Early breast cancer diagnostics based on hierarchical machine learning classification for mammography images

M. Saeed Darweesh, Mostafa Adel, Ahmed Anwar, Omar Farag, Ahmed Kotb, Mohamed Adel, Ayman Tawfik & Hassan Mostafa |

Published online: 09 Sep 2021.

Submit your article to this journal ✎

Article views: 85

View related articles ✎

View Crossmark data ✎

cogent
engineering

CrossMark

*Corresponding author: M. Saeed Darweesh, Electronics and Computer Engineering, Nile University School of Engineering and Applied Sciences, 26th of July Corridor, Al Sheikh Zayed City, Giza, Egypt
E-mail: mdarweesh@nu.edu.eg

COMPUTER SCIENCE | RESEARCH ARTICLE

Early breast cancer diagnostics based on hierarchical machine learning classification for mammography images

M. Saeed Darweesh[1,2]*, Mostafa Adel[3], Ahmed Anwar[4], Omar Farag[4], Ahmed Kotb[4], Mohamed Adel[5], Ayman Tawfik[6] and Hassan Mostafa[47]

Abstract: Breast cancer constitutes a significant threat to women's health and is considered the second leading cause of their death. Breast cancer is a result of abnormal behavior in the functionality of the normal breast cells. Therefore, breast cells tend to grow uncontrollably, forming a tumor that can be felt like a breast lump. Early diagnosis of breast cancer is proved to reduce the risks of death by providing a better chance of identifying a suitable treatment. Machine learning and artificial intelligence play a key role in healthcare systems by assisting physicians in diagnosing early, better, and treating various diseases. For achieving the early detection of breast cancer, this paper proposes a Machine Learning-based two-level top-down hierarchical approach for breast cancer detection and classification into three classes: normal, benign, and malignant, using the Mammographic Image Analysis Society (MIAS) mammography dataset. Different data preprocessing techniques are applied before using feature extraction techniques and machine learning algorithms for classification. The first classification stage which distinguishes

ABOUT THE AUTHOR

M. Saeed Darweesh is a (M'17 - SM'20) received his Ph.D. (with honors) in Electronics and Electrical Communications Engineering from the Faculty of Engineering, Cairo University, Giza, Egypt, in 2017. Currently, he is a full-time Assistant Professor in the ECE program, School of Engineering and Applied Sciences, Nile University, Egypt. Besides, Dr. Darweesh is also an expert at the Phi Science Institute.

Dr. Darweesh is a PI of different research projects funded by the Information Technology Industry Development Agency (ITIDA), Co-PI, and a Research Associate in seven research projects funded by different agencies like the Science and Technology Development Fund (STDF), National Telecom Regulatory Authority (NTRA), and the Academy of Scientific Research and Technology (ASRT). His research interests focus on Artificial Intelligence, Narrow-Band IoT, Autonomous Driving Vehicle to Vehicle Systems, Wireless Communications, Biomedical Engineering (EEG Seizure Detection, Sleepiness Detection using EEG, and Breast Cancer Classification), and Data Compression.

M. Saeed Darweesh

PUBLIC INTEREST STATEMENT

Machine Learning and Artificial Intelligence play a key role in healthcare systems, in which Machine Learning-based approaches can assist physicians to better diagnose and treat various diseases. This paper proposes a Machine Learning-based approach for breast cancer detection and classification into three classes: normal, benign and malignant. A two-level top-down hierarchical classification approach is proposed. The proposed approach is based on two consecutive stages utilizing two classifiers. The first classification stage distinguishes between normal and abnormal cases, followed by the second classification stage, which classifies the abnormal cases into benign or malignant cases.

cogent · oa

between normal and abnormal cases is comprised of Gray Level Co-occurrence Matrix (GLCM) as a feature extraction technique and random forest as a classifier, followed by the second classification stage which classifies the abnormal cases into benign or malignant cases and is comprised of Local Binary Patterns (LBP) as a feature extraction technique and random forest as a classifier. The classification accuracy for the first stage is 97% and an F1-score of 0.98 and 0.97 for normal and abnormal classes. While for the second stage, the classification accuracy is 75% and an F1-score of 0.76 and 0.74 for benign and malignant classes. The overall hierarchical classification system achieves a classification accuracy of 85%, Matthews correlation coefficient (MCC) of 0.76, and F1-score of 0.98, 0.7, and 0.74 for normal, benign, and malignant test cases.

Subjects: Machine Learning - Design; Technology; Image Processing

Keywords: Breast cancer; Mammography; Machine learning; Gray Level Co-Occurrence Matrix (GLCM); Local Binary Patterns (LBP); Random forest

1. Introduction

Breast cancer is a consequence of abnormal behavior in the normal breast cells' functionality that causes a disturbance in the ordinary properties of the breast cells. As a result, breast cells tend to grow in an uncontrollable manner forming either a benign or malignant tumor (Warburg, 1956). Breast, lung, and colorectal cancer are considered the three most popular types of cancers among women, respectively. These three cancer types represent about 50% of all cancer cases in women, and breast cancer constitutes about 30% of all the cancer cases (Siegel et al., 2016). The morbidity rate for breast cancer is about 14.7% such that nearly half a million patients diagnosed with breast cancer are dead, with an increase of almost 1.7 million patients diagnosed with breast cancer annually (Han et al., 2017). By 2025, it is predicted that the annual cancer cases (all races, all sites) will be increased among men by 26.1% and by 23.7% among women (Siegel et al., 2020).

Breast cancer is widely spread among women aging more than 40 years, especially those ranging between 60 and 79 years old (Han et al., 2017). Fortunately, the likelihood of women's death cases at a premature stage is about 3% (Gbenga et al., 2017). According to (Unger-Saldaña, 2014), the incidence rates of breast cancer are higher in the most developed countries. However, death rates are significantly higher in developing countries. In Egypt, and according to Baheya Foundation (Bahya Hospital, 2018), there is a probability that 1 out of each 8 Egyptian women would be diagnosed with breast cancer in their lifetime. Besides, breast cancer represents about 17.6% of all cancer patients in Egypt. Nevertheless, early diagnosis of breast cancer could effectively increase survival rates. The probability that a woman could be totally cured of breast cancer at its first and second stages could reach up to 98% and 93%, respectively. However, these survival rates are supposed to drop significantly to 73% and 22% when the diagnosis occurs at the third and fourth stages, respectively.

Computer-aided diagnostic tools have been recently utilized for the sake of early detection of female breast cancer. A computer program is used to determine any abnormal part of the breast. The most common techniques used for this purpose are Magnetic Resonance Imaging (MRI) and X-ray mammography. MRI uses radio waves and magnetic fields to produce high-quality (2D or 3D) images of the breast (Cheung & Donlon, 2016). Mammography uses X-rays with low energy to create images of what is inside the breast. The mammography diagnostic procedure is applied as follows: Firstly, the breast is initially pressed using two parallel plates. This compression decreases the area that X-rays can penetrate so the output images could be more accurate. Following that, top to bottom and angle-sided view images are collected. This process usually takes time ranging from 15 to 20 minutes. Mammography can be classified according to its purpose into two types,

which are screening mammography and diagnostic mammography (EE et al., 2005). Detecting breast cancer could be cost-effective using mammography images. Computer-aided diagnostic tools are capable of speeding up this process and improving the overall accuracy by helping the radiologists as artificial intelligence (AI) systems are capable of surpassing human experts in breast cancer prediction (McKinney et al., 2020).

Aside from mammography and MRI, some imaging techniques depend mainly on ultrasound waves. The most popular method is ultrasound elastography, which makes use of the fact that breast cancer tissues are stiffer than normal breast tissues, where the stiffening process begins in an early stage of cancer (Zhou et al., 2017). The elastography technologies could be divided into strain elastography and shear wave elastography (Barr, 2012). Some invasive techniques are used to extract high-quality images for the breast. Fine needle biopsy (FNB) is considered the most famous among them. FNB is carried out by obtaining a sample directly from the tumor. This sample is then exposed to microscopic examination for image extraction (Kowal et al., 2013). However, all these imaging techniques are not enough for breast cancer identification. There is always a significant probability for false positives, which may lead to unneeded surgical involvement (Becker et al., 2017).

Machine learning and data mining techniques are recently utilized to assist breast cancer diagnostics, in which accuracy is a crucial factor. Machine learning is based on developing a complex statistical and mathematical model that can effectively learn from data. This model can then find the hyperplanes that can effectively separate between different classes and can be used afterward for predicting the state of the breast cancer patient. Applying machine learning algorithms is always preceded by image processing and feature extraction algorithms for dataset construction. Machine learning techniques can efficiently extract important predictive features from complex and noisy datasets (Kourou et al., 2015). The most important advantage of using machine learning in breast cancer diagnostics is that by increasing the dataset size (more patient samples), the chance that the model can effectively learn from data increases. Subsequently, the accuracy of examinations could be significantly increased (Ahmad et al., 2013).

In (Adel et al., 2019), a support vector machine (SVM) with radial basis function (RBF) is used to differentiate between benign and malignant breast lesions in combined elastogram and B-mode images.

The contributions of this work can be summarized as follows:

• Two-level hierarchical system for breast cancer detection and classification into normal, benign, and malignant cases.
• Two different feature extraction techniques are implemented based on the classification level. Gray Level Co-Occurrence Matrix (GLCM) is used for the first level (Normal/Abnormal), and Local Binary Patterns (LBP) is used for the second level (Benign/Malignant).
• Balanced test data for the two classification levels is performed, which leads to unbiased results.
• We have found that GLCM and Random Forest are the best combinations of a feature extractor and a classifier for the first stage, yielding a classification accuracy of 97%, and an F1-score of 0.98 and 0.97 for the normal and abnormal classes, respectively. While LBP and Random Forest are the best combinations of a feature extractor and a classifier for the second stage, yielding a classification accuracy of 75%, and F1-score of 0.76 and 0.74 for the benign and malignant classes, respectively.
• The overall hierarchical classification system achieves a classification accuracy of 85%, Matthews correlation coefficient (MCC) of 0.76, and F1-score of 0.98, 0.7, and 0.74 for normal, benign, and malignant test cases.

The objective and aim of this study are introducing a Machine Learning-based breast cancer diagnosis system to detect and classify normal from abnormal breasts and benign from malignant tumors. The study shows the strengths and weaknesses of the proposed system through a detailed classification report including precision, recall, and F1-score, which helps physicians understand when to rely on the system with higher confidence. Moreover, the study shows that differentiating benign and malignant tumors is a difficult task and proposes LBP as a feature extraction technique for this task, which provides better results than GLCM, which differentiates normal from abnormal breasts.

In this paper, a hierarchical-based classification approach is used with the Mammographic Image Analysis Society (MIAS) dataset. The paper is organized as follows: Section I discusses the related work done to tackle the breast cancer detection and classification problem. Section III describes the materials and methods used in the proposed approach of classification. The results and discussion are given in Section IV. Finally, the whole work is concluded in Section V.

2. Literature review

The main problem in breast cancer detection is that the tumor size is very small with respect to the size of the mammographic image. For example, an image might be 4000 $\times$ 3000 pixels, and the tumor might be in the range of millimeters which represents around 30 $\times$ 30 pixels for a pixel resolution of 50 µm per pixel. Hence, there is always a need for data with tumor labels.

The most popular data sets with tumor labels are the MIAS dataset and the Digital Database for Screening Mammography (DDSM). Many researchers have used the MIAS dataset such as in (Bektas et al., 2018). First, the authors preprocess the images to remove doctor annotations, followed by three filters for noise reduction. The main goal is to identify whether there is a tumor or not, and if a tumor is found, whether it is malignant or benign. The authors use three feature extraction techniques: histogram of oriented gradients (HOG), LBP, and GLCM, and compare them. Other useful transforms could be used in feature extraction from images (e.g., proposed in (Abdulhussain & Mahmmod, 2021) (Abdul-Hadi et al., 2020), and (Mahmmod et al., 2020)). Features are extracted after preprocessing the images. A two-stage classification approach is used. The first stage identifies the tumors with a maximum accuracy of 65%, and the second stage classifies whether the tumor (if exists) is malignant or benign with a maximum accuracy of 65%. The authors have not provided an F1 score, precision, or recall, although these performance metrics are essential to determine the model reliability and to determine the classes that result in degrading the accuracy through the confusion matrix as well as to perform dataset balancing if needed.

In (Charan et al., 2018), the authors resize the images to overcome the problem of small tumor size concerning the whole image and then perform morphological operations for noise removal. Then, a convolutional neural network (CNN) model is trained in a two-stage approach. The first stage classifies the image as normal or abnormal breast. The second stage detects six classes of findings: asymmetry, calcification, speculated masses, circumscribed masses, architectural distortion, and miscellaneous. The model is trained and tested on an unbalanced data set. The achieved classification accuracy is 65%. It is important to note that the provided evaluation metric is the average accuracy only, and no F1-score, precision, or recall are evaluated.

In (Saraswathi et al., 2016), the authors apply curvelet transform for feature extraction followed by particle swarm optimization (PSO) as a dimensionality reduction technique. The SVM is used as a classifier, and an accuracy of 92% has been achieved on 182 selected samples from the MIAS database. Sensitivity and specificity evaluation metrics are provided. In (Saraswathi et al., 2016), there are no given justifications on how the selection of these 182 samples has been performed.

Wavelet transform as a feature extraction technique is used in (Abirami et al., 2016). Both multi-layer perceptron (MLP) and radial basis function (RBF) neural networks are used for classification.

cogent ·· engineering

However, this work discusses only the classification between normal and abnormal breasts and does not discuss whether the abnormality is benign or cancerous. The achieved accuracy is 95.5% of Haar wavelet transform and MLP, with specificity and sensitivity of 0.95 and 0.96, respectively.

In (Setiawan et al., 2015), the authors use the Law's Texture Energy Measure (LAWS) technique to extract secondary features from the images. First, the Region of Interest (ROI) is extracted using the labels in the dataset, followed by feature extraction using the LAWS feature extraction technique. A hierarchical classification approach is used. Firstly, the classification between normal and abnormal breasts is carried out, followed by another classification between benign and malignant cases if the first stage classification is abnormal. The accuracy, specificity, and sensitivity for classifying normal and abnormal cases are 93.9%, 100%, and 91%, respectively. The accuracy, specificity, and sensitivity of classifying benign and malignant cases are 83.3%, 88%, and 80%, respectively.

In (Kashif, 2020), the authors have applied a 2D median filter to remove the noises from the images of the MIAS dataset. The dataset is divided into two classes: normal and abnormal (Benign and Malignant). They have used two methods for image segmentation to select the ROI from the images. The first method is the edge-based technique which differentiates the region of the image by locating edges. The second method is by dividing the image into small blocks using region-based segmentation. The segmented image is then represented into five features: the radius of the highlighted region, the entropy of the highlighted region, the smoothness of the highlighted region, the mean texture of the highlighted region, and texture-based features. Multiple machine learning models have been trained on the dataset, but the best model was SVM, achieving 90%, 100%, 90%, 0%, and 95% on the accuracy, recall, precision, specificity, and f-score, respectively.

In (Shi et al., 2019), using the MIAS dataset, the authors have applied segmentation for skin-air boundary and breast region boundary by combining vertical and horizontal gradient magnitude weight arrays. Segmented pectoral muscle images are reshaped as 200×200 in size in the first step. Then, 32, 32, and 64 convolutional cores with a kernel of 3×3 size are used in each layer to extract deep features of input mammography images. Finally, the output of four BI-RADS density classes is driven after a flattened layer followed by a dense layer. This CNN has achieved an accuracy of 83.6%.

In (Saber et al., 2021), the authors have applied a 2D median filter with a kernel size of 3×3 to aid in removing the noise from the images of the MIAS dataset. Classical histogram equalization has been applied in order to strengthen the contrast of the original image to make the image anomalies more visible. Removing non-breast regions was done by morphological analysis by applying image opening, image closing, white top hat, black top hat, and Mathematical Morphology. These operations are followed by threshold-based segmentation. All images are resized to 244×244 to fit the pre-trained models input size. Various pre-trained deep learning

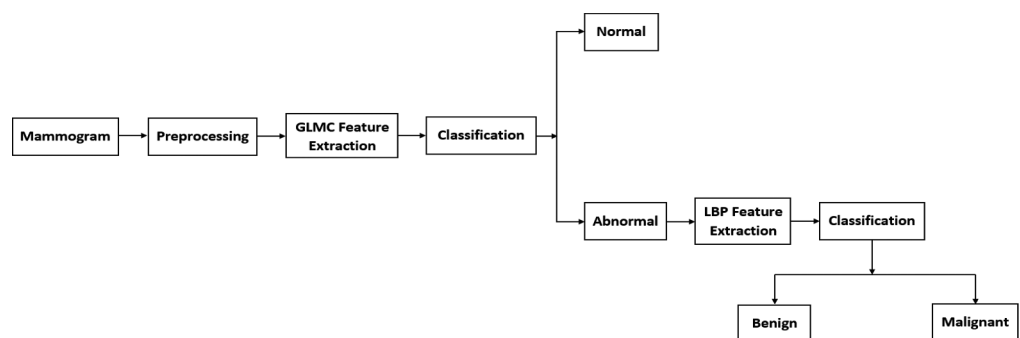**Figure 1. Proposed breast cancer classification approach.**

cogent ·· engineering

Figure 2. Normal, benign and malignant samples from MIAS dataset.

(a) Normal　　　(b) Benign　　　(c) Malignant
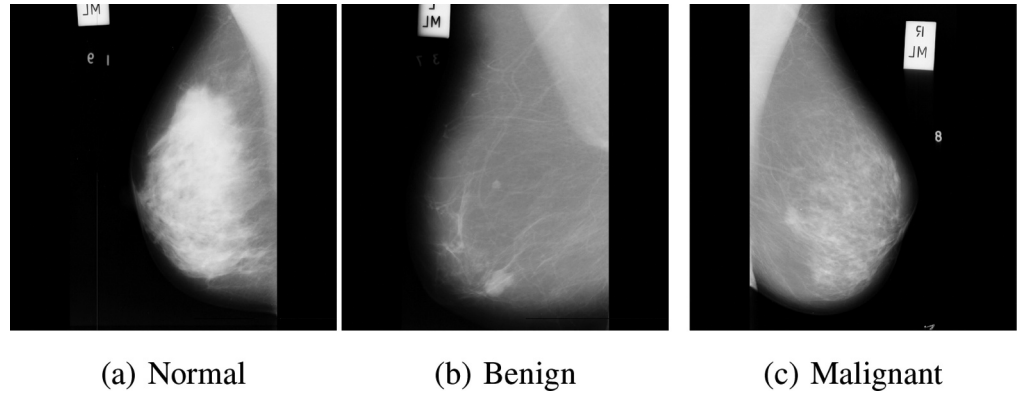
Table 1. First stage training and test sets distribution		
Class	**Training samples**	**Test samples**
Normal	187	20
Abnormal	91	20

Table 2. Second stage training and test sets distribution		
Class	**Training samples**	**Test samples**
Benign	53	10
Malignant	38	10

Figure 3. The effect of applying histogram equalization on a sample from MIAS dataset.

(a) Before histogram equalization　　　(b) After histogram equalization
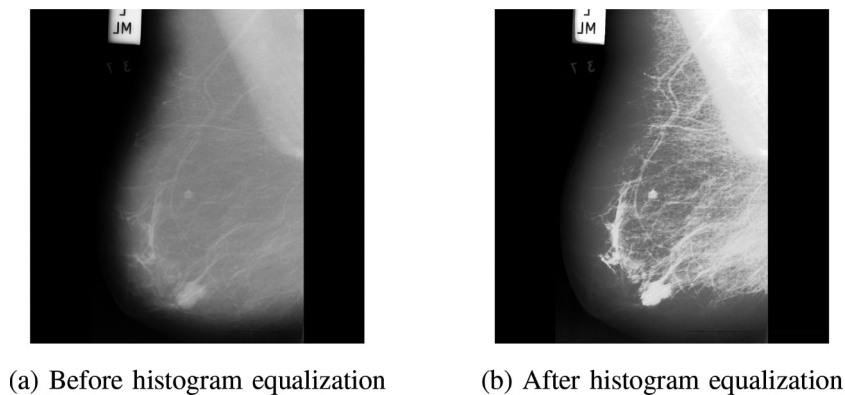
models have been used as a feature extractor, but the best model was VGG 16 followed by SVM classifier. Achieving accuracy, sensitivity, specificity, precision, and f-score on benign images 99.31%, 99%, 99.3%, 97%, and 98%, respectively, on malignant images 98.62%, 95.6%, 99.1%, 95.6%, and 96%, respectively, and on normal images 98.96%, 98.9%, 99%, 99%, and 99%, respectively, with a total accuracy of 98.87%.

In (Yu et al., 2020), the authors have divided the MIAS dataset into two classes Normal and Abnormal (Benign and Malignant). They have designed a deep fusion learning system for

cogent··engineering

mammographic image classification. They have applied a median filter to remove the noise from the images in addition to contrast limited adaptive histogram equalization for image enhancement. ROI image is then resized to 120 $\times$ 120 pixels. Small patches are randomly cropped of size 72 $\times$ 72 pixels on each ROI. They have collected 500 small patches for each normal ROI and 2,000 small patches for each abnormal ROI. VGG 16 model has been used with the idea of concatenating global average pooling layers to form a longer one and connect them to the batch normalization layer. Achieving recall, precision, and f1-score on normal class 87.80%, 94.74%, and 91.14%, respectively, while on abnormal class 91.30%, 80.77%, and 85.71%, respectively, with a total accuracy of 89.06%.

3. Materials and methods

The proposed system is a hierarchical-based classifier. Initially, an image is classified as normal or abnormal (benign or malignant), followed by a benign or malignant classification if the first-stage classification is abnormal.

Figure 1 shows an overview of the proposed classification system. First, the input image is preprocessed, followed by a feature extraction process using two sets of features based on GLCM and LBP. GLCM features are used in the first stage to differentiate between normal and abnormal breasts, as GLCM expresses the frequency of occurrence of a combination of pixel pair values. Next, it is used to compare the relation between pixel intensity and the neighborhood's intensity (Sreehari Sastry et al., 2012). Accordingly, GLCM can capture large diversities of normal breasts for benign and malignant breasts.

For the classification between benign and malignant, LBP features are used due to the capability of LBP to detect patterns regardless of variations of grey intensity (Öztürk & Akdemir, 2018). This helps in detecting tumor patterns, whether malignant or benign.

   (A) *Dataset*

A digital mammogram database obtained from the MIAS has been used in this study (MIAS database, 2021). The mammograms in the dataset initially consist of 322 samples of digital mammogram images, categorized into three categories: normal, benign, and malignant, where 207 cases are diagnosed as normal, 63 as benign, and 48 as malignant. The images are of the same size 1024 $\times$ 1024 and are centered in the matrix. The dataset provides information about each images, such as the class of abnormalities present (Calcification, Well-defined/circumscribed masses, Speculated masses, Other, ill-defined masses, Architectural distortion, Asymmetry, Normal). The coordinates of the center of the tumor and the radius of the circle that delimits the tumor. This information helps in the process of segmentation and extraction of the tumor. Figure 2 shows three normal, benign, and malignant samples from the MIAS dataset.

The dataset is split into training and testing portions. A fair test is to have a test set with an equal number of samples for each class in every classification stage, as the dataset is skewed towards the normal class. In the first classification stage, the benign and malignant samples are grouped as the abnormal class. The test set for the first stage comprises 20 samples as normal and 20 samples as abnormal. The test set for the second stage comprises 10 samples as benign and 10 samples as malignant. Tables 1 and Tables 2 show the training and testing set distribution for the first and second classification stages.

   (A) *Preprocessing*

   (1) *Histogram Equalization*

Histogram equalization (Pizer et al., 1987) is performed to enhance the details of the image. This technique spreads out the most frequent intensity values, resulting in a better distribution on the

histogram. Consequently, the areas of lower local contrast gain a higher contrast. Figure 3 shows an image before and after histogram equalization from the MIAS dataset.

(1) *Unlabeled Data Removal*

Removing any unlabeled or mislabeled data is essential for a better classification process before feature extraction and training. It has been observed that some samples have no labels for the coordinates of the tumor. This issue has been observed in four samples leading to their removal. Therefore, the dataset is reduced from 322 to 318 samples.

(1) *ROI Extraction*

Breast cancer classification using the whole image is difficult, as the tumors constitute a small region of the image. Consequently, it is important to crop the image focusing on the ROI where the tumor is present. ROI extraction is carried out using the provided data in the MIAS dataset, which coordinates the center of the tumor and the radius of the circle that encloses the tumor. This procedure helps in removing unnecessary information such as the annotations and pictorial muscles. For the normal cases, the ROI is extracted from the center of the image with a radius equal to the radii's average found in the abnormally provided ROI annotations.

(A)*Feature Extraction*

The main purpose of feature extraction is to capture the important characteristics of the image (Arjun Kumar Joginipelly, 2014) (Arjun Joginipelly et al., 2012).

(1) *Gray Level Co-Occurrence Matrix (GLCM)*

GLCM was first introduced in (Haralick et al., 1973) as a way of computing textural features based on gray tone spatial dependencies. These dependencies can capture different distributions of pixel values, which can detect important features like edges and smoothness. GLCM can be considered as a histogram of pixel-level pairs, i.e., it tells us how many times the pixel level pair (x, y) (where x and y are pixel levels values) appears. For any image, the GLCM matrix can be constructed as a two-dimensional matrix of the image's pixel-level pairs.

To construct pixel-level pairs, it is necessary to define two parameters: d (distance between them) and A (angle between them). As the distance, d is increased, features from distant pixels are captured, whereas, the less the value of d, features from adjacent pixels are captured.

Figure 4. Illustration of varying size neighborhood by changing *p* and *r* values to construct the LBP feature vector.

$p = 4$ $p = 8$ $p = 4$
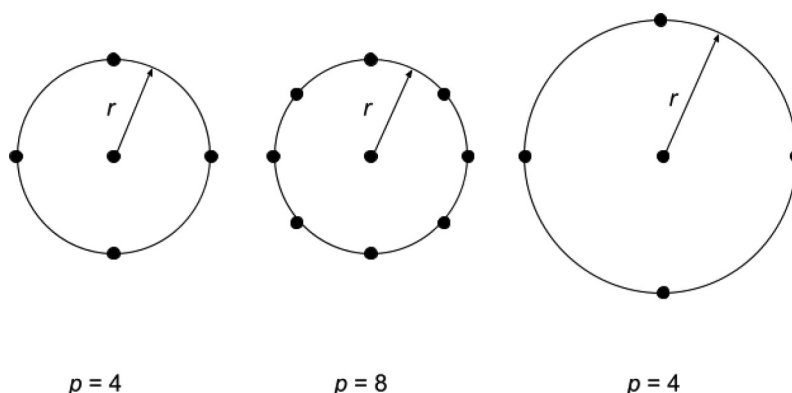
The angle defines the direction of motion to construct pairs. Typically, it is (0, 90, 180, 270 degrees). Each angle represents a different direction to form pairs. An isotropic GLCM is defined as making the matrix for each angle, followed by obtaining the matrices average. Following are the nine different extracted features:

(1) Correlation: It defines how each pixel is correlated with its neighbor.
(2) Contrast: It defines the difference in contrast between each pixel and its neighbor.
(3) Angular Second Moment: It measures the textural uniformity of an image (Gong et al., 1992).
(4) Like contrast, but the motion is linear instead of moving exponentially from a pixel.
(5) Inverse Difference Moment: It measures image homogeneity.
(6) Entropy: It measures the disorder of an image.
(7) Sum Average: It measures the mean of the grey level sum distribution of the image.
(8) Cluster Prominence: It reflects the level of asymmetry in an image. When it is high, the image is less symmetric.
(9) Cluster Shade: It measures the skewness of the GLCM matrix.
   (1) *Local Binary Pattern (LBP)*

LBP is a texture descriptor popularized by the performed work in (Ojala et al., 2002). In the LBP algorithm, a local representation of the texture is constructed by comparing each pixel with its surrounding neighbors. For each pixel in the image, a neighborhood of size r is selected to compute the LBP for this pixel. For example, suppose that the thresholding of a center pixel against the surrounding eight pixels is computed; the computation is carried out such that if the center pixel is equal to or greater than the neighbor pixel, the value is set to one. Otherwise, the value is set to zero. This thresholding creates the LBP code. The LBP value for the center pixel is calculated by flattening the neighboring eight pixels after thresholding, starting from the upper right pixel and rotating clockwise or counter-clockwise consistently across all the data samples, followed by converting the resulting binary 8-bit number into a decimal number. This process is carried out for every pixel in the image. The final feature vector is a histogram ranging from 0 to 255 for a 3 × 3 neighborhood in this example. The advantage of this algorithm is that it captures fine-grained details; however, it cannot capture details at varying scales. To accommodate for this problem, two hyper-parameters are introduced in (Ojala et al., 2002) to address the problem of variable neighborhood sizes:

• p: The number of points in a circularly symmetric neighborhood.
• r: The radius of a circle allowing to consider different scales.

Figure 4 demonstrates how a pixel's neighborhood varies by changing the values of r and p. The value of r represents the distance between the center pixel and the neighboring pixels, and the value of p represents the number of the neighboring pixels. If the left neighborhood has a value of $r = R$ and $p = P$, the middle neighborhood has a value of $r = R$ and $p>P$, and the right neighborhood has a value of $r>R$ and $p = P$.

The resulting histogram is of dimension $p + 2$, in which $p + 1$ represents the uniform patterns, and the added term represents the non-uniform patterns. The values of p and r are chosen after iterations to get the best features for classification. In the first classification stage, the values of p and r are 16 and 5, respectively. The values of p and r for the second stage are 10 and 3, respectively. This hyper-parameter selection results in several features of 18 and 12 for the first and second stages, respectively.

   (1) *Data Standardization*

⚛ cogent · engineering

Data standardization is considered an essential step in data preprocessing when the values of the various extracted features are significantly different in magnitudes and with other units. These different values result in a negative impact on the machine learning algorithms and might influence their performance.

Data standardization is used to statistically normalize the features to follow a Gaussian normal distribution with zero mean and unit variance. Consequently, treating all the data skewness as well as decreasing the number of data outliers. Data standardization could also be seen in the significant declination of time consumed in the training of the machine learning algorithms.

The data standardization is carried out using the basic Gaussian normal distribution given in (1), where Z represents transformed sample value, x represents the original sample value, μ represents the mean of the feature, and σ represents the standard deviation of the feature (Abdulhussain et al., 2020).

$$Z = \frac{x - \mu}{\sigma} \tag{1}$$

(A)*Evaluation Metrics*

In binary and multi-class classification, especially when the problem is related to health care, it is crucial not to depend only on the classifier's average accuracy as an evaluation metric but to assess the classifier's performance differently from other evaluation metrics. This leads to a better understanding of the classifier's behavior and evaluates each dataset class's performance. This study uses precision, recall, F1-score, MCC, and accuracy as evaluation metrics for the problem of breast cancer classification.

In the following equations, *TP* indicates true positives, *TN* indicates true negatives, *FP* indicates false positives, and *FN* indicates false negatives.

Table 3. First stage classification report					
Class	**Precision**	**Recall**	**F1 score**	**MCC**	**Accuracy**
Normal	0.95	1.00	0.98	2*0.95	2*97%
Abnormal	1.00	0.95	0.97		

Figure 5. First stage confusion matrix.

cogent··engineering

Table 4. Second stage classification report					
Class	**Precision**	**Recall**	**F1 score**	**MCC**	**Accuracy**
Benign	0.73	0.80	0.76	2*0.5	2*75%
Malignant	0.78	0.70	0.74		

Table 5. Hierarchical system classification report					
Class	**Precision**	**Recall**	**F1 score**	**MCC**	**Accuracy**
Normal	0.95	1.00	0.98	0.76	85%
Benign	0.70	0.70	0.70		
Malignant	0.78	0.70	0.74		

(1) *Precision*

Precision is the proportion of true positive samples out of all classified samples as positive. It shows the ability of the classifier not to label a negative sample as positive. (2) shows the formula used to calculate the precision (Abdulhussain et al., 2020).

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

(1) *Recall*

The recall is the number of true positive samples that are correctly classified. It shows the ability of the classifier to classify all positive samples correctly. (3) shows the formula used to calculate recall.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

(1) *F1-score*

Figure 6. Second stage confusion matrix.

Figure 7. Hierarchical system confusion matrix.

Table 6. Hierarchical system processing time			
Stage	**Operation**	**Number of samples**	**Processing time**
1	Feature extraction	40	125.02 s
1	Classification	40	7.64 ms
2	Feature extraction	19	53.51 ms
2	Classification	19	7.14 ms

F1-score is the harmonic mean of precision and recall. (4) shows the formula used to calculate F1-score.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

(1) *Matthews Correlation Coefficient (MCC)*

MCC is used as a measure of the quality of binary classification. It is a balanced measure that considers true and false positives and negatives, even if the class sizes are different. The MCC ranges between −1 and +1. An MCC of +1 indicates a perfect prediction, 0 indicates a random prediction, while −1 indicates that the predictions totally disagree with the true labels. (5) shows the formula used to calculate the MCC.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{5}$$

(1) *Accuracy*

Accuracy is a measure of the correctly classified samples. (6) shows the formula used to calculate the accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

4. Results and discussion

This section discusses the results obtained in the first and second stages of the hierarchical classification system and the overall results after combining the two stages.

Table 7. Comparison between the proposed work and other related studies

Study	(Bektas et al., 2018) Random Forest			(Charan et al., 2018) Discrete Krawtchouk			(Saraswathi et al., 2016) Swarm Intelligence			(Abirami et al., 2016) Neural Network			(Setiawan et al., 2015) Neural Network			(Kashif, 2020) SVM			(Shi et al., 2019) CNN			(Saber et al., 2021) VGG16 + SVM			(Yu et al., 2020) VGG16			Proposed		
Class	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M
Precision	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0.9	-	-	-	0.99	0.97	0.956	0.9474		0.8077	0.95	0.70	0.78
Recall	-	-	-	-	-	-	-	-	-	-	-	-	0.91	-	0.8	-	0	1.00	-	-	-	0.989	0.99	0.956	0.878		0.913	1.00	0.70	0.70
F1-score	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.95	-	-	-	0.99	0.98	0.96	0.9114		0.8571	0.98	0.70	0.74
MCC	-			-			-			-			-			-			-			-			-			0.76		
Accuracy	65%			65%			92%			95.5%			83.8%			90%			83.6%			98.87%			89.06%			85%		

Table 8. Improvement ratio between the proposed work and other related studies

Study	(Bektas et al., 2018) Random Forest			(Charan et al., 2018) Discrete Krawtchouk			(Saraswathi et al., 2016) Swarm Intelligence			(Abirami et al., 2016) Neural Network			(Setiawan et al., 2015) Neural Network			(Kashif, 2020) SVM			(Shi et al., 2019) CNN			(Saber et al., 2021) VGG16 + SVM			(Yu et al., 2020) VGG16		
Class	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M	N	B	M
Precision	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.11	-	-	-	0.95	0.72	0.81	1.002		1.23
Recall	-	-	-	-	-	-	-	-	-	-	-	-	1.09	-	0.875	INF		0.95	-	-	-	1.01	0.7	0.73	1.13		1.04
F1-score	-	-	-	-	-	-	-	-	-	-	-	-	1.09	-	0.875	-		1.02	-	-	-	0.98	0.71	0.77	1.07		1.13
Accuracy	1.3			1.3			0. 92			0.89			1.01			0.94			1.01			0.85			0.95		

Evaluation metrics such as precision, recall, F1-score, MCC, and accuracy are used to evaluate the classifier's performance. In addition, the MCC evaluation metric is used to assess the quality of the binary classification.

(A)*First Classification Stage (Normal/Abnormal)*

The first classification stage differentiates between normal and abnormal cases. Different classifiers such as SVM, k-nearest neighbors (KNN), XGBoost, Decision Tree, Random Forest, and Naive Bayes are tested as the first stage classifier. GLCM and LBP feature extraction algorithms are performed to extract features for the first stage classification. It has been observed that the best performing classifier is Random Forest, and the best feature extraction algorithm is GLCM.

Table 3 shows the classification report of the first stage based on Random Forest as a classifier combined with GLCM for feature extraction. The achieved precision, recall, and F1-score for the normal class are 0.95, 1.00, and 0.98, respectively. For the abnormal class, the achieved precision, recall, and F1-score are 1.00, 0.95, and 0.97, respectively. Thus, the MCC and accuracy for the first classification stage are 0.95% and 97%, respectively.

Figure 5 demonstrates the confusion matrix for the first classification stage. The Random Forest classifier successfully classified all the normal cases correctly and misclassified one abnormal case as normal.

(A)*Second Classification Stage (Benign/Malignant)*

The second classification stage differentiates between benign and malignant cases. Different classifiers such as SVM, KNN, XGBoost, Decision Tree, Random Forest, and Naive Bayes are tested as the second stage classifier. GLCM and LBP feature extraction algorithms are performed to extract features for the second classification stage. It has been observed that the best performing classifier is Random Forest, and the best feature extraction algorithm is LBP.

Table 4 shows the second stage classification report based on Random Forest as a classifier in combination with LBP for feature extraction. The achieved precision, recall, and F1-score for the benign class are 0.73, 0.80, and 0.76, respectively. The achieved precision, recall, and F1-score for the malignant class are 0.78, 0.70, and 0.74, respectively. The MCC and the accuracy for the second classification stage are 0.5% and 75%, respectively.

It is evident that classifying a tumor, whether benign or malignant, is more complicated than classifying an ROI into whether it is normal or abnormal, as the achieved accuracy for the first and second stages are 97% and 75%, respectively.

Figure 6 demonstrates the confusion matrix for the second classification stage. The Random Forest classifier successfully classified eight benign cases correctly and misclassified two benign cases as malignant. Seven samples are correctly classified as malignant for the malignant class, while three samples are misclassified as benign.

(A)*Hierarchical Classification*

The hierarchical classification system is composed of the best feature extraction algorithm and a machine learning classifier for the first and second classification stages based on the results mentioned above in a hierarchical structure. The whole test set contains 40 images in which 20 samples are labeled as normal, and the other 20 samples are labeled as abnormal. The 20 abnormal samples are further labeled as 10 benign samples and 10 malignant samples. Thus, the whole test set passes through the first classification stage to classify the images as normal or

abnormal. The first stage classifier's samples classified as abnormal are further passed to the second classification stage to be classified as benign or malignant.

Table 5 Hierarchical system classification of the hierarchical classification system based on Random Forest as a classifier, combined with GLCM and LBP for feature extraction in the first and second stages, respectively. The achieved precision, recall, and F1-score for the normal class are 0.95, 1.00, and 0.98, respectively. For the benign class, the achieved precision, recall, and F1-score are 0.70, 0.70, and 0.70, respectively. Finally, the achieved precision, recall, and F1-score for the malignant class are 0.78, 0.70, and 0.74, respectively. Thus, the MCC and accuracy for the overall hierarchical classification system are 0.76% and 85%, respectively.

Obviously, the first classification stage results are better than that of the second classification stage, with an F1-score of 0.98 and 0.97 for the normal and abnormal classes, respectively. This indicates that the proposed solution can be used to assist physicians in classifying mammography images into normal and abnormal with high confidence. Moreover, the results show that classifying an abnormal sample into benign or malignant is a relatively difficult task. This could be due to some similarities between benign and malignant cases. The extracted features do not represent all the important distinctions between the two classes, resulting in a lower F1-score than the first classification stage.

Figure 7 demonstrates the confusion matrix for the overall hierarchical classification system. All the normal cases are classified correctly. For the benign class, seven samples are classified correctly, two samples are classified as malignant, and one sample is classified as benign. The malignant class has seven correct predictions, with three misclassified samples as benign.

(A)*Processing Time*

The processing time is calculated for each classification stage in terms of the time taken by the feature extraction and classification processes. Forty images are processed in the first stage, and 19 abnormal images are processed in the second stage. The processing time is calculated on Intel(R) Xeon(R) CPU @ 2.20 GHz provided by the Google Collaboratory. Table 6 shows the processing time taken by the feature extraction and classification processes in both the first and second stages of the hierarchical system. It has been observed that the GLCM feature extraction technique takes a relatively longer processing time of 125.02 s for 40 images. The LBP feature extraction technique takes 53.51 ms for 19 images. The classification process takes between 7.64 ms and 7.14 ms for 40 and 19 images for the first and second stage classifiers.

(A)*Comparison*

The proposed work is compared with other related studies investigating breast cancer detection and classification using the MIAS dataset. Tables 7 and Tables 8 show the comparison and improvement ratio in results between the studies and the proposed approach. The studies are compared according to precision, recall, F1-score, MCC, and accuracy. It is important to note that some studies include the results between normal and abnormal cases only. These results are compared to our first stage results in Table 8.

The proposed approach achieves higher accuracy than (Bektas et al., 2018) (Charan et al., 2018), (Shi et al., 2019) and (Setiawan et al., 2015). The reported accuracy in (Saraswathi et al., 2016) and (Abirami et al., 2016) is higher than in the proposed work. However, these studies do not include the results of other evaluation metrics such as precision, F1-score, and MCC, giving a better explanation for the classifier's performance on different classes. The proposed approach outperforms (Kashif, 2020) in the precision of abnormal, the recall of normal, and F1-score of abnormal while (Kashif, 2020) exceeds our approach in the accuracy. The proposed approach exceeds (Saber et al., 2021) only on the recall of the normal. While (Yu et al., 2020) outperforms our approach in terms of accuracy only. On the other hand, the proposed approach outperforms it on the other

evaluation metrics. A detailed classification report is provided in the proposed work, which explains the strengths and weaknesses of the classification system. This can help physicians to know when to rely on the proposed solution as an assistant with higher confidence.

5. Conclusion

In this paper, breast cancer detection and classification using machine learning are performed on the MIAS dataset using a hierarchical system. In the first classification stage (Normal/Abnormal) it achieves an accuracy of 97%, and the second classification stage (Benign/Malignant) it achieves an accuracy of 75%. The whole hierarchical classification system achieves an accuracy of 85%. This paper shows that differentiating benign from malignant tumors is a difficult task compared to differentiating normal from abnormal breasts. It has been observed that LBP provides better results than GLCM as a feature extraction technique for the second classification stage. The proposed system can be used to assess physicians in diagnosing breast cancer, and the provided classification reports highlights the strengths and weaknesses of the system. More research can be done to enhance the results of the second classification stage.

Author details
M. Saeed Darweesh[1,2]
E-mail: mdarweesh@nu.edu.eg
ORCID ID: http://orcid.org/0000-0001-9684-3325
Mostafa Adel[3]
E-mail: mostafaadel610@gmail.com
Ahmed Anwar[4]
E-mail: ahmed.anwar.saeed@gmail.com
Omar Farag[4]
E-mail: Omar16farag16@gmail.com
Ahmed Kotb[4]
E-mail: ak.ahmedkotb10@gmail.com
Mohamed Adel[5]
E-mail: madel.badwi@pg.cu.edu.eg
Ayman Tawfik[6]
E-mail: a.tawfik@ajman.ac.ae
ORCID ID: http://orcid.org/0000-0003-1438-6541
Hassan Mostafa[47]
E-mail: hassanmostafahassan@gmail.com
[1] Wireless Intelligent Networks Center (WINC), Nile University, Giza, Egypt.
[2] School of Engineering and Applied Sciences, Nile University, Giza, Egypt.
[3] Electronics and Communications Engineering Department, Cairo University, Giza, Egypt.
[4] Department of Electronics Program, German University in Cairo (Guc), Cairo, Egypt.
[5] Department of Engineering Applications of Laser, National Institute of Laser Enhanced Science, Cairo University, Egypt.
[6] Department of Electrical and Computer Engineering, Ajman University, Ajman, UAE.
[7] University of Science and Technology, Nanotechnology and Nanoelectronics Program, Zewail City of Science and Technology, Giza, Egypt.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Citation information
Cite this article as: Early breast cancer diagnostics based on hierarchical machine learning classification for mammography images, M. Saeed Darweesh, Mostafa Adel, Ahmed Anwar, Omar Farag, Ahmed Kotb, Mohamed Adel, Ayman Tawfik & Hassan Mostafa, *Cogent Engineering* (2021), 8: 1968324.

References
Abdul-Hadi, A. M., Abdulhussain, S. H., Mahmmod, B. M., & Pham, D. On the computational aspects of Charlier polynomials. (2020). *Cogent Engineering* Ed. by Duc Pham, *7*(1), 1763553. https://doi.org/10.1080/23311916.2020.1763553

Abdulhussain, S., Al-Haddad, S. A. R., Saripan, M. I., Mahmmod, B. M., & Hussien, A. (2020). Fast Temporal video segmentation based on Krawtchouk-Tchebichef moments. *IEEE Access, 8,* 72347–72359. https://doi.org/10.1109/ACCESS.2020.2987870

Abdulhussain, S. H., & Mahmmod, B. M. (2021). Fast and efficient recursive algorithm of Meixner polynomials. *Journal of Real-Time Image Processing.* https://doi.org/10.1007/s11554-021-01093-z

Abirami, C., Harikumar, R., & Chakravarthy, S. R. S. (2016). Performance analysis and detection of micro calcification in digital mammograms using wavelet features. In *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 2327–2331), Chennai, India. IEEE.

Adel, M., Kotb, A., Farag, O., M. Saeed Darweesh, & Mostafa, H. (2019). Breast cancer diagnosis using image processing and machine learning for elastography images. In *2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST)* (pp. 1–4), Thessaloniki, Greece. IEEE.

Ahmad, L. G., Eshlaghy, A. T., Poorebrahimi, A., Ebrahimi, M., and Razavi, A. R. (2013). Using three machine learning techniques for predicting breast cancer recurrence. *Journal of Medical Informatics, 4*(124), 3. https://doi.org/10.4172/2157-7420.1000124

Bahya Hospital. Retrieved July 19, 2018, from https://www.baheya.org

Barr, R. G. (2012). Sonographic breast elastography: A primer. *Journal of Ultrasound in Medicine, 31*(5), 773–783. https://doi.org/10.7863/jum.2012.31.5.773

Becker, A. S., Marcon, M., Ghafoor, S., Wurnig, M. C., Frauenfelder, T., & Boss, A. (2017). Deep learning in mammography: Diagnostic accuracy of a multipurpose image analysis software in the

detection of breast cancer. *Investigative Radiology*, *52*(7), 434–440. https://doi.org/10.1097/RLI.0000000000000358

Bektas, B., Emre ,İ., Kartal, E., & Gulsecen, S. (2018). "Classification of mammography images by machine learning techniques". In *2018 3rd International Conference on Computer Science and Engineering (UBMK)* (pp. 580–585). IEEE.

Charan, S., Khan, M. J., & Khurshid, K. (2018). "Breast cancer detection in mammograms using convolutional neural network". In: *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)* (pp. 1–5), Sarajevo, Bosnia, and Herzegovina. IEEE.

Cheung, P., & Donlon, ca. 2016. MRI upgrades at radiology associates. San Luis Obispo: California Polytechnic State University.

EE, P., JE, J., & Db, P. 2005. *SavingWomen's lives: Strategies for improving breast cancer detection and diagnosis* Institute of Medicine (US), National Research Council (US) Committee on New Approaches to Early Detection, and Diagnosis of Breast Cancer. National Academies Press (US.

Gbenga, D. E., Christopher, N., & Yetunde, D. C. (2017). Performance comparison of machine learning techniques for breast cancer detection. *Nova*, *6*(1), 1–8. https://doi.org/10.20286/nova-jeas-060105

Gong, P., Marceau, D. J., & Howarth, P. J. (1992). A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data. *Remote Sensing of Environment*, *40*(2), 137–151. https://doi.org/10.1016/0034-4257(92)90011-8

Han, Z., Haralick, R., Shanmugam, K.,& Dinstein (2017). Breast cancer multi-classification from histopathological images with structured deep learning model. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1109/TSMC.1973.4309314

Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *6*(6), 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Joginipelly, A. et al. (2012). Efficient FPGA implementation of steerable Gaussian smoothers. In *Proceedings of the 2012 44th Southeastern Symposium on System Theory (SSST)* (pp. 78–82). IEEE,Jacksonville, FL, USA https://doi.org/10.1109/SSST.2012.6195131.

Joginipelly, A. K. 2014. *Efficient FPGA architectures for separable filters and logarithmic multipliers and automation of fish feature extraction using Gabor filters*. University of New Orleans Theses and Dissertations.

Kashif, M. (2020). Chapter 6 - Application of machine learning and image processing for detection of breast cancer. In M. D. Lytras, A. Sarirete, et al (Eds.), *Innovation in health informatics* ISBN: 978-0-12-819043-2. (pp. 145–162). Next Gen Tech Driven Personalized MedSmart Healthcare. Academic Press.

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, *13*, 8–17. https://doi.org/10.1016/j.csbj.2014.11.005

Kowal, M., Filipczuk, P., Obuchowicz, A., Korbicz, J., & Monczak, R. (2013). Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. *Computers in Biology and Medicine*, *43*(10), 1563–1572. https://doi.org/10.1016/j.compbiomed.2013.08.003

Mahmmod, B. M., Abdul-Hadi, A. M., Abdulhussain, S. H., & Hussien, A. (2020). On computational aspects of Krawtchouk polynomials for high orders. *Journal of Imaging*, *6*(8), 81. https://doi.org/10.3390/jimaging6080081

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G. S., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F. J., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C. J., King, D., Ledsam, J. R., & Shetty, S. (2020, January). International evaluation of an AI system for breast cancer screening. *Nature*, *577*(7788), 89–94. https://doi.org/10.1038/s41586-019-1799-6

MIAS database. Retrieved January 1, 2021. http://peipa.essex.ac.uk/info/mias.html

Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(7), 971–987. https://doi.org/10.1109/TPAMI.2002.1017623

Öztürk, Ş., & Akdemir, B. (2018). Application of feature extraction and classification methods for histopathological image using GLCM, LBP, LBGLCM, GLRLM and SFTA. *Procedia Computer Science*, *132*, 40–46. https://doi.org/10.1016/j.procs.2018.05.057

Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., Ter Haar Romeny, B., Zimmerman, J. B., & Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, *39*(3), 355–368. https://doi.org/10.1016/S0734-189X(87)80186-X.

Saber, A., Sakr, M., Abo-Seida, O. M., Keshk, A., & Chen, H. (2021). A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique. *IEEE Access*, *9*, 71194–71209. https://doi.org/10.1109/ACCESS.2021.3079204

Saraswathi, D., Dharani, D., & Srinivasan, E. (2016). "An efficient feature extraction technique for breast cancer diagnosis using curvelet transform and swarm intelligence". In: *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 441–445), Chennai, India. IEEE.

Setiawan, A. S., Wesley, J., Purnama, Y., & Purnama, Y. (2015). Mammogram classification using law's texture energy measure and neural networks. *Procedia Computer Science*, *59*, 92–97. https://doi.org/10.1016/j.procs.2015.07.341

Shi, P., Wu, C., Zhong ,J., & Wang, H. (2019). Deep learning from small dataset for BI-RADS density classification of mammography images. In *2019 10th International Conference on Information Technology in Medicine and Education (ITME), Qingdao, China* (pp. 102–109). https://doi.org/10.1109/ITME.2019.00034.

Siegel, R. L., Miller, K., Sauer, A., Fedewa, S., Butterly, L., Anderson, J., Cercek, A., Smith, R., & Jemal, A. (2020). *Colorectal cancer statistics, 2020*. CA: a cancer journal for clinicians.

Siegel, R. L., Miller, K. D., & Jemal, A. (2016). Cancer statistics, 2016. *CA: A Cancer Journal for Clinicians*, *66*(1), 7–30. https://doi.org/10.3322/caac.21332

Sreehari Sastry, S., Vindhya Kumari, Nageswara Rao, C., Mallika, K., Lakshminarayana, S., & Tiong, H. (2012). Transition temperatures of thermotropic liquid crystals from the local binary gray level cooccurrence matrix. *Advances in Condensed Matter Physics*. https://doi.org/10.1155/2012/527065

Unger-Saldaña, K. (2014). Challenges to the early diagnosis and treatment of breast cancer in developing

countries. *World Journal of Clinical Oncology, 5*(3), 465. https://doi.org/10.5306/wjco.v5.i3.465

Warburg, O. (1956). On the origin of cancer cells. *Science, 123*(3191), 309–314. https://doi.org/10.1126/science.123.3191.309

Yu, X., Pang, W., Xu, Q., & Liang, M. (2020, September). Mammographic image classification with deep fusion

learning. *Scientific Reports, 10* (1), 14361. https://doi.org/10.1038/s41598-020-71431-x

Zhou, B.-G., Wang, D., Ren, W., Li, X., He, Y., Liu, B., Wang, Q., Chen, S., Alizad, A., & Xu, H. (2017). Value of shear wave arrival time contour display in shear wave elastography for breast masses diagnosis. *Scientific Reports, 7*(1), 1–9. https://doi.org/10.1038/s41598-017-07389-0

Cogent Engineering (ISSN: 2331-1916) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com