JOURNAL OF SOFTWARE ENGINEERING & INTELLIGENT SYSTEMS
ISSN 2518-8739
31st December 2018, Volume 3, Issue 3, JSEIS, CADMEI Copyright © 2016-2018
www.jseis.org

EFFICIENT FEATURES REDUCTION FOR BREAST CANCER IMAGES CLASSIFICATION

[1]M. H. SAAD, [2]MOHAMED S. EL_TOKHY

[1]Radiation Engineering Dept., NCRRT, 11787, Nasr City, Cairo, Atomic Energy Authority, Egypt
[2]Engineering Department, Nuclear Research Center, Atomic Energy Authority, P.No. 13759, Inshas, Egypt
Email: [1]m.hassansaad@gmail.com

OPEN ACCESS

Academic Editor: Muhammad Sajid Khan, APCOMS, Pakistan.

Published: 31st December, 2018

Funding: The author has no funds.

Article Type: Research Article

ABSTRACT

Reducing the classification process time, for identifying male tissues in medical images, has always been a challenge. In this paper, an efficient approach is presented that improves the classification time of breast cancer mammogram images. First, a preprocessing phase is applied for background correction and noise removal using Wave Atom transform and a universal cascaded filter, respectively. Then, the segmentation process is carried out through Particle Swarm Optimization, Hidden Markov Random Field Model with its Expectation-Maximization algorithm, and Kirsch's templates. Finally, features extraction and classification are employed. The proposed method is used to efficiently reduce the amount of data that represent the statistically based textural features of the image. Experimental results show that the employment of the selected preprocessing algorithms leads to a better classification accuracy than that of the commonly known digital filters. Moreover, the proposed method improves the classification time by using only six, out of twenty-two, features that mostly contains the image information. Finally, classification procedures using Naive Bayes (NB) is applied depending on the reduced features. The accuracy of the obtained results is compared with literature and good results are achieved.

Keywords : signal and image processing; mammography; image classification; breast cancer; particle swarm optimization; hidden Markov random field;

1. INTRODUCTION

The interpretation and analysis of medical images represent an important and exciting part of computer vision and pattern recognition. Signal and image processing techniques have been successfully applied on medical data as a tool for screening and improving diagnosis [1]. Developing a computer-aided diagnosis (CAD) system for cancer diseases, such as breast cancer, to assist physicians in hospitals is becoming of high importance and priority for many researchers and clinical centers. It is a complex process to develop a computer vision system to perform such tasks [2]. Breast cancer is the most common cancer among women [3]. It is today one of the most common diseases and a lot of effort has been put in prevention through different types of screening processes [2]. It is the second leading cause of cancer death in women [1, 4] after lung cancer. It can be most effectively treated when detected in its early stage. Mammography is used a lot for early detection and screening of breast cancer in past decades [3]. However, reading mammography is a labor-intensive job for radiologists, who cannot always provide consistent results between readings. The readings depend on training, experience, and subjective criteria [3]. In fact, radiologists miss about ten percent of all malignant tumors in mammograms, and most of the missed tumors are in dense breasts [3]. On the other hand, mammography causes a high false positive rate [3].

Computer aided diagnosis (CAD) algorithms rely on breast segmentation as a first preprocessing step [2]. To perform breast segmentation as well as other image enhancements or region detections, many image-processing methods have been presented. Accurate breast segmentation is important in both scanned film mammograms and digital mammograms but is performed with a different degree of difficulty according to the capturing conditions. Scanned images usually suffer from much more artifacts than digital images, which are usually segmented by detector calibration during capturing. Artifacts that are usually found in scanned film mammograms are orientation tags, light leakages caused by non-uniform sensitivity and thickness of film and imperfection of scanning process itself [2]. There have been many attempts to develop an optimal breast segmentation algorithm [2]. The objective of a typical breast cancer CAD system is to detect and evaluate various cancerous mammograms automatically [5]. To improve the accuracy and efficiency of digital mammogram interpretation, CAD systems have been developed [5]. Computer-aided diagnosis (CAD) systems can help radiologists in interpreting sonography for mass detection and classification. The combination of CAD scheme and experts' knowledge

Corresponding Author: M. H. Saad
Email Address: m.hassansaad@gmail.com

would greatly improve the detection and classification accuracy. This paper focuses on developing a novel CAD system for automatic mass detection and classification using radiation therapy images in [3]. It is evolves different steps identified in Figure 1.

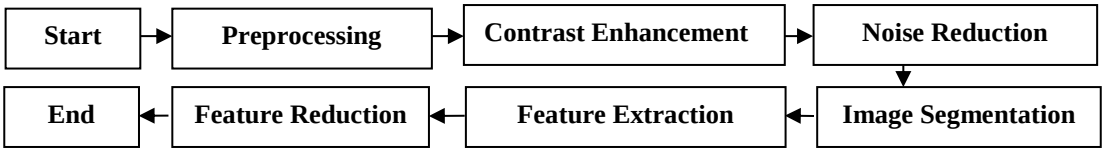

Figure. 1 General process flow of automated mass detection using breast radiotherapy images systems

2. BREAST CANCER DETECTION ALGORITHMS

Mammograms are images difficult to interpret. Therefore, a preprocessing phase of images is necessary to improve the quality of the images and make the feature extraction phase more reliable. This section introduces the preprocessing techniques before the feature extraction stage. The preprocessing stage consists of two main phases, which are used together. The first phase involves the removal of background information and unwanted parts from the image, while the second phase deals with enhancing the contrast of suspicious areas in the image [6].

It is well known that breast cancer image often contains plenty of artifacts and noises due to the complex imaging environment and imaging principle, such as low contrast. They greatly degrade the performance of conventional segmentation methods. A preprocessing procedure for noise reduction was required [4]. Image preprocessing suppresses noise and enhances the contrast between the suspicious areas and breast tissue background. Here, we preprocess the region of interests (ROIs) of the radiation therapy images, which were provided by the radiologists [3].

2.1 Image preprocessing using wave atom transform

Wave atoms (WA) is multi-scale geometric transform that have revealed themselves quite useful over the past few years in such diverse fields as image processing, seismic imaging, numerical analysis, and analysis of partial differential equations [7, 8].

Wave atom was proposed as an efficient representation of oscillating patterns and textures in a large class of images. Like fast discrete curvelet transform (FDCT). It is implemented in the frequency-domain by applying windowing and wrapping operations [9]. Unlike fast discrete cosine transform (FDCT), the bandwidth of the frequency subbands of WA is not dyadic ally related across the scales. From the fundamentals of the WA implementation, a simple procedure is derived to determine the cut-off frequencies of the low pass filters [9]. In addition, the WA functions are isotropic in space and consist of oscillations oriented in a single direction. They achieve sparse representation of oscillating patterns in images because of the property that the wavelength of oscillation of any WA function is proportional to the square of its diameter [9]. This implies that if the scale, which is proportional to the diameter, increases dyadically. Then, the bandwidth of the frequency subband, which is inversely proportional to the wavelength, must decrease by powers of four [9]. This ‘‘parabolic scaling’’ is implemented by tiling the frequency-domain according to a certain scheme shown in Fig. 2 by using the Villemoes’ wavelet packets [9].

Background correction algorithm for breast cancer image using WA is depicted in Fig. 3. In what follows, the exposition is restricted to two dimensions, and write $x= (x_1, x_2)$, although this restriction is not essential. The convention for the 2D Fourier transform is [10, 11]

$$\overline{f}(\omega) = \int e^{-ix.\omega} f(x) dx \tag{1}$$

$$f(x) = \frac{1}{(2\pi)^2} \int e^{ix\omega} \overline{f}(\omega) d\omega \tag{2}$$

The wave atoms is written as $\varphi\mu(x)$, with subscript $\mu=(j,m,n)=(j,m1,m2,n1,n2)$. All five quantities j, m1, m2, n1, n2 are integer-valued and index a point $(x\mu,\omega\mu)$ in phase-space, as [10, 11]

$$x_\mu = 2^{-j} n \tag{3}$$

$$\omega_n = \pi 2^j m \tag{4}$$

$$C_1 2^j \le \max_{i=1,2} |m_i| \le C_2 2^j \quad (5)$$

where C1, C2 are two positive constants left unspecified for convenience, but whose values will be implied by the specifics of the implementation. Heuristically, the position vector xμ is the center of φμ(x) and the wave vector ωμ determines the centers of both bumps of $\overline{\phi}_\mu(\omega)$ as ±ωμ. The range of m needs to be further reduced to m2>0 (or m2=0 and m1>0) to account for the central symmetry of the Fourier transform of real-valued functions about the origin in ω. Some further restriction on n (cutoff in space) and j (cutoff in scale), are of course necessary in practice, but not for the description of a frame of L2(^{2}). Then, WAs needs to obey a localization condition around the phase-space point (xμ, ωμ).

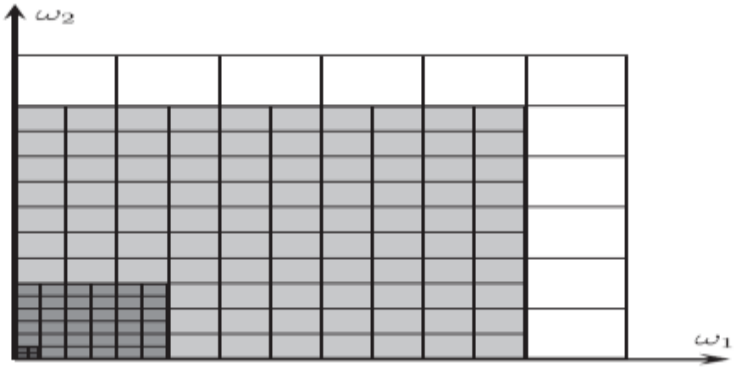

Figure. 2 Frequency-domain tiling corresponding to WA

- **Input Breast Mammogram Image**
- **Image Resizing Process**
- **Perform Wave Atom Transform**
  - **Select Type of Transform (Complex, Directional, Orthobasis)**
  - **Specify the Type of Frequency Partition**
  - **Extract Absolute Coefficients**
- **Perform Inverse Wave Atom Transform**
- **Image Reconstruction**

Figure. 3 Background correction of breast cancer image using WA

2.2 The universal denoising filter

Image filtering is used for removing noise, sharpening contrast, or highlighting contours in the breast cancer images. The proposed filter is a cascaded spatial filter based on median fitter, Wiener filter and Coiflet Wavelets [12]. Its edge-preserving nature makes it useful in cases where edge blurring is undesirable. This filter is designed for minimizing the Mean Square Error (MSE) between the filtered image and the original image. It can be applied to the image adaptively, tailoring itself to the local image variance. When the variance is large, it performs little smoothing, otherwise it performs more smoothing. Since the intensities of the pixels in the mass area are distributed in a Gaussian-like distribution, the Wiener filter is selected for filtering breast cancer images of welds. Median filtering is used to remove the high frequency components in the mammogram image, where it is used to remove the noise without disturbing the edges. For each pixel a window of pixels are extracted and the median value is calculated for that window. The intensity value of the center pixel value is replaced with the median value. This procedure is done for all the pixels in the image to smoothen the mammogram image. The window size that we use it in median filtering for enhancing the image give us another benefit and strength the relation between pixels in the same window and that causing more strength to mammogram layers.

2.3 Image segmentation

Segmentation is the process of dividing an image into non-overlapping regions, such that each region is homogeneous but the union of any two neighboring regions is inhomogeneous [3]. The image segmentation method refers to the major step in image processing. The inputs are images. However, outputs are the attributes extracted from those images. Segmentation divides image into its constituent regions or objects. The level to

which segmentation is carried out depends upon the problem being solved i.e. segmentation should stop when the objects of interest in an application have been isolated. Image segmentation refers to the decomposition of a scene into its components [13]. It plays an important role in image analysis. The goal of segmentation is to isolate the regions of interest depending on the problem and its characteristics. Many applications of image analysis need to obtain the regions of interest before the analysis can start. Therefore, the need of an efficient segmentation method has always been required [13]. Segmentation algorithms for gray images are generally based on two basic properties of image intensity values, discontinuity and similarity. In the first category, the approach is to partition an image based on abrupt changes in intensity, such as edges in an image. The principle approaches in the second category are based on partitioning image into regions that is similar according to a set of predefined criteria. Thresholding, region growing, region splitting and merging are examples of the methods in this category [13]. Breast segmentation consists of breast border contour extraction, pectoral muscle extraction and nipple identification. Therefore, we have discussed different image segmentation techniques and their role in image analysis.

a) Image segmentation using particle swarm optimization

The breast cancer image is segmented using particle swarm optimization (PSO). This algorithm accepts any type of image with multiple intensity levels; grayscale or/and color. This algorithm uses different methods for segmentation. These methods are particle swarm optimization (PSO), darwinian PSO (DPSO), fractional order DPSO (FO-DPSO) and exhaustive methods based on the image histogram shape. This algorithm is shown in Fig. 4.

- **Input Background Corrected Breast Cancer Image**
- **Parameters Definition**
  - **Number of Thresholds**
  - **Number of Iterations of the PSO Algorithm**
  - **Individual and Social Weight**
  - **Predefined PSO Population for Multi-Segmentation**
- **Select Segmentation Level and Method to Perform Segmentation**
- **Determine Velocities of Particles**
- **Identify**
  - **Current population of the swarm**
  - **Maximum and minimum population**
  - **Population from new swarms**
  - **Current number of swarms**
  - **Maximum and minimum number of swarms**
  - **Number of iterations of the PSO algorithm**
  - **Start and end of swarm**
  - **Save the best particle from all swarms**
  - **Evaluate each swarm**
  - **Create a new particle**
- **Identify and Eliminate Worst Particle**
- **Output the Segmented Image**

Figure. 4 Segmentation of breast cancer image using PSO method

b) Image Segmentation Using Hidden Markov Random Field Model and Its Expectation-Maximization (HMRF-EM) Algorithm

Markov random fields (MRFs) have been widely used for computer vision problems, such as image segmentation, surface reconstruction and depth inference [14]. Image segmentation can be thought of as a labelling problem where the task is to label the image pixels as belonging to one of several groups. Particle selection amounts to a binary filtering operation, in that pixels are marked according to whether they belong to the particle or the background. Hence, such a segmentation could be thought of as labelling of the image field with labels picked up from a binary set. Here, the rationale for image segmentation based on the use of hidden Markov random field models and its expectation maximization is discussed [15]. Given an image $y = (y_1, \ldots, y_N)$ where each y_i is the intensity of a pixel, and it is necessary to infer a configuration of labels $x = (x_1, \ldots, x_N)$ where $x_i \in L$ and L is the set of all possible labels. In a binary segmentation problem, $L = \{0, 1\}$. The breast cancer image segmentation algorithm using HMRF-EM is depicted in Figure. 5 [14].

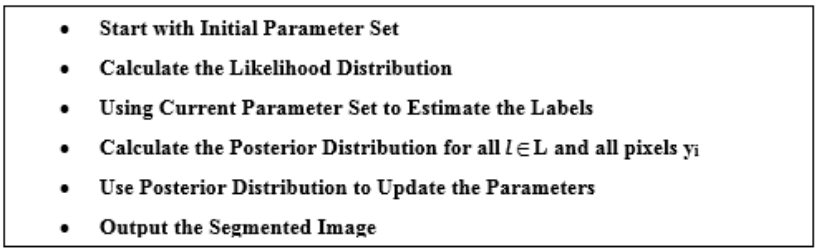

Figure. 5 Segmentation of breast cancer image using HMRF-EM

3. AUTOMATIC DETECTION AND SUPPRESSION OF PECTORAL MUSCLE USING KIRSCH'S TEMPLATES

Accurate segmentation is essential to obtain a correct computer aided diagnosis. Intensity based method may produce erroneous result when applied to dense structure, abnormal masses, fibroglandular disc. Since, it has the nearly same opacity. So, it is important to detect the pectoral muscle and isolate it from the region of interest (ROI). Usually pectoral muscles are homogenous areas. The pectoral muscle has a slightly higher intensity compared to the rest of the breast tissue and appear in upper left corner of medio-lateral oblique (MLO) view of mammogram [13]. In majority of cases, pectoral muscles are consisted of brighter pixels than surrounding fat tissue and are usually easy to distinguish from the surrounding tissue. In some cases, parts of pectoral muscle or the entire pectoral muscle are not visible [2]. Reasons for that can be low overall contrast of the image or imperfection in mammogram capturing process. These cases present the most problems for automatic detection algorithms. To overcome that problem we have applied Kirsch's templates on the extracted rectangle, in order to suppress the pectoral muscle.

Spatial filtering of the input breast cancer image is done with the Kirsch's templates in different orientations. Then, it is followed by thresholding and results in the extracted pectoral muscle. Using this Kirsch's templates, pectoral muscle that is an unwanted part of the mammogram will be eliminated. By superimposing the eliminated pectoral region on the original mammogram. The breast image containing the breast tissues only (ROI) is provided. It will be helpful for further investigation of mammogram properly [13]. Therefore, it is summarized in to different steps as shown in Fig. 6.

Figure. 6 Pectoral muscle suppression using Kirsch's templates

4. IMAGE CLASSIFICATION

Here, we are concerned with image classification. There are several multi-class classifiers that can be employed but the ones selected must be able to cope with non-linearity and eventually high dimensionality. Features are extracted from both normal and abnormal images using texture features as depicted in Fig. 7.

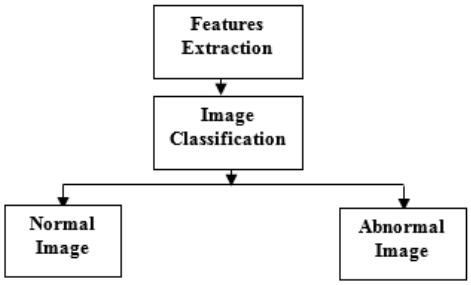

Figure. 7 Flowchart of image classification process to identify normal and abnormal breast

4.1 Textural feature

Texture is one of the important characteristics used in identifying objects or regions of interest in an image. This section describes some easily computable textural features based on graytone spatial dependencies [16]. The

concept of tone is based on the varying shades of gray of resolution cells in a mammogram image, while texture is concerned with the spatial (statistical) distribution of gray tones. Texture and tone are not independent concepts. They are always present in an image, although one property can dominate the other at times [16]. The basic intuitively perceived relationships between tone and texture are the following. When a small-area patch of an image has little variation of discrete gray tone, the dominant property of that area is tone. When a small-area patch has a wide variation of features of discrete gray tone, the dominant property of that area is texture. Crucial to this distinction are the size of the small-area patch, the relative sizes of the discrete features and the number of distinguishable discrete features [16].

One important property of tone-texture is the spatial pattern of the resolution cells composing each discrete tonal feature. A fine texture result when there is no spatial pattern and the gray-tone variation between features is wide. As the spatial pattern becomes more definite and involves more and more resolution cells, a coarser texture result [16]. Discrete tonal features are really quite fuzzy, since they do not necessarily stand out as entities by themselves. Therefore, the texture analysis is concerned with more general or macroscopic concepts than discrete tonal features [16].

The procedure for obtaining the textural features of an image is based on the assumption that the texture information on an image, I, is contained in the overall spatial relationship. More specifically, the texture information is assumed to be adequately specified by a set of gray-tone spatial-dependence matrices. These are computed for various angular relationships and distances between neighboring resolution cell pairs on the image. All of textural features are derived from these angular nearest-neighbor gray-tone spatial-dependence matrices [16].

a) Textural feature extraction

Textural features are based on co-occurrence matrices of the texture information. All textural features are derived from the spatial gray-level dependence (SGLD) matrices, which are two-dimensional histograms. An element of the SGLD$_\theta$ matrix $P(i, j,d, \theta)$ is defined as the joint probability of the gray levels i and j separated by distance, d, and along direction, θ. In order to simplify the computational complexity, the values of θ are often given as 0º, 450º, 90º, and 135º. The distance, d, is often defined as the Manhattan or city block distance. Since, $\theta = \arctan[(y2 - y1)/(x2 - x1)]$, the element $P(i, j,d, \theta)$ of the SGLD$_\theta$ matrix can be expressed as [16, 17]:

$$P(i, j, d, 0^o) = \left\| \{(x_1, y_1)(x_2, y_2) \text{ where } |x_2 - x_1| = d, y_2 - y_1 = 0\} \right\|$$

$$P(i, j, d, 45^o) = \left\| \{(x_1, y_1)(x_2, y_2) \text{ where } (x_2 - x_1 = d, y_2 - y_1 = d) \text{ or } (x_2 - x_1 = -d, y_2 - y_1 = -d)\} \right\|$$

$$P(i, j, d, 90^o) = \left\| \{(x_1, y_1)(x_2, y_2) \text{ where } x_2 - x_1 = 0, |y_2 - y_1| = d\} \right\|$$

$$P(i, j, d, 135^o) = \left\| \{(x_1, y_1)(x_2, y_2) \text{ where } (x_2 - x_1 = d, y_2 - y_1 = -d) \text{ or } (x_2 - x_1 = -d, y_2 - y_1 = d)\} \right\|$$

where $i = I(x_1, y_1)$, $j = I(x_2, y_2)$, $I(x, y)$ denotes the intensity value of pixel (x, y) and $\|S\|$ is the cardinality of set, S.

Textural features can be extracted from SGLD matrices with different distances, d, and directions, θ. Given a distance, d, four SGLD matrices can be calculated corresponding to 0°, 45°, 90° and 135°, respectively. A set of four values for each of the 14 measures listed in Appendix A are produced. For each measure, we can compute the mean and range of the set of four values. Therefore, a set of 28 textural features is extracted from these four matrices for a given distance, d. There are five d values (d = 1, 2, 4, 8 and 16). Hence, totally 140 textual features are extracted.

b) Feature reduction of textural features

A simple program code was written to compare between the extracted features from normal and abnormal breast cancer images as shown in Fig. 8. Features are extracted from more than 40 images by texture features. The average of these features are calculated and compared for normal and abnormal images.

- **Input Normal and Abnormal Preprocessed Images**
- **Extract the Features from the Normal and Abnormal Images**
- **Take Average of Each Feature for Both Normal and Abnormal Images**
- **If** $\phi_i - \lambda_i > \frac{1}{2}\max(\phi_i, \lambda_i)$, **Where** ϕ_i **Denotes the Features of Abnormal Image and** λ_i **Denote the Features of Normal Image**
- **Select This Feature for Comparison**
- **Otherwise, Reject This Feature and Search for Other Features**
- **Summarize the Accepted Features for Comparison and Image Classification**

Figure. 8 Features reduction algorithm based on textures feature extraction

c) Classification using Naive Bayes

Naïve Bayes (NB) classifier is a probabilistic classifier that gives a probability distribution of set of all classes as well as given an input rather than providing a single output. Naive Bayes Nearest Neighbor has been proposed as a powerful, learning-free, non-parametric approach for object classification. Its good performance is mainly due to the avoidance of a vector quantization step, and the use of image-to-class comparisons, yielding good generalization. The advantage of using Naive Bayes Classifier is simplicity and converges quicker than logistic regression. Moreover, it works well for small amount of training to calculate the classification parameters [18, 19]. Each image (*Q)* is classified as belonging to class ($\hat{C}$) according to

$$\hat{C} = \arg\max_{C} p(C|Q) \tag{6}$$

Assuming a uniform prior over classes and applying Bayes' rule,

$$\hat{C} = \arg\max_{C} \log(p(C|Q)) \tag{7}$$

The assumption of independence of the descriptors d_i found in image Q gives

$$\hat{C} = \arg\max_{C} \left[\log(\prod_{i=1}^{n} p(d_i|C)) \right] \tag{8}$$

$$\hat{C} = \arg\max_{C} \left[\sum_{i=1}^{n} \log(p(d_i|C)) \right] \tag{9}$$

Next, approximating $p(d_i|C)$ in Equation 9 by a Parzen window estimator, with kernel K, gives

$$\hat{p}(d_i|C)) = \frac{1}{L}\sum_{j=1}^{L} K(d_i - d_j^C) \tag{10}$$

where there are L descriptors in the training set for class C and d_j^C is the j-th nearest descriptor in class C. This can be further approximated by using only the r nearest neighbors as

$$\hat{p}_r(d_i|C)) = \frac{1}{L}\sum_{j=1}^{r} K(d_i - d_j^C) \tag{11}$$

Also, NBNN takes this to the extreme by using only the single nearest neighbor $(NN_C(d_i))$:

$$\hat{p}_1(d_i|C)) = \frac{1}{L} K(d_i - NN_C(d_i)) \tag{12}$$

Choosing a Gaussian kernel for K and substituting Equation 12 (the single nearest neighbor approximation of $p(d_i|C)$ into Equation 9 (the sum of log probabilities) gives:

$$\hat{C} = \arg\max_{C} \left[\sum_{i=1}^{n} \log \frac{1}{L} e^{-\frac{1}{2\sigma^2}\|d_i - NN_C(d_i)\|^2} \right] \quad (13)$$

$$\hat{C} = \arg\max_{C} \left[\sum_{i=1}^{n} \|d_i - NN_C(d_i)\|^2 \right] \quad (14)$$

Equation 14 is the NBNN classification rule: find the class with the minimum distance from the query image.

5. RESULTS AND DISCUSSION

5.1 Results of image preprocessing using wave atom transform

A preprocessing step improves the total results of detecting the suspicious segments. It is very important in order to limit the search for abnormalities without undue influence from background of the mammogram. A preprocessing procedure for reducing the speckles and preserving the boundaries was performed before the segmentation. Image preprocessing is necessary to remove the background, enhance the image if it is dark, and normalize the image. So, every image has the same maximal intensity. The preprocessing module is used to remove the background and normalize all the line images to the same maximum gray level. Image reconstruction using wave atom transform with both 1536 and 512 coefficients is depicted in Fig. 9(a) and (b), respectively. Also, comparison between them in terms of MSE and PSNR is illustrated in Table 1. We noticed that the best results are achieved with smaller number of coefficients.

Wave atom transform at different scales for spatial domain at μ=(2,3,2) and μ=(3,5,3) is shown in Fig. 10 (a) and Fig. 10 (b), respectively. However, wave atom transform at different scales for frequency domain at μ=(2,3,2) and μ=(3,5,3) is depicted in Fig. 10 (c) and Fig. 10 (d), respectively.

Table. 1 PSNR and RMS for breast cancer image using WA with 1536 and 512 coefficients

Number of Coefficients	MSE	PSNR
Wave Atom with 1536	0.010054	68.1075
Wave Atom with 512	0.0092352	68.4763

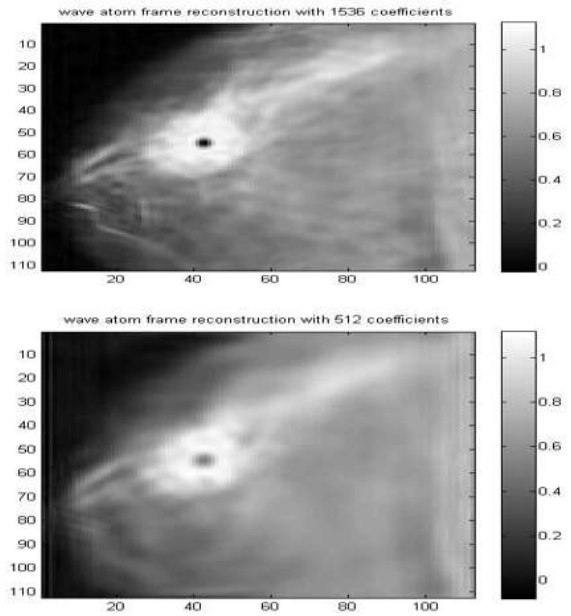

Figure. 9 Image reconstruction using wave atom transform with a) 1536 coefficients and b) 512 coefficients

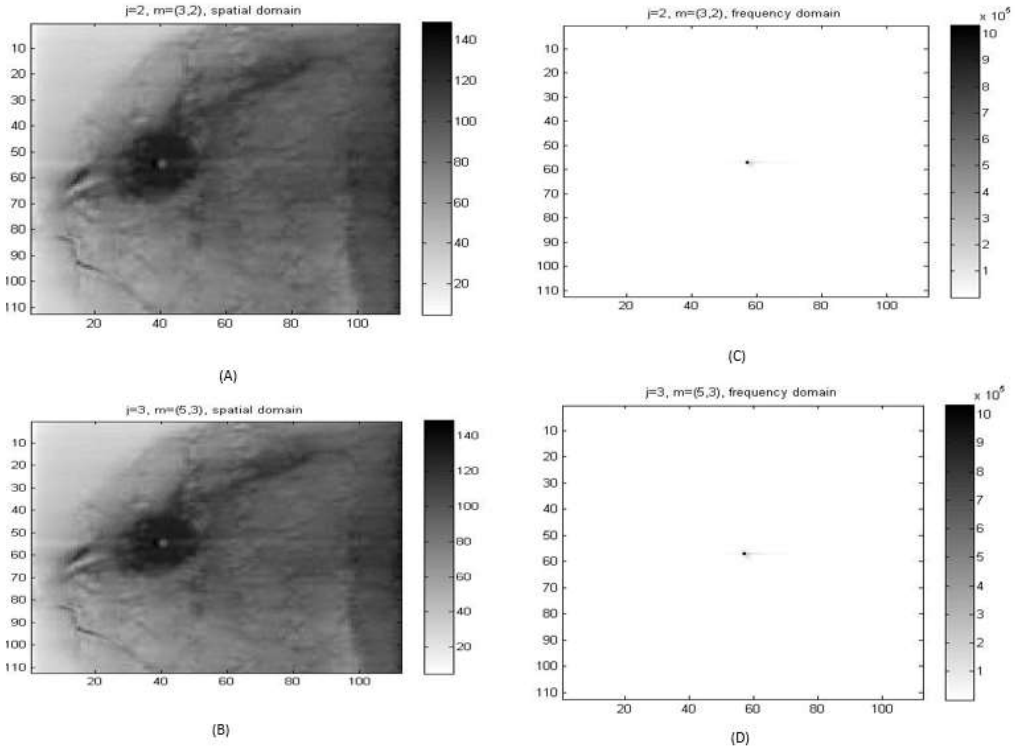

Figure. 10 Wave atom transform at different scales for spatial domain a) μ=(2,3,2) b) μ=(3,5,3) and frequency domain c) μ=(2,3,2) d) μ=(3,5,3)

5.2 Image denoising results using the universal filter

In this subsection, we are interested with removing the existing overall noise. Filters are routinely designed to search for specific shapes and at the same time manage noise, reduce aliasing and subtract background fluctuations [16]. Therefore, we are concerned with de-noising the breast cancer image using the proposed filter. Figures 11(a)–2(d) show the original image and filtered image results, respectively. Comparison between these filters in terms of peak signal to noise ratio (PSNR) and RMS is illustrated in Table 2. The obtained results confirm that the proposed universal denoising filter proved more efficiency compared with the traditional filters and achieves higher PSNR and low MSE.

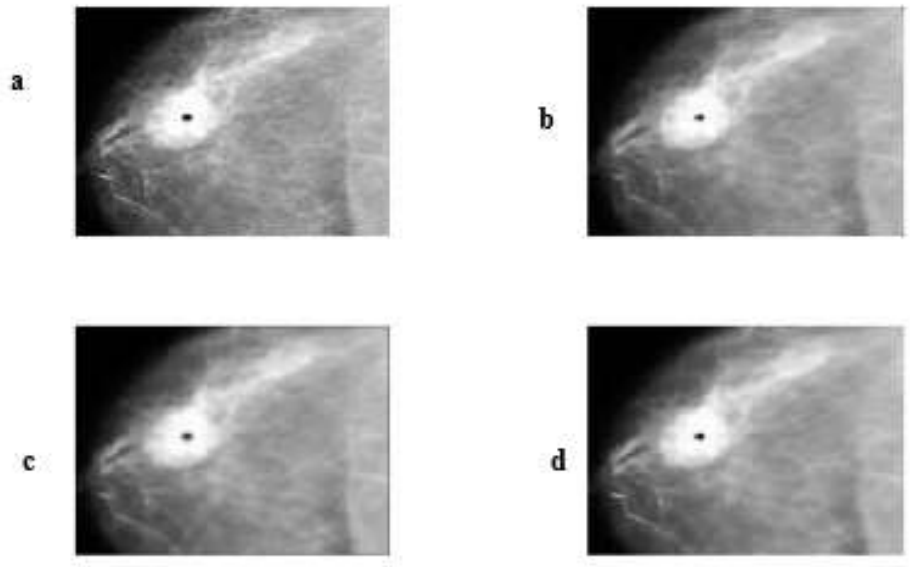

Figure. 11 Filtered images using a) Original image b) median filter, c) Wiener filter and d) Proposed filter

Table. 2 Evaluation of breast cancer image de-noising using median, Wiener and proposed filters

	RMS	PSNR
Median filter	**15.3273**	**24.4215**
Wiener filter	**15.6482**	**24.2415**
proposed filter	**10.7555**	**41.0404**

5.3 Image enhancement results

Some preprocessing may take place like noise reduction and contrast enhancement (histogram equalization) to assist the breast mass segmentation. Breast cancer image enhancement is effectively used for better image interpretation. The image digitization process may produce small contrast images containing some granularities. Image enhancement is required for improving the contrast between the image background and the abnormal areas and for removing any noise resulting from the digitization process. Figure 12 shows the effect of the histogram equalization at different orders. Histogram equalization with order 1 achieves better results.

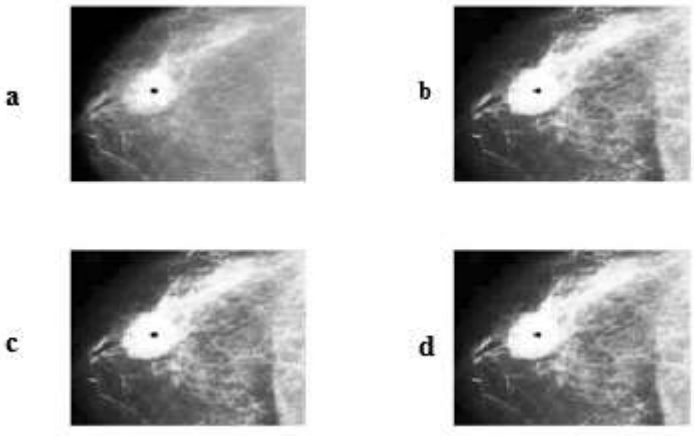

Figure. 12 Image enhancement using histogram equalization a) Original image b) Enhanced image with order 1, c) Enhanced image with order 2 d) Enhanced image with order 3

5.4 Image segmentation results using PSO method

The original breast cancer is depicted in Fig. 13 (a). The segmented image based on level 4 is shown in Fig. 13 (b). However, the segmented image is illustrated in Fig. 13 (c). Furthermore, this algorithm is evaluated in terms of intensity that maximizes between class variance, fitness between class variance and run time as depicted in Table 3.

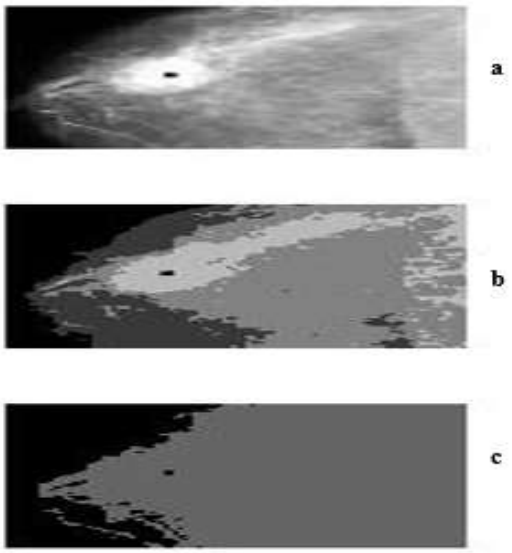

Figure. 13 Image segmentation using PSO method a) Original image b) Segmented image with level four c) Segmented Image

Table. 3 Evaluation of breast cancer image segmentation using PSO method

Intensity	Fitness	Run Time (Sec)
[102 101 101]	3.1684×10^3	7.3542

5.5 Image segmentation results using HMRF-EM

The initial labels obtained by the k-means algorithm are not smooth enough, have morphological holes, and do not preserve the canny edges. The original image is depicted in Fig. 14(a). However, the Gaussian blurred image is shown in Fig 14(b). The initial labels obtained by k-means, where k = 5 is shown in Fig. 14(c), and the final labels obtained by HMRF-EM algorithm is depicted in Fig. 14(d). The total posterior energy in each iteration of the EM algorithm is illustrated in Figure. 15.

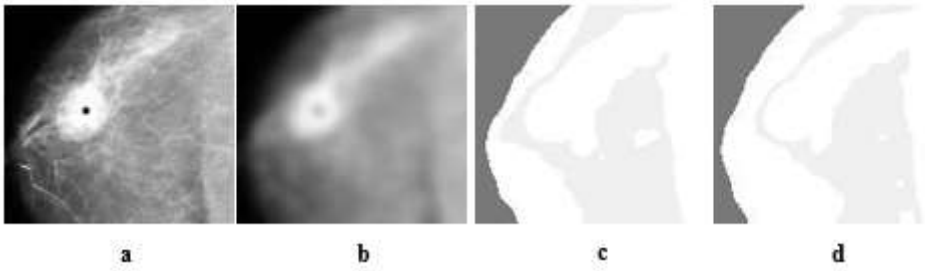

Figure. 14 Edge-prior-preserving image segmentation results (a) Original image (b) Gaussian blurred image. (c) Initial labels obtained by k-means, where k = 5 and (d) Final labels obtained by HMRF-EM algorithm.

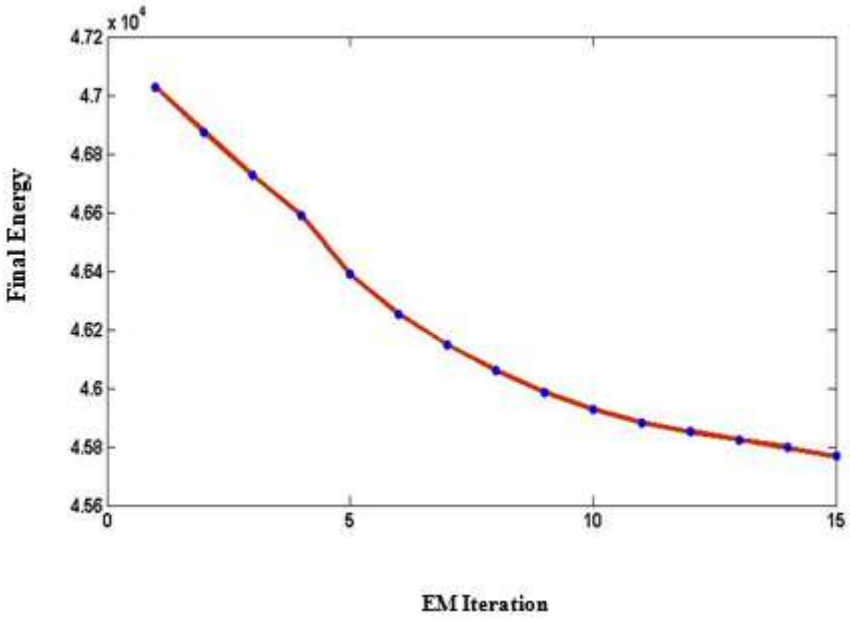

Figure. 15 Total posterior energy in each iteration of the EM algorithm

5.6 Automatic detection and suppression of pectoral muscle results using Kirsch's templates

Removal of the pectoral muscle is one of the important preprocessing. The main problem with pectoral muscle detection is low visibility of the pectoral muscle in certain types of breasts, usually higher density breasts, where there is almost no intensity variation between the pectoral muscle and the breast tissue [2]. Therefore, it is almost impossible to achieve accurate segmentation of that visually almost invisible part of the muscle. Therefore, kirsch's templates method was used. The original breast cancer is depicted in Fig. 16 (a). However, the pectoral muscle extraction is shown in Fig. 16 (b).

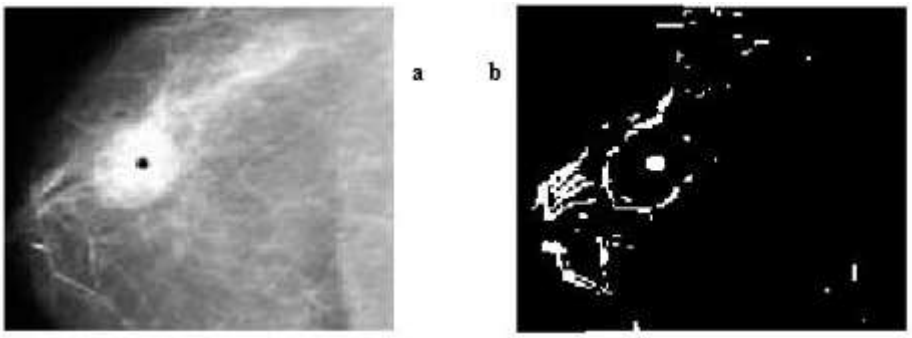

Figure. 16 Pectoral muscle extraction a) Original image b) Extracted pectoral muscle

5.7 Feature reduction results

After extraction of features from both normal and abnormal breast images, a program code is implemented to compare these features with each other. More than 22 features are extracted for both cases and tested by the proposed code. The obtained result confirms that 6 features are used to distinguish normal and abnormal breast images. These features are; autocorrelation, contrast, cluster, sum variance, difference variance, and information measure of correlation1. However, 16 features introduce similar results. Therefore, they not valid to differentiate normal and abnormal breast images as depicted in Table 4.

Table. 4 Features reduction for normal and abnormal breast under Mammograms

Feature	Accepted Feature	Feature	Rejected Feature
Autocorrelation	√	Correlation	x
Contrast	√	External Correlation	x
Cluster	√	Cluster Prominence	x
Sum Variance	√	Dissimilarity	x
Difference Variance	√	Energy	x
Information Measure of Correlation1	√	Entropy	x
		Homogeneity	x
		External Homogeneity	x
		Maximum Probability	x
		Sum of Squares Variance	x
		Sum Average	x
		Sum Entropy	x
		Difference Entropy	x
		Information Measure of Correlation2	x
		Inverse Difference Normalized	x
		Inverse Difference Moment Normalized	x

5.8 Classification results

The proposed approach has been tested using Digital Database for Screening Mammography (DDSM) at the University of South Florida [20]. DDSM is an infrastructure resource for the mammography image analysis research community. The purpose of the DDSM is to make it possible for researchers to rigorously compare the performance of different image analysis techniques. To evaluate the performance of the proposed system, the following these different measures were used

$$Accuracy(\%) = \frac{TP + TN}{TP + FN + TN + FP} \times 100 \quad (15)$$

$$Sensitivity(\%) = \frac{TP}{TP + FN} \times 100 \quad (16)$$

$$Specificity(\%) = \frac{TN}{FP + TN} \times 100 \quad (17)$$

where TP, TN, FN and FP stands for true positive, true negative, false positive and false negative respectively. Table 5 shows the performance measure of the proposed system w.r.t the others.

Table 6 presents the computation time of the proposed system. The computation time is another important factor to evaluate the classifier. For each 512 ×512 image, the averaged computation time on feature extraction, feature reduction, and NB classification is 0.097 s, 0.0889 s, and 0.00691 s, respectively.

Table. 5 Performance measures of the proposed system against the previous

Method	Accuracy	Sensitivity	Specifity
D.Wang, L.Shi and P.A.Heng [5]	91.4%	88.7%	89.4%
D.C.Pereira, R.P.Ramos and M.Z.Nascimento [21]	92%	89.1%	90.5%
B.Verma, P.McLeod and A.Klevansky [22]	90.5%	83.2%	85.8%
R.Rouhi, M.Jafari, S.Kasaei and P.Keshavarzian [23]	93.85%	92.3%	93.6%
Y.Li, H.Chen, G.K.Rohde, C.Yao and L.Cheng [24]	85.96%	80.4%	83.5%
S. Ergin, O. Kilinc [25]	78%	66.3%	69.1%
Zemmal, Nawel, Nabiha Azizi, and Mokhtar Sellami [26]	93%	92.1%	93.1%
Proposed CAD system	**95.7%**	**96.4%**	**96.9%**

Table. 6 Time performance

Steps	Consuming time
Feature Extraction	0.097 s
Feature Reduction	0.0889 s
NB Classification	0.00691 s

6. CONCLUSION

Automatic detection of cancer in early detection diagnosis is the main goal of this paper. This is done through image processing methodology. Therefore, different algorithms are presented in different related steps. The first step is preprocessing algorithm that done through Wave Atom (WA) transforms. Then, image de-noising using digital filters are accomplished. From the obtained results, the proposed filter gives the best PSNR and low MSE. The next step concerns image enhancement. An algorithm using generalized histogram equalization method is studied. Better enhancement quality is obtained. Also, image segmentation is performed using two different algorithms. These algorithms are image segmentation using Particle Swarm Optimization (PSO) and image segmentation using hidden Markov random field model and it's Expectation-Maximization (HMRF-EM). The pectoral muscle was extracted from a breast cancer image using Kirsch's templates. The breast cancer image features are extracted using textural features. These features are used for comparison purposes. Features reduction was performed using a program code that was implemented. It is based on the average value of both normal and abnormal image. We notice that the extracted features reduced from 22 to 6 only. Therefore, better execution time is obtained for image classification.

Appendix A

28 textural features are suggested which can be extracted from each of the gray-tone spatial-dependence matrices. The following equations define these features [16-17].

One should notice that $p(i, j)$ is the $(i, j)^{th}$ entry in a normalized SGLD matrix, which is equal to $P(i, j)/R$, where R is a normalizing constant and n is the number of distinct gray levels of the image, $p_x(i)^{th}$ entry in the

marginal-probability matrix obtained by summing the rows of $p(i,j)=\sum_{j=1}^{N_g}P(i,j)$, N_g denotes the number of distinct gray levels in the quantized image. The fourteen measurements are [16-17]:

1) Angular Second Moment.

$$f_1=\sum_i\sum_j\left(p(i,j)\right)^2 \tag{A}$$

2) Contrast:

$$f_2=\sum_{n=0}^{N_g-1}n^2\left(\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}p(i,j)\right) \tag{B}$$

3) Correlation:

$$f_3=\frac{\sum_i\sum_j(ij)\,p(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y} \tag{C}$$

4) Sum of Squares: Variance

$$f_4=\sum_i\sum_j(i-\mu)^2\,p(i,j) \tag{D}$$

5) Inverse Difference Moment

$$f_5=\sum_i\sum_j\frac{1}{1+(i-j)^2}\,p(i,j) \tag{E}$$

6) Sum Average:

$$f_6=\sum_{i=2}^{2N_g}ip_{x+y}(i) \tag{F}$$

7) Sum Variance:

$$f_7=\sum_{i=2}^{2N_g}(i-f_8)^2\,p_{x+y}(i) \tag{G}$$

8) Sum Entropy:2

$$f_8=-\sum_{i=2}^{2N_g}p_{x+y}(i)\log\left\{p_{x+y}(i)\right\} \tag{H}$$

9) Entropy:

$$f_9=-\sum_i\sum_j p(i,j)\log\left[p(i,j)\right] \tag{I}$$

10) Difference Variance:

$$f_{10}=Variance\ of\ p_{x-y} \tag{J}$$

11) Difference Entropy:

$$f_{11}=-\sum_{i=0}^{N_g-1}p_{x-y}(i)\log\left[p_{x-y}(i)\right] \tag{K}$$

12), 13) Information Measures of Correlation:

$$f_{12} = \frac{HXY - HXY1}{Max(HX, HY)} \quad \text{(L)}$$

$$f_{13} = \sqrt{\left(1 - e^{\left[-2.0(HXY2 - HXY)\right]}\right)} \quad \text{(M)}$$

$$HXY = -\sum_i \sum_j p(i,j) \log\left(p(i,j)\right) \quad \text{(N)}$$

where HX and HY are entropies of p_x, and p_y, and

$$HXY1 = -\sum_i \sum_j p(i,j) \log\left[p_x(i)\, p_y(j)\right] \quad \text{(O)}$$

$$HXY2 = -\sum_i \sum_j p_x(i)\, p_y(j) \log\left[p_x(i)\, p_y(j)\right] \quad \text{(P)}$$

14) Maximal Correlation Coefficient:

$$f_{14} = \sqrt{Second \text{ Largest EigenValue of Q}} \quad \text{(Q)}$$

where $$Q(i,j) = \sum_k \frac{p(i,k)\, p(j,k)}{p_x(i)\, p_y(k)} \quad \text{(R)}$$

REFERENCES

1. Borchartt, T.B., et al., *Breast thermography from an image processing viewpoint: A survey.* Signal Processing, 2013. 93(10): p. 2785-2803.
2. Mustra, M. and M. Grgic, *Robust automatic breast and pectoral muscle segmentation from scanned mammograms.* Signal Processing, 2013. 93(10): p. 2817-2827.
3. Shi, X., et al., *Detection and classification of masses in breast ultrasound images.* Digital signal processing, 2010. 20(3): p. 824-836.
4. Huang, Q.-H., et al., *A robust graph-based segmentation method for breast tumors in ultrasound images.* Ultrasonics, 2012. 52(2): p. 266-275.
5. Wang, D., L. Shi, and P.A. Heng, *Automatic detection of breast cancers in mammograms using structured support vector machines.* Neurocomputing, 2009. 72(13-15): p. 3296-3302.
6. Mousa, R., Q. Munib, and A. Moussa, *Breast cancer diagnosis system based on wavelet analysis and fuzzy-neural.* Expert Systems with Applications, 2005. 28(4): p. 713-723.
7. Demanet, L. and L. Ying. *Curvelets and wave atoms for mirror-extended images*. in *Wavelets XII*. 2007: International Society for Optics and Photonics.
8. Demanet, L., *Curvelets, wave atoms, and wave equations*. 2006, California Institute of Technology.
9. Palakkal, S. and K. Prabhu, *Poisson image denoising using fast discrete curvelet transform and wave atom.* Signal Processing, 2012. 92(9): p. 2002-2017.
10. Demanet, L. and L. Ying, *Wave atoms and time upscaling of wave equations.* Numerische Mathematik, 2009. 113(1): p. 1-71.
11. Demanet, L. and L. Ying, *Wave atoms and sparsity of oscillatory patterns.* Applied and Computational Harmonic Analysis, 2007. 23(3): p. 368-387.
12. Patidar, P., et al., *Image de-noising by various filters for different noise*. International journal of computer applications, 2010. 9(4).
13. Maitra, I.K., S. Nag, and S.K. Bandyopadhyay, *Technique for preprocessing of digital mammogram.* Computer methods and programs in biomedicine, 2012. 107(2): p. 175-188.
14. Wang, Q., *HMRF-EM-image: implementation of the hidden markov random field model and its expectation-maximization algorithm.* arXiv preprint arXiv:1207.3510, 2012.
15. Singh, V., D.C. Marinescu, and T.S. Baker, *Image segmentation for automatic particle identification in electron micrographs based on hidden Markov random field models and expectation maximization.* Journal of Structural Biology, 2004. 145(1-2): p. 123-141.

16. Li, X.-Z., S. Williams, and M.J. Bottema, *Background intensity independent texture features for assessing breast cancer risk in screening mammograms.* Pattern Recognition Letters, 2013. 34(9): p. 1053-1062.
17. Haralick, R.M. and K. Shanmugam, *Textural features for image classification.* IEEE Transactions on systems, man, and cybernetics, 1973(6): p. 610-621.
18. Timofte, R., T. Tuytelaars, and L. Van Gool. *Naive bayes image classification: beyond nearest neighbors*. in *Asian Conference on Computer Vision*. 2012: Springer.
19. McCann, S. and D.G. Lowe. *Local naive bayes nearest neighbor for image classification*. in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. 2012: IEEE.
20. Rose, C., Daniele Turi, and Alan Williams. *Digital database for screening mammography*. Available from: http://marathon.csee.usf.edu/Mammography/Database.html.
21. Pereira, D.C., R.P. Ramos, and M.Z. Do Nascimento, *Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm.* Computer methods and programs in biomedicine, 2014. 114(1): p. 88-101.
22. Verma, B., P. McLeod, and A. Klevansky, *Classification of benign and malignant patterns in digital mammograms for the diagnosis of breast cancer.* Expert Systems with Applications, 2010. 37(4): p. 3344-3351.
23. Rouhi, R., et al., *Benign and malignant breast tumors classification based on region growing and CNN segmentation.* Expert Systems with Applications, 2015. 42(3): p. 990-1002.
24. Li, Y., et al., *Texton analysis for mass classification in mammograms.* Pattern Recognition Letters, 2015. 52: p. 87-93.
25. Ergin, S. and O. Kilinc, *A new feature extraction framework based on wavelets for breast cancer diagnosis.* Computers in biology and medicine, 2014. 51: p. 171-182.
26. Zemmal, N., N. Azizi, and M. Sellami. *CAD system for classification of mammographic abnormalities using transductive semi supervised learning algorithm and heterogeneous features*. in *Programming and Systems (ISPS), 2015 12th International Symposium on*. 2015: IEEE.

AUTHORS PROFILE