



Segmentation and classification of breast cancer using novel deep learning architecture

S. Ramesh¹ · S. Sasikala² · S. Gomathi² · V. Geetha² · V. Anbumani²

Received: 16 October 2020 / Accepted: 29 March 2022 / Published online: 25 May 2022
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

Abstract

Breast cancer is one of the most frequent cancers in women, and it has a higher mortality rate than other cancers. As a result, early detection is critical. In computer-assisted disease diagnosis, accurate segmentation of the region of interest is a vital concept. The segmentation techniques have been widely used by doctors and physicians to locate the pathology, identify the abnormality, compute the tissue volume, analyze the anatomical structures, and provide treatment. Cancer diagnostic efficiency is based on two aspects: The precision value associated with the segmentation and calculation of the tumor area and the accuracy of the features extracted from the images to categorize the benign or malignant tumors. A novel deep-learning architecture for tumor segmentation is therefore proposed in this study, and machine learning algorithms are used to categorize benign or malignant tumors. The segmentation results improve the decision-making capability of the physicians to identify whether a tumor is malignant or not and normally, the machine learning techniques need expert annotation and pathology reports to identify this. This challenge is overcome in this work with the help of the GoogLeNet architecture used for segmentation. The segmentation results are then offered to the Support Vector Machine, Decision Tree, Random Forest, and Naïve Bayes classifier to improve their efficiency. Our work has provided better results in terms of accuracy, Jaccard and dice coefficient, sensitivity, and specificity compared to conventional architectures. The proposed model offers an accuracy score of 99.12% which is relatively higher than the other techniques. A 3.78% accuracy improvement is noticed by the proposed model against the AlexNet classifier and the actual increase is 4.61% on average when compared to the existing techniques.

Keywords Breast cancer segmentation · Breast cancer classification · MIAS · Deep learning · Machine learning

1 Introduction

Breast cancer is ranked among the first foremost causes of female cancer worldwide. It mostly occurs when cancer cells grow in the breast tissues. A tumor is a mass of abnormal tissues, and there are two types of breast cancer tumors (normal and malignant). There are numerous methods to diagnose breast cancer, like breast self-examination (BSE), mammography or imaging, clinical breast examination (CBE), and surgery. The most reliable method for breast cancer screening and diagnosis is a mammogram, which can identify 85 to 90% of all breast cancers. Masses and calcifications are the most commonly reported anomalies that signify breast cancer.

A mass examined on a mammogram may be either benign or malignant based on its shape. Benign tumors typically have oval or circular forms, whereas malignant

✉ S. Ramesh
rameshsundars.sse@saveetha.com; rameshbe04@gmail.com

S. Sasikala
sasikalas.ece@kongu.edu

S. Gomathi
sgomathi.ece@kongu.edu

V. Geetha
vgeetha.ece@kongu.edu

V. Anbumani
anbumani.ece@kongu.ac.in

¹ Department of Artificial intelligence and Machine learning, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India

² Department of ECE, Kongu Engineering College, Perundurai, Erode, Tamilnadu, India

tumors have a somewhat circular form with a sharp or uneven outer surface. Benign or noncancerous tumors consist of fibroadenomas, breast hematomas, and cysts. A malignant or cancerous tumor mostly in the breast is a collection of abnormal and uncontrolled growth of breast tissue [1]. Therefore, early diagnosis is very important. The biggest challenge is to determine the location of the tumor and the level of severity. A mammogram image is used in a variety of image processing methods to diagnose pathology. These methods are as follows: I. Image enhancement, II. Segmentation, III. Feature extraction and IV. Classification.

Figure 1 demonstrates the overall process flow for the proposed approach. Image enhancement is the first step. Image enhancement was performed by Contrast Limited Adaptive Histogram Equalization (CLAHE) methodology [60]. The second step is segmentation, and a new novel deep-learning-based model is suggested for segmentation.

It extracts the affected area from the enhanced image. The third step is the extraction feature. A gray-level co-occurrence matrix (GLCM) and shape strategies are utilized to extract the valuable features from the segmented image. Finally, by using machine learning approaches including decision tree, random forest, SVM [2], and Naive Bayes [3], it is used to determine whether it is normal or abnormal and benign or malignant.

1.1 Image enhancement

Image enhancement in a mammogram is a way of adjusting mammogram images to improve their brightness and decrease the noise seen to assist radiologists in identifying abnormalities. CLAHE methodology is used in this paper for image enhancement.

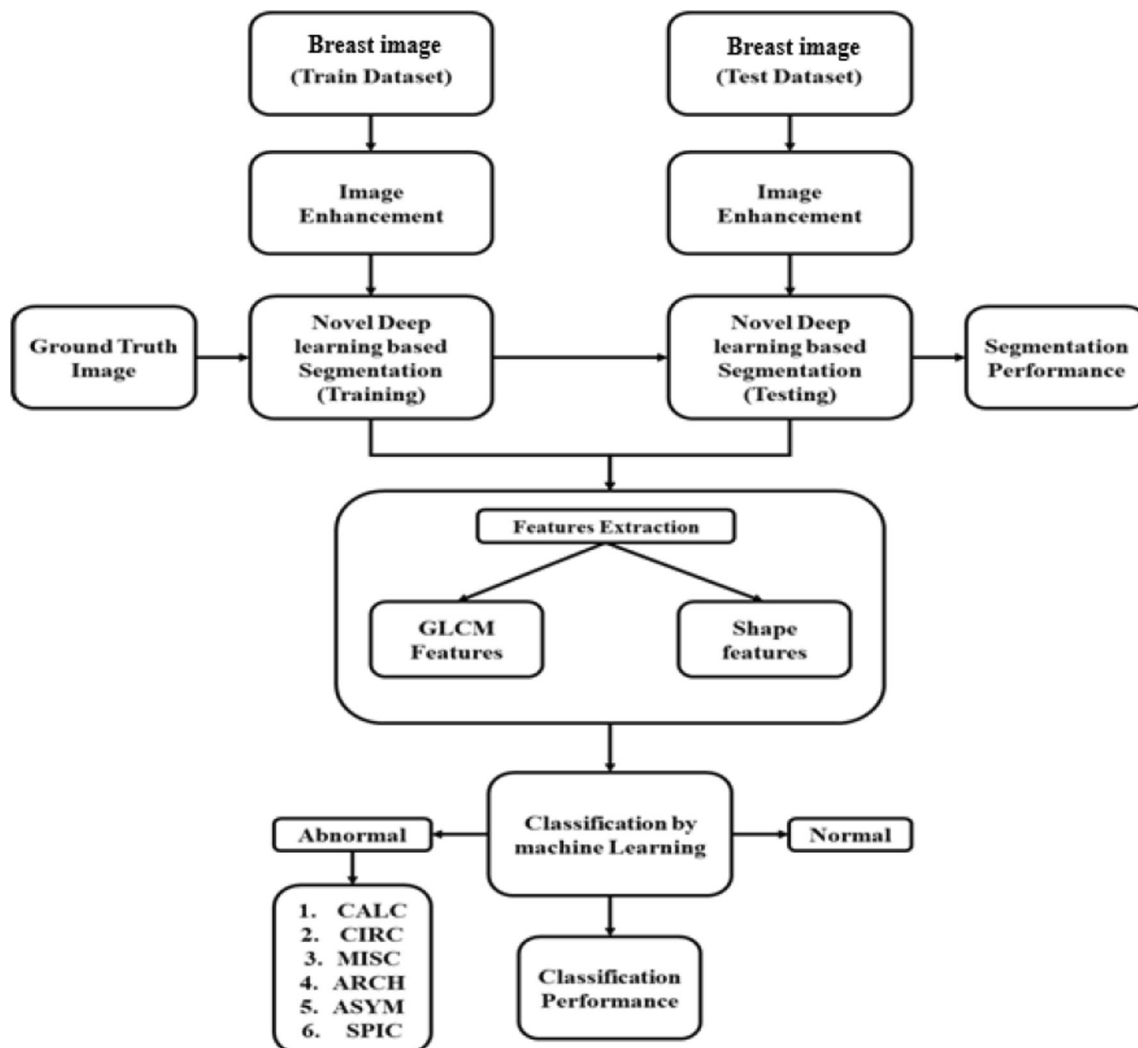


Fig. 1 Overall process flow for the proposed approach

1.2 Segmentation

Recently, many methodologies have been proposed for image segmentation [4]. This study focuses on semantic segmentation. A novel deep learning-based architecture is proposed for segmentation based on GoogLeNet architecture. The GoogLeNet was introduced by Google research which was known by the name Inception v1 in 2014 [5]. This structure won the best ILSVRC 2014 image classification challenge. The error rate was substantially lower relative to previous Alexnet, ZF-Net winners, and slightly lower than VGG. This model uses methods like 1×1 convolution in the center of design and global average pooling to develop a deeper structure.

1.3 Features extraction

Feature extraction acquires more specific information from the segmented image. Commonly shape and GLCM techniques are used for feature extraction [6].

1.4 Classification

Classification generally has 3 types, namely machine learning [7], deep learning [8], and neural network [9]. Here, machine learning algorithms are used for severity level classification.

The remainder of the study is organized as shown below: The related work studies are addressed in Sect. 2. Details of the proposed technique are addressed in Sect. 3. The outcomes and descriptions of the study are provided in Sect. 4. Subsequently, the conclusions are set out in Sect. 5.

2 Review of related works

In the past, most analyses were used to enhance mammogram images and numerous spatial and frequency-domain methods have been explored [10]. A comparative [11] review on digital mammography imaging enhancement mechanisms, including wavelet-based improvement, CLAHE, unsharp masking, and the morphological operator was provided. In digital mammography images, methods have been developed for both regional contrast enhancement and background texture reduction [12, 12]. The CLAHE [14] method is the most widely utilized methodology that results in an improvement of the contrast around medical images.

Previous researches suggested various methods of mammogram image segmentation, including region-growing segmentation techniques [15], contour-based methods

[16, 16], cluster methods [17], threshold-based methods [18], watershed-based techniques [19], and deep learning related techniques [20, 48–54]. There could be a risk of segmentation if an incorrect threshold value is used in threshold-based segmentation [21]. Even several other hybrid variants of the clustering method have been suggested to achieve the best results [22–24]. However, it is difficult to select a number of clusters in k-means and centroids in FCM for cluster-based approaches. Other methods provide low-performance results, except for the deep-learning technique.

A few other complex, efficient architectures were also addressed in the following papers [25–29, 55–57]. Severity level classification done by many proposed methods included SVM [30, 31], decision tree [32], naïve Bayes [33, 34], random forest [31], hybrid version [35], PNN [36], and deep learning [37, 37]. Yiqiu Shen et al. [55] presented a weakly supervised localization technique for high-resolution breast cancer images. Several authors have also used metaheuristic algorithms [58, 59] to enhance classification performance. However, these techniques suffer from high computational time, low training efficiency, manual processing, and low accuracy. To overcome these problems, a novel methodology is presented in this paper and it is explained in detail in the subsequent sections.

3 Methodology

Figure 1 demonstrates the overall process flow for the proposed approach. The proposed model has four stages: Image Enhancement, Segmentation, Feature Extraction, and Classification. Image Enhancement is the first step and it is performed by CLAHE methodology [60]. The second step is segmentation, where a novel deep learning-based architecture is used. It extracts the affected area from the enhanced image. The third step is the feature extraction where the gray-level co-occurrence matrix and the shape strategies are utilized to extract the valuable features from the segmented image. Finally, by using machine learning approaches including decision tree [39], random forest [40], SVM [2], and Naive Bayes [3], the features are classified as benign and malignant.

3.1 Dataset details

The MIAS dataset [47] contains mammography scan images along with their labels. The mammogram images are centered in the matrix and the size of each image is 1024×1024 pixels. In most of the images, calcifications are present in the centers and the radii are mainly applied in clusters rather than individual calcification. In certain

Table 1 MIAS dataset description

| Column name | Description | Categorization |
|-------------|--|--|
| 1 | MIAS Database reference number | – |
| 2 | Background tissue characteristics | F-Fatty G-Fatty-Glandular D-Dense-Glandular |
| 3 | Abnormality Classes | CALC- Calcifications ASYM- Asymmetry CIRC- Circumscribed masses SPIC-Spiculated masses ARCH- Architectural distortion MISC(Other)-ill-defined masses NORM-Normal |
| 4 | Severity of abnormality | Benign and Malignant |
| 5,6 | X, Y | Coordinates of abnormality center |
| 7 | The approximate radius of the circle surrounding the abnormality in pixels | – |

cases, the calcifications are dispersed entirely to the image instead of concentrating on a single site alone. For these cases, both the central location and radii are considered unnecessary and omitted. A detailed description of the MIAS dataset is presented in Table 1. The dataset consists of a total of 322 images where 70% of images are used for training and the remaining 30% of images are used for testing.

3.2 Enhancement

Image enhancement is the digital image adjustment process so that the performance can be more convenient for further image analysis like segmentation. CLAHE is one of the widely used image processing techniques to improve the prediction accuracy by enhancing the regions of the tiny veins which are often ignored during contrast enhancement. The contrast enhancement process of the standard Histogram Equalization (HE) is limited by the CLAHE technique which performs an operation similar to noise

improvement. The main aim of using the CLAHE technique is to limit the noise that occurs during contrast enhancement which serves as a major hurdle for medical images. A histogram is sliced at a certain threshold level and then the formula is implemented. It is an adaptive histogram equalization approach [41, 41], in which the contrast of an image is boosted by implementing CLAHE to limited data sections called tiles instead of the entire image. To patch the desired outcome in adjacent tiles, bilinear interpolation is used. Contrast can be limited within the same region, thus avoiding noise amplification [14].

Contrast enhancement is mainly a slope function that interrelates the intensity values of the input image to generate the desired output image intensities. When the slope of the relating function is controlled, the contrast value is minimized. The height of a histogram at a particular intensity mainly represents the contrast enhancement. Contrast enhancement is mainly controlled by limiting the slope value and clipping the histogram height. The CLAHE algorithm mainly limits the contrast via a clip limit.

The clipping limit (CL) of the CLAHE algorithm is mainly shown in Eq. (1)

$$CL = \left[\frac{\phi}{GS} \right] + \left[\alpha \cdot \left(\psi - \left[\frac{\psi}{GS} \right] \right) \right] \quad (1)$$

The controllable threshold value of our proposed methodology is explained in Eq. (2)

$$CL = \frac{GT}{80} \quad (2)$$

In Eqs. (1) and (2), GS is the grayscale value, ψ is the pixel population of each block, GT is the global threshold, and the clip factor is represented as α .

3.3 Segmentation

3.3.1 GoogleNet layer description

Google's research resulted in the creation of GoogLeNet, which was described as inception V1 in 2014 [5]. This structure won the best ILSVRC 2014 image classification challenge. The error rate was substantially lower relative to previous Alexnet, ZF-Net winners, and slightly lower than VGG. This model uses methods like 1×1 convolution in the center of design and global average pooling to develop a deeper structure.

GoogLeNet is a deep convolutional, 22-layer wide, neural network. It is a pre-trained architecture that uses a places365 or ImageNet dataset for evaluation. The network trained on ImageNet categories which comprise 1000 types of objects, such as a mouse, keyboard, pencils, and various

Table 2 Novel GoogLeNet architecture after improvement

| S. No | Layers | Output size | Number of kernels | Kernel size | Depth |
|-------|----------------------|-------------|-------------------|-------------|-------|
| 1 | Input image | 224*224*3 | – | – | 0 |
| 2 | Convolution | 224*224*64 | 64 | 7*7*7 | 1 |
| 3 | Convolution | 224*224*64 | – | – | 2 |
| 4 | ReLU | Activation | – | – | 1 |
| 5 | convolution | 112*112*64 | – | – | 2 |
| 6 | Max pooling | 56*56*64 | – | 3*3*3 | 0 |
| 7 | Convolution | 56*56*192 | 192 | – | 2 |
| 8 | Max pooling | 28*28*192 | – | 3*3*3 | 0 |
| 9 | Inception | 28*28*256 | – | – | 2 |
| 10 | Inception | 28*28*480 | – | – | 2 |
| 11 | Max pooling | 14*14*480 | – | 3*3*3 | 0 |
| 12 | Inception | 14*14*512 | – | – | 2 |
| 13 | Inception | 14*14*512 | – | – | 2 |
| 14 | Inception | 14*14*512 | – | – | 2 |
| 15 | Inception | 14*14*512 | – | – | 2 |
| 16 | Inception | 14*14*832 | – | – | 2 |
| 17 | Max pooling | 7*7*832 | – | 3*3*3 | 0 |
| 18 | Inception | 7*7*832 | – | – | 2 |
| 19 | Inception | 7*7*1024 | – | – | 2 |
| 20 | Fully connected | 4096 | – | – | 0 |
| 21 | Fully connected | 4096 | – | – | 0 |
| 22 | Pixel classification | 4*1*1 | – | – | 0 |

animals. Places365 is closely related to the ImageNet-trained network but categorizes images through 365 different categories of places, such as ground, park, lobby, and runway. Transfer learning is a potential way in GoogLeNet to train another image dataset.

The first convolutional layer in the GoogLeNet utilizes a patch size of 7×7 which is relatively higher than the remaining patches used in the network and the main purpose of this layer is to minimize the size of the input image without losing the spatial features. The size of the input image is reduced by a factor of four when it reaches the second convolutional layer and before reaching the inception module it is reduced by a factor of eight. These processes mainly generate a larger number of feature maps. The second convolutional layer uses a 1×1 convolutional block with a depth of 2. The main aim of this 1×1 convolutional block is dimensionality reduction which decreases the number of operations done by different layers thus reducing the computational burden.

The nine inception modules used in the GoogLeNet are one of its crucial layers. The inception module's main functionality is to reduce the computational cost associated with dimensionality reduction by identifying the features in varying scales via the use of convolution operators with different filters. Two max-pooling layers are placed in between some inception modules and the main use of the max-pooling layer is to downsample the input when it is

propagated to different layers. The downsampling process mainly reduces the height and width of the input data. In this way, the computational burden that exists between different inception modules is reduced. The mean value of every feature map is taken by the average pooling layer present at the last inception module where the input height and width are reduced to 1×1 . To prevent the overfitting of the network, a dropout layer is used which randomly minimizes the number of interconnected neurons with a neural network. The linear layer mainly comprises 1000 hidden units in which each layer represents the image class of the ImageNet dataset. The last layer is the softmax layer which utilizes the softmax function to derive the probability distribution of the input vector. The softmax function vector is a set of values whose probability sums up to 1.

3.3.2 Novel deep learning architecture

The pre-trained GoogLeNet architecture cannot be directly applied for segmentation tasks since its output layer is mainly designed for performing classification. This arises the need for modifying the pre-trained architecture for segmentation purposes by using transfer learning. Transfer learning reuses an already trained model for another task to solve a similar task of the same sort. In this way, the training time of the model is reduced and it also offers increased performance for a small training set. The model

Table 3 GLCM features and their corresponding equations

| S. No | GLCM features | Equation |
|-------|--------------------------------------|--|
| 1 | Entropy | $En = \sum_m \sum_n BS_z(m, n) \log BS_z(m, n)$ |
| 2 | Energy | $E = \sum_m \sum_n BS_z(m, n)$ |
| 3 | Contrast | $C = \sum_m \sum_n m - n ^2 BS_z(m, n)$ |
| 4 | Correlation | $Corr = \frac{\sum_m \sum_n (m - \mu_m)(n - \mu_n) BS_z(m, n)}{\sigma_m \sigma_n}$ |
| 5 | Homogeneity | $H = \sum_m \sum_n \frac{1}{1 + (m - n)^2} BS_z(m, n)$ |
| 6 | Variance | $Var = \sum_m \sum_n (m - n)^2 BS_z(m, n)$ |
| 7 | Inverse difference moment | $IDM = \sum_{m, n} \frac{BS_z(m, n)}{1 + m - n }$ |
| 8 | Dissimilarity | $D = \sum_{m, n=0}^{n-1} BS_z(m, n) m - n $ |
| 9 | Cluster prominence | $CP = \sum_m \sum_n (m + n - \mu_m - \mu_n)^4 BS_z(m, n)$ |
| 10 | Cluster shade | $CS = \sum_m \sum_n (m + n - \mu_m - \mu_n)^3 BS_z(m, n)$ |
| 11 | Difference variance | $DV = \text{Variance of } BS_z(m, n)$ |
| 12 | Difference entropy | $DE = -\sum_{m=0}^{N_g} BS_z(m, n) (i) \log \{BS_{z_{x-y}}(m)\}$ |
| 13 | Information measure of correlation 1 | $IMC1 = \sum_{m, n} \frac{HMN - HNM1}{\max\{HM, HN\}}$ |
| 14 | Information measure of correlation 2 | $IMC2 = (1 - \exp[-2.0(HMN2 - HMN)])^{1/2}$ $HMN = -\sum_m \sum_n BS_z(m, n) \log BS_z(m, n)$ |
| 15 | Sum entropy | $SE = \sum_{m=2}^{2N_g} BS_z(m) \log \{BS_{z_{m+n}}(m)\}$ |
| 16 | Sum average | $SA = \sum_{m=2}^{2N_g} m BS_{z_{m+n}}(m)$ |
| 17 | Sum variance | $SV = \sum_{m=2}^{2N_g} (i - SE^2) BS_{z_{m+n}}(m)$ |
| 18 | Sum of square variance | $SSV = -\sum_m \sum_n (m - \mu)^2 BS_z(m, n)$ |
| 19 | Maximum probability | $MP = \max(BS_z(m, n))$ |

weights used for the previous problem by the same architecture are transferred for the novel architecture by slightly modifying them to suit the novel dataset.

In this paper, we provide a novel architecture model for segmentation which is based on GoogLeNet (Table 2). The main difference between our model and GoogLeNet is the top and bottom few layers changes. In GoogLeNet last 3 layer contains the dropout Softmax and output classification layer. Instead of these 3 layers, we add fully connected layers and a pixel classification layer. This pixel classification layer provides classification output for each pixel on the image. So pixel classification plays a major role in semantic segmentation. Semantic segmentation is a process of connecting the image pixels to its class label and it is a technique that classifies the image at a pixel level. In this way, an accurate level of the image is obtained and it can help the system to understand what is exactly present in the image via computer vision and enhance accuracy. By taking into consideration pixel-wise loss and in-network-up

sampling, the fully connected network is employed for dense prediction.

Table 2 presents the novel GoogLeNet segmentation architecture details regarding the number of layers used, output size, number of kernels, kernel size, and depth. The architecture used in this paper is 26 layers deep. Initially, there are normal convolutional layers followed by blocks of inception layers and max-pooling layers as shown in Table 2. Both the convolution and inception modules use a ReLU (Rectified Linear Unit) activation function. The pixel classification layer in the last provides a class label for each image pixel processed using the GoogLeNet architecture and the undefined pixel labels are ignored during training. In the training phase, the enhanced image and the ground truth image are given to the new model. Using the MIAS dataset, the proposed model is trained. During the testing phase, the images were segmented with the help of our model.

Table 4 Shape features and their corresponding equation

| S. No | Shape features | Equation description | Parameter description |
|-------|---------------------|--|--|
| 1 | Area (A) | The area is specified as the pixel density corresponding to the region | – |
| 2 | Asymmetry | $Z = \frac{\Delta C_{area}}{CA}$; | Z-asymmetry CA- calcification area, ΔC_{area} is the area between the two calcification halves |
| 3 | Perimeter | $\rho = \sum_{m=1}^{N-1} D_m = \sum_{m=1}^{N-1} BL_m - BL_{m+1} $ | ρ -Calcification perimeter value D- Distance BL- Boundary Length |
| 4 | Border irregularity | $S = \frac{ab}{2\pi(a^2+b^2)} \frac{\rho}{\Delta A}$ | a-major axis b-minor axis S- Input image ρ -Calcification perimeter value ΔA - related area |
| 5 | Compactness | $S = \frac{4\pi\rho^2}{\Delta A}$ | ρ -Calcification perimeter value ΔA - related area |

3.4 Features extraction

3.4.1 Gray level co-occurrence matrix (GLCM)

The GLCM [43] is a strategy for extracting statistical second-order texture features from the segmented image. The differences in the complex image textures can be analyzed via the GLCM matrices. The variances are usually caused by differences in the relative arrangement of pixels at different intensities. The GLCM [61, 62] overcomes this problem by differentiating the spatial relationship of pixels. The 2×2 GLCM is represented by the Black and White GLCM (BWglcm) in three directions namely 45 diagonal upper-left-to-bottom right, vertical down, and horizontal left-to-right). The input image taken is rectangular and has M columns and N rows. In grey level of each pixel is quantized to G levels.

For the G quantized grey levels ($Q_M = \{0, 1, 2, 3, \dots, G - 1\}$), the columns and rows are represented as $L_M = \{1, 2, 3, \dots, L_M\}$ and $L_N = \{1, 2, 3, \dots, L_N\}$. $L_N \times L_M$ is the set of pixels that can be transformed into the row-column description. The input image $I : L_M \times L_N$ is taken as a function that allocates some grey level value Q for every pair of pixels or individual pixels in the coordinates. The texture-related information is expressed using a matrix of related frequencies named $BS_x(m, n)$. Here, m and n represent the adjacent gray level pixels separated by a distance α . These gray-level co-occurrence frequencies are a function that

comprises the adjacent pixel distance and angular relationship. This paper utilizes a total of 19 GLCM-based texture features along with their equations as shown in Table 3. The (m, n) th normalized GLCM entry is represented as $BS_x(m, n)$. The mean(μ) and standard deviation (σ) for the (m, n) th normalized GLCM entry are computed as shown in the below equations:

$$\begin{aligned} \mu_M &= \sum_m \sum_n m \cdot BS_x(m, n) \\ \mu_N &= \sum_m \sum_n n \cdot BS_x(m, n) \end{aligned} \tag{3}$$

$$\begin{aligned} \sigma_M &= \sum_m \sum_n (m - \mu_M)^2 \cdot BS_x(m, n); \\ \sigma_N &= \sum_m \sum_n (n - \mu_N)^2 \cdot BS_x(m, n) \end{aligned} \tag{4}$$

The co-occurrence matrix in GLCM is represented as $BS_x(m, n)$ which is the frequency of a reference matrix R and a component c is present in an angle-distance metric(α) where the value m has n near it. The reference matrix R is the image and (m, n) is the pixel differences. In the GLCM, the value of each component (m, n) equals the number of times the pixel m is associated with n using an angle-distance metric. Based on the grayscale intensity values, the number of rows and columns are identified. Since an image contains grayscale values ranging from 0–255, the GLCM output comprises more than 256 rows. Every GLCM are represented using a 2×2 matrix since the segmentation results in a binary image where the pixel value either falls

in 0 or 1. The working of the GLCM is represented using the matrix A as follows:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \tag{5}$$

If the adjacent pixel distance is 1, then the angle is represented as 0° . When the focus is moved from left to right, the frequency value is calculated, which indicates whether a certain component is the same or different, and it is expressed as follows:

$$B_{1,90} = \begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix} \tag{6}$$

The sum of the elements in the matrix is 12 due to the presence of 12 locations in the matrix and each component is located at the right. The value of the elements $B_x, b_{x1,2}$, and $b_{x2,1}$ is zero due to the absence of number zero elements with a 1 to its right ($b_{x1,2}$) and number one elements with a 0 to the right ($b_{x2,1}$). From the actual $i \times j$ binary image (B_{img}), the GLCM value is computed as follows:

$$2B_{img}(m, n) - B_{img}(m + x, n + y) = Z(m, n) \tag{7}$$

$$GLCM = \begin{bmatrix} \sum_{n=0}^{u-y-1} \sum_{m=0}^{v-x-1} \varphi(Z(m, n), 1) & \sum_{n=0}^{u-y-1} \sum_{m=0}^{v-x-1} \varphi(Z(m, n), 2) \\ \sum_{n=0}^{u-y-1} \sum_{m=0}^{v-x-1} \varphi(Z(m, n), -1) & \sum_{n=0}^{u-y-1} \sum_{m=0}^{v-x-1} \varphi(Z(m, n), 0) \end{bmatrix} \tag{8}$$

$$\text{Where } \varphi(Z(m, n), h) = \begin{cases} 1, & Z(m, n) = h \\ 0, & \text{else} \end{cases} \tag{9}$$

The given GLCM formula is for a binary image. The values x and y determine the angle-distance metric(α). When the value of y is positive and x is zero, then it represents the down vertical distance. The diagonal direction has equal x and y values and the left to right horizontal direction has a positive y value and a zero x value. For an image of interest, there are eight scales and it is represented as $x, y \in \{2^0, 2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7\}$. Table 3 presents the different texture elements of BS_x .

The relative intensity measure is known as the contrast which is computed between a pixel and its adjacent value α . The contrast value is 0 for a non-zero element due to the presence of a single non-zero element in the GLCM matrix. The proportion of values on the GLCM diagonal versus the proportion of values of the diagonal is measured by homogeneity. The homogeneity value lies in the range [0,1] and the pixel intensity value in a specific distance equals the reference pixel and the diagonal matrix value is 1 for every pixel. Energy is a matrix normalization type

that measures the orderliness of the image and its value falls in the range fall in [0,1]. The value 1 mainly represents the constant image and energy is also interrelated with entropy. The correlation value mainly measures the correlation present between the reference pixel and the pixel in distance α . The mean value measures the average of gray levels in the image and the variance is a measure of heterogeneity where the variance increases when a difference is noted in the grey level values. The dissimilarity value is similar to contrast where the weights of the components increase in a linear fashion.

3.4.2 Shape features

Shape features extract the shapes from the segmented image and a set of 5 shape features is suggested in this paper. Table 4 represents the shape features and their corresponding equations. If the shape of the calcification is normal, the ratio is 1, or else it is near zero.

3.5 Breast cancer classification

The normal, malignant, and benign classes of the mammography images are classified using the Naïve Bayes (NB) [3], Decision tree (DT)[39], Support Vector Machine (SVM) [2], and Random Forest (RF) [40] classifiers. A brief description of each classifier is shown below:

Decision Tree(DT): The decision tree is a supervised machine learning classifier where the data is continuously partitioned based on a certain parameter. The nodes represent the features extracted for the breast cancer classification problem and the edges represent the outcome of the test by interconnecting the next node or leaf. The classification result (benign, malignant, or normal) is present in the leaf node.

Random Forest (RF): It offers multiple trained decision tree classifiers for the testing stage which makes it always preferred over the conventional Decision tree. The correct input features need to be obtained to act as the nodes. There is an N number of decision tree classifiers and the features obtained from the input image are sent through every Decision Tree to obtain the class labels. At last, a bagging technique is applied to the result obtained from the trees in the previous step.

Support Vector Machine (SVM): SVM is a machine learning classifier that provides high accuracy with less computational power. The main aim of the SVM is to find the hyperplane in an N -dimensional space to uniquely classify the data points. Where N represents the number of features and the hyperplane is a decision boundary used to classify the data points. The hyperplane with a maximum

Table 5 Performance results

| S. No | Abnormality | Input Image | Enhanced Image | Segmented image |
|-------|--|-------------|----------------|-----------------|
| 1 | Calcification (Benign (Normal)) | | | |
| 2 | Well-defined/ circumscribed masses (Malignant) | | | |
| 3 | Speculated masses (Malignant) | | | |
| 4 | Other, ill- defined masses (Malignant) | | | |
| 5 | Architectural distortion (Benign (Normal)) | | | |
| 6 | Asymmetry (Malignant) | | | |
| 7 | Normal | | | |

margin can distinctively separate the classes with high accuracy.

Naïve Bayes (NB): To classify the breast cancer classes, a probabilistic machine learning model known as Naïve Bayes is used. It is formulated using the Bayes theorem.

$$P(R|M) = \frac{P(M|R)P(R)}{P(M)} \tag{10}$$

The above equation is used to derive the Bayes theorem. If the incident M happened means one can easily find the probability value of R . If R is the number of malignant cases, then M is the disease progression. Here M is the hypothesis and R is the evidence. The features are mostly independent, and one feature does not rely on the other feature in any way. Hence, it is known as Naïve. If m is the malignant class that classifies whether a patient is subjected to breast cancer or not. The value $R = r_1, r_2, r_3, \dots, r_n$ represents the list of the input features. After expanding the Naïve Bayes Rule by substituting the values for R, we get

$$P(m|r_1, r_2, \dots, r_n) = \frac{P(r_1|m)P(r_2|m) \dots P(r_n|m)P(m)}{P(r_1)P(r_2) \dots P(r_n)} \tag{11}$$

$$P(m|r_1, r_2, \dots, r_n) \subset P(m) \prod_{i=1}^n P(r_i|m) \tag{12}$$

Here, the class variable(m) makes two predictions: yes or no. The main aim is to find the class m with maximum probability.

$$m = \arg \max_m P(m) \prod_{i=1}^n P(r_i|m) \tag{13}$$

By using the above equation, one can make a classification by taking the predictions.

4 Experiments results and discussions

The experiments are conducted on an Intel Core I9-10,850 K 3.60 GHz processor with 32 GB memory and 1 TB storage. The Matlab programming language is used to implement this model. Table 5 represents the original input image, enhanced image, and segmented image results for various abnormalities, respectively. The performance metrics used are namely Dice coefficient, Jaccard Index, Accuracy, Sensitivity, and Specificity. The dice coefficient is mainly used to compute the similarity between two sets. It is defined as two times the area of M and N divided by the sum of the areas M and N . Jaccard Index is mainly used to compute the similarity and diversity of sample image sets. They find the similarity of the finite size of samples by taking a ratio of intersections over the union.

$$Dice(M, N) = \frac{2|M.N|}{|M + N|} \tag{14}$$

$$Jaccard(M, N) = \frac{|M.N|}{|M| + |N| - |M.N|} \tag{15}$$

where M and N are binary vectors of equal length with values of 1 and 0, respectively. The value one indicates that an element is present in the set whereas a value 0 indicates an absence of the element in the set. $|M.N|$ represents an inner product of M and N where M and N represent the true positive values. Sensitivity ($S1$) identifies the percentage of pixels in the diseased area that is accurately segmented as abnormal masses. It is computed using the following formula:

$$S1 = \frac{X_1}{X_1 + Y_2} \tag{16}$$

Specificity ($S2$) is the percentage of normal tissues correctly segmented by the model.

$$S2 = \frac{X_2}{X_2 + Y_1} \tag{17}$$

$$Accuracy = \frac{X_1 + X_2}{X_1 + X_2 + Y_1 + Y_2} \tag{18}$$

The true positive (X_1) represents the abnormal tissue correctly segmented as abnormal and the true negative(X_2) represents the normal masses segmented as normal. False-positive (Y_1) represents the normal tissue incorrectly segmented as abnormal and False-negative(Y_2) represents the abnormal masses incorrectly segmented as normal.

Figure 2 shows a comparison of proposed and existing models such as modified Xception [63], modified AlexNet

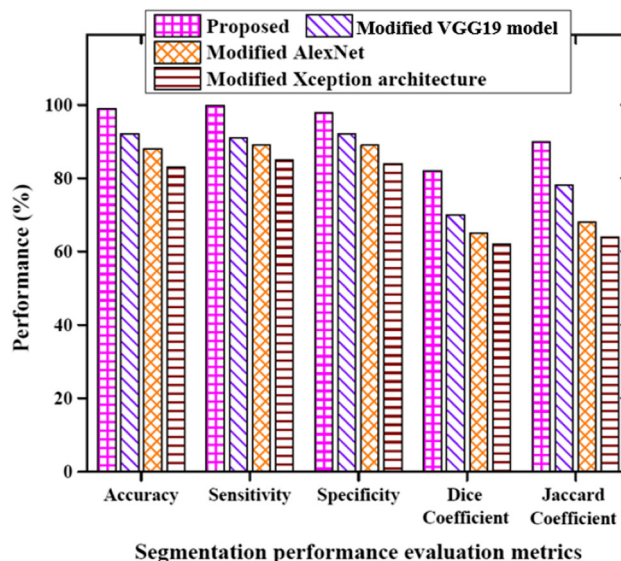


Fig. 2 Segmentation performance comparison between proposed, VGG16 [44], AlexNet [45] and CNN [46]

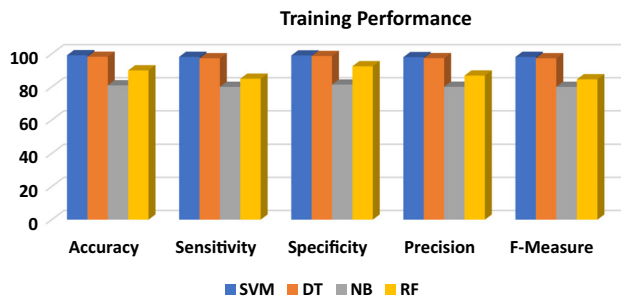


Fig. 3 Classification training performance comparison between various ML models

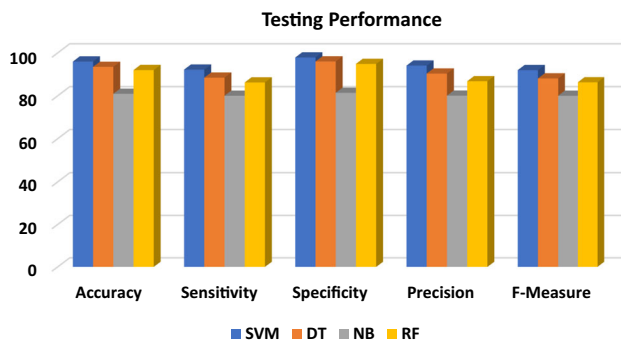


Fig. 4 Classification testing performance comparison between various ML models

[64], and modified VGG-19 [65] in terms of segmentation performance. In the modified Xception model, the multi-level features that are acquired from different convolutional layers are fed into a Multilayer perceptron Network (MLP) for training. In the modified AlexNet architecture, the multiclass Support Vector Machine (SVM) layer is used instead of the normal classification layer. In the modified VGG-19 model, the authors replaced the global pooling layer in the final block instead of the max-pooling layer.

The pre-trained CNN architectures are trained for 100 epochs with a learning rate of 0.001 using a stochastic gradient descent algorithm. By comparing the output image to the actual segmented image acquired from the radiologist, the segmentation performance is assessed. During the training process, a snapshot of the CNN model is taken for each epoch, and the model with the highest dice coefficient value is chosen as the winner. The proposed model offers an accuracy, sensitivity, specificity, Dice coefficient, and Jaccard coefficient score of 99.12%, 99.89%, 98.45%, 82.15%, and 89.11% which is relatively higher than the other techniques.

According to segmentation performance analysis, when comparing the proposed model with the existing models, the accuracy increased approximately 4.61%, the sensitivity increased 5.43%, the precision increased 4.76%, the

Dice coefficient increased 13.4% and the Jaccard coefficient increased 17.1%. The experimental results obtained show that the proposed methodology offers significant performance and outperforms other conventional methodologies when evaluated in terms of different performance evaluation metrics.

Figures 3 and 4 represent the classification training and testing performance comparison between various machine learning approaches and they are self-explanatory. From the analysis of the results, the SVM classifier provides the best classification performance in terms of accuracy, sensitivity, precision, and F-measure. The SVM achieves higher performance and surpasses the DT, NB, and RF techniques when evaluated using the MIAS dataset. The SVM mainly offers higher performance due to the detailed calcification segmentation results offered by the GoogLeNet architecture. The GoogLeNet architecture offers improved performance by the shape features extracted via precise lesion segmentation. Even though the performance of the SVM is dependent on the segmentation results, it doesn't need additional time and effort because no manual intervention is utilized here for segmentation. Thus we can conclude that with the help of the transfer learning approach, the training and testing efficiency has been improved significantly.

5 Conclusion

According to information released by the World Health Organization (WHO), breast cancer is another of the most common cancers among women. Mammography is the most effective tool for the early detection of this type of cancer. Mammography can detect cancer in the breast ten years before it manifests. We employed segmentation to assess the breast tumor, which aids doctors in determining the volume of the tumor and results in more effective treatment. In this study, we proposed a GoogLeNet architecture for breast cancer segmentation. From the analysis of the results, the proposed model provides the best segmentation results performance while comparing it with the existing model. To determine whether it is normal or abnormal, benign or cancerous, machine-learning approaches are used. When compared to other methods, SVM gives better performance results, according to classification training and test performance analyses. The proposed method has a segmentation accuracy of 99.12%, which helps to improve the classification performance of various machine learning architectures. As a result, when applied to the medical field, our proposed methodology has proven to be very beneficial.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent For this type of study informed consent is not required.

References

- Blackwell M, Nikou C, DiGioia AM, Kanade T (2000) An image overlay system for medical data visualization. *Med Image Anal* 4(1):67–72
- Chinnu A (2015) MRI brain tumor classification using SVM and histogram based image segmentation. *Int J Comput Sci Inf Technol* 6(2):1505–1508
- Amrane M, Oukid S, Gagaoua I, Ensari T (2018) Breast cancer classification using machine learning. In: 2018 electric electronics, computer science, biomedical engineering's meeting (EBBT). IEEE
- Amrane M, Oukid S, Gagaoua I, Ensari T (2018) Image segmentation techniques: a survey. *Image* 5(04):1–4
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9.
- Kamalakaran J, Babu MR (2018) Classification of breast abnormality using decision tree based on GLCM features in mammograms. *Int J Comput Aid Eng Technol* 10(5):504–512
- Pranckevičius T, Marcinkevičius V (2017) Comparison of naive bayes, random forest, decision tree, support vector machines, and logistic regression classifiers for text reviews classification. *Baltic J Modern Comput* 5(2):221
- Perez L, Wang J (2017) The effectiveness of data augmentation in image classification using deep learning. [arXiv:1712.04621](https://arxiv.org/abs/1712.04621)
- Sridhar D, Krishna IM (2013) Brain tumor classification using discrete cosine transform and probabilistic neural network. In: 2013 international conference on signal processing, image processing & pattern recognition. IEEE
- Sivakumari CY (2013) Comparison of diverse enhancement techniques for breast mammograms. *Int J Adv Res Comput Sci Manage Stud* 1(7):400–407
- Srivastava S, Sharma N, Singh SK, Srivastava R (2014) A combined approach for the enhancement and segmentation of mammograms using modified fuzzy C-means method in wavelet domain. *J Med Phys/Assoc Med Phys India* 39(3):169
- Stoji CT, Reljin I, Reljin B (2005) Local contrast enhancement in digital mammography by using mathematical morphology. In: International symposium on signals, circuits and systems. ISSCS 2005. IEEE
- Hanumantharaju MC, Gopalakrishna MT (2014) Review of mammogram enhancement techniques for detecting breast cancer. *Int J Comput Appl* 975:8887
- Sahu S, Singh AK, Ghreera SP, Elhoseny M (2019) An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. *Opt Laser Technol* 110:87–98
- Li X, Yang C, Wu S (2016) Automatic segmentation algorithm of breast ultrasound image based on improved level set algorithm. In 2016 IEEE international conference on signal and image processing (ICSIP), IEEE
- Prabhakar T, Poonguzhali S (2017) Automatic detection and classification of benign and malignant lesions in breast ultrasound images using texture morphological and fractal features. In: 2017 10th biomedical engineering international conference (BMEI-CON). IEEE
- Dominguez AR, Nandi AK (2008) Detection of masses in mammograms via statistically based enhancement, multilevel-thresholding segmentation, and region selection. *Comput Med Imaging Graph* 32(4):304–315
- Sahar M, Nugroho HA, Ardiyanto I, Choridah L (2016) Automated detection of breast cancer lesions using adaptive thresholding and morphological operation. In: 2016 International conference on information technology systems and innovation (ICITSI). IEEE.
- Husain RA, Zayed AS, Ahmed WM, Elhaji HS (2015) Image segmentation with improved watershed algorithm using radial bases function neural networks. In: 2015 16th International conference on sciences and techniques of automatic control and computer engineering (STA). IEEE
- Román KL, Ocaña MI, Urzelai NL, Ballester MÁ, Oliver IM (2020) Medical image segmentation using deep learning In: Deep learning in healthcare. Springer. pp 17–31
- Moftah HM, Ahmad TA, Al-Shammari ET, Ghali NI, Hassanien AE, Shoman M (2014) Adaptive k-means clustering algorithm for MR breast image segmentation. *Neural Comput Appl* 24(7):1917–1928
- Subramani T (2019) Brain tumor segmentation based on a hybrid clustering technique. California State University, Northridge
- Menon N, Ramakrishnan R (2015) Brain tumor segmentation in MRI images using unsupervised artificial bee colony algorithm and FCM clustering. In: 2015 international conference on communications and signal processing (ICCSP). IEEE
- Khalifa I, Youssif A, Youssry H (2012) MRI brain image segmentation based on wavelet and FCM algorithm. *Int J Comput Appl* 47(16)
- Rouhi R, Jafari M, Kasaei S, Keshavarzian P (2015) Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Syst Appl* 42(3):990–1002
- Oh KT, Lee S, Lee H, Yun M, Yoo SK (2020) Semantic segmentation of white matter in FDG-PET using generative adversarial network. *J Digit Imag*, 1–10
- Jafari M, Auer D, Francis S, Garibaldi J, Chen X (2020) DRU-Net: an efficient deep convolutional neural network for medical image segmentation. In: 2020 IEEE 17th international symposium on biomedical imaging (ISBI). IEEE
- Myronenko A (2018) 3D MRI brain tumor segmentation using autoencoder regularization. In: International MICCAI brainlesion Workshop. Springer
- Cui S, Mao L, Jiang J, Liu C, Xiong S (2018) Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network. *J Healthc Eng* (2018)
- Sewak M, Vaidya P, Chan CC, Duan ZH (2007) SVM approach to breast cancer classification. In: Second international multi-symposiums on computer and computational sciences (IMSCCS 2007). IEEE
- Vijayarajeswari R, Parthasarathy P, Vivekanandan S, Basha AA (2019) Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform. *Measurement* 146:800–805
- Venkatesan E, Velmurugan T (2015) Performance analysis of decision tree algorithms for breast cancer classification. *Indian J Sci Technol* 8(29):1–8
- Saritas MM, Yasar A (2019) Performance analysis of ANN and Naive Bayes classification algorithm for data classification. *Int J Intell Syst Appl Eng* 7(2):88–91

34. Dumitru D (2009) Prediction of recurrent events in breast cancer using the Naive Bayesian classification. *Ann Univ Craiova-Math Comput Sci Ser* 36(2):92–96
35. Virmani J, Dey N, Kumar V (2016) PCA-PNN and PCA-SVM based CAD systems for breast density classification. In: *Applications of intelligent optimization in biology and medicine*. Springer, 159–180
36. Azar AT, El-Said SA (2013) Probabilistic neural network for breast cancer classification. *Neural Comput Appl* 23(6):1737–1751
37. Bayramoglu N, Kannala J, Heikkilä J (2016) Deep learning for magnification independent breast cancer histopathology image classification. In: *2016 23rd international conference on pattern recognition (ICPR)*. IEEE
38. Zhou Y, Xu J, Liu Q, Li C, Liu Z, Wang M, Zheng H, Wang S (2018) A radiomics approach with CNN for shear-wave elastography breast tumor classification. *IEEE Trans Biomed Eng* 65(9):1935–1942
39. Naik J, Patel S (2014) Tumor detection and classification using decision tree in brain MRI. *Int J Comput Sci Netw Security* 14(6):87
40. Nguyen C, Wang Y, Nguyen HN (2013) Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic
41. Sheet SSM, Tan TS, As'ari MA, Hitam WHW, Sia JS (2021) Retinal disease identification using upgraded CLAHE filter and transfer convolution neural network. *ICT Express*
42. Patel VK, Uvaid S, Suthar A (2012) Mammogram of breast cancer detection based using image enhancement algorithm. *Int J Emerg Technol Adv Eng* 2(2012):143–147
43. Soh LK, Tsatsoulis C (1999) Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Trans Geosci Remote Sens* 37(2):780–795
44. Geng L, Zhang S, Tong J, Xiao Z (2019) Lung segmentation method with dilated convolution based on VGG-16 network. *Comput Assist Surg* 24(sup2):27–33
45. Zhang YD, Govindaraj VV, Tang C, Zhu W, Sun J (2019) High performance multiple sclerosis classification by data augmentation and AlexNet transfer learning model. *J Med Imag Health Informatics* 9(9):2012–2021
46. Wahab N, Khan A, Lee YS (2019) Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathological images. *Microscopy* 68(3):216–233
47. Suckling JP (1994) The mammographic image analysis society digital mammogram database. *Digital Mammo*, pp 375–386
48. Chen Z, Yunjie Y, Jiabin J, Pierre B (2020) Deep learning based cell imaging with electrical impedance tomography. In: *2020 IEEE international instrumentation and measurement technology conference (I2MTC)*, pp 1–6. IEEE
49. Vinu S (2016) An efficient threshold prediction scheme for wavelet based ECG signal noise reduction using variable step size firefly algorithm. *Int J Intell Eng Syst* 9(3):117–126
50. Vinu S (2019) Optimal task assignment in mobile cloud computing by queue based ant-bee algorithm. *Wirel Pers Commun* 104(1):173–197
51. Sundararaj V, Muthukumar S, Kumar RS (2018) An optimal cluster formation based energy efficient dynamic scheduling hybrid MAC protocol for heavy traffic load in wireless sensor networks. *Comput Secur* 77:277–288
52. Sundararaj V (2019) Optimised denoising scheme via opposition-based self-adaptive learning PSO algorithm for wavelet-based ECG signal noise reduction. *Int J Biomed Eng Technol* 31(4):325
53. Jose J, Gautam N, Tiwari M, Tiwari T, Suresh A, Sundararaj V, Rejeesh MR (2021) An image quality enhancement scheme employing adolescent identity search algorithm in the NSSST domain for multimodal medical image fusion. *Biomed Signal Process Control* 66:102480
54. Rejeesh MR (2019) Interest point based face recognition using adaptive neuro fuzzy inference system. *Multimed Tools Appl* 78(16):22691–22710
55. Shen, Yiqiu, Nan Wu, Jason Phang, Jungkyu Park, Kangning Liu, Sudarshini Tyagi, Laura Heacock et al. “An interpretable classifier for high-resolution breast cancer screening images utilizing weakly supervised localization.” *Medical image analysis* 68 (2021): 101908.
56. Prasath AS, Vasuki RS (2021) Breast cancer detection using mark rcnn segmentation and ensemble classification with feature extraction. *Indian J Comput Sci Eng* 12(1):239–245
57. Gautam C, Pratik KM, Aruna T, Bharat R, Hari MP, Shuihua W, Muhammad T (2020) Alzheimer's Disease Neuroimaging Initiative. Minimum variance-embedded deep kernel regularized least squares method for one-class classification and its applications to biomedical data. *Neural Netw* 123:191–216
58. Ahmad F, Nor AMI, Zakaria H, Siti NS (2013) A genetic algorithm-based multi-objective optimization of an artificial neural network classifier for breast cancer diagnosis.”. *Neural Comput Appl* 23(5):1427–1435
59. Aličković E, Subasi A (2017) Breast cancer diagnosis using GA feature selection and Rotation Forest. *Neural Comput Appl* 28(4):753–763
60. Reza AM (2004) Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *J VLSI Signal Process Syst Signal Image Video Technol* 38(1):35–44
61. Honeycutt CE, Plotnick R (2008) Image analysis techniques and gray-level co-occurrence matrices (GLCM) for calculating bioturbation indices and characterizing biogenic sedimentary structures. *Comput Geosci* 34(11):1461–1472
62. Park Y, Guldmann JM (2020) Measuring continuous landscape patterns with Gray-Level Co-Occurrence Matrix (GLCM) indices: an alternative to patch metrics? *Ecol Indic* 109:105802
63. Kassani SH, Kassani PH, Khazaeinezhad R, Wesolowski MJ, Schneider KA, Deters R, (2019) Diabetic retinopathy classification using a modified xception architecture. In: *2019 IEEE international symposium on signal processing and information technology (ISSPIT)*. IEEE, pp 1–6
64. Hosny KM, Kassem MA, Foad MM (2019) Classification of skin lesions using transfer learning and augmentation with Alexnet. *PLoS ONE* 14(5):e0217293
65. Ramesh MJ (2021) Feature extraction of ultrasound prostate image using modified Vgg-19 transfer learning. *Turk J Comput Math Educ (TURCOMAT)* 12(10):7597–7606

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.