Breast Cancer Detection Using Mammogram Images with Improved Multi-Fractal Dimension Approach and Feature Fusion

Dilovan Asaad Zebari [1], **Dheyaa Ahmed Ibrahim** [2], **Diyar Qader Zeebaree** [3], **Mazin Abed Mohammed** [4], **Habibollah Haron** [5], **Nechirvan Asaad Zebari** [3], **Robertas Damaševičius** [6,*] and **Rytis Maskeliūnas** [7]

1   Department of Computer Science, College of Science, Nawroz University, Duhok 42001, Iraq; dilovan.majeed@nawroz.edu.krd
2   Communications Engineering Techniques Department, Information Technology Collage, Imam Ja'afar Al-Sadiq University, Baghdad 10064, Iraq; Dheyaa.ibrahim@sadiq.edu.iq
3   Information Technology Department, Duhok Technical Institute, Duhok Polytechnic University, Duhok 42001, Iraq; dqszeebaree@dpu.edu.krd (D.Q.Z.); nechirvan720@gmail.com (N.A.Z.)
4   Information Systems Department, College of Computer Science and Information Technology, University of Anbar, Ramadi 31001, Iraq; mazinalshujeary@uoanbar.edu.iq
5   School of Computing, Faculty of Engineering, University Teknologi Malaysia, Johor Bahru 81310, Malaysia; habib@utm.my
6   Faculty of Applied Mathematics, Silesian University of Technology, 44-100 Gliwice, Poland
7   Department of Multimedia Engineering, Kaunas University of Technology, 51368 Kaunas, Lithuania; rytis.maskeliunas@ktu.lt
*   Correspondence: robertas.damasevicius@polsl.pl

Abstract: Breast cancer detection using mammogram images at an early stage is an important step in disease diagnostics. We propose a new method for the classification of benign or malignant breast cancer from mammogram images. Hybrid thresholding and the machine learning method are used to derive the region of interest (ROI). The derived ROI is then separated into five different blocks. The wavelet transform is applied to suppress noise from each produced block based on BayesShrink soft thresholding by capturing high and low frequencies within different sub-bands. An improved fractal dimension (FD) approach, called multi-FD (M-FD), is proposed to extract multiple features from each denoised block. The number of features extracted is then reduced by a genetic algorithm. Five classifiers are trained and used with the artificial neural network (ANN) to classify the extracted features from each block. Lastly, the fusion process is performed on the results of five blocks to obtain the final decision. The proposed approach is tested and evaluated on four benchmark mammogram image datasets (MIAS, DDSM, INbreast, and BCDR). We present the results of single- and double-dataset evaluations. Only one dataset is used for training and testing in the single-dataset evaluation, whereas two datasets (one for training, and one for testing) are used in the double-dataset evaluation. The experiment results show that the proposed method yields better results on the INbreast dataset in the single-dataset evaluation, whilst better results are obtained on the remaining datasets in the double-dataset evaluation. The proposed approach outperforms other state-of-the-art models on the Mini-MIAS dataset.

Keywords: breast cancer detection; mammogram images; multi-fractal dimension; feature fusion

1. Introduction

Global deaths have been caused by noncommunicable diseases (NCDs). Among NCDs, cancer is a primary contributor to the increase in mortality rate and the main impediment to improving the life span of humans around the world in the 21st century. About 2.1 million women have been diagnosed with breast cancer worldwide since 2018. Breast cancer is fatal, the second most frequently diagnosed disease, and the most common

malignancy that causes deaths in women among all cancers. Mastectomy can be avoided by early detection of cancer [1].

Disease prevention remains a major challenge because the cause of breast cancer is still unknown. However, effectively diagnosing breast cancer at an early stage can increase the possibility of total recovery. Early detection of breast cancer using mammography [2] and other imaging modalities such as ultrasound [3], magnetic resonance imaging (MRI) [4], and thermal images [5] can help reduce the mortality rate and the probability of recurrence due to the early detection of benign and malignant breast cancer masses with the progress of mammogram imaging. However, expert radiologists are still missing a significant proportion of abnormalities in the early stage of cancer. Given many images available for each radiologist, the main problem is the identification of lesions or suspicious regions in X-ray images. In the early detection of breast cancer, mammography is widely used to detect presymptomatic cancer cells, which in turn helps to prevent the advancement of cancer into critical stage through administering suitable treatments [6]. The latest advancements in mammography include breast tomosynthesis, computer-aided detection, and digital mammography [7].

When it comes to radiographic image interpretation, the evaluation and organization of image data are required to arrive at a diagnosis. These procedures can be difficult to perform due to the variability of breast parenchyma and structural noise caused by the masking of dense tissue [7]. These elements can cause cancer lesions to be buried or ignored. Manual diagnostic involves multiple subjective decisions with increased variations between and within observers that can lead to serious errors and health implications [8]. Clinical professionals have strongly argued for intolerance toward 'false alarms' [9]. Therefore, the development of a computer-aided diagnosis (CAD) system as a decision support tool adds value in terms of helping the radiologist to reduce false-positive and false-negative cases. However, the radiologist experience is still the guide and the final decision maker because ground truth for the training stage is reflected in the radiologist labels [10,11].

Furthermore, manually interpreting medical images in large volumes consumes a significant amount of time, is a monotonous process, and is vulnerable to mistakes and biases due to the nature of human judgments. The workload of radiologists can be reduced with computerized analysis supported by artificial intelligence techniques, which can help detect tumors and malignant tissues [12]. The structure and characteristics of breast abnormalities make the detection of abnormalities challenging. Consequently, CAD systems were developed to identify breast malignancies and assist medical professionals in the efficient interpretation of medical images with increased accuracy and speed [13,14]. The role of CAD systems is to solve the challenge of interpreting mammogram images for the effective diagnosis of cancer [15,16]. The confidence level of the final decision for the detection of breast cancer is a significant advantage of CAD systems when expert radiologists use the CAD as a second opinion in the classification of breast cancer, as well as to improve the sensitivity of the diagnosis [17].

CAD systems generally perform image preprocessing, image segmentation, feature extraction, feature selection, and classification. The two main modules for detecting breast mass in the CAD system are detection of regions of interest (ROIs) and suspicious region identification based on segmentation and classification modules, which classify identified ROIs into benign or malignant categories [18,19]. A major phase that can strongly impact the classification rate is feature extraction [20,21]. Thus, this study aims to develop an efficient CAD system to detect breast cancer masses in mammogram images. Relevant features must be extracted effectively to enhance the efficiency and precision of CAD systems. Additionally, the success of accurate mammogram image retrieval is based on the extracted features. Parameters such as texture, shape, and color are typically used to characterize images. However, mammogram images contain useless regions, such as dark background and bright interesting regions with various colorless shapes located in the foreground. Therefore, texture is a convenient descriptor in mammogram images for the detection of masses. Consequently, many studies have used texture characteristics, such as

the gray aurora matrix, Gabor characteristics, gray-level cooccurrence matrices (GLCMs), gray-level run-length matrices (GLRLMs), grey-level difference matrices (GLDMs), and gray neighbor matrix, for retrieval [22].

The following are the main contributions of this study:

- Wavelet transform (WT) based on BayesShrink soft thresholding is applied to suppress noise from each produced block by capturing high-frequency information and processing it to reduce noise that might decrease accuracy.
- A new mechanism is used to produce multi-fractal dimension, called multi-FD technique. The traditional FD method has been enhanced by a new thresholding technique to produce features that can help identify the risk of malignant cancer.
- Different threshold values are produced and then placed on each block to create different binary images to extract multi-FD features from each block. To compute the values of thresholds that are distributed uniformly and to determine the range of grayscale, the interval's ROI intensity is calculated. Based on that, sensitive threshold values are produced to obtain different binary images.
- A fusion technique is used to fuse different classifiers trained to classify the features of each block. The fusion process uses the majority vote for decision-making.

The rest of the organization of this paper is as follows: Section 2 presents related work of the study. Section 3 presents the proposed M-FD model and is divided into several subsections. In Sections 4 and 5, the experimental results are discussed in detail with significance and the limitations of the study are presented, respectively. Lastly, the conclusions and future work of this study are presented in Section 6.

2. Related Work

Feature extraction is crucial in the CAD system, and a variety of methods are introduced to distinguish between malignant and benign masses. Some studies have used shape features for mass classification from mammogram images such as mean intensity, perimeter, diameter, centroid, and area [23]. The neural network (NN) with many hidden nodes and additional features that are fed to the classifier obtained high accuracy for only 40 Mini-MIAS images. A mammogram classification using the Hough transform to assess the parameters of the shape from its boundary points was proposed in [24]. It is used to insulate features of specific shapes in mammogram images. This work has been evaluated on only 95 mammogram images based on the support vector machine (SVM) classifier.

Some studies have reviewed the contributions of texture to the risk assessment for each density. Patches of mammogram images consist of directionally oriented ligaments, ducts, texture features of the image due to its fibroglandular tissues, and blood vessels. Patches of texture features based on mammogram images can be classified into statistical [25], local pattern histogram [26], directional [27], and transform-based [28]. Statistical features, such as entropy, mean, variance, kurtosis, contrast, correlation, energy, coarseness, standard deviation, direction, linearity, regularity, skewness, and roughness, can be derived from five sets of statistical texture features, namely GLCM, GLRLM, Tamura features, first-order statistics (FOS), and GLDM. The dynamic cross-propagation algorithm (DCPA) for mammogram classification was introduced in [29]. Adaptive median filter was used for noise reduction, and superior ROI (SROI) was used for ROI segmentation. Some statistical features, such as mean, entropy, standard deviation, variance, correlation, skewness, kurtosis, homogeneity, contrast, and entropy value, were extracted from the ROI and used to discover the loss function in the image micro and then compared with the training model. This work has been tested on DDSM datasets.

Patil and Biradar [30] utilized an optimized hybrid classifier based on a convolutional neural network (CNN) and a recurrent neural network (RNN). Median filter is used to reduce noise; then, region growing based on a hybrid metaheuristic technique is used for tumor segmentation. The features of GLCM and GRLM are extracted from the segmented tumor. The statistical features extracted are used as input to RNN and the segmented binary image is used as input to CNN. The work obtained higher results than conventional

methods when the AND operation of the output of two classifiers yields the overall diagnostic accuracy. The study [31] reinforces the statistical characteristics of GLCM and GLRLM to improve the differentiation of the masses with the help of super-resolution reconstruction. This study obtained an accuracy of 96.7% on the MIAS dataset. The classification of breast density was proposed in [32]. Breast glandular tissue was segmented by performing a new threshold technique. Mean, skewness, and kurtosis features were obtained by implementing the GLCM texture feature. SVM and extreme learning machine (ELM) were used to classify mixed feature vectors. SVM obtained an accuracy of 96.19% and 96.35% for the MIAS and DDSM databases, respectively.

Some studies have exploited the local distribution of textural features for mammogram classification, including histogram of gradients and local configure pattern (LCP) [33], local ternary pattern (LTP) [34], and local quinary patterns (LQP) [35]. The study of [36] proposes a local photometric attribute (LPA) based on a local texture feature to classify mammograms into benign and malignant. Local information has been measured through the optical density using LPA; since the region of the background has been suppressed, it can help in providing mass lesion details. The performance of this method has been evaluated based on receiver operating characteristic (ROC) and accuracy. For Mini-MIAS dataset it obtained 0.94 and 86.90%, whereas for DDSM it obtained 0.89 and 80.76%, respectively.

Ghasemzadeh et al. [37] present an effective technique for mammogram classification by exploiting texture feature-based directional transform. The feature vector is obtained by using the Gabor wavelet transform. Machine learning techniques are used in the decision-making stage. Finally, the performance was tested using the DDSM database and obtained 0.939, 0.951, and 0.92 in terms of accuracy, sensitivity, and specificity, respectively. Haar wavelet decompositions are used to extract texture features by [38]. Points and corners of interest are detected using the speeded-up robust feature (SURF) and minimum eigenvalue algorithm (MinEigenAlg). The subset features of the selected subset using a filter and embedded methods as feature selection techniques are used to train a random forest (RF) binary classifier. This study was tested using 260 images from the BCDR database; it obtained an accuracy of 97.31% and 88.46% for normal/abnormal and benign/malignant, respectively. Eventually, the discrete cosine transform (DCT) is used to extract texture characteristics [28].

Some studies have used statistical and directional transform-based texture feature extraction. Discrete WT (DWT) and GLCM were used for the feature extraction-based fusion method [39,40]. SURF, Gabor filter, and GLCM were used for feature extraction. ELM was proposed for the identification of microcalcification (MC) [41]. Furthermore, some studies used shape and texture characteristics for mammogram classification. In [42], shape features, FOS, and GLCM of statistical texture features with some intensity-based features are extracted, obtaining 91.42% accuracy for benign or suspicious classification.

Sheba and Gladston Raj [43] extracted shape, texture, and gray level features from mammogram images from the MIAS dataset, and optimal features were selected using the regression tree (CART). With feed-forward NNs, the classification was carried out and an accuracy of 96% was obtained. The varied regions of masses were analyzed using concave geometry (alpha shapes) and geostatistical methods [44]. SVM was applied to evaluate each feature extraction, and the study produced a detection rate of 97.30% and 91.63% in MIAS and DDSM datasets, respectively. Several characteristics, namely intensity, margin, texture, and shape characteristics, were used with SVM to classify breast cancer into benign or malignant categories [45]. Researchers obtained an accuracy of 91.37% and 93.22% in the MIAS and DDSM datasets, respectively. Many studies have exploited the integration between different features of shape and texture to improve the CAD system for the classification of breast cancer [46–48].

The existing works have some limitations which are used as challenges in this work. Lesion classification in breast cancer detection is considered a key challenge. Mammogram images influenced by noise can lead to the provision of imprecise data, which reduces the performance of the method. Reduced performance is considered an important issue in

classification. Mammogram images comprise complex textures, which are a substantial problem in lesion detection. Failure to preserve complete information and obtain satisfactory performance in terms of accuracy for incomplete data is a very serious problem that can result in inappropriate decision-making. Overall, the literature showed that result interpretation is still based on heuristics and various databases with a small set of images for training and testing stages and validation. In addition, performance factors, including feature extraction and classification, have influenced the results of various methods. These limitations indicate that improvements can still be achieved.

The literature review showed that many feature extraction techniques have been proposed. However, the fractal dimension method for extraction features has not been sufficiently evaluated and tested to classify mammogram images. Thus, a new texture feature extraction technique is proposed in this study and the results are compared with those of previous studies under the same dataset.

3. Proposed Methodology

3.1. Outline of Proposed Methodology

Digital images contain a massive amount of data, and their interpretation requires efficient image processing and analysis solutions. In this model, an effective and reliable extraction of significant features, which are relevant medical information, is necessary for automatic diagnosis. Reliable and effective feature extraction is dependent on image quality. However, image quality cannot be ensured in the capture phase. Thus, enhancing the quality of the image is a fundamental task. Analyzing the extracted features is a major diagnostic goal that needs advanced classification tools and machine learning methods.

This section presents the proposed model for feature extraction and classification tools in detail. The segmentation task has been performed successfully based on the thresholding value and machine learning techniques. The wavelet-based enhancement feature extraction model (EWBFEM) is then applied after ROI extraction to remove the noise before feature extraction. The following stage extracts relevant features from enhanced ROIs. Additional features extracted from ROIs for automatic diagnosis differ from features manually extracted by expert radiologists. This work concentrates on texture features rather than measurement features. This study also uses the genetic algorithm (GA) to select relevant features. Lastly, the extracted features are fed into a classifier to obtain a diagnostic score/decision. As shown in Figure 1, the focus is on producing powerful features that can help identify the risk of breast cancer at an early stage.

3.2. Image Segmentation

Mammogram images suffer from several unwanted regions, including background, pectoral muscle, and different image noises and artifacts, which cause poor image quality. The undesirable presence of unwanted regions, noises, and artifacts can adversely affect textural features. Moreover, mammogram images do not provide optimal results in the case of extracting texture features and classification. Thus, the segmentation task is important because it limits abnormalities to the relevant breast region without interference from the background of the image. The cancerous region or ROI should be extracted from the image to obtain better classification results.

In this study, two phases were performed for the extraction of ROI from mammogram images. We used a threshold technique to estimate the background of the image, including artifacts [49]. The threshold method used various features, including entropy, mean, and median, which are calculated based on subregions for each pixel according to the local information. We obtained the binary image from the grayscale image using this technique. Masking was used to retrieve original pixel values of the breast region and help select the cancer area, as well as the pervasion out of the cancer cells.
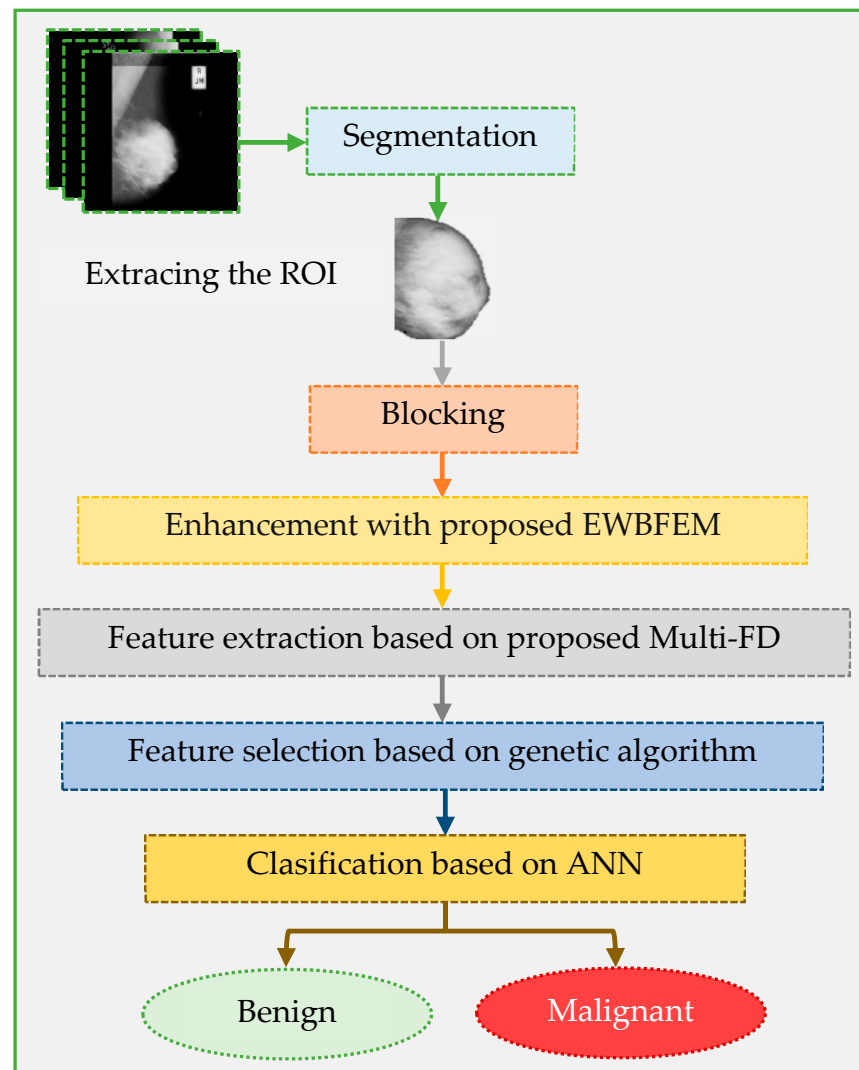
Figure 1. Proposed automated breast cancer recognition model.

The multi-segmentation stage is used to extract the whole ROI and pectoral muscle. The first stage depends on the threshold method to extract the whole ROI. However, a remaining area, called the pectoral muscle, can affect the next stage of CAD processing. This stage is difficult due to the overlap between the ROI and the unwanted ROI (pectoral muscles). Then, the second stage involves a trainable model based on histogram of oriented gradients (HOG) texture features and the NN classifier used to isolate and extract the ROI. In this stage, we have built the trainable model by splitting the data into testing and training. The single split testing strategy has been used by randomly selecting 35% from each class (for each dataset) to form a testing sample, whilst the remaining 65% of each class form a training sample. In the training stage, we have selected blocks from the pectoral muscle out of the pectoral muscle. The HOG features were extracted for each training sample and used as input vector with the labels (pectoral muscle or not) in the NN classifier. In the testing stage, each region extracted from the initial segmentation has been divided into blocks, and the HOG features extracted from each block have been used to classify it as pectoral muscle or not. The cropping technique is then applied to datasets with failed segmentation to derive the ROI for further processing. The cropping technique is dependent on the center and radius values, which are determined by the expert radiologists in the dataset. ROI extraction in abnormal cases is performed depending on the values of the center and radius, whereas in normal cases it is performed in randomly selected locations [50,51].

3.3. Image Enhancement

The visual signal distortion of mammogram images is considered a major problem of diagnosis due to the coherent nature of the transferred wave. This distortion is called noise. Images can be affected by either adaptive or multiplicative noise [46]:

$$Additive\ noise: w(x,\ y) = s(x,\ y) + n(x,y) \tag{1}$$

$$Multiplicative\ noise: w(x,\ y) = s(x,\ y) \times n(x,y) \tag{2}$$

where the original image is represented by $s(x,y)$, and noise introduced in the image is indicated by $n(x,y)$, resulting in the corrupted image $w(x,y)$. The position of pixels is represented by (x,y). The concept of denoising images is shown in Figure 2. The denoising technique has been applied to form the denoised image $z(x,y)$.
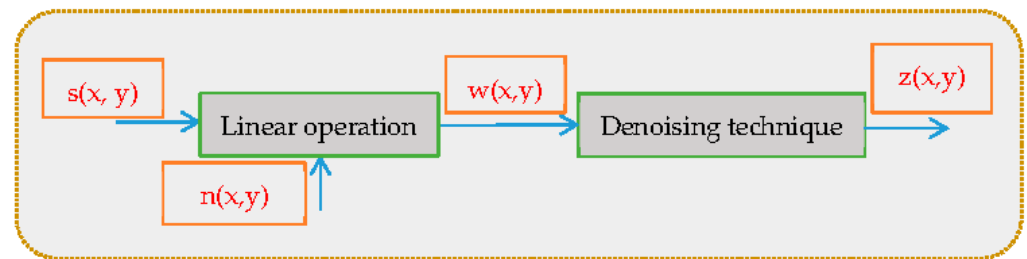

Figure 2. The concept of noise and denoising process.

Random noise that affects mammogram images during data acquisition can lead to corrupted images. Thus, image processing filters must be implemented to enhance and eliminate existing noise. In this study, the EWBFEM model is proposed to eliminate noise and enhance mammogram images. EWBFEM can reduce the noise of images by removing high-frequency bands and enhance the mammogram based on the threshold value. The BayesShrink technique is used to minimize the Bayesian risk via soft thresholding, which is a smoothness-adaptive strategy that depends on sub-bands. The thresholding operator in the decomposition of WT is performed at each resolution band. The t_B of the Bayes threshold is defined in Equation (3).

$$t_B = \sigma^2 / \sigma_s^2, \tag{3}$$

where σ^2 represents variance of noise and σ_s^2 represents the signal variance without noise. Using Equation (4), σ_w^2 can be calculated. The signal variance σ_s^2 can be calculated using Equation (5). Using σ^2 and σ_s^2, the Bayes threshold has been calculated as in Equation (5).

$$\sigma_w^2 = \frac{1}{n^2} \sum_{x,y=1}^{n} w^2(x,y), \tag{4}$$

$$\sigma_s = \sqrt{max(\sigma_w^2 - \sigma^2, 0)}, \tag{5}$$

High-frequency bands are thresholded within one level of decomposition. The HH, LH, and HL bands are contained in the high-frequency mammogram component. The EWBFEM uses four different filters under the same threshold value. First, only the HH band is denoised based on the threshold. Second, the HH–LH bands are denoised using the threshold value. Third, the HH–HL band is denoised using the same threshold value. Lastly, all bands containing high frequencies are denoised (LH–HH–HL).

3.4. Fractal Dimension

The complex geometry of A can be characterized using Hausdorff–Besicovitch (HB) number or FD of a bounded set A in R'. The set is a fractal set when the topological dimension is strictly smaller than the HB dimension [52]. FD can be estimated via the

notion of self-similarity. The bounded set A in the Euclidean n-space is self-similar if A is the union of N and distinct copies are scaled up by itself or scaled down by the ratio r. Equation (6) can be used to estimate FD using the differential box counting (DBC) method.

$$D = \lim_{\epsilon \to 0} \frac{log \; log \; N(\epsilon)}{log \; log \; \epsilon^{-1}} \tag{6}$$

Feature extraction is an essential phase in ROI classification, and it can highly influence the rate of classification. Texture is the most appropriate descriptor for mass detection in mammogram images. This study focuses on providing a robust method for feature extraction. Therefore, the fractal dimension feature has been exploited to extract significant features based on the multi-fractal dimension (M-FD).

3.5. Proposed Multi-Fractal Dimension

The segmentation and enhancement stages are followed by extraction of texture features from the mammogram image to identify the risk of mammogram breast cancer at an early stage. The mammogram is divided into five blocks (EB1 to EB5) and each block is used as input to the machine learning model to obtain an accurate diagnosis and improve the features of the fractal dimension. The blocking process divides the image into four equal squares depending on the size of the image. We then crop the fifth square by tapping the center of each square. Blocks are used instead of the entire image to (1) increase the number of fractal dimensions and select the optimal information to identify the risk of malignancy and (2) capture additional specific information and identify the optimal sensitive threshold that can capture as much as edges. Each block is used as input to the enhancement algorithm. A vector of fractal dimension characteristic of each block (FD1 to FD5) is extracted and used as input to the classifier. This process, followed by the fusion model, is applied to fuse the five classifier decisions (ANN1 to ANN5) and obtain a single final decision. Blocks are primarily used to identify specific information from different areas in the image and arrive at a decision based on this information. The final decision remains unaffected, despite having two blocks in the image with the wrong decision. This idea helps to avoid areas that include nonimportant information.

The proposed fusion model is illustrated in Figure 3. Decision level fusion is used to fuse the decision of the classifiers into a final decision by combining the classification decision outcomes generated by different classifiers. Our proposed method uses the idea of voting to identify the final decision. That is, our final decision indicates benign if the tumor reaches three or more benign votes; otherwise, the final decision is malignant.
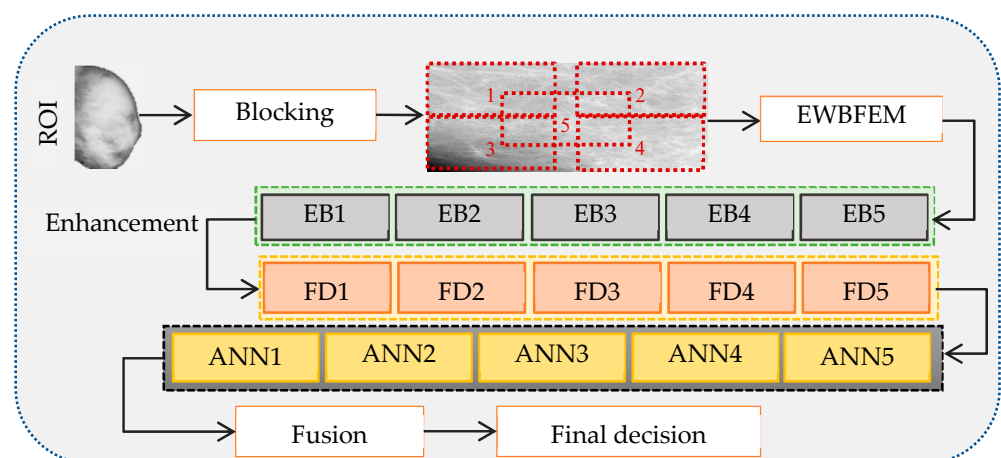

Figure 3. Multi-fractal fusion model based on blocks.

3.5.1. Training Stage

In this stage, the derived ROI is grouped into five blocks and the noise from each generated block is removed using EWBFEM. The features are extracted from each enhanced block after grouping and enhancing blocks from mammogram images. The first block is selected, and binary images are produced from the selected block based on thresholding values (Figure 4). Traditional histogram is used to compute the threshold value.
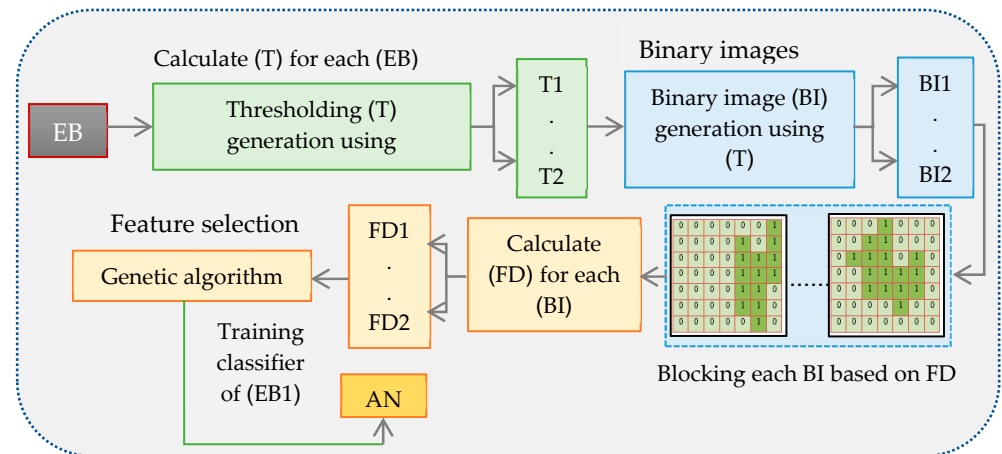
Figure 4. Training stage of feature extraction for each enhanced block.

The first technique can display pixel numbers for each gray value, which is within the scope of the image grayscale. The image histogram demonstrates the number of pixels for each gray value in the image range of 0 to 255. Common gray values within any image are represented by crests within the histogram that are commonly composed of nearly uniform regions. A low number of common values in the histogram is represented by valleys. The vacant districts inside the histogram show that the gray values are absent from the pixels inside the image containing these unfilled locales. Transformation from grey to binary images is possible when thresholding techniques are used. The threshold value can be determined in any threshold technique. The interval intensity amongst the proposed threshold values can be computed to determine uniformly distributed threshold values. The following algorithm demonstrates the binarization steps of the multi-thresholding image that uses the histogram technique.

The threshold value was used in Algorithm 1 to generate a binary image. The histogram presents the range of greyscale values, and the algorithm generates the number of thresholds using the interval value. Equation (7) is used to measure the interval S.

$$S = Max - Min/(N+1) \tag{7}$$

The S value will be calculated using maximum and minimum values of the color image in the first step. The threshold value will be calculated using the threshold counter (R) ($S = 6$) and minimum values of the color image at the second step. The five thresholds are 16, 32, 48, 64, and 80.

Algorithm 1 Calculate the threshold value:

Step 1: Calculate the intensity interval (S) between the intensity thresholds.
Step 2: Calculate thresholds
*   For R = 1 to N*
*   T (R) = Min + S · R*
*   End*

Lastly, the selected block from the ROI was transferred to binary images. After converting the image into a binary image with values 0 and 1, regions with the pixel value

of 1 correspond to regions that contain the information of possible masses. However, normal breast tissue corresponds to regions with the pixel value of 0. Twenty-five different binary images were produced from the selected block after producing 25 thresholds and then applying them to the selected block. Chosen object pixels were acquired depending on binary images. The binary image will be divided into blocks based on FD. FD of the existing object in the binary image is calculated using Equation (6). Box-counting algorithm can be used to assess FD when the binary image is used as the object representation. The first step is to divide the image into a grid containing squares with the size $\epsilon \times \epsilon$. The number of squares of size $\epsilon \times \epsilon$ N(ϵ) comprising at least one object pixel is then determined. The last step is plotting ε versus N on a log–log plot, whereby least squares fitting is applied to get the straight line. The slant of this line is the FD fractal measurement.

The proposed method can effectively obtain the M-FD based on various binary images. As illustrated in Figure 4, ROI was divided into five blocks. Twenty-five different threshold values are produced for each block based on Algorithm 1. Twenty-five images are generated for each block, depending on the produced threshold values. The box-counting technique is used for calculating the FD for each binary image. Based on this, 25 FDs will be produced for each block. GA is then used as the feature selection technique to reduce the number of FDs to 10. Selected FDs are fed to the ANN classifier to obtain the final decision for the first block.

3.5.2. Testing Stage

Figure 5 illustrates the steps required in this stage. The segmentation step is the same method presented in the training stage to derive the ROI. Five different blocks are generated from the ROI before noise removal. A model based on the blocking technique is built to (1) obtain specific and local information from the image; (2) acquire a binary image for FD using the threshold technique, and thus use blocks to determine the suitable threshold for extracting effective texture structures; and (3) build classifiers by using information from different regions of the image. The final decision depends on all trained classifiers.
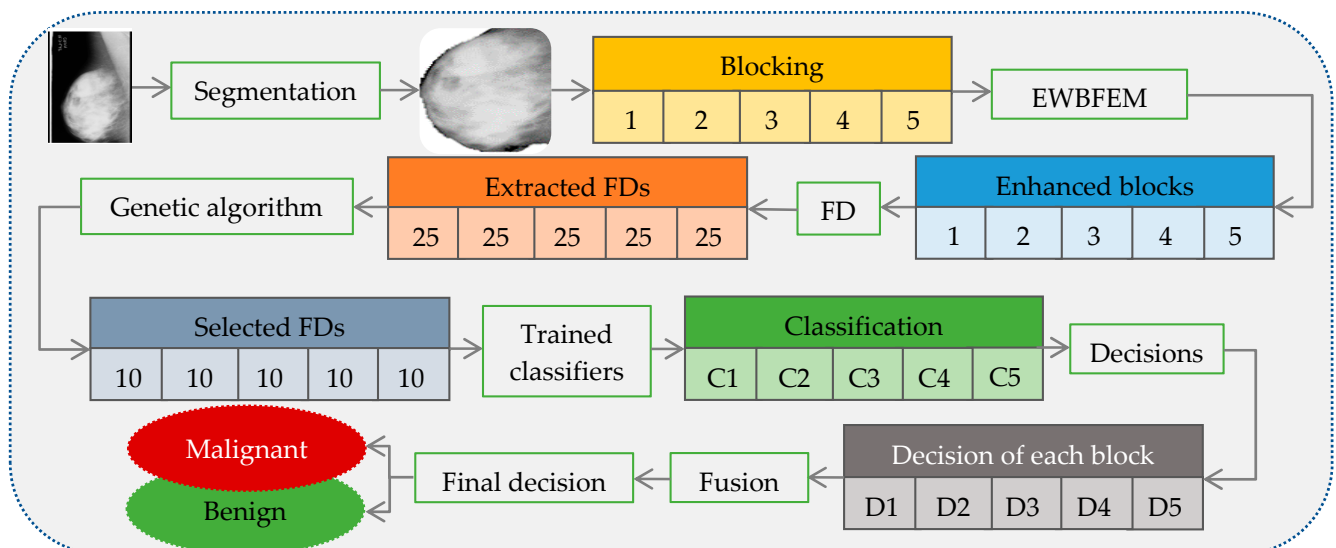
Figure 5. Testing stage of the proposed models.

The proposed EWBFEM is applied to produce blocks in the enhancement step. This step, which plays a crucial role in extracting powerful features, is performed separately on each block to remove noise prior to feature extraction. Texture features are extracted from each block based on FD after grouping and enhancing each ROI block in the feature extraction step. This step extracts 25 texture features from each block (total features = 125) according to threshold value generation, binary image generation, and FD calculation for

each binary image step in the training stage. The genetic algorithm is used as the feature selection method in the next step to reduce the number of features and achieve the two objectives of accelerating the process and reducing the overlapping between classes.

The model created in the training stage can classify breast cancer into benign or malignant cases in the classification step. The ANN classifier is trained into five classifiers; each block contains its own classifier. Classifiers 1 to 5 are trained for blocks 1 to 5 and are denoted C1 to C5, respectively. Each block is tested based on its classifier, and each classifier produces a binary decision of 0 or 1 in the testing stage. D1 to D5 represent the first to fifth binary decisions, respectively.

The fusion process is conducted based on produced binary decisions to decide whether the selected ROI is malignant or benign, as shown in Figure 5. The fusion stage can determine the breast cancer class of the ROI based on the results of five blocks as the final decision after testing all blocks. The confidence level that links diagnosis decisions is reflected in the evaluation of the risk of malignancy. Improvement in the confidence level of diagnosis through the fusion process leads to increasingly accurate diagnosis decisions. Occasional conflicts that occur between classifiers can lead to uncertain cases based on the fusion process; hence, such cases require further investigation. The majority vote fusion technique is used to obtain the final decision considering the five decisions of each ROI. The final decision for each ROI is subsequently determined using majority voting. Five different decision-based trained classifiers are fused to obtain a single decision at the final decision level.

3.6. Feature Selection

Genetic algorithm (GA) is a search technique that depends on the basics of genetics and natural selection. GA is considered one of the most effective and efficient global optimization techniques. Optimization problems can be solved by means of the GA, by using the object function $f(x)$ where N-dimensional vector of parameters of optimization is indicated by $x = [x_1, x_2, \ldots, x_n]$. The GA in contrast to the traditional searching technique is based on the population of candidate solutions. The size of the population is a parameter that should be selected by the user and affects the scalability and performance of the GA.

In this study, GA is used as a feature selection method to reduce the number of features extracted from the ROI of mammogram images. The main purpose of the process is to reduce the overlapping between cancer classes that are benign and malignant, as well as to speed up the classification process. In this study, from each block, 25 features were extracted, for a total of 125 features, and then we used GA to select the most relevant 10 features from each block; as a result, the most relevant 50 features were selected.

3.7. Classification

After extracting the features from the ROIs of mammograms, these features are utilized as input or fed to the classifier to classify the ROI as benign or malignant. This procedure is considered as data mapping to predefined classes. In this study, the ANN classifier is used as it can handle high-dimensional feature vectors. The back-propagation ANN has one input layer, two hidden layers (include 10 nodes for each layer), and one output layer. The ANN learning rate is determined as 0.01, using Trainbr (Bayesian regularization backpropagation) and Sigmoid as training and activation functions, respectively. The training stage is a mapping between the input and the output values. Using a learning algorithm (e.g., generalized delta rule), the mapping procedure is accomplished by adjusting the weight value. The error for the training phase is defined as follows:

$$E_p = \frac{1}{2} \sum_j \left(t_{pj} - o_{pj} \right)^2, \tag{8}$$

where the index p corresponds to single input vector, and the vectors t_p and o_p are the target and observed output vectors, respectively, according to the p.

The error function has been used to calculate the difference between the observed and the target vectors. The error can be minimized by changing the weight w. To minimize that error, a gradient descent in E has been implemented by the generalized delta rule in Equation (9), which is written as Equation (10):

$$\Delta_p w_{ji} \propto -\frac{\partial E_p}{\partial w_{ji}},\tag{9}$$

$$\Delta_p w_{ji} = \eta \delta_{pj} o_{pi},\tag{10}$$

The output of node j due to input p is $o_{pj} = f_j\left(net_{pj}\right)$ with activation function f. For a sigmoid activation function, $f(z) = 1/(1 - e^{-x})$. Then, δ_{pj} is rewritten as Equation (12):

$$\delta_{pj} = \begin{cases} \left(t_{pj} - o_{pj}\right) f'\left(net_{pj}\right) & \textit{for an output node} \\ f'\left(net_{pj}\right) \sum_k \delta_{pk} w_{kj} & \textit{for a hidden node} \end{cases}\tag{11}$$

$$\delta_{pj} = \begin{cases} \left(t_{pj} - o_{pj}\right) o_{pj}\left(1 - o_{pj}\right) & \textit{for an output node } o_{pj} \\ \left(1 - o_{pj}\right) \sum_k \delta_{pk} w_{kj} & \textit{for a hidden node} \end{cases}\tag{12}$$

where $net_{pj} = \sum_k w_{ji} o_{pi} + \theta_j$ is the total input to node j including a bias term θ_j and the parameter η is the learning rate.

Finally, additional momentum is added to the learning equation, resulting in (13).

$$\delta_{pj} = \begin{cases} \left(t_{pj} - o_{pj}\right) f'\left(net_{pj}\right) & \textit{for an output node} \\ f'\left(net_{pj}\right) \sum_k \delta_{pk} w_{kj} & \textit{for a hidden node} \end{cases},\tag{13}$$

$$w_{ji}[n+1] = w_{ji}[n] + \Delta_p w_{ji}\,[n],\tag{14}$$

where μ is the momentum rate at each iteration, and the weights are modified as in Equation (14). After the ANN is adjusted with specific weight values in the training phase, the trained ANN can be used to classify unknown samples.

3.8. Performance Evaluation

The cross-validation strategy is carried out to improve the robustness and validity of our performance evaluation and determine the result of each dataset. The performance metrics, namely sensitivity (Sn), specificity (Sp), accuracy (AC), and F-measure, are used to evaluate the classification efficacy of this study.

The sensitivity (true positive rate) measure (Equation (15)) is the proportion of the number of correct powerful texture signs for classification, which enhances and segments breast cancer mammogram images found in the dataset and heavily relies on false negatives. The ratio of identified positive cases to actual positive cases increases and the rate of false negatives decreases when sensitivity increases. The specificity (true negative rate) measure (Equation (16)) denotes the proportion of the number of correct powerful texture signs for classification that enhances and segments breast cancer mammogram images to the total number of real images found in the dataset. Accuracy (classification rate) (Equation (17)) denotes the correctness of the proposed detection method. Accuracy indicates the correctness of the predicted evaluation results. Finally, the F-measure is calculated using Equation (18).

$$Sensitivity = \frac{TP}{TP + FN} \times 100\%\tag{15}$$

$$Specificity = \frac{TN}{TN + FP} \times 100\%\tag{16}$$

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \times 100\%\tag{17}$$

$$F - measure = \frac{2TP}{2TP + FN + FP} \times 100\% \tag{18}$$

Here, true positives (TP) are correctly identified disease cases, false positives (FP) are incorrectly recognized disease cases, true negatives (TN) are correctly recognized healthy cases, and false negatives (FN) are incorrectly recognized healthy cases.

A receiver operating characteristic (ROC) graph is used as a representation of the comparative trade-off between costs (false positives) and benefits (true positives). A point on the ROC curve represents sensitivity and specificity corresponding to a decision. The following equations show the way of calculating the FP rate and the TP rate:

$$FP\ Rate = FP/(FP + TN) \tag{19}$$

$$TP\ Rate = TP/(TP + FN) \tag{20}$$

4. Experimental Results

4.1. Image Acquisition

In this study, two datasets are used to compare the proposed approach with existing methods. Using images from the mammogram database obtained from a sensible population is necessary in the evaluation of the mammogram classification performance approaches. The testing of such schemes is carried out based on specific protocols that guide the division of the mammogram dataset into testing and training sets for classification. Several mammogram databases can be used to compare results during experiments. In this study four widely used and publicly available mammogram databases, namely the Mini Mammographic Image Analysis Society (Mini-MIAS) [53], the Digital Database for Mammography Screening (DDSM) [54], INbreast [55] and the Breast Cancer Digital Repository (BCDR) [56], are used.

The Mini-MIAS database is made up of 322 mammograms and is freely available to the public for scientific research. The 322 mammogram images were collected from 161 pairs of mediolateral oblique (MLO) views from both right and left views. These mammogram images were obtained from a film screen imaging process conducted by a national breast screening program in the UK. This database contains two major categories, abnormal and normal mammograms. It contains 207 benign and 115 malignant mammograms; the total is equal to 322. However, depending on both MLO and craniocaudal (CC) mammogram views, the DDSM contains 2620 cases. It contains normal, benign, and malignant categories with proven pathological information. The mammogram images in this dataset were obtained from different sources, including the Washington University of St. Louis School of Medicine and the Wake Forest University School of Medicine. Figure 6 shows samples from both databases.
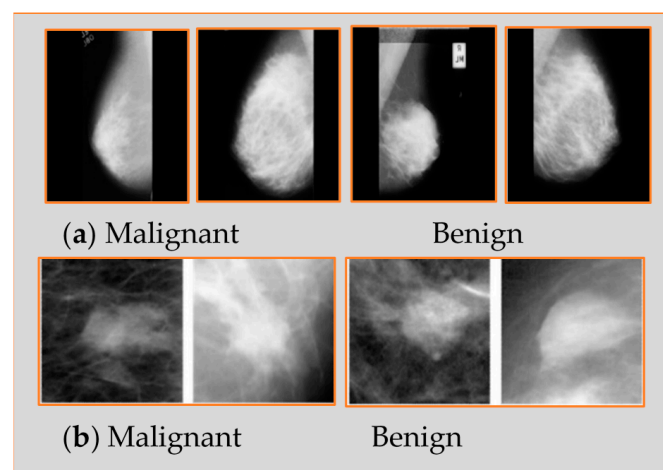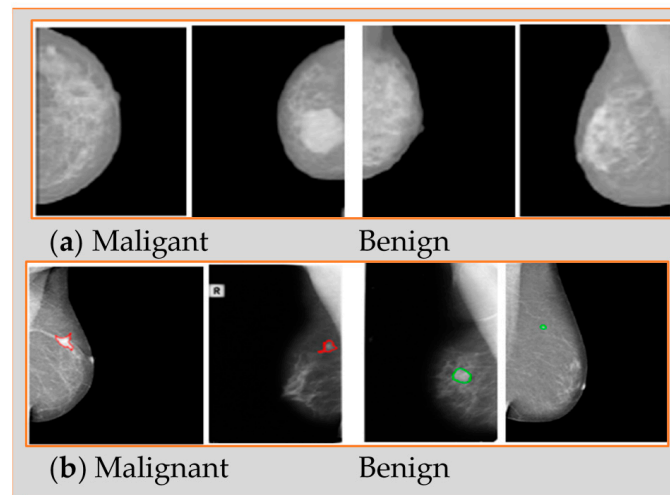

Figure 6. Mammogram samples: (**a**) Mini-MIAS database [49]; (**b**) DDSM database [50].

The INbreast dataset has images from 115 patients (cases). From each of 90 patients, 4 breast images of both breasts (right and left) were collected, while 50 breast images were collected from 25 mastectomy patients. Therefore, 410 normal, benign, and malignant cases of mammogram images were collected, including MLO and CC views.

The Breast Cancer Digital Repository (BCDR) database contains 1125 images from both views MLO and CC totaling 3703 mammogram images with 720×1168-pixel resolution. Figure 7 shows samples from both databases.

Figure 7. Mammogram samples: (**a**) INbreast database [51]; (**b**) BCDR database [52].

This study takes digitalized mammogram images collected from Mini-MIAS, DDSM, INbreast, and BCDR datasets. A total of 316 and 981 mammogram images are collected from both the Mini-MIAS and DDSM databases, respectively. Mini-MIAS has 206 benign and 110 malignant whereas DDSM has 479 benign and 502 malignant. INbreast has a total of 200 more mammogram images, including 127 benign and 73 malignant. Finally, 736 more images are taken from the BCDR database, which include 426 benign and 310 malignant. As a result, a total of 2233 breast cancer mammograms are collected from different databases, including 1238 benign and 995 malignant, to evaluate and validate the proposed methodology.

The ground truth labels have been collected by experts, and they diagnose those cases (labeled them) using experience as well as using different tests including the mammogram scan. Table 1 shows the distribution of mammogram images from the used datasets. During evaluation, images of each dataset are classified into two sets, namely training and testing sets.

Table 1. Distribution of used benign and malignant mammogram images.

Database	Total No. of Images		Training		Testing	
	Benign	**Malignant**	**Benign**	**Malignant**	**Benign**	**Malignant**
Mini-MIAS	206	110	162	88	44	22
			44	22	162	88
DDSM	479	502	319	335	160	167
			160	167	319	335
INbreast	127	73	101	58	26	15
			26	15	101	58
BCDR	426	310	284	207	142	103
			142	103	284	207

To assess the adequacy of the proposed classification method through identification, in particular, it should be determined whether the automated texture features extracted from the breast cancer case can likewise prompt great characterization results if a breast malignancy case indicates a malignant or benign case. All images have been segmented based on the segmentation method described in Section 3.2, whereas the DDSM dataset has been segmented using the cropping technique. The Mini-MIAS dataset used as an input for the segmentation method resulted in 206 out of 207 images successfully segmented from benign cases and 110 out of 115 images segmented correctly from the malignant cases. More so, INbreast and BCDR are segmented as whole images correctly.

The proposed methodology was validated through a number of different experiments. The experiments were performed with MATLAB (2020b) with a Core-i7 processor, 32 GB RAM, and Windows 10 operating system.

4.2. Segmentation Results

To evaluate the proposed segmentation method, we have used the ROI to specify the important information that can help to identify the abnormal cases. To achieve the evaluation goal, we have investigated the textures of the extracted regions and tested them by using ANN. Moreover, to highlight the effectiveness of the extracted region, we have compared it with the texture features extracted from the whole image.

Figure 8 shows the result of the features of the traditional fractal dimension extracted from complete images and the ROI extracted from segmented images. The comparison of results of the entire image and ROI showed that the segmentation stage is very important in dealing with specific information (ROI) whilst ignoring the rest of the image.
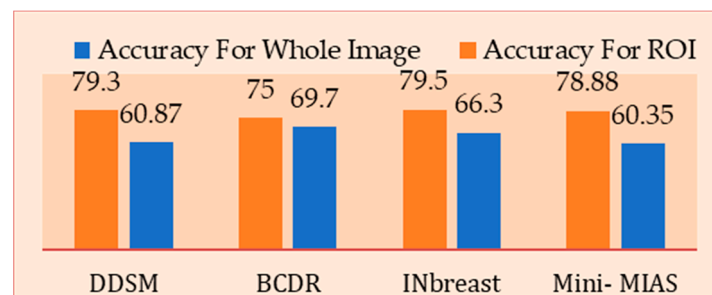
Figure 8. Accuracy of the traditional FD dimension for both segmented and non-segmented images.

4.3. Denoising Results

The proposed models were evaluated and compared by peak signal-to-noise ratio (PSNR), signal-to-noise ratio (SNR), and mean square ratio (MSE) using mammogram images (see an example in Figure 9), as shown in Table 2. The optimal mammogram is generated when the HH, LH, and HL bands are denoised using the threshold.

(a)   (b)
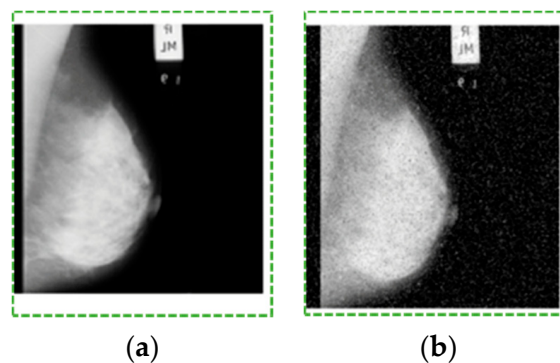
Figure 9. An example of mammogram images: (**a**) ground truth; (**b**) noisy image.

Table 2. Mammogram denoising results based on EWBFEM.

Filters	MSE	PSNR (db)	SNR (db)
HH	0.026	65.99	47.62
HH–HL	0.026	65.13	47.76
HH–HL–LH	0.028	69.95	47.57
HH–LH	0.27	64.32	47.64

4.4. Classification Results

This section presents the classification results of the proposed method using the M-FD technique for extracting features and the ANN classifier for classifying breast cancer mammogram images as benign or malignant. The multilayer perceptron, which was used as the ANN classifier to identify malignant cases from benign ones, is based on the generalized least mean square (LMS) rule that uses the gradient search method to minimize the average difference between the output and the target value of the neural network. Three layers, namely input, hidden, and output layers, were used to build the ANN. The input and hidden layers include the same number of nodes (125 nodes). ANN works to change weights during the training stage, and we reach a minimum error close to 0.1.

We extract 25 features from each ROI block and 125 features from all ROI blocks using the proposed M-FD technique. This finding indicated that all extracted features are not significant for classification. The set of extracted features may present some correlated features because features were extracted from each block and blocks were classified separately. More so, all extracted features do not contribute to feeding the ANN classifier to the overall performance evaluation of the proposed model. Therefore, we initially omit correlated features, use GA as the feature selection technique, and select relevant features from each block. The ANN classifier was trained for each block to classify mammogram cases. Parameter settings of the ANN classifier present a value learning rate, momentum, and number of hidden nodes of 0.05, 0.5, and 8, respectively. We obtained five results for the entire ROI based on the five trained ANN classifiers in this stage. Finally, the fusion process was performed on the obtained results to achieve the final decision.

4.4.1. Performance Evaluation of Single Dataset

Two experiments were conducted to improve the robustness of our performance evaluation. The first experiment is the single-dataset evaluation that uses stratified leave-out-one cross-validation, that is, two images are selected (one from each class) for testing and the remaining images are used for preparing at every cycle of the cross-validation. Separate midpoints of general accuracy, sensitivity, specificity, and F-measure account for the example dataset after performing the forget-about-one cross-approval method. The process is repeated for the entire dataset to limit irregular impacts of the example, and the entire dataset is used for training and testing. This experiment was separately carried out on Mini-MIAS, DDSM, INbreast, and BCDR datasets, thereby indicating that training and testing dataset samples are selected from a specific database.

Table 3 shows the performance analysis of the proposed model using four different datasets. In the literature, it has been discovered that traditional FD and local binary pattern (LBP) features are efficient and strong features used in classification. Therefore, to show the efficacy of the proposed model, traditional FD and LBP are extracted from each dataset and evaluated from the same dataset. The main two reasons behind this step are to show which feature can obtain higher results for mammogram classification as well as observe the efficacy of the proposed M-FD for each dataset. The extracted features are fed to the ANN classifier to classify and discover the cancer subtype as benign or malignant. Furthermore, this section evaluates the proposed model using each database separately, which means that training and testing datasets are taken from the same database.

Table 3. First experiment of mammogram classification results as benign or malignant using different databases.

Database	Feature	Class		No. of Images				Percentage (%)			
		BN	MA	TN	TP	FN	FP	AC	Sn	Sp	F-Measure
Mini-MIAS	LBP			73	177	37	29	79.11	82.71	71.56	84.28
	FD	206	110	76	177	34	29	80.06	83.88	72.38	84.89
	Proposed M-FD			103	201	7	5	96.2	96.63	95.37	97.1
DDSM	LBP			397	380	105	99	79.2	78.35	80.04	78.83
	FD	479	502	401	393	101	86	80.93	79.55	82.34	80.78
	Proposed M-FD			494	473	8	6	98.57	98.33	98.8	98.54
INbreast	LBP			50	115	23	12	82.5	83.33	80.64	86.79
	FD	127	73	52	119	21	8	85.5	85	86.66	89.13
	Proposed M-FD			71	127	2	0	99	98.44	100	99.21
BCDR	LBP			208	329	102	97	72.96	76.33	68.19	76.77
	FD	426	310	211	333	99	93	73.91	77.08	69.4	77.62
	Proposed M-FD			304	416	6	10	97.82	98.57	96.81	98.11

A total of 162 benign and 88 malignant samples are selected for training and 44 benign and 22 malignant samples are taken for testing from the Mini-MIAS dataset. Table 3 presents the different results obtained with different features (LBP, FD, and proposed M-FD) for benign versus malignant cases on four datasets. The results showed that the traditional FD outperforms LBP for each dataset of mammogram images. The M-FD model is then proposed to extract features from mammograms for breast cancer classification. The obtained results demonstrated that the proposed M-FD model outperforms the traditional FD and LBP features on the four datasets. The Mini-MIAS dataset obtained an accuracy, sensitivity, specificity, and F-measure of 96.2%, 96.63%, 95.37%, and 97.1%, respectively. A total of 319 benign and 335 malignant samples from the DDSM dataset are used for training whilst 160 benign and 167 malignant samples are used for testing in the first round of evaluation. Samples used in testing were then used for training, and all 981 trained samples from the first round were used for testing the DDSM dataset. The DDSM dataset obtained an accuracy, sensitivity, specificity, and F-measure of 98.57%, 98.33%, 98.8%, and 98.54%, respectively. A total of 127 benign and 73 malignant samples from the INbreast dataset were used for testing whilst 101 benign and 58 malignant samples were used for training in the first round of evaluation. The testing stage used 26 benign and 15 malignant samples. The testing samples were subsequently used for training the remaining samples from the INbreast dataset. This evaluation obtains an accuracy, sensitivity, specificity, and F-measure of 99%, 98.44%, 100%, and 99.21%, respectively. Lastly, this study evaluates 736 additional samples from the BCDR dataset. A total of 491 (284 benign and 204 malignant) samples are used in the training stage whilst 245 (142 benign and 103 malignant) samples are used for testing. Testing samples are subsequently used in the training stage, and training samples are applied in the testing stage. The classification accuracy, sensitivity, specificity, and F-measure of the BCDR dataset were 97.82%, 98.57%, 96.81%, and 98.11% respectively. The proposed M-FD model obtains an excellent F-measure on the Mini-MIAS dataset, high specificity on DDSM and INbreast datasets, and excellent sensitivity on the BCDR dataset. The evaluation results showed that the proposed M-FD model obtains the best accuracy, sensitivity, and specificity on the INbreast dataset amongst the four datasets.

4.4.2. Performance Evaluation Using Different Datasets

The performance of a single testing process using a double dataset is evaluated in this section. The double dataset (two different databases) was used in a single evaluation. The performance evaluation is measured by taking training samples from one dataset and testing samples from another dataset. The main reason behind this evaluation strategy is to ensure that the proposed M-FD model can work using different datasets. Traditional

FD and LBP features are also used in the double-dataset evaluation. The proposed M-FD model outperformed both traditional FD and LBP features in all evaluations when Mini-MIAS, DDSM, INbreast, and BCDR datasets were used in the training stage, as shown in Tables 4–7, respectively.

Table 4. Performance evaluation obtained on the test data for benign or malignant classification using Mini-MIAS as training data.

Database		Feature	Class		No. of Images				Percentage (%)			
TR	TS		BN	MA	TN	TP	FN	FP	Acc	Sn	Sp	F-Measure
Mini-MIAS	DDSM	LBP	479	502	398	384	104	95	79.71	78.68	80.73	78.9
		FD			405	397	97	82	81.75	80.36	83.16	81.6
		Proposed M-FD			501	479	1	0	99.89	99.79	100	99.89
	INbreast	LBP	127	73	49	113	24	14	81	82.48	77.77	85.6
		FD			50	116	23	11	83	83.45	81.96	87.21
		Proposed M-FD			68	126	5	1	97	96.18	98.55	97.67
	BCDR	LBP	426	310	211	334	99	92	74.04	77.13	69.63	77.76
		FD			215	338	95	88	75.13	78.06	70.95	78.69
		Proposed M-FD			308	425	2	1	99.59	99.53	99.67	99.64

Table 5. Performance evaluation obtained on the test data for benign or malignant classification using DDSM as training data.

Database		Feature	Class		No. of Images				Percentage (%)			
TR	TS		BN	MA	TN	TP	FN	FP	Ac	Sn	Sp	F-Measure
DDSM	Mini-MIAS	LBP	206	110	79	187	31	19	84.17	85.77	80.61	88.2
		FD			84	188	26	18	86.07	87.85	82.35	89.52
		Proposed M-FD			109	204	1	2	99.05	99.51	98.19	99.27
	INbreast	LBP	127	73	50	115	23	12	82.5	83.33	80.64	86.79
		FD			51	116	22	11	83.5	84.05	82.25	87.54
		Proposed M-FD			71	125	2	2	98	98.42	97.26	98.42
	BCDR	LBP	426	310	231	349	79	77	78.8	81.54	75	81.73
		FD			240	352	70	74	80.43	83.41	76.43	83.01
		Proposed M-FD			303	414	7	12	97.41	98.33	96.19	97.75

Table 6. Performance evaluation obtained on the test data for benign or malignant classification using INbreast as training data.

Database		Feature	Class		No. of Images				Percentage (%)			
TR	TS		BN	MA	TN	TP	FN	FP	Ac	Sn	Sp	F-Measure
INbreast	Mini-MIAS	LBP	206	110	84	193	26	13	87.65	88.12	86.59	90.82
		FD			89	196	21	10	90.10	90.32	89.89	92.67
		Proposed M-FD			110	205	0	1	99.68	100	99.09	99.75
	DDSM	LBP	479	502	392	380	110	99	78.69	77.55	79.83	78.43
		FD			396	390	106	89	80.12	78.62	81.64	80
		Proposed M-FD			492	471	10	8	98.16	97.92	98.4	98.12
	BCDR	LBP	426	310	236	356	74	70	80.43	82.79	77.12	83.17
		FD			246	357	64	69	81.92	84.79	78.09	84.29
		Proposed M-FD			307	419	3	7	98.64	99.28	97.77	98.82

Table 7. Performance evaluation obtained on the test data for benign or malignant classification using BCDR as training data.

Database		Feature	Class		No. of Images				Percentage (%)			
TR	TS		BN	MA	TN	TP	FN	FP	Ac	Sn	Sp	F-Measure
BCDR	Mini-MIAS	LBP	206	110	83	194	25	14	83.61	88.58	85.56	90.86
		FD			87	196	23	10	89.55	89.49	89.69	92.23
		Proposed M-FD			110	206	0	0	100	100	100	100
	DDSM	LBP	479	502	394	382	108	93	79.1	77.95	80.9	79.17
		FD			399	391	103	88	80.53	79.14	81.93	80.36
		Proposed M-FD			497	479	5	0	99.49	98.96	100	99.48
	INbreast	LBP	127	73	50	113	23	14	81.5	83.08	78.12	85.93
		FD			50	114	23	13	82	83.21	79.36	86.36
		Proposed M-FD			69	124	4	3	96.5	96.87	95.83	97.25

Table 4 shows the performance analysis for benign or malignant classification of the proposed M-FD method with traditional FD and traditional LBP. In this evaluation, the Mini-MIAS dataset has been used in the training stage whereas other datasets have been used for the testing stage. From Mini-MIAS a total of 250 samples including 162 benign and 88 malignant are taken in the training stage. A total of 981, 200, and 736 samples from DDSM, INbreast, and BCDR, respectively, are taken for testing. It is seen from the results that the DDSM obtained 99.89% for accuracy, 99.79% sensitivity, 100% specificity, and 99.89% F-measure. Moreover, BCDR obtained 99.59%, 99.53%, 99.67%, and 99.64% for accuracy, sensitivity, specificity, and F-measure. Results of DDSM and BCDR outperformed the previous evaluation (single-dataset evaluation) in this evaluation. DDSM and BCDR achieved higher results in terms of specificity. However, the INbreast dataset obtained higher results in the previous evaluation (single-dataset evaluation) by achieving 97% accuracy, 96.18% sensitivity, 98.55% specificity, and 97.67% F-measure.

Table 5 shows performance for benign and malignant classification when DDSM is used as a training dataset. A total of 654 samples from DDSM are taken for training including 319 benign and 335 malignant. However, 316 samples from Mini-MIAS, 200 samples from INbreast, and 736 samples from BCDR are used in the testing stage. It is observed that Mini-MIAS obtained higher results from the single-dataset evaluation, achieving 99.05% accuracy, 99.51% sensitivity, 98.19% specificity, and 99.27% F-measure. Moreover, in this evaluation, BCDR obtained higher results from single-dataset evaluation as well for accuracy 97.41%, sensitivity 98.33%, specificity 96.16%, and F-measure 97.75% whereas the INbreast dataset still outperformed in single-dataset evaluation by obtaining 98%, 98.42%, 97.26%, and 98.42% for accuracy, sensitivity, specificity, and F-measure, respectively.

Table 6 shows that in the INbreast dataset used in the training stage, a total of 159 samples with 101 benign and 58 malignant are taken to train the proposed model. Table 6 shows that 316 samples are taken from Mini-MIAS for testing, the proposed M-FD achieved 99.68% accuracy, 100% sensitivity, 99.09% specificity, and 99.75% F-measure. From the DDSM dataset, 981 samples are taken for testing, and the classification accuracy and F-measure obtained were 98.16% and 98.12%, respectively while DDSM obtained 97.92% and 98.4% classification sensitivity and specificity, respectively. Furthermore, a total of 736 samples are taken from BCDR to test the proposed model. The performance of the classification evaluation achieved 98.64% and 99.28% for accuracy and sensitivity while obtaining 97.77% and 98.82% for specificity and F-measure, respectively.

Finally, this study evaluates the proposed M-FD by taking the BCDR dataset as training data. A total of 491 samples including 284 benign and 207 malignant are taken in the training stage. Table 7 depicts the result achieved with the proposed M-FD on the Mini-MIAS, DDSM, and INbreast datasets. M-FD yields better evaluation performance by achieving 100% accuracy, sensitivity, specificity, and F-measure. The classification accuracies for DDSM and INbreast datasets are 99.49% and 96.5%, respectively. The testing

datasets obtained 98.96% and 96.87% and 100% and 95.83% for sensitivity and specificity, respectively. More so, F-measure obtained was 99.48% for DDSM and 97.25% for INbreast.

Two different experiments were conducted to evaluate the performance of the proposed M-FD method, as illustrated above. Two different traditional methods (FD and LBP) were used for feature extraction with the same datasets. Extracted features based on the traditional FD, traditional LBP, and the proposed M-FD were then fed to the ANN classifier. This evaluation and comparison showed the effectiveness of the proposed method. The results clearly showed that the traditional FD produces a better result than the traditional LBP. We then enhanced the traditional FD, and the results showed that the proposed M-FD model increases the result and improves the confidence of the system.

The results obtained for four datasets are depicted in Tables 3–7. Table 3 shows that higher results are obtained on the INbreast dataset in the single-dataset evaluation (INbreast was used in training and testing) whilst other datasets achieve higher results in the double-dataset evaluation. Table 4 presents that DDSM and BCDR datasets obtain high results when Mini-MIAS is used as the training dataset. Table 6 shows that the Mini-MIAS obtains a sensitivity of 100% when the INbreast dataset is used for training. The proposed M-FD obtained optimal accuracy, sensitivity, and F-measure of 99.68%, 100%, and 99.75%, respectively, for Mini-MIAS when INbreast was used as the training dataset. The best specificity, 100%, was achieved by INbreast in single-dataset evaluation and DDSM when using BCDR in training. Table 7 demonstrates that the DDSM dataset obtained a specificity of 100% when the BCDR dataset was used for training. Table 7 also indicates that the Mini-MIAS dataset achieves optimal results when the BCDR dataset is used in training by obtaining a performance classification of 100% for accuracy, sensitivity, specificity, and F-measure.

We used the ROC plot to evaluate the power of the texture feature (M-FD) and performance when different texture features are used. The ROC curve that visualizes the performance of the classifier is commonly used in medical image decision-making and is increasingly used in malignant identification research. Figure 10 shows the performance of traditional fractal dimension features for the Mini-MIAS dataset. Figure 11 presents the performance of the proposed fractal dimension for the Mini-MIAS dataset. The area under the curve (AUC) and the performance of ANN increased in training, validation, and testing.
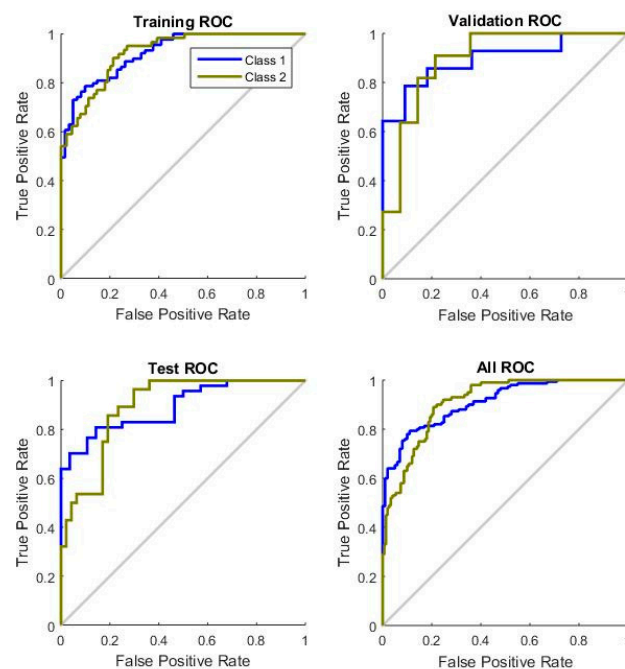
Figure 10. ROC for training, validation, and testing of the proposed model using traditional FD features.
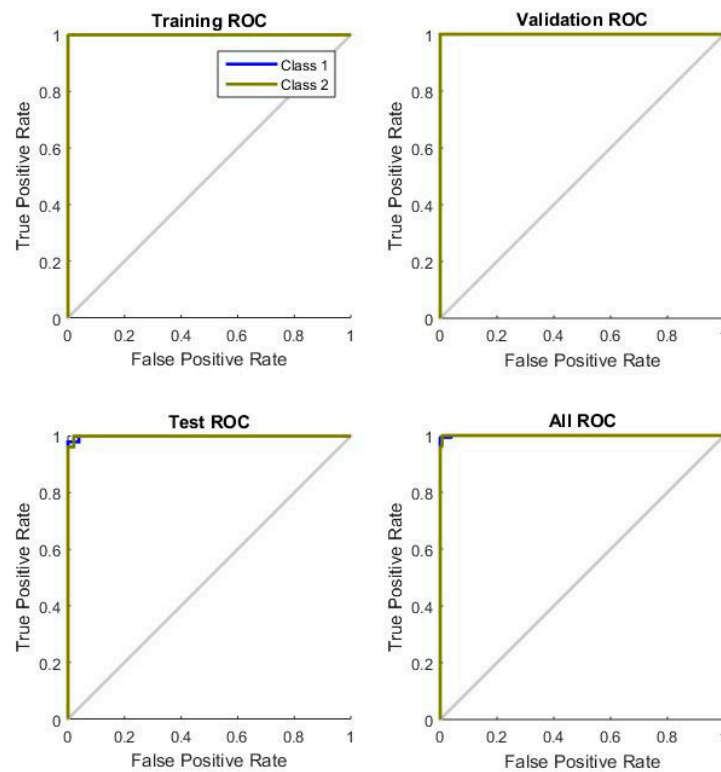
Figure 11. ROC for the training, validation, and testing of the proposed model using new M-FD features.

4.5. Comparision with Other Techniques

This section compares the performance of the proposed method with the models known from the literature. Most of the previous models were evaluated on Mini-MIAS and DDSM datasets, whereas two CAD models have been evaluated on INbreast and BCDR datasets. The average performance values obtained for the literature and proposed models are summarized in Tables 8–10. The proposed model yields better results for performance evaluation in terms of accuracy, sensitivity, specificity, and F-measure when compared to traditional FD and LBP features. Table 8 presents a comparative study between our proposed M-FD + GA + ANN and several studies in the literature for the Mini-MIAS dataset. Our proposed model yields a 100% performance result in terms of all used matrices in the context of classification using 316 images out of 322 from the Mini-MIAS database. The previous study [47] obtained 100% sensitivity, whereas the work of [57] used hybrid LB-GLCM+LPQ texture features, and fewer images were used for the evaluation in both works. Moreover, we also found that the study of [57] gives 100% accuracy where this work used LWT+PCA (32 features), and 119 mammogram images from Mini-MIAS were evaluated. However, our proposed model exceeds all presented results in Table 8.

Table 9 presents a comparative study between our proposed model and previous works for the DDSM dataset. We discovered that the proposed model yields better results compared to the traditional FD and LBP features. The previous study [46] obtained 100% sensitivity, and the previous study [58] gives 100% accuracy in the Mini-MIAS dataset. Table 9 depicts that our proposed model yields better sensitivity and accuracy compared to previous studies [57,58] in the DDSM dataset. In the same trend, the previous study [63] gives 100% sensitivity and our proposed model gives 99.79% in the DDSM dataset due to using a small set of mammogram images. DDSM holds a high number of images, while the previous study of [63] used only 250 images in the evaluation.

Table 8. Performance comparison using Mini-MIAS dataset. Best values are shown in bold.

Classification Methods	Total No. of Images	Classifier	Feature	Ac (%)	Sn (%)	Sp (%)	F-Measure (%)
[45]	84	ANN	GLCM	61	100	77	-
[50]	107	SVM	Contourlet + FOA (45 features)	97.5	95.81	97.65	-
[57]	109	DT	Hybrid LB-GLCM + LPQ (texture)	98.2	100	97	-
[58]	119	MFO-ELM	LWT + PCA (32 features)	100	99.58	99.82	-
[59]	60	ELM	Gray level and textural features, 32	96.02	96.29	94.32	-
[60]	322	MLP	GLCM (texture)	93.37	94.18	92.43	-
[61]	126	ANN	HOG, GLCM (texture)	94	90	96.67	-
[62]	107	WC-SSA	Block-based discrete wavelet packet transforms	99.28	99.44	99	99.44
Proposed M-FD model	316	ANN	LBP FD M_FD	87.65 90.1 100	88.58 90.32 100	86.59 89.89 100	90.86 92.67 100

Table 9. Performance comparison with respect to DDSM dataset. Best values are shown in bold.

Classification Methods	Total No. of Images	Classifier	Feature	Ac (%)	Sn (%)	Sp (%)	F-Measure (%)
[50]	981	SVM	Contourlet + FOA (45 features)	98.18	-	-	-
[57]	72	Ensemble	Hybrid LB-GLCM + LPQ (texture)	93.1	95	88	-
[58]	981	MFO-ELM	LWT + PCA (32 features)	98.8	96.2	97.66	-
[59]	300	ELM	Gray level and textural features, 32	95.73	94.88	97.16	-
[60]	300	FGMM	Intensity, shape, texture	93	90	96	-
[62]	981	WC-SSA	Block-based discrete wavelet packet transforms	99.63	99.77	99.37	99.45
[63]	250	ANN	GLCM, GLRLM, wavelet (50 texture and shape)	98.9	100	97.8	-
[64]	287	ANN	GLRL, GLSZ, NGTD, LTEM (13 histogram, texture, 18 shape)	96.2	97.6	95.2	-
[65]	2781	AdaBoost	SIFT, GIST, HOG, LBP, VGG16, ResNet, and DenseNet (shape, texture, and deep learning)	90.91	82.96	98.38	-
Proposed M-FD model	981	ANN	LBP FD M_FD	79.71 81.75 99.89	78.68 80.36 99.79	80.9 83.16 100	79.17 81.6 99.89

Table 10. Performance comparison with respect to INBreast and BCDR datasets.

Classification Methods	Dataset	Total No. of Images	Classifier	Feature	Ac (%)	Sn (%)	Sp (%)	F-Measure (%)
[65]	INbreast	387	AdaBoost	SIFT, GIST, HOG, LBP, VGG16, ResNet, and DenseNet (shape, texture, and deep learning)	87.93	57.20	97.73	-
Proposed M-FD model	INbreast	200	ANN	LBP	82.5	83.33	80.64	86.79
				FD	85.5	85	86.66	89.13
				M_FD	99	98.44	100	99.21
[62]	BCDR	736	WC-SSA	block-based discrete wavelet packet transform	99.6	100	99.35	99.7
Proposed M-FD model	BCDR	736	ANN	LBP	80.43	82.79	77.12	83.17
				FD	81.93	84.79	78.09	84.29
				M_FD	99.59	99.53	99.67	99.64

Moreover, because sensitivity deals with positive cases, when the TP rate is low it means the sensitivity rate is high. This means that the sensitivity is used to identify the risk of benign; thus, when the number of benign images is low, the sensitivity rate will be high. On contrary, for the DDSM dataset, our proposed model exceeds all other previous studies in terms of all used matrices with using a high number of images in our evaluation.

In Table 10, we present a comparative study between our proposed model and two recent previous studies in INbreast and BCDR datasets. Our proposed model outperformed in terms of accuracy, sensitivity, specificity, and F-measure compared to [65] despite using texture, shape, and deep learning features. However, the previous study of [65] used 387 images, whereas we evaluate 200 images in the INbreast dataset. Furthermore, the proposed model yields better performance in terms of specificity whereas the study [62] has better results in terms of accuracy, sensitivity, and F-measure.

The proposed model also provides better evaluation performance when compared to the study [62] in Mini-MIAS and DDSM datasets. Our proposed model gives accuracy, sensitivity, and F-measure of 99.59%, 99.53%, and 99.64%, respectively, while [62] obtained 99.6%, 100%, and 99.7%, respectively. This means that the previous study of [62] outperformed in terms of accuracy, sensitivity, and F-measure slightly compared to our proposed model. In contrast, our proposed study outperformed in terms of specificity by obtaining 99.67% while the previous study yields 99.35% (see Table 10).

As a result, it is observed that our proposed model yields state-of-the-art evaluation performance compared to several previous studies in the Mini-MIAS, DDSM, INbreast, and BCDR. Therefore, our proposed model can work with different datasets efficiently while outperforming the previous study of [61] slightly on the BCDR dataset.

Table 11 compares the proposed method with deep learning methods investigated in [66]. DenseNet and DenseNet-II models were compared with our proposed model. We have used the same models for the Mini-MIAS, DDSM, INbreast, and BCDR datasets to show the effectiveness of deep learning models compared with the proposed model. The proposed model has achieved more effective results when compared to the models used in [66] because the proposed method needs less data for training. Deep learning techniques require a considerable amount of training data to build an acceptable model that can identify the risk of malignancy at an early stage. In addition, the balance between the two classes is an important point that may affect the prediction of deep learning techniques. Training a model with imbalanced data may obtain high weight for one class that can reduce the quality of the sensitivity and specificity of the system. Table 11 shows that applying DenseNet and DenseNet-II on DDSM and BCDR datasets achieved better results

than on Mini-MIAS and INbreast datasets, when we applied the same models, because DDSM and BCDR datasets contain more images than Mini-MIAS and INbreast datasets.

Table 11. Comparison of the proposed method with DenseNet and DenseNet-II models.

Dataset	Method	Performance (%)		
		Ac	**Sn**	**Sp**
Mini-MIAS	DenseNet	90.21	89.84	90.62
	DenseNet-II	91.3	90.42	91.83
	Proposed M-FD model	100	100	100
DDSM	DenseNet	95.08	94.6	96.1
	DenseNet-II	96.1	95.91	96.57
	Proposed M-FD model	99.89	99.79	100
INbreast	DenseNet	87.77	86	88.3
	DenseNet-II	89.22	89.07	88.28
	Proposed M-FD model	99	98.44	100
BCDR	DenseNet	98.3	97.8	98.72
	DenseNet-II	98.9	99.8	98.5
	Proposed M-FD model	99.59	99.53	99.67

5. Discussion

As shown by the experiments, our proposed model provides a better evaluation performance in accuracy, sensitivity, specificity, and the F-measure compared to several previous studies on the four different datasets. Our proposed model includes algorithms such as WT, thresholding, morphological operations, ML, M-FD + GA, and ANN that provide several features. WT is effective in dividing images into different sub-bands and separates high frequencies from low frequencies, which can be helpful in noise capturing and reducing. The thresholding technique has been used due to the remarkable intensity variation between foreground tissues and the background of mammograms. Morphological operations have been used to remove small objects that remain in the binary images. Due to the homogeneity between the pectoral muscle and ROI, an ML technique was built based on the histogram of the oriented gradient feature with neural network classifiers to determine the region of the pectoral muscle and ROI. Based on our evaluation and comparison between FD and LBP, this study has motivated us to improve FD and propose M-FD for feature extraction, selecting the most relevant feature from the extracted features that GA has employed. Class discrimination information is maintained by the selected features that help in proper breast cancer classification on mammogram images [67]. As a result, the proposed technique is an efficient technique to discover subtypes of breast cancer in mammogram images, which has been supported by the evaluation of the proposed technique on four different datasets.

However, the proposed method presents some limitations. Preprocessing mainly aims to enhance the texture and obtain powerful texture features from the image. The application of the proposed enhancement model removed some important information that may help identify the risk of malignancy. Thus, threshold selection is considered an important step in denoising-based image enhancement. Selection should be performed accurately and carefully because noise will still exist in the image when a small threshold is selected, whereas details of the image will be destroyed whilst producing artifacts and blurs when a large threshold is selected. Hence, a suitable threshold value must be determined to address this limitation and avoid denoising-based overfitting and underfitting in the proposed model. Moreover, low data contrast affects the feature extraction stage because there is some overlap between texture features of benign and malignant samples in this

stage. The use of texture features alone is insufficient to identify abnormalities and describe cases. Therefore, combining texture features with other features, such as geometry features, is reasonable.

6. Conclusions

The appropriate extraction of characteristics depends on the techniques used and the image itself. However, the extracted features should be robust or constant in terms of scale, translation, illumination, and orientation and partially consistent in terms of noises and artifacts. This study proposes an efficient model for the detection of breast cancer in mammogram images. The main contribution of this study is the development of a new model for feature extraction in mammogram images. The M-FD model is proposed as an improvement of the FD technique. The EWBFEM model based on WT is applied to ROIs for noise reduction before feature extraction. The GA is used to select relevant extracted features from ROIs. Extracted texture features can capture the essential differences amongst classes of objects (benign or malignant). The selected features are then fed to the trained ANN to capture the essential differences between benign and malignant samples. A series of experiments were conducted using four different databases. The first experiment was conducted using a single dataset for training and testing. The required training time is 237–277 s, and 0.3 s is required for testing each case. The second experiment was performed using one dataset for training and another dataset for testing. The proposed model required 324–367 s to train the proposed model and 0.3 s to test one case.

The performance evaluation in this study is efficient but still can be improved in the future. A suitable or strong threshold value can be used to remove noise and preserve the quality of the image with increasing PSNR. Furthermore, the proposed model can be used along with different deep learning methods to provide better performance for the classification of breast cancer on mammogram images. Overall, the continuous development of systems is very important to improve and contribute to the research on breast cancer, create algorithms that can aid experts, and reduce the examination time subjectively. Therefore, the proposed method shows great potential and will be investigated further.

Author Contributions: Conceptualization and methodology, D.A.Z. and D.A.I.; supervision M.A.M. and H.H.; validation D.Q.Z. and N.A.Z., writing—original draft preparation D.A.Z. and D.A.I.; writing—review and editing, R.D. and R.M.; funding acquisition, R.D. and R.M. All authors have read and agreed to the published version of the manuscript.

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

References

1. Zebari, D.A.; Ibrahim, D.A.; Zeebaree, D.Q.; Haron, H.; Salih, M.S.; Damaševičius, R.; Mohammed, M.A. Systematic review of computing approaches for breast cancer detection-based computer aided diagnosis using mammogram images. *Appl. Artif. Intell.* **2021**, 1–47. [CrossRef]
2. Hussein, I.J.; Burhanuddin, M.A.; Mohammed, M.A.; Benameur, N.; Maashi, M.S.; Maashi, M.S. Fully-automatic identification of gynaecological abnormality using a new adaptive frequency filter and histogram of oriented gradients (HOG). *Expert Syst.* **2021**, e12789. [CrossRef]
3. Irfan, R.; Almazroi, A.; Rauf, H.; Damaševičius, R.; Nasr, E.; Abdelgawad, A. dilated semantic segmentation for breast ultrasonic lesion detection using parallel feature fusion. *Diagnostics* **2021**, *11*, 1212. [CrossRef] [PubMed]
4. Kadry, S.; Damasevicius, R.; Taniar, D.; Rajinikanth, V.; Lawal, I.A. Extraction of tumour in breast mri using joint thresholding and segmentation—A study. In Proceedings of the 2021 Seventh International Conference on Bio Signals, Images and Instrumentation (ICBSII), Chennai, India, 25–27 March 2021. [CrossRef]
5. Rajinikanth, V.; Kadry, S.; Taniar, D.; Damasevicius, R.; Rauf, H.T. Breast-cancer detection using thermal images with marine-predators-algorithm selected features. In Proceedings of the 2021 Seventh International Conference on Bio Signals, Images and Instrumentation (ICBSII), Chennai, India, 25–27 March 2021. [CrossRef]
6. Hussein, I.J.; Burhanuddin, M.; Mohammed, M.A.; Elhoseny, M.; Garcia-Zapirain, B.; Maashi, M.S.; Maashi, M.S. Fully automatic segmentation of gynaecological abnormality using a new viola-jones model. *Comput. Mater. Contin.* **2021**, *66*, 3161–3182. [CrossRef]
7. Holland, K.; Van Gils, C.H.; Mann, R.M.; Karssemeijer, N. Quantification of masking risk in screening mammography with volumetric breast density maps. *Breast Cancer Res. Treat.* **2017**, *162*, 541–548. [CrossRef] [PubMed]
8. Brady, A.P. Error and discrepancy in radiology: Inevitable or avoidable? *Insights Imaging* **2017**, *8*, 171–182. [CrossRef]
9. Olliaro, P.; Torreele, E. Managing the risks of making the wrong diagnosis: First, do no harm. *Int. J. Infect. Dis.* **2021**, *106*, 382–385. [CrossRef]
10. Ramadan, S.Z. Methods used in computer-aided diagnosis for breast cancer detection using mammograms: A review. *J. Healthc. Eng.* **2020**, *2020*, 9162464. [CrossRef] [PubMed]
11. Sorantin, E.; Grasser, M.G.; Hemmelmayr, A.; Tschauner, S.; Hrzic, F.; Weiss, V.; Lacekova, J.; Holzinger, A. The augmented radiologist: Artificial intelligence in the practice of radiology. In *Pediatric Radiology*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2021. [CrossRef]
12. Hosny, A.; Parmar, C.; Quackenbush, J.; Schwartz, L.H.; Aerts, H.J.W.L. Artificial intelligence in radiology. *Nat. Rev. Cancer* **2018**, *18*, 500–510. [CrossRef]
13. Liew, X.; Hameed, N.; Clos, J. A review of computer-aided expert systems for breast cancer diagnosis. *Cancers* **2021**, *13*, 2764. [CrossRef]
14. Yassin, N.I.; Omran, S.; Houby, E.M.; Allam, H. Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review. *Comput. Methods Programs Biomed.* **2018**, *156*, 25–45. [CrossRef]
15. Diniz, J.O.B.; Diniz, P.H.B.; Valente, T.L.A.; Silva, A.C.; Paiva, A.; Gattass, M. Detection of mass regions in mammograms by bilateral analysis adapted to breast density using similarity indexes and convolutional neural networks. *Comput. Methods Programs Biomed.* **2018**, *156*, 191–207. [CrossRef]
16. Hong, B.-W.; Sohn, B.-S. Segmentation of Regions of Interest in Mammograms in a Topographic Approach. *IEEE Trans. Inf. Technol. Biomed.* **2009**, *14*, 129–139. [CrossRef]
17. Chan, H.-P.; Samala, R.K.; Hadjiiski, L.M. CAD and AI for breast cancer—Recent development and challenges. *Br. J. Radiol.* **2020**, *93*, 20190580. [CrossRef] [PubMed]
18. Zeebaree, D.Q.; Haron, H.; Abdulazeez, A.M.; Zebari, D.A. Machine learning and region growing for breast cancer segmentation. In Proceedings of the 2019 International Conference on Advanced Science and Engineering (ICOASE), Zakho-Duhok, Iraq, 2–4 April 2019; pp. 88–93.
19. Zeebaree, D.Q.; AbdulAzeez, A.M.; Zebari, D.A.; Haron, H.; Hamed, H.N.A. Multi-level fusion in ultrasound for cancer detection based on uniform LBP features. *Comput. Mater. Contin.* **2021**, *66*, 3363–3382. [CrossRef]
20. Khan, M.; Sharif, M.; Akram, T.; Damaševičius, R.; Maskeliūnas, R. Skin lesion segmentation and multiclass classification using deep learning features and improved moth flame optimization. *Diagnostics* **2021**, *11*, 811. [CrossRef]
21. Maqsood, S.; Damaševičius, R.; Maskeliūnas, R. Hemorrhage detection based on 3D CNN deep learning framework and feature fusion for evaluating retinal abnormality in diabetic patients. *Sensors* **2021**, *21*, 3865. [CrossRef]
22. Zemouri, R.; Devalland, C.; Valmary-Degano, S.; Zerhouni, N. Neural network: A future in pathology? *Ann. De Pathol.* **2019**, *39*, 119–129. [CrossRef] [PubMed]
23. Azli, E.A.; Huddin, A.B.; Ibrahim, M.F.; Samad, S.A. Mass classification in mammograms using neural network. In Proceedings of the 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI), Langkawi, Malaysia, 25–27 November 2017; pp. 1–5.
24. Vijayarajeswari, R.; Parthasarathy, P.; Vivekanandan, S.; Basha, A.A. Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform. *Measurement* **2019**, *146*, 800–805. [CrossRef]

25. Al Rasyid, M.B.; Arnia, F.; Munadi, K. Histogram statistics and GLCM features of breast thermograms for early cancer detection. In Proceedings of the 2018 International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI-NCON), Chiang Rai, Thailand, 25–28 February 2018; pp. 120–124.

26. Abdel-Nasser, M.; Rashwan, H.A.; Puig, D.; Moreno, A. Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern. *Expert Syst. Appl.* **2015**, *42*, 9499–9511. [CrossRef]

27. Gedik, N. A new feature extraction method based on multi-resolution representations of mammograms. *Appl. Soft Comput.* **2016**, *44*, 128–133. [CrossRef]

28. Laadjel, M.; Al-Maadeed, S.; Bouridane, A. Combining Fisher locality preserving projections and passband DCT for efficient palmprint recognition. *Neurocomputing* **2015**, *152*, 179–189. [CrossRef]

29. Sakthi, S.; Balasubramanie, P. Dynamic cross propagation algorithm based detection of micro calcification in digital mammogram. *J. Ambient. Intell. Humaniz. Comput.* **2021**, *12*, 5877–5894. [CrossRef]

30. Patil, R.S.; Biradar, N. Automated mammogram breast cancer detection using the optimized combination of convolutional and recurrent neural network. *Evol. Intell.* **2020**, *14*, 1459–1474. [CrossRef]

31. Boudraa, S.; Melouah, A.; Merouani, H.F. Improving mass discrimination in mammogram-CAD system using texture information and super-resolution reconstruction. *Evol. Syst.* **2020**, *11*, 697–706. [CrossRef]

32. Gong, X.; Yang, Z.; Wang, D.; Qi, Y.; Guo, Y.; Ma, Y. Breast density analysis based on glandular tissue segmentation and mixed feature extraction. *Multimed. Tools Appl.* **2019**, *78*, 31185–31214. [CrossRef]

33. Ergin, S.; Kilinc, O. A new feature extraction framework based on wavelets for breast cancer diagnosis. *Comput. Biol. Med.* **2014**, *51*, 171–182. [CrossRef]

34. Muramatsu, C.; Hara, T.; Endo, T.; Fujita, H. Breast mass classification on mammograms using radial local ternary patterns. *Comput. Biol. Med.* **2016**, *72*, 43–53. [CrossRef] [PubMed]

35. Rampun, A.; Scotney, B.W.; Morrow, P.J.; Wang, H.; Winder, J. Breast density classification using local quinary patterns with various neighbourhood topologies. *J. Imaging* **2018**, *4*, 14. [CrossRef]

36. Rabidas, R.; Arif, W. Characterization of mammographic masses based on local photometric attributes. *Multimed. Tools Appl.* **2020**, *79*, 21967–21985. [CrossRef]

37. Ghasemzadeh, A.; Azad, S.S.; Esmaeili, E. Breast cancer detection based on Gabor-wavelet transform and machine learning methods. *Int. J. Mach. Learn. Cybern.* **2019**, *10*, 1603–1612. [CrossRef]

38. Fanizzi, A.; Basile, T.M.A.; Losurdo, L.; Bellotti, R.; Bottigli, U.; Dentamaro, R.; Didonna, V.; Fausto, A.; Massafra, R.; Moschetta, M.; et al. A machine learning approach on multiscale texture analysis for breast microcalcification diagnosis. *BMC Bioinform.* **2020**, *21*, 91. [CrossRef]

39. Mohanty, F.; Rup, S.; Dash, B.; Majhi, B.; Swamy, M.N.S. Digital mammogram classification using 2D-BDWT and GLCM features with FOA-based feature selection approach. *Neural Comput. Appl.* **2019**, *32*, 7029–7043. [CrossRef]

40. Shen, L.; He, M.; Shen, N.; Yousefi, N.; Wang, C.; Liu, G. Optimal breast tumor diagnosis using discrete wavelet transform and deep belief network based on improved sunflower optimization method. *Biomed. Signal Process. Control* **2020**, *60*, 101953. [CrossRef]

41. Melekoodappattu, J.G.; Subbian, P.S. A hybridized ELM for automatic micro calcification detection in mammogram images based on multi-scale features. *J. Med. Syst.* **2019**, *43*, 183. [CrossRef] [PubMed]

42. Loizidou, K.; Skouroumouni, G.; Nikolaou, C.; Pitris, C. An automated breast micro-calcification detection and classification technique using temporal subtraction of mammograms. *IEEE Access* **2020**, *8*, 52785–52795. [CrossRef]

43. Sheba, K.; Raj, S.G. An approach for automatic lesion detection in mammograms. *Cogent Eng.* **2018**, *5*, 1444320. [CrossRef]

44. Junior, G.B.; da Rocha, S.V.; de Almeida, J.D.; de Paiva, A.C.; Silva, A.C.; Gattass, M. Breast cancer detection in mammography using spatial diversity, geostatistics, and concave geometry. *Multimed. Tools Appl.* **2018**, *78*, 13005–13031. [CrossRef]

45. Pezeshki, H.; Rastgarpour, M.; Sharifi, A.; Yazdani, S. Extraction of spiculated parts of mammogram tumors to improve accuracy of classification. *Multimed. Tools Appl.* **2019**, *78*, 19979–20003. [CrossRef]

46. Virmani, J.; Agarwal, R. Effect of despeckle filtering on classification of breast tumors using ultrasound images. *Biocybern. Biomed. Eng.* **2019**, *39*, 536–560. [CrossRef]

47. Mabrouk, M.S.; Afify, H.M.; Marzouk, S.Y. Fully automated computer-aided diagnosis system for micro calcifications cancer based on improved mammographic image techniques. *Ain Shams Eng. J.* **2019**, *10*, 517–527. [CrossRef]

48. Pérez-Benito, F.J.; Signol, F.; Pérez-Cortés, J.-C.; Pollán, M.; Pérez-Gómez, B.; Salas-Trejo, D.; Casals, M.; Martínez, I.; Llobet, R. Global parenchymal texture features based on histograms of oriented gradients improve cancer development risk estimation from healthy breasts. *Comput. Methods Programs Biomed.* **2019**, *177*, 123–132. [CrossRef]

49. Zebari, D.A.; Zeebaree, D.Q.; Abdulazeez, A.M.; Haron, H.; Hamed, H.N.A. Improved threshold based and trainable fully automated segmentation for breast cancer boundary and pectoral muscle in mammogram images. *IEEE Access* **2020**, *8*, 203097–203116. [CrossRef]

50. Mohanty, F.; Rup, S.; Dash, B.; Majhi, B.; Swamy, M.N.S. Mammogram classification using contourlet features with forest optimization-based feature selection approach. *Multimed. Tools Appl.* **2019**, *78*, 12805–12834. [CrossRef]

51. Zemouri, R.; Omri, N.; Morello, B.; Devalland, C.; Arnould, L.; Zerhouni, N.; Fnaiech, F. Constructive deep neural network for breast cancer diagnosis. *IFAC-PapersOnLine* **2018**, *51*, 98–103. [CrossRef]

52. Mandelbrot, B.B. *The Fractal Geometry of Nature*; W.H. Freeman and Co.: New York, NY, USA, 1982; pp. 394–397.

53. Suckling, J.P. The mammographic image analysis society digital mammogram database exerpta medica. *Int. Congr. Ser.* **1994**, *1069*, 375–378.

54. Heath, M.; Bowyer, K.; Kopans, D.; Kegelmeyer, P.; Moore, R.; Chang, K.; Munishkumaran, S. The digital database for screening mammography. In *Fifth International Workshop on Digital Mammography*; Yaffe, M.J., Ed.; Medical Physics Publishing: Madison, WI, USA, 2001; pp. 212–218.

55. Moreira, I.C.; Amaral, I.; Domingues, I.; Cardoso, A.; Cardoso, M.J.; Cardoso, J. INbreast: Toward a full-field digital mammographic database. *Acad. Radiol.* **2012**, *19*, 236–248. [CrossRef]

56. Moura, D.C.; López, M.A.G.; Cunha, P.; De Posada, N.G.; Pollan, R.R.; Ramos, I.; Loureiro, J.P.; Moreira, I.C.; De Araújo, B.M.F.; Fernandes, T.C. Benchmarking datasets for breast cancer computer-aided diagnosis (CADX). In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8258, pp. 326–333.

57. Sadad, T.; Munir, A.; Saba, T.; Hussain, A. Fuzzy C-means and region growing based classification of tumor from mammograms using hybrid texture feature. *J. Comput. Sci.* **2018**, *29*, 34–45. [CrossRef]

58. Muduli, D.; Dash, R.; Majhi, B. Automated breast cancer detection in digital mammograms: A moth flame optimization based ELM approach. *Biomed. Signal Process. Control* **2020**, *59*, 101912. [CrossRef]

59. Xie, W.; Li, Y.; Ma, Y. Breast mass classification in digital mammography based on extreme learning machine. *Neurocomputing* **2016**, *173*, 930–941. [CrossRef]

60. Pashoutan, S.; Shokouhi, S.B.; Pashoutan, M. Automatic breast tumor classification using a level set method and feature extraction in mammography. In Proceedings of the 2017 24th National and 2nd International Iranian Conference on Biomedical Engineering (ICBME), Tehran, Iran, 30 November–1 December 2017; pp. 1–6.

61. Tatikonda, K.C.; Bhuma, C.M.; Samayamantula, S.K. The analysis of digital mammograms using HOG and GLCM features. In Proceedings of the 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 10–12 July 2018; pp. 1–7.

62. Mohanty, F.; Rup, S.; Dash, B.; Majhi, B.; Swamy, M. An improved scheme for digital mammogram classification using weighted chaotic salp swarm algorithm-based kernel extreme learning machine. *Appl. Soft Comput.* **2020**, *91*, 106266. [CrossRef]

63. Mohamed, B.A.; Salem, N.M. Automatic classification of masses from digital mammograms. In Proceedings of the 2018 35th National Radio Science Conference (NRSC), Cairo, Egypt, 20–22 March 2018; pp. 495–502.

64. Uthoff, J.; Sieren, J.C. Information theory optimization based feature selection in breast mammography lesion classification. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI), Washington, DC, USA, 4–7 April 2018; pp. 817–821.

65. Zhang, H.; Wu, R.; Yuan, T.; Jiang, Z.; Huang, S.; Wu, J.; Hua, J.; Niu, Z.; Ji, D. DE-Ada*: A novel model for breast mass classification using cross-modal pathological semantic mining and organic integration of multi-feature fusions. *Inf. Sci.* **2020**, *539*, 461–486. [CrossRef]

66. Li, H.; Zhuang, S.; Li, D.-A.; Zhao, J.; Ma, Y. Benign and malignant classification of mammogram images based on deep learning. *Biomed. Signal Process. Control* **2019**, *51*, 347–354. [CrossRef]

67. Lahoura, V.; Singh, H.; Aggarwal, A.; Sharma, B.; Mohammed, M.A.; Damaševičius, R.; Kadry, S.; Cengiz, K. Cloud computing-based framework for breast cancer diagnosis using extreme learning machine. *Diagnostics* **2021**, *11*, 241. [CrossRef]