	$L = 0.467 - 0.423 = 0.044$
BIMF5	$L = 0.591 - 0.552 = 0.039$
RSIDUE	$L = 0.441 - 0.422 = 0.019$
Reconstructed Image	$L = 0.612 - 0.724 = 0.112$
Original Image	$L = 0.602 - 0.771 = 0.169$

The observations taken from the two series of healthy and pathological mammograms that were taken into consideration are represented in space of reduced attributes associated with the most discriminating, as shown by an analysis of the results shown in Figures 8 and 9 and Table 3, and to provide a visual overview of the discriminatory power of this subset of attributes, we note that:

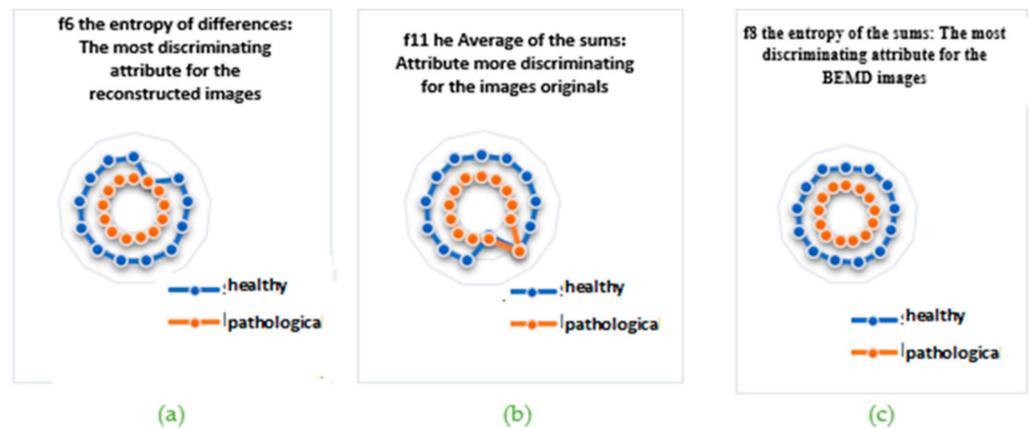


Figure 8. Projection of observations extracted from healthy and pathological mammography images obtained from the most discriminating BIMFs level and reconstructed and original images: (a) the categorization results obtained for healthy and cancerous images for the images reconstructed after decomposition. (b) the categorization results obtained for healthy and cancerous images for the originals images; (c) the categorization results obtained for healthy and cancerous images the images decomposed by the BEMD.

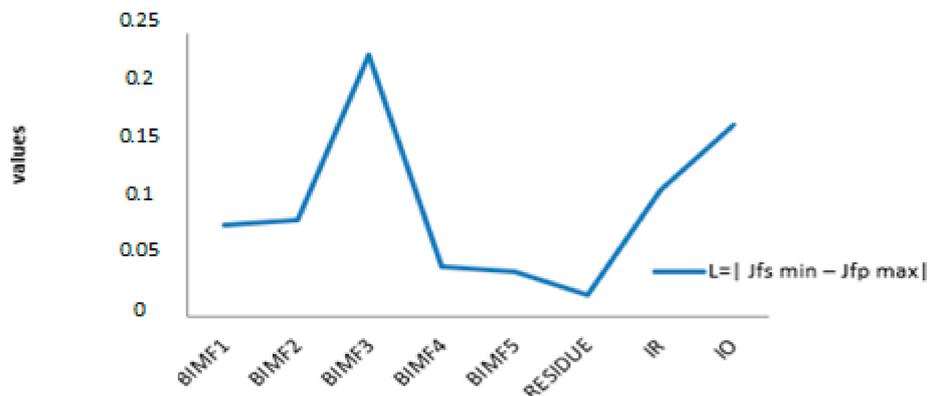


Figure 9. The interval between the minimum value of Jf healthy images (Jfs min) and the maximum value of Jf des pathology images (Jfp max).

For images decomposed by BEMD, the interval $L = |J_{fs\ min} - J_{fp\ max}|$ is larger than that of the original images and the reconstructed images, according to Figure 8c. The point clouds associated with each class of images, both pathological and healthy, are projected into the reduced attribute space and are thought of as form sets that are well separated from one another. The inter-class inertia maximum, and the trace J criterion is to be maximized with respect to the other levels. These point clouds are then divided into two subgroups: one for the pathological images and the other for the healthy images. Since this instance allows for the differentiation of these two groups, a discriminating threshold for mammography pictures may be established. In the reconstructed images and the original images Figure 6, The images that are classified as pathological and healthy overlap, the measure of intra-class inertia is minimal, the point clouds corresponding to each class are close together, and the variation intervals of this parameter for all the images (pathological and healthy) overlap. These two examples demonstrate that it is not possible to distinguish between the populations of cancer patients and healthy.

We use our database as our learning base, then, the non-linear SVM classifier is trained on our learning base. The evaluation is done through a ROC analysis. Another parameter from this curve provides a quantifiable value, this is the area under the ROC curve (AUROC). It is between 0 and 1 and a value close to 1 implies better precision. The classification's performance is assessed on the basis of prediction in terms of sensitivity and specificity defined by:

$$\text{sensitivity} = \frac{VP}{VP + FN} \quad \text{specificity} = \frac{VN}{VN + FP} \tag{20}$$

Alongside: FP stands for "false positives", VP represents "true negatives", FN stands "false negatives", and VN represents "true negatives".

By observing Figure 10 representing the ROC curves and the corresponding Table 4 of the AUROC values, we find that the most discriminating attribute that we defined by our method (SMDA-BEMD), is the most discriminating one (AUROC = 0.991) is higher compared to other attribute extraction based on SMDA- Original Image, (SMDA-Reconstructed Image after decomposition by BEMD, their AUROC = 0.83, AUROC = 0.942 respectively.

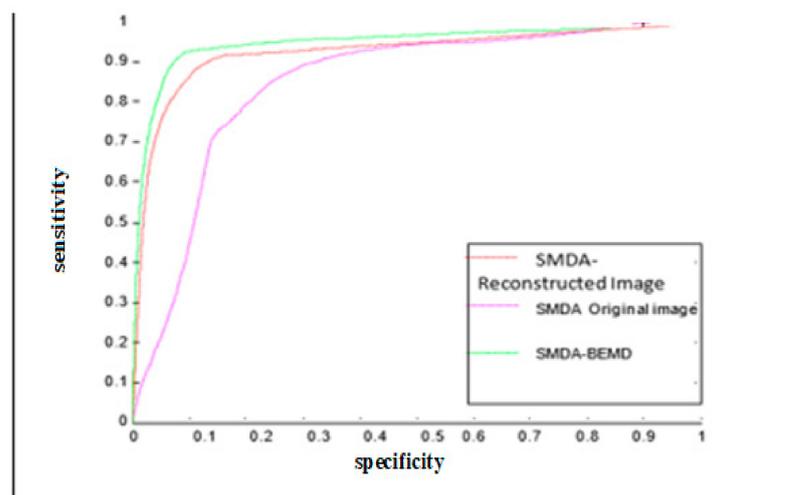


Figure 10. ROC curve comparison using SVM between the SMDA-BEMD, SMDA-Reconstructed Image, and the SMDA-Original.

Table 4. AUROC values for SMDA-BEMD, SMDA-Reconstructed Image and SMDA- Original Image.

Method	SMDA-BEMD	SMDA-Reconstructed Image	SMDA-Original Image
AUROC	0.991	0.942	0.83

According to the Figure 11 by observing that the classification results obtained are given in the form of point clouds one for the healthy and the other for the pathologies. For Figure 11c we observe that there is a difference in its ability to distinguish healthy populations from cancerous populations, these point clouds are then divided into two subgroups.: one for pathological images and the other for healthy images, In the reconstructed images and the original images Figure 11a,b, the images classified as pathological and healthy overlap, The point clouds that belong to every class are near each other, and the parameter's variation intervals overlap for all photos, whether diseased and healthy. These two methods demonstrate that the distinction between cancer and healthy patient populations is poor compared to the SMDA-BEMD method.

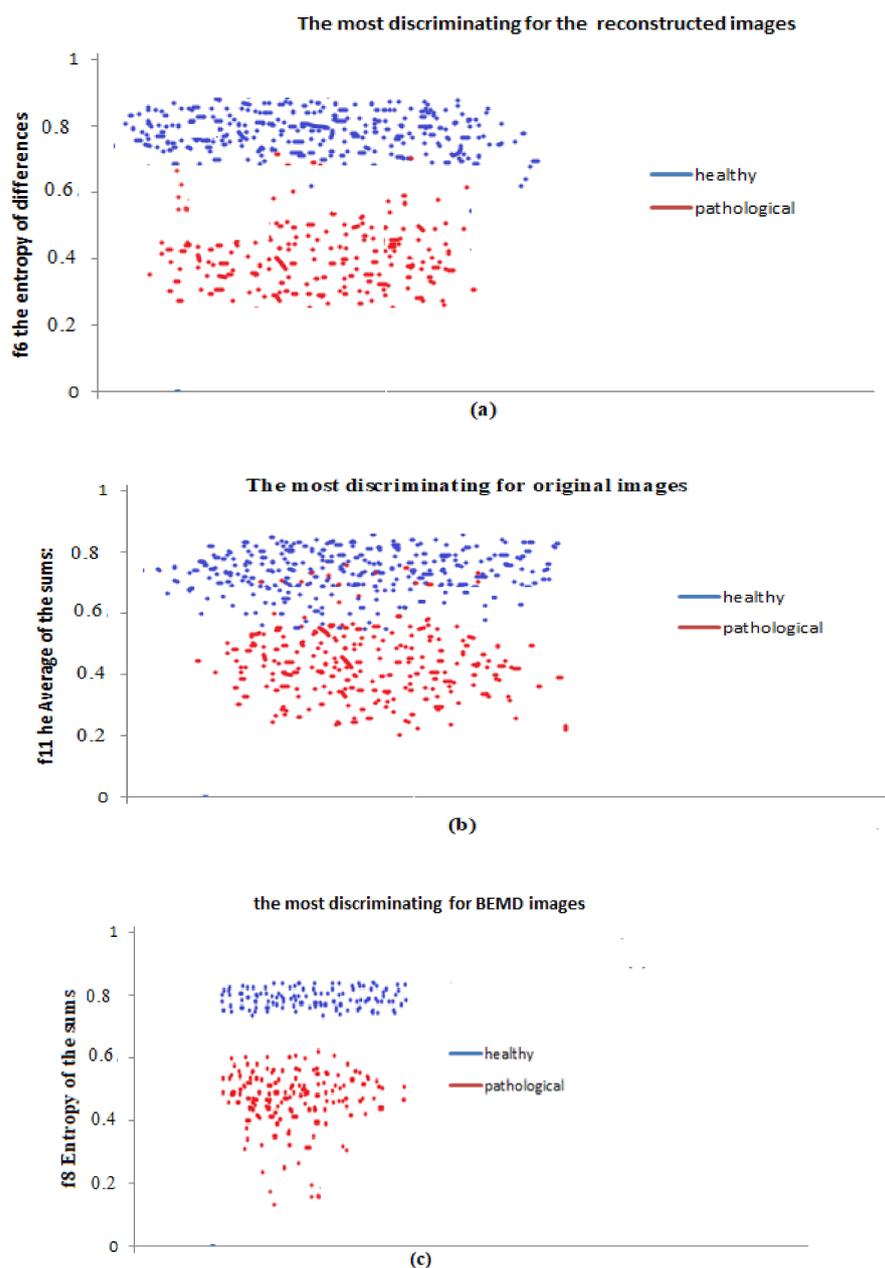


Figure 11. The results obtained from the classifications of mammographic images of the breast in the form of point clouds: one for healthy (blue) and the other for cancerous (red): (a) SMDA-Reconstructed Image, (b) SMDA- Original Images, (c) proposed methodology (SMDA-BEMD).

From Figure 11c, we notice that for the SMDA-BEMD method, we can define a discrimination threshold which can help detect pathological images, it could be integrated into diagnostic workflows, and it is defined by:

$$\text{threshold} = \frac{(\text{MAXJfp} + \text{MINJfs})}{2} \tag{21}$$

with:

- MINJfs is the minimum value of Jfs min healthy BIMFs and residue
- MAXJfp is the maximum value of Jfp max des pathology BIMFs and residue.

If the discriminating power Jf of one of the BIMFS of the processed mammography image is lower than the threshold then we have a malignant mammogram.

This threshold can help radiologists evaluate breast images, by providing a percentage risk of malignancy of a lesion. It can also reduce reading time and improve detection.

Following these comparisons, we will now see the advantage of the proposed methodology compared to other methods defined in existence, by observing the Figure 12 representing the ROC curves and Table 5 corresponding to the AUROC values for the proposed methodologies is also for the other existence methods in the literature.

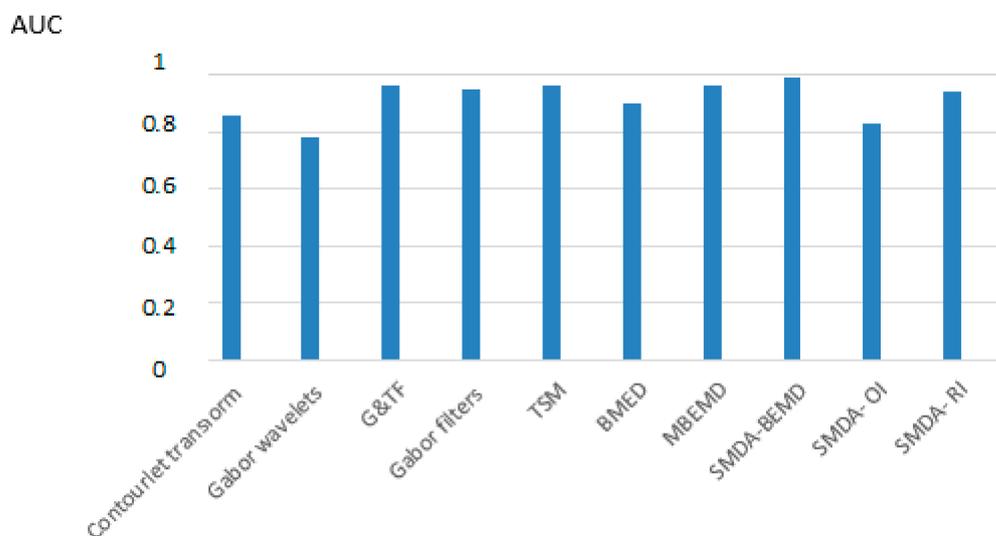


Figure 12. Quantifying the effectiveness of classification by measuring the value of the area under the ROC curve (AUC) of the existing method with the proposed methodology.

Table 5. Comparison between the different existing methods and the most discriminative feature extraction provided.

Feature Extraction Methods	The Classification Rate (%)	AUC
Contourlet transform [20]	87	-
Gabor wavelets [19]	78	0.78
Geometry and texture features (G&TF) [16]	94	0.9615
Gabor filters [15]	93.95	0.948
tructural similarity mapping (TSM) [14]	94.57	0.98
BMED [21]	90	0.90
MBEMD [21]	96.2	0.966
Proposed (SMDA-BEMD)	98.6	0.991
SMDA- Original Image (SMDA-OI)	82.3	0.83
SMDA-Reconstructed Image after decomposition by BEMD(SMDA-RI)	92.8	0.942

Table 5 and Figure 12 comparative results show that the following are the primary issues with the feature extraction techniques that are available in the literature:

- In terms of mass classification accuracy, methods utilizing gray level characteristics performed poorly.
- The mass cannot be categorized as cancerous or benign based just on its morphology. Its sole purpose is to differentiate the bulk area from healthy breast tissue.
- Because they rely on filtering schemes or basis functions, the majority of texture feature extraction approaches are non-adaptive and based on transformation methods.
- The vast set of features produced by previous feature extraction methods necessitates feature selection or dimensionality reduction in our approach.

In our methodology the Selection of the Most Discriminating Attributes the images pre-processed by BEMD (SMDA-BEMD).

We are based on the iterative selection process to determine the most discriminating characteristics is a strong point of study, because it allows to determine the essential characteristics that contribute significantly to the classification of mammography images by dimensionality reduction of the set of features and the selection of the most discriminating characteristics of these descriptors which must be carried out on multiresolution decomposition techniques called BEMD for texture extraction of these characteristics, Unlike the Fourier and wavelet transforms, which employ fundamental functions, BEMD is a decomposition technique that breaks down an image into a collection of two-dimensional intrinsic mode functions, or modes (BIMF). these extracted BIMF are the strongest Characteristics of mammography images elements.

The iterative selection method to determine the most discriminating features is a strong point of the study, as it allows the identification of key attributes that contribute significantly to the classification of mammography images and adds a valuable dimension. This involves choosing the most pertinent characteristics from the original dataset using a new Reduced set.

The results obtained are robust and very encouraging, this indicates that the system operates generally and consistently when the use of mammographic images is broken down and also the use of a set of data from the Reference Center for Reproductive Health of Kenitra -Morocco (CRSRKM), which includes normal and abnormal mammo-grams, adds credibility to the study results and ensures their relevance to real-world clinical applications, The experiments of our method are rapid.

6. Conclusions and Outlook

Breast cancer is becoming the most deadly illness for women, it is well-known. This has led researchers to develop various computer-aided diagnosis techniques to identify breast lumps before they develop into life-threatening malignancies and to identify breast tumors before they develop into lethal cancer since breast cancer is becoming the most serious cancer in women. In this paper, we suggest a novel strategy. for characterizing and differentiating medical breast mammogram images decomposed by bidimensional empirical multimodal decomposition "BEMD" according to the use of co-occurring GLCM matrices from which sets of attributes are extracted, especially assessment because of a process of iterative selection. These characteristics enable the distinction between pathological breast cancer and healthy breast mammography images. The advantage of BEMD decomposition is its intrinsic adaptability to the image. By iterating the process on BIMFs at each level of decomposition, we obtain multi-scale information, and hidden textures are better identified. To show the performance of our method, we made use of a database of medical mammograms from the Kenitra-Morocco Reproductive Health Reference Center (CRSRKM), which includes both abnormal and normal breast mammograms cases. With the perspective of establishing a certain number of studies to further refine this methodology in the future, this proposal calls for the utilization of alternative, more effective analytic methods. Furthermore, we suggest enhancing the characterization stage by incorporating other descriptors that can enhance the accuracy of the breast mass classification.

Author Contributions: Methodology, F.G.; Formal analysis, A.B.; Supervision, F.A. and K.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study did not require ethical approval.

Informed Consent Statement: We collected data on the disease through a research request authorized by the competent administration. Additionally, respecting patient privacy is important, which is why we used the information under an anonymous name.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. De Oliveira, J.E.; de Albuquerque Araújo, A.; Deserno, T.M. Content-based image retrieval applied to BI-RADS tissue classification inscreening mammography. *World J. Radial.* **2011**, *3*, 24–31. [[CrossRef](#)] [[PubMed](#)]
2. Trayes, K.P.; Cokenakes, S.E.H. Breast Cancer Treatment. *Am. Fam. Physician* **2021**, *104*, 171–178. [[PubMed](#)]
3. Liu, C.; Cheng, S.; Chen, C.; Qiao, M.; Zhang, W.; Shah, A.; Bai, W.; Arcucci, R. M-FLAG: Medical Vision-Language Pre-training with Frozen Language Models and Latent Space Geometry Optimization. *Comput. Vis. Pattern Recognit.* **2023**, *14220*, 637–647. [[CrossRef](#)]
4. Wan, Z.; Liu, C.; Zhang, M.; Fu, J.; Wang, B.; Cheng, S.; Ma, L.; Quilodrán-Casas, C.; Arcucci, R. Med-UniC: Unifying Cross-Lingual Medical Vision-Language Pre-Training by Diminishing Bias. *Comput. Sci.* **2024**, *36*. [[CrossRef](#)]
5. Brunelle, F.; Brunelle, P. Intelligence artificielle et imagerie médicale: Définition, état des lieux et perspectives Artificial intelligence and medical imaging: Definition, state of the art and perspectives. *Bull. L'Académie Natl. Médecine* **2019**, *203*, 683–687. [[CrossRef](#)]
6. Gautherot, M.; Yepremian, S.; Bretzner, M.; Jacques, T.; Hutt, A.; Pruvo, J.P.; Kuchcinski, G.; Lopes, R. 15 minutes pour comprendre et évaluer un logiciel d'intelligence artificielle appliquée à l'imagerie médicale (15 minutes to understand and assess an artificial intelligence-based software in medical imaging). *J. D'Imag. Diagn. Interv.* **2021**, *4*, 167–171.
7. Nithya, R.; Shanthi, B. Comparative study on feature extraction method for breast cancer classification. *J. Theor. Appl. Inf. Technol.* **2011**, *12*, 220–226.
8. Eltonsy, N.; Tourassi, G.; Elmaghraby, A. A concentric morphology model for the detection of masses in mammography. *IEEE Trans. Med. Imaging* **2007**, *26*, 880–889. [[CrossRef](#)] [[PubMed](#)]
9. Delogu, P.; Fantacci, M.E.; Kasae, P.; Retico, A. Characterization of mammographic masses using a gradient-based segmentation algorithm and a neural classifier. *Comput. Biol. Med.* **2007**, *37*, 1479–1491. [[CrossRef](#)]
10. Kim, D.H.; Lee, S.H.; Ro, Y.M. Mass type specific sparse representation for mass classification in computer-aided detection on mammograms. *Biomed. Eng. Online* **2012**, *12*, S3. [[CrossRef](#)]
11. Tralic, D.; Bozek, J.; Grgic, S. Shape analysis and classification of masses in mammographic images using neural networks. In Proceedings of the 18th International Conference on Systems, Signals and Image Processing, Sarajevo, Bosnia & Herzegovina, 16–18 June 2011.
12. Bouguila, T.; Elguebaly, N. Bayesian approach for the classification of mammographic images using neural networks. In Proceedings of the 2013 Inter-national Conference on Developments in eSystems Engineering (DeSE), Abu Dhabi, United Arab Emirates, 16–18 December 2013. [[CrossRef](#)]
13. Cascio, D.; Fauci, F.; Magro, R.; Raso, G.; Bellotti, R.; De Carlo, F.; Tangaro, S.; De Nunzio, G.; Quarta, M.; Forni, G.; et al. Mammogram segmentation by contour searching and mass lesions classification with neural network. *IEEE Trans. Nucl. Sci.* **2006**, *53*, 2827–2833. [[CrossRef](#)]
14. Rabidas, R.; Midya, A.; Chakraborty, J. Neighborhood structural Similarity mapping for the classification of masses in mammo-grams. *IEEE J. Biomed. Health Inf.* **2018**, *22*, 826–834. [[CrossRef](#)]
15. Khan, S.; Hussain, M.; Aboalsamh, H.; Mathkour, H.; Bebis, G.; Zakariah, M. Optimized Gabor features for mass classification in mammography. *Appl. Soft Comput.* **2016**, *44*, 267–280. [[CrossRef](#)]
16. Liu, X.; Tang, J. Mass classification in mammograms using selected Geometry and texture features, and a new SVM-based feature selection method. *IEEE Syst. J.* **2014**, *8*, 910–920. [[CrossRef](#)]
17. Mudigonda, N.; Rangayyan, R.; Desautels, J. Gradient and texture analysis for the classification of mammographic masses. *IEEE Trans. Med. Imaging* **2000**, *19*, 1032–1043. [[PubMed](#)]
18. Mohanty, A.K.; Senapati, M.R.; Beberta, S.; Lenka, S.K. Texture-based features for classification of mammograms using decision tree. *Neural Comput. Appl.* **2012**, *23*, 1011. [[CrossRef](#)]
19. Ioan, B.; Alexandru, G. Directional features for automatic tumor classification of mammogram images. *Biomed. Signal Process. Control.* **2011**, *6*, 370–378.
20. Moayedi, F.; Azimifar, Z.; Boostani, R.; Katebi, S. Contourlet-based mammography mass classification using the SVM family. *Comput. Biol. Med.* **2010**, *40*, 373–383. [[CrossRef](#)] [[PubMed](#)]

21. Nagarajan, V.; Britto, E.C.; Veeraputhiran, S.M. Feature extraction based on empirical mode decomposition for automatic mass classification of mammogram images. *Med. Nov. Technol. Devices* **2019**, *1*, 100004. [[CrossRef](#)]
22. Nunes, J.C.; Bouaoune, Y.; Delechelle, E.; Niang, O.; Bunel, P. Image analysis by bidimensional Empirical Mode Decomposition. *Image Vis. Comput.* **2003**, *21*, 1019–1026. [[CrossRef](#)]
23. Zhang, B.; Zhang, C.; Wu, J.; Liu, H. medical image fusion method based on energy classification of BEMD components. *Optik* **2014**, *125*, 146–153. [[CrossRef](#)]
24. Haralick, R.; Shannungan, K.; Dinstein, I. Textural Features For Image Classification. *IEEE Syst. Man. Cybern.* **1973**, *3*, 610–621. [[CrossRef](#)]
25. Vandenbroucke, N.; Macaire, L.; Postaire, J.G. Color image segmentation by supervised pixel classification in a color texture feature space. Application to soccer image segmentation. In Proceedings of the 15th International Conference on Pattern Recognition (ICPR'00), Barcelona, Spain, 3–8 September 2000; Volume 3, pp. 625–628.
26. Jain, A.; Zongker, D. Feature selection: Evaluation, application and small sample performance. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *19*, 153–158. [[CrossRef](#)]
27. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis. *Proc. R. Soc. Lond.* **1998**, *454*, 903–995. [[CrossRef](#)]
28. Coughlin, K.; Tung, K.K. Chapter 10. Empirical Mode Decomposition of Climate Variability in the Atmospheric. In *Hilbert-Huang Transform: Introduction and Applications*; Huang, N., Shen, S., Eds.; World Scientific Publishing: Singapore, 2004.
29. Addai, A. *Bridge Sensor Data Analysis Using the Hilbert-Huang Transform*; Rapport technique; University of Illinois-Urbana Champaign: Champaign, IL, USA, 2003.
30. Feldman, M.; Seibold, S. Damage diagnosis of rotors: Application of Hilbert transform and multi-hypothesis testing. *J. Vib. Control.* **1999**, *5*, 421–442. [[CrossRef](#)]
31. Boudra, A.O.; Pollet, C.; Cexus, J.C.; Saidi, Z. Caractérisation des fonds marins par décomposition modale empirique. In Proceedings of the Colloque GRETSI 05, Louvain-La-Neuve, Belgium, 6–9 September 2005; pp. 559–562.
32. Huang, W.; Shen, Z.; Huang, N.E.; Fung, Y.C. Engineering analysis of biological variables: An example of blood pressure over 1 day. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4816–4821. [[CrossRef](#)]
33. Zhou, Y.; Li, H. Adaptive noise reduction method for DSPI fringes based on bi-dimensional ensemble empirical mode decomposition. *Opt. Express* **2011**, *19*, 18207–18215. [[CrossRef](#)]
34. Linderhed, A. Compression by image empirical mode decomposition. In Proceedings of the IEEE International Conference on Image Processing 2005, Genoa, Italy, 11–14 September 2005; Volume 51.
35. Liu, Z.; Peng, S. Boundary processing of bidimensional EMD using texture synthesis. *IEEE Signal Process. Lett.* **2004**, *12*, 33–36.
36. He, Z. Multivariate gray model-based BEMD for hyperspectral image classification. *IEEE Trans. Instrum. Meas.* **2013**, *62*, 889–904. [[CrossRef](#)]
37. Benkuider, A.; Aarab, A. Content Based Image Retrieval using the Generalized Gamma Density to model BEMD's IMF. *J. Comput.* **2011**, *6*, 1168–1174. [[CrossRef](#)]
38. Benkuider, A.; Aarab, A. A New Scheme for Watermarking Images based on the BEMD. *ICGST Int. J. Graph. Vis. Image Process.* **2011**, *11*, 9–17.
39. Philipp, S. *Analyse de Texture Applique aux Radiographies Industrielles*. Ph.D. Dissertation, These de University P. et M. Curie, Paris VI, Paris, France, 1988.
40. Linderhed, A. Image empirical mode decomposition: A new tool for image processing. *Adv. Adapt. Data Anal.* **2009**, *1*, 265–294. [[CrossRef](#)]
41. Rilling, G.; Flandrin, P.; Gonçalves, P. On empirical mode decomposition and its algorithms. In Proceedings of the IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, Grado, Italy, 1 June 2003.
42. Flandrin, P.; Rilling, G.; Gonçalves, P. Empirical mode decomposition as a filter bank. *IEEE Signal Process. Lett.* **2004**, *11*, 112–114. [[CrossRef](#)]
43. Beucher, S. Geodesic reconstruction, saddle zones and hierarchical segmentation. *Image Anal. Stereol.* **2001**, *20*, 137–141. [[CrossRef](#)]
44. Beucher, S.; Meyer, F. The morphological approach to segmentation: The watershed transformation. In *Mathematical Morphology in Image Processing*; Dougherty, E., Ed.; Marcel Dekker, Inc.: New York, NY, USA, 1993; pp. 433–481.
45. Hardy, R.L. Multiquadratic equations of topography and other irregular surfaces. *J. Geophys. Res.* **1971**, *76*, 1905–1915. [[CrossRef](#)]
46. Chen, P.C.; Pavlidis, T. Segmentation by texture using a co-occurrence matrix and a splitand-merge algorithm. *Comput. Graph. Image Process.* **1979**, *10*, 172–182. [[CrossRef](#)]
47. Marceau, D.J.; Howarth, P.J.; Dubois, J.; Gratton, D.J. Evaluation of the grey-level co-occurrence matrix method for land-cover classification using spot imager. *IEEE Trans. Geosci. Remote Sens.* **1990**, *28*, 513–519. [[CrossRef](#)]
48. Kovalev, V.; Petrou, M. Multidimensional co-occurrence matrices for object recognition and matching. *Graph. Models Image Process.* **1996**, *58*, 187–197. [[CrossRef](#)]
49. Kim, J.K.; Park, H.W. Statistical texture features for detection of microcalcifications in digitized mammograms. *IEEE Trans. Med. Imaging* **1999**, *18*, 231–238.
50. Valkealahti, K.; Oja, E. Reduced multidimensional co-occurrence histograms in texture classification. *IEEE Trans Pattern Anal. Mach. Intell.* **1998**, *20*, 90–94. [[CrossRef](#)]

51. Webel, J.; Gola, J.; Britz, D.; Mücklich, F. A new analysis approach based on Haralick texture features for the characterization of microstructure on the example of low-alloy steels. *Materials* **2018**, *144*, 584–596.
52. Haddon, J.F.; Boyce, J.F. Co-occurrence matrices for image analysis. *IEEE Electron. Commun. Eng. J.* **1993**, *5*, 71–83. [[CrossRef](#)]
53. Chekouo, T.; Mohammed, S.; Rao, A. A Bayesian 2D functional linear model for gray-level co-occurrence matrices in texture analysis of lower grade gliomas. *NeuroImage Clin.* **2020**, *28*, 102–437. [[CrossRef](#)] [[PubMed](#)]
54. Chen, L.; Bentley, P.; Mori, K.; Misawa, K.; Fujiwa, M.; Rueckert, D. Self-supervised learning for medical image analysis using image conte restoration. *Med. Image Anal.* **2019**, *58*, 101539. [[CrossRef](#)]
55. Kalina, J.; Matonoha, C. A sparse pair-preserving centroid-based supervised learning method for high-dimensional biomedical data or images. *Biocybern. Biomed. Eng.* **2020**, *40*, 774–786. [[CrossRef](#)]
56. Kudo, M.; Sklansky, J. Comparison of algorithms that select feature for pattern classifiers. *Pattern Recognit.* **2000**, *33*, 25–41. [[CrossRef](#)]
57. Zhu, Z.; Ong, Y.S.; Dash, M. Wrapper-filter feature selection algorithm using a mimetic framework. *IEEE Trans. Syst. Man Cybern.* **2007**, *37*, 70–76. [[CrossRef](#)]
58. Della Pietra, S.; Della Pietra, V.; Lafferty, J. Inducing features of random fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **1997**, *19*, 380–393. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.