Multimedia Tools and Applications
https://doi.org/10.1007/s11042-018-6560-x

# A deep convolutional neural network for the early detection of breast carcinoma with respect to hyper-parameter tuning

Saranyaraj D.[1] · Manikandan M.[1] · Maheswari S.[1]

Received: 16 May 2018 / Revised: 23 July 2018 / Accepted: 15 August 2018
Published online: 29 August 2018


**Abstract**
Medical image processing needs attention towards accurate analysis rate which directly implies on the treatment. This paper focuses on the mammogram image analysis for early prediction of breast cancer (Screening) and reduce the mortality rate rather than using an invasive diagnosis technique. To classify the mammogram images a novel network called Deep Convolutional neural network (DCNN) is utilized in which multi-layer perceptron is used in the fully connected layer to accurately classify the mammogram images as three classes benign, malignant and normal. Before classifying the breast cancer, image pre-processing and feature extraction plays a major role in preserving the useful information and extracting the desired features. The Bilateral filter with a vector grid computing is used as the noise reduction filter to preserve the edge information which is essential in differentiating the masses and the dense tissue. The Features like Area, Radius, Perimeter and smoothness are extracted to train the network and to detect the malignant tumor stating if the patient is positive or negative with the cancer. Five stages have been proposed and implemented such as: (a) Crop and resize of the original mammogram; (b) De-Noising the DDSM (Digital Database for Screening Mammography) image to preserve the edge information. (c) Train the proposed DCNN model using the features extracted, (d) Classifying the DDSM images (e) Evaluating the performance using hyper parameter tuning of the proposed system. Unstinted Observations are made to justify the listed findings and by comparing the proposed outline with the help of the literature about the several in-use image classification models. A confusion matrix is drawn with the classes based on: Those with Benign, Malignant and normal tissues. The results are discussed (benchmarked) to show that fine-tuning of the final layers or the entire network parameters leads in achieving 96.23% of overall test accuracy and 97.46% of Average Classification Accuracy.

**Keywords** Breast cancer detection · Deep learning · Deep convolutional neural network · Hyper-parameter tuning

✉ Saranyaraj D.
saranya.elec6691@gmail.com

Extended author information available on the last page of the article

## 1 Introduction

The most perilous and common disease suffered majorly by women in the world is found by the team from American Cancer Society (ACS) [14], who gave a statistic stating that, 1/8th and 1/12th are predicted to be incident with the breast melanoma in their lifetime [2]. 19% of the cancer mortality is expected from the study made by the European community representing 24% of cancer incidence [13]. Similarly, 16.5% of mortality rate from breast cancer is estimated in the Tunisian country [38]. The major age group among which one-fourth of all the breast cancer in women has occurred and diagnosed to be between 40 and 49. Proper awareness and treatments along with mammography screening and diagnosis have declined the mortality rate even though the incidence has increased over the years [51].

The early stage of prediction helps in aiding the treatment and reducing the mortality rate. According to ACS [14], the following changes in the breast denotes the breast cancer proliferation [14]:

1. A feel of mass or pain in the breast
2. Skin dryness/flakiness or pigmentation
3. Layer dimpling/itching
4. Breast inflammation
5. The structural change in the breast
6. Leakage of nipple with the pulling effect.

The evolution of the technical era has risen computerized diagnostic system as the main source of a screening method for analysing mammogram images to classify breast cancer. Features play a vital role in diagnosing the medical images, thus it is used to train the machine learning classifiers. Features are the deciding factor for reducing false positive and false negative breast cancer classification. The artificial neural network is simultaneously booming in the computer vision application in the field of medical image investigations.

Computer Aided Detection (CAD) has also leaned on such technology for the diagnosis of the medical images. Extracting features in other hand is not an easy job which lays enormous challenges to the researchers from medicine to computer vision. Generally, CAD makes use of defined features based on the knowledge gathered from the prior art and expert guidance. For more predictability and precise information on the classification, Automatic feature extraction is required, which is obtained by self-training the layers. This recent evolution is known to be Deep Learning Technique.

### 1.1 Motivation

The development in the computer processors involving huge memory capacity and faster buses has paved a reasonable and inculcating path towards medical imaging. The most facilitating algorithm towards health care industry is said to be machine learning/Artificial intelligence that helps the medical experts to process and analyse the medical outcomes accurately and fast. Preserving the important features and information from a medical image is the most critical factor to be considered while analysing the medical images.

Several digital image processing algorithms have played its role in assisting the experts to analyse the mammograms which greatly reduced the mortality rate to 25% [51]. The image pre-processing, segmentation and feature extraction plays avital role in analysing the mammogram images aiding the early detection of cancer.

Generic visual recognition tasks have drawn the attention of many researchers for which the deep learning algorithm called deep convolutional neural networks [31] has been utilized producing an outstanding performance [50]. The features extracted by the CNN framework consists of several hidden layers to extract the input images for high-level feature representations. These advances and advantages of the deep CNN has motivated us to research on the medical images classification using CNN framework.

### 1.2 Problem statement

he cause of the breast cancer is not accurately known so far, the prevention suggestion also cannot be termed as the prominent solution, these being the factors to be considered the early stage detection of breast cancer is essential for the timely treatment [2].

The statistics by ACS shows that by 2030 the mortality rate increases drastically, for which the early detection of breast cancer greatly influences the reduction in mortality rate. Hence accurate screening of the breast mammogram images is essential for the successful treatment. Based on the study made by [13], the breast cancer detection seems to be difficult and challenging due to the following reasons involving dense masses,

1. Contrast and brightness is poor masking the prominent features
2. varies in geometrical and textural features
3. Overlapping of the masses
4. difficult to differentiate between masses and dense tissues.
5. Inconsistent tissue background.

There are few screening abnormalities to be considered while analysing the mammogram images which are listed as,

1. False negative and false positive rates are more
2. More the screening, More the radiation is.

### 1.3 Contribution/novelty

The main contribution of this research is said to address all the problems mentioned in latter sections as below,

1. The best suiting image processing technique to reducing the noise and for preserving the useful information
2. The most efficient classification method for classifying the cancer
3. The automated segmentation algorithm for edge preservation
4. The appropriate feature extraction algorithm for extracting almost every prominent feature.
5. The best suitable network supporting the aforementioned points and to reduce the computational time and cost.
6. The Algorithm to yield good accuracy rate.

The novelty of this research focuses on classifying the mammogram images using Deep convolutional neural network which combines both the Deep neural network and convolutional neural network algorithm. This consists of deep layers of input, Multi hidden layers of input and output layer processing with multi-layer perceptron and back propagation (MLP-BP) algorithm based on multiple convolution layer (Multi hidden layers) to capture the feature maps from regional connectivity via the weighted filter, and pooling layers to reduce data size, at last a fully connected NN is added as the classifier that acts as MLP-BP.

The classification of the DDSM dataset as benign (Dense tissue), malignant (Cancer tissue) and normal is performed using a Deep Convolutional neural network. To optimize the proposed classifier, hyperparameter tunning are explored by benchmarking classifiers with different factors termed as hidden layers, kernel size, filters, activation function etc. Consequently, iterating upon the proposed learning model through a combination of statistical and machine learning theory, domain knowledge of the obtained data, and trial and error few results are yield.

The paper is organized in the following sequence for better understanding. (i) Related work, (ii) Contemplated work (Pre-processing, Architecture designing), (iii) Observation and findings (Hyperparameter tunning) (iv) Experimental results (Result and discussion).

## 2 Related work

Some previously completed research has aided the progress of this paper in directing the proposed methods and augmenting the understanding of CNN image classification.

Bhattacherjee, A.et al., in 2016 [7] performed a histopathological breast cancer analysis with 32 features involving nuclei detection for the breast cancer detection. The author used back propagation neural network to train the network using the weights obtained from the features extracted. This approach facilitates the increase accuracy rate in early prediction of breast cancer.

Lan, K., et. Al., in 2018 [33] discussed about various data mining and deep learning methods in bioinformatics. The authors gave a very detailed discussion on how deep learning influences the medical image analysis in a huge way. The authors gave a summarization on various optimized neural networks in deep learning which greatly influenced the research work performed in this paper. From the paper CNN and SAE contributes more in medical image reconstruction, Medical image auto segmentation, classification and detection of cancer.

K. Peterson et al., in 2014 [41] used deep stacked auto-encoders to segment breast tissue and to calculate breast density score using multi-scale features and convolutional neural network (CNN) models.

X. S. Zhang in 2014 [60] used CNN to characterize micro-calcifications as a representational strategy. The closest work was developed in [23] to classify malignant/benign breast lesions, which uses an adaptive deep convolutional network.

Three different architectures of CNNs to locate masses in mammography images was used by M. G. Ertosun and D. L. Rubin, 2015 [12]. The images from DDSM database were grouped as validation, training and testing consisting of total 2420 images, in the ratio 8:1:1. They also used cropping, conversion, rotation, flipping and scaling techniques to get an augmented training set to improve the generalization ability of the framework. The experiment was divided into two stages: the first is to classify the image containing the mass and the second will be localization of masses in the images.

John Arevalo et al. in 2015 [4], obtained 86% of the Receiver Operating Characteristic (ROC) curve area by classifying mammography mass lesions using a CNN as a feature extractor and an SVM as a classifier. The dataset considered by the authors was from BCDR database. This dataset was comprised of 426 mass lesions and 310 with cancerous lesions all together 736 images. The data augmentation was achieved by flipping and rotating the images.

Jiao et al. in 2016 [24], used DDSM database to classify malignant and benign using the CNN as feature extractor and SVM as classifier. The images were normalized and whitened. On the other hand, the CNN was trained with a subset of ImageNet and the features to perform the classification were extracted from two different layers of the CNN.

Mohsin Jadoon et al. in 2017 [21], obtained an accuracy 83.74% by using the CNN curvelet method in MIAS database.

Three main aspects are being focused by the researchers to improve the performance aspect of deep learning models. ResNet, GoogleNet and VGG models have been in use, which shows the large number of hidden layers go on better than the smaller networks [20]. The independent execution of each layer is also possible with the larger networks and the result is obtained by combining all the responses from each layer. This being advantages the CNN will be utilized more efficiently.

The CAD algorithm retrospective evaluation of cancer cases which consumes more time and it is not cost effective for a clinical setting. The alternative of this will be by analysing the mammography images from the mammography centre, which turns out to be more time consuming and expensive. CNN being an automatic classifier is cost efficient and produces accurate results. Hence the proposed system is developed using a LeNet-5 architecture as a deep CNN algorithm. Table 1 shows various problem statement which has been addressed by the researchers in detecting the breast cancer using image processing techniques.

## 3 Contemplated framework

The proposed deep CNN based classification framework is described as image pre-processing, LeNet framework, LeNet training and feature extraction & classification as shown in Fig. 1. The.

### 3.1 Image pre-processing

Image pre-processing is the most important step involved to obtain the desired features and good classification rate. In the proposed system, the DDSM database is chosen. The images from DDSM contains variation, in contrast and brightness; basically, some induced noises are also present. To minimize such an effect which adversely affects the deep CNN process the normalization of images are to be performed by taking the difference between the minimal intensity values of the images. Considering D1 as the minimum intensity value and D2 as the maximum intensity value, the difference between them will give the resultant intensity as G1 = D2 - D1. To freeze the uniformity in the normalized images the images are resized to $200 \times 200$ which is further used for training. Before resizing the images, the dimension was $195 \times 365$ which increases the cost to train more parameters, whereas even smaller dimension might lead to the reduction in the feature information. To minimize the loss and computational cost the image has been resized into $200 \times 200$. The training is performed using the whole image without ROI operation [41] which would affect the classification performance.

**Table 1** A summary on the related work on detection of breast cancer

| S.NO | YEAR | et al. | TITLE | MOTIVATION | LIMITATION | CONCLUSION |
|---|---|---|---|---|---|---|
| 1 | JAN 2018 | 1. A. Elmoufidi, K. El Fahssi,<br>2. S. Jai-andaloussi,<br>3. A. Sekkaki,<br>4. Q. Gwenole<br>5. M. Lamard | Anomaly classification in digital mammography based on multiple-instance learning [11] | 1. Anomaly classification using multiple layer instance<br>2. Performs pixel level segmentation and image level classification<br>3. Automatic ROI | 1. Microcalcification being the import subfactor influencing malignant stage of a cancer is nor analyzed.<br>2. The computational analyses is tedious | 1. Detection sensitivity obtained is 95.6%<br>2. Concentrates on lesions |
| 2 | NOV 2017 | 1. Gustavo Carneiro,<br>2.Jacinto Nascimento,<br>3. Andrew P. Bradley | Automated Analysis of Unregistered Multi-View Mammograms with Deep Learning [10] | 1. Analysis of unregistered mammographic view dentification<br>2. comparison of the medical images with the pretrained Deep neural network model | 1. The accuracy cannot be relatively compared with the accuracy of ImageNet outcome as the image database to detect breast cancer has high feature dependency. | 1. Medical images can be diagnosed using the pretrained Deep neural network model<br>2. The accuracy would be equivalent to the accuracy obtained for the ImageNet database. |
| 3 | JULY 2016 | 1. Gwenol'e Quellec.<br>2. Mathieu Lamard.<br>3. Michel Cozic, Gouenou Coatrieux.<br>4. Guy Cazuguel | Multiple-Instance Learning for Anomaly Detection in Digital Mammography [43] | 1. Multiple layer instance Paradigm<br>2. MIL algorithms<br>3. Mammography on Breast Cancer.<br>4. Irregularity identification on surface investigation. | 1. Early adjustments of tissues can't be caught by an emphatically regulated classifier | 1. Highlights got from the identification of masses and Microcalcifications.<br>2. Multiple-Instance Learning (MIL) calculation is utilized to perceive anomalous areas and guide local to a global labeling ("ordinary examination record" or "strange examination record") utilizing a proposed mammography aggregator. |
| 4 | AUGUST 2018 | 1. Kamal, S.,<br>2.Dey, N.,<br>3. Nimmy, S. F.,<br>4. Ripon, S. H.,<br>5. Ali, N. Y.,<br>6. Ashour, A. S | Evolutionary framework for coding area selection from cancer data [26] | 1. Cancer DNA data analysis<br>2. Four different machine learning techniques are used to analyze such as support vector machine (SVM), principal component analysis (PCA) technique, neural mapping skyline filtering (NMSF) and Fisher's discriminant analysis (FDA). | 1. The enormous data analysis take more time<br>2. Evidential reasoning becomes more c omplicated. | 1. Simulation and computational analysis reduces cost<br>2. The large-scale cancer infected DNA data was efficiently supported by the neural mapping skyline filtering. |

**Table 1** (continued)

| S.NO | YEAR | et al. | TITLE | MOTIVATION | LIMITATION | CONCLUSION |
|---|---|---|---|---|---|---|
| | | | | 3. To estimate coding are | | |
| 5 | JULY 2016 | 1. Maxine Tan.<br>2. Bin Zheng.<br>3. Joseph K. Leader.<br>4. David Gur | Association Between Changes in Mammographic Image Features and Risk for Near-Term Breast Cancer Development [57] | 1. Prediction of near-term breast cancer risk.<br>2. A measurable assessment of bilateral mammographic image feature variations in a series of negative full-field digital mammography (FFDM) images. | 1. The computational time is more.<br>2. The parameters considered to analyze the near-term risk is not termed to be sufficient. | 1. Using a cross validation method an SVM model was developed.<br>2. The Mutation, Age, and Density of the breast were considered as a factor influencing the near-term risk |
| 6 | MAR 2012 | 1. N. Selvarasu<br>2.Alamelu Nachaippan<br>3.N.M. Nandhitha | Image Processing Techniques and Neural Networks for Automated Cancer Analysis from Breast Thermographs- A Review [47] | 1. Computer Aided analysis of breast thermographs have been developed to interpret the abnormality.<br>2. Artificial Neural Networks (ANNs) are used for classifying abnormality based on its severity.<br>3. Kurtosis, Difference in mean, Difference in Variance are the parameters considered. | 1. The masses overlapping is not detected efficiently.<br>2. The edge segmentation used in this paper does not contain sufficient information. | 1. Asymmetry analysis can be done by measuring skewness, kurtosis, difference in mean and standard deviation or variance between the right and left halves in turn to plot a histogram<br>3. Boundaries are expected to be obtained using Hough Transform and gradient based Hough transform.<br>4. Canny detector, snake transforms are to be utilized to identify abnormalities.<br>5. Artificial Neural Networks such as Radial Basis Function Networks and Back propagation Networks can be used for abnormality classification and severity detection. |
| 7 | Sep 2004 | 1. J. Koay.<br>2. C. Herry.<br>3. M. Frize. | Analysis of Breast Thermography with an Artificial Neural Network [30] | 1. Thermal Image Processing on Breast Cancer Detection.<br>2. The ROI is performed on the contralateral breasts of 19 Patients. | 1. The cost of procedure is high.<br>2. The computation time seems to be more.<br>3. Validation of the results needs refinement. | 1. ROI segmentation was performed using top-hat and bottom-hat morphological filtering operations, thereby emphasizing the contours and the edges in the images. Edges were then detected using a Canny edge detector. |

**Table 1** (continued)

| S.NO | YEAR | et al. | TITLE | MOTIVATION | LIMITATION | CONCLUSION |
|---|---|---|---|---|---|---|
| | | | | 3.SemiAutomated segmentation of the Breasts was Used. | | 2. A backpropagation (BP) neural network was used to analyze the infrared breast images.<br>3. The five parameters chosen for further analysis were mean, standard deviation, skewness, kurtosis and heat content.<br>4. Two back propagation training algorithms were compared for suitability: the Levenberg-Marquardt (trainlm) and the Resilient Backpropagation (trainrp).<br>5. The Levenberg- Marquardt (LM) algorithm was chosen as it generally yields the fastest convergence and the lowest mean square error (MSE), especially for relatively small networks with The transfer function for all neurons was the tan-sigmoid function. |

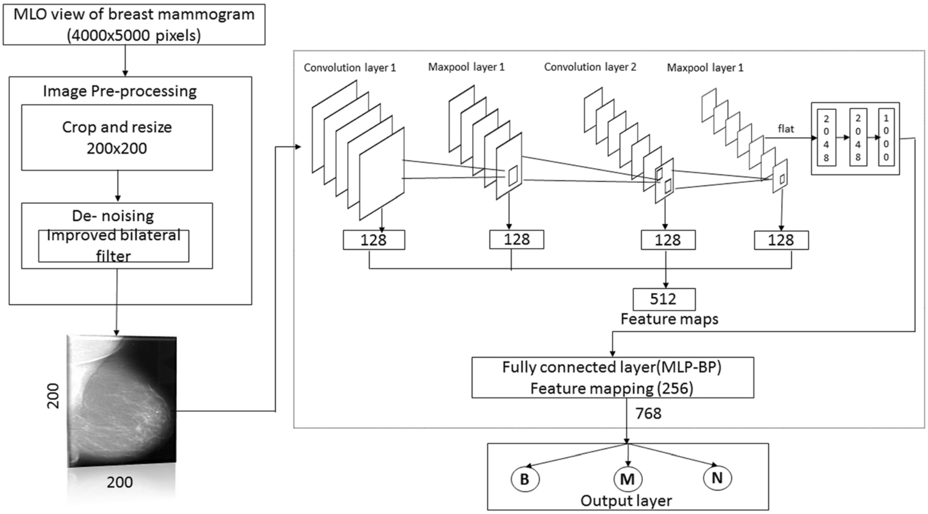


**Fig. 1** Contemplated Framework: Deep neural network and convolutional network as deep Convolutional network

The resized image is then processed with the bilateral filter to preserve the edges considering both the spatial and the range. The modified bilateral filter is performed as in Fig. 2.

Algorithm 1: Modified bilateral filter with vector valued grid.

---

1. Input: 2D image I

---

The trajectory valued grid
$V^- : S_1 \, x \, R_1 \rightarrow R_1{}^2$ so that

$$V^-(Q_a, Q_b, t) = \begin{cases} \left(I(Q_a, Q_b),\ 1\right) & if\ t = I(Q_a, Q_b) \\ (0,0) & else \end{cases}$$

2. Perform a gaussian smoothing of $V^-$, for each component independently
$GB\left[V^-\right](Q_a, Q_b, t) = G_{\sigma s\,\sigma r} * V^-(Q_a, Q_b, t)$, Where $G_{\sigma s1\,\sigma r1}$ is a 3D gaussian with $\sigma s1$ as parameter along the 2 spatial dimensions & $\sigma r1$ along the range dimensions.
3. Extracting the result: for a pixel P with initial intensity $I_Q$, $we\ denote\ (\widetilde{XI}, \tilde{X})$ the value at position $(Q_a, Q_b, I_Q)$. The result of the bilateral filter is
$$BF\left[I_Q\right] = \left(\widetilde{XI} | \tilde{X}\right)$$

---

### 3.2 LeNet framework

Selecting a proper network architecture to implement CNN is very important in deep neural network [9]. The network chosen for the proposed work is LeNet-5 architecture with 5 layers shown in Fig. 3 Convolutional and max pooling layers alternatively makes the first four layers connecting to the fully connected layer.

### 3.3 Convolutional layer

Considering the nth layer where $M^n$ denotes the mth layer feature maps. The feature map is symbolized as $s_k^n$ (k = 1, 2,.., $M^n$). The designed convolutional layer is categorized by an array of 2-D filters $Z_{lk}^n$ associating the ith feature map $s_l^{n-1}$ . In the (n − 1) th layer with the kth feature map $s_k^n$ in the nth layer and the bias $b_l$. The convolution operation is applied between all the input feature map location with the feature detector as a filter to predict the desired feature. $s_k^n$ is obtained after, $s_l^{n-1}$ (l = 1, 2,..., $M^{n-1}$). Input features with the respective filter $Z_{lk}^n$ is convolved.

The bias $b_k^n$. is summed up with the results and further also appended nonlinear activation function φ(·) is applied to an element-wise manner. Mathematically, between feature maps of two successive layers the full-complete connection scheme is used, the feature maps of the lth layer can be expressed as follows:

$$s_k^n = \varnothing\left(\sum_{l=1}^{M^{n-1}} s_l^{n-1} * Z_{lk}^n + b_k^n\right), k = 1, 2, \ldots M^n \tag{1}$$

* represents the convolution operation.

### 3.4 Pooling layer

The computation of the CNN can be reduced by down sampling the feature maps using the pooling layer. This action captures the images without making any further changes. Max-pooling is advantageous for the proposed framework, it ruled over the average pooling since it facilitates in choosing the maximum activation [17].

### 3.5 Classification layer

Classification happens in the fully connected layer that links all the output obtained from the previous layer and feeding it as input. The proposed system consists of a single, fully connected layer, F5 depicted in the Fig. 2. The input to the fifth layer is the fully connected layer obtained from the fourth layer F4 as feature maps, where Z8 and b8 is the given parameter. F5 layers consist of x number of neurons w.r.t. x classes of masses and give out the probability to be $\hat{t} = \left[\hat{t}_1, \hat{t}_2, \ldots, \hat{t}_x\right] \epsilon\, Q^x$ Via softmax regression as follows:

$$s_8 = Z_8 s_7 + b_{8,}\ \ s_8 \epsilon\, Q^x \tag{2}$$

$$\hat{t}_k = \frac{\exp\left(s_k^8\right)}{\sum_{l=1}^{x} \exp\left(s_l^8\right)}, k = 1, 2, \ldots, x \tag{3}$$

Where $\hat{t}_k$ is the output probability of kth neuron. The feature map from spatial resolution declines w.r.t. the features extracted in a consecutive manner from a layer to layer when a mammogram image is fed forward and propagated via the network. The spatial convolution and pooling is responsible for the extraction, and the spatial information extraction from the feature maps.

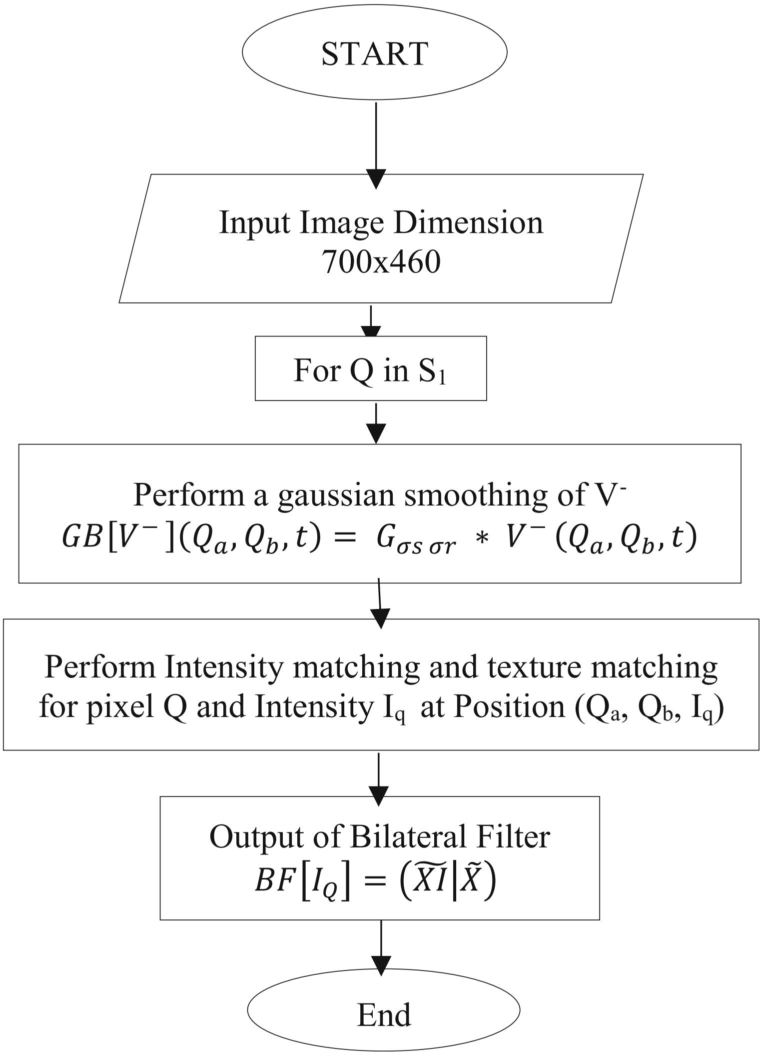


**Fig. 2** Denoising technique: Improved bilateral filter with the vector grid computation

### 3.6 Fully connected layer

The F5, fully connected layer used in the proposed framework consists of a multilayer perceptron algorithm that uses the SoftMax activation function in the output layer. The fully connected layer is responsible for linking all the neurons from the next and the previous layer respectively. The prime purpose of it is to classify various classes using the features of the input image based on the training data set.

Combinations of those from Convolutional and pooling layers extracted features might yield better. The total probability obtained from the fully connected layer is one. The SoftMax is used as activation function in the fully connected layer's output.

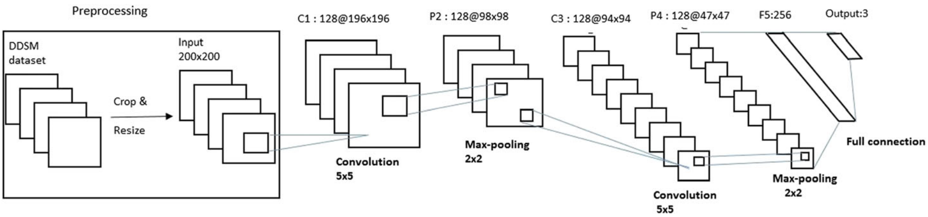


**Fig. 3** Mammogram image classification architecture using the deep CNN classification system

The SoftMax function sums to 1 by combining the vector of arbitrary real value scores into a vector value between 0 and 1.

These layers have been designed based on the Algorithm 1 Shown in Table 2.

### 3.7 LeNet training

Setting the training parameters of a network is very important, in order to produce a faster solution which is in due to the non-convex nature of the deep CNN cost function.

In the proposed framework, the weights and biases $\{Z^n, b^n\}$ of several distinct fully connected layers and convolutional layers are considered, where $n = 1, 3, 5, 7, 8$. The output probability vector $\hat{t} = [\hat{t}_1, \hat{t}_2, \ldots, \hat{t}_x]$ and the binary class label vector t = [t1, t2,..., tx] are minimized and utilized to train the network which is obtained as cross-entropy.

$$E(t, \hat{t}) = -\sum_{k=1}^{x} t_k \log \hat{t}_k \tag{4}$$

**Table 2** Algorithm 2: breast cancer detection using deep CNN

| |
|---|
| **Input:** DDSM dataset- 2D images<br>**Output:** CNN parameters S, predicted labels $\hat{t}$<br>**Training:**<br>**for** File **in** Listing do<br>Resize image data set<br>Create a matrix to store all flattened images<br>Initialize hypermeters such as batch size, the number of epochs,<br>the number of filters, kernel size,<br>number of classes.<br>Split dataset into training and testing set.<br>Convert class vectors to binary class matrices<br>Define CNN architecture<br>Train the defined architecture<br>end for<br>**Testing:**<br>**For** I **in** range ( ) do<br>Re-define CNN architecture<br>Test and validate the defined architecture<br>Classify the classes and generate confusion matrix<br>Calculate precision, recall, f1-score and accuracy<br>Load weights and save weights<br>Visualize accuracy and losses.<br>End for |

**Table 3** Statistical feature analysis using ROC for malignant diagnosis (100 samples)

| MARKER | AUC | SE.AUC | LOWERLIMIT | UPPERLIMIT | Z | *p* VALUE |
|---|---|---|---|---|---|---|
| Radius.mean | 0.93752 | 0.01046 | 0.91702 | 0.95801 | 41.83856 | 0 |
| Texture.mean | 0.77582 | 0.01973 | 0.73715 | 0.8145 | 13.9769 | 0 |
| Perimeter.mean | 0.9469 | 0.00928 | 0.92871 | 0.96509 | 48.1501 | 0 |
| Area.mean | 0.93832 | 0.01043 | 0.91787 | 0.95877 | 42.00747 | 0 |
| Smoothness.mean | 0.72204 | 0.02127 | 0.68036 | 0.76372 | 10.44103 | 0 |

A uniform distribution between (−c, c) with $c = \sqrt{\frac{6}{(fan_{out}+fan_{in})}}$ incurs the weights. The number of inputs and outputs to a neuron is $fan_{out}$ and $fan_{in}$ . The stochastic gradient descent is considered to update trainable parameters periodically [12] before which a mini batch is replaced with the cost function of training images which are evaluated with high training rate. Momentum is incorporated to speed up the learning to smoothen the network coverage (faster) and directions of gradient descent [34]. Higher the stabilization of the training error rate, the learning process turns to be finer. Classifying the error rates of validation set and training set, the training ends at some high point of the epochs.

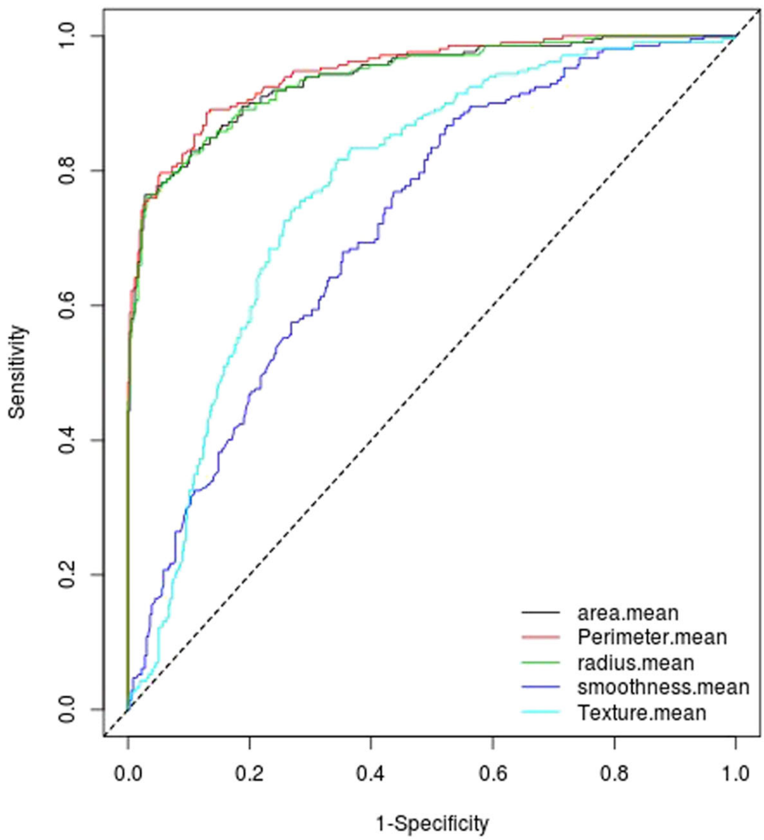


**Fig. 4** ROC Analysis for a malignant diagnosis (100 samples)

**Table 4** Comparison between activation functions

| Activation function | Accuracy (on test set) | | Test duration (in seconds) |
|---|---|---|---|
| | Overall Accuracy (%) | ACA (%) | |
| Sigmoid | 35.85 | 47.24 | 24.868 |
| Tanh | 62.26 | 72.06 | 8.481 |
| Rectifier | 96.23 | 97.46 | 6.069 |

### 3.8 Feature extraction and classification

The pre-processing applied is used to implement the test image to classify. The probability of each class is obtained by forward propagating through the network. The number of layers in CNN is increased to obtain the robustness after the stability of the training process [35, 46]. The maximum output probability over the average of all the probabilities is depicted as the predicted class.

Table 3 and Fig. 4 gives the statistical analysis of the malignant cases (100 samples) from the diagnosed images which reflects in the early detection of breast cancer. The parameters (Features) influencing the early prediction are Area, Radius, Perimeter and smoothness. The mean of those features is calculated, and a ROC is analysed. These features extracted are utilized for training the network and performing classification in Fully connected layer of CNN-MLP.

## 4 Observations and findings

This section elaborates the observations obtained, from applying the proposed frameworks by varying several hyperparameter on the mammogram images from DDSM dataset and hence tabulated for further understanding.

### 4.1 Network design and training

In [35, 46] the base work and guidance for designing a network using hyperparameter is discussed based on which the proposed system is designed and trained to measure the performance. Tuning of the hyperparameters in a CNN framework plays a major role in deciding the training process and designing a network accordingly. And this is verified based on a trial-and-error approach.

The Hyperparameters in the CNN framework is listed as 1) training relevant: initial learning rate, mini-batch size, number of epochs and momentum coefficients controlling the training process sequentially; 2) Model relevant: pooling method, the size of a kernel, number of filters convolutional layer need, the size of the pooling region and the number of neurons for the fully connected layer.

**Table 5** The system requirement/framework

| S.No | Operating system | Processor | RAM | Framework | Language |
|---|---|---|---|---|---|
| 1. | 64 bit, Windows 8.1 | Intel (R) Core (TM) i7 CPU | 20GB | Theano, TensorFlow, Keras | Python |

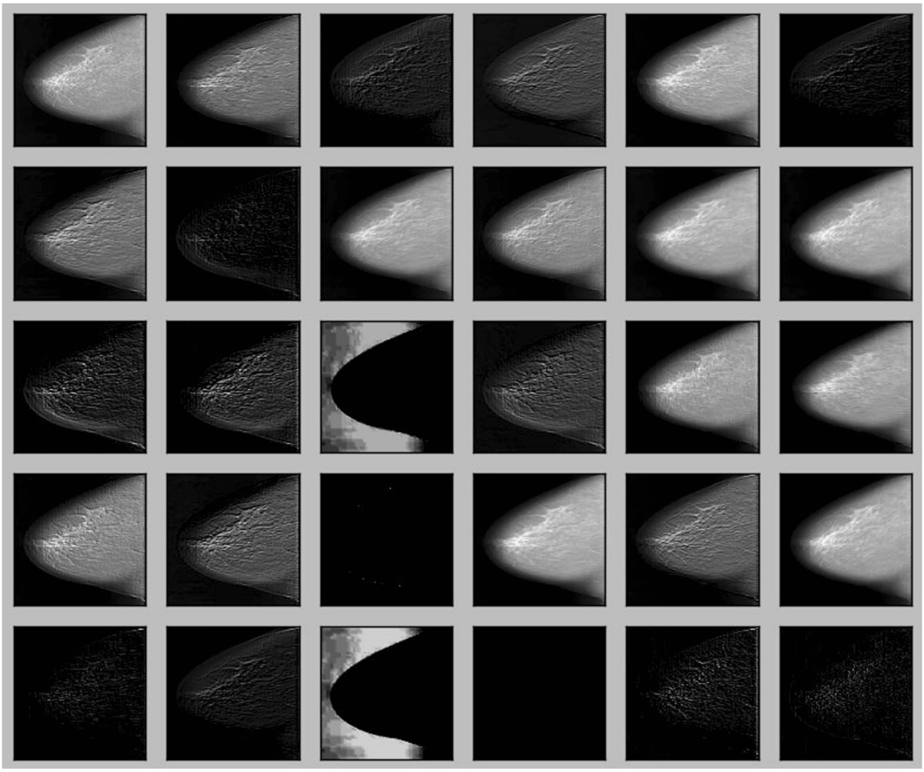

**Fig. 5** Foreground, Background and whole image mask of one image from the database

**Table 6** Model- relevant hyper parameters obtained

| Layer | Layer type | Hyper Parameter |
|---|---|---|
| Input | Input | Image size W x W: 200 × 200 |
| C1 | Convolution | Filter size k1 xk1:5 × 5<br>Number of filtersn1: 128<br>Stride: 1, Padding: 0<br>Activation function:<br>φ (x) = max (0, x) |
| P2 | Pooling | Pooling region size k2 x k2: 2 × 2<br>Pooling method: max-pooling<br>Stride: 2 |
| C3 | Convolution | Filter size k3 xk3:5 × 5<br>Number of filtersn2: 128<br>Stride: 1, Padding: 0<br>Activation function:<br>φ (x) = max (0, x) |
| P4 | Pooling | Pooling region size k4 x k4: 2 × 2<br>Pooling method: max-pooling<br>Stride: 2 |
| R1 | Regularization | Regularization method: Dropout<br>Dropout size p1: 0.5 |
| F5 | Full connection | Neurons number n3: 256<br>Activation function:<br>φ (x) = max (0, x) |
| R2 | Regularization | Regularization method: Dropout<br>Dropout size p2: 0.5 |

**Table 7** Training- relevant hyper parameters obtained

| Hyperparameter | Initial Learning rate | Batch size | Number of class | Number of epochs |
|---|---|---|---|---|
| Value | 0.01 | 32 | 3 | 20 |

The framework is analysed for different activation functions to observe the performance betterment if the appropriate activation function is used. Those are i) sigmoid $\varphi(\mathrm{u}) = \frac{1}{1+e^{-u}}$, ii) Hyperbolic tangent $\varphi(u) = 1.7159 \tanh(2/3)u$, and iii) Rectifier $\varphi(u) = \max(0, u)$. While comparing these activation functions, rectifiers that are also known as ReLU (Rectified Linear Units) is preferable, because it results in the neural network training several times faster with significant accuracy, this is shown in Table 4. ReLU is a non-saturating activation function which consists layers of neuron. Without affecting the receptive fields of Convolutional layer, it increases the nonlinear properties of the decision function and of the overall network.

It is found that the tangent function and sigmoid function gives poor performance for the proposed image classification and network learning respectively. The layer at the bottom takes a long time because of its non-symmetry around zero and to saturate top hidden layer quickly [6]. This effect leads to the poor classification performance, prevents from feature learning.

Table 5 details on the system requirement/Framework to perform this entire research. The language used is python which is fast to implement and flexible for the models CNN and DNN. Since high resolution images are processed and needs high computational setup 20 GB RAM and i7 Processor is used.

Sigmoid function on the other hand performs worse in the proposed learning network. The hyperparameters relevant to the training aspect can significantly affect by the convergence of the cost function, learning speed and classification purpose of the network.

### 4.2 Impact of foreground masks

To improve classification performance, images have to be filtered from background noise and hence image masks can be used to perform. Image classification tasks don't make use of the image mask in order to sketch out the mask that depends on labour intensive with high domain specific knowledge. This proposed experiment includes the following three cases.

Primarily, the whole images are used to train and test CNN without using masks which are called "CNN-whole-image". Then the images are segmented with image mask into

**Table 8** Evaluation of different network design

| Network design | Accuracy (on test set) | | Test duration (in seconds) |
|---|---|---|---|
| | Overall Accuracy (%) | ACA (%) | |
| Decrease W from 200 to 100 | 90.57 | 93.54 | 1.457 |
| Increase W from 200 to 300 | 92.45 | 93.65 | 12.402 |
| Decrease n1 from 128 to 64 | 94.34 | 95.23 | 3.478 |
| Decrease n1 and n2 from 128 to 64 | 88.68 | 90.85 | 2.672 |
| Decrease k1 and k3 from 5 to 3 | 88.68 | 92.23 | 3.528 |
| Increase k1 and k3 from 5 to 7 | 92.45 | 94.85 | 8.772 |
| Decrease n3 from 256 to 128 | 90.57 | 93.54 | 5.954 |
| Our design with Table 1 | 96.23 | 97.46 | 6.069 |

background and foreground. In the second case foreground is considered for training and testing the Framework. This is called "CNN-FG-Mask." And the third case considers the background for training and testing the CNN, that is called as "CNN-BG-Mask". As a result, CNN-whole-image out performs CNN-FG-Mask (86.07 vs 85.19%, and 96.23 vs 97.46%). This result shows that for images, the outside region not only contains the noise, but also have useful information to distinguish different classes. In this case of image classification, the noise occupies background and further noise is removed with masks to improve the classification

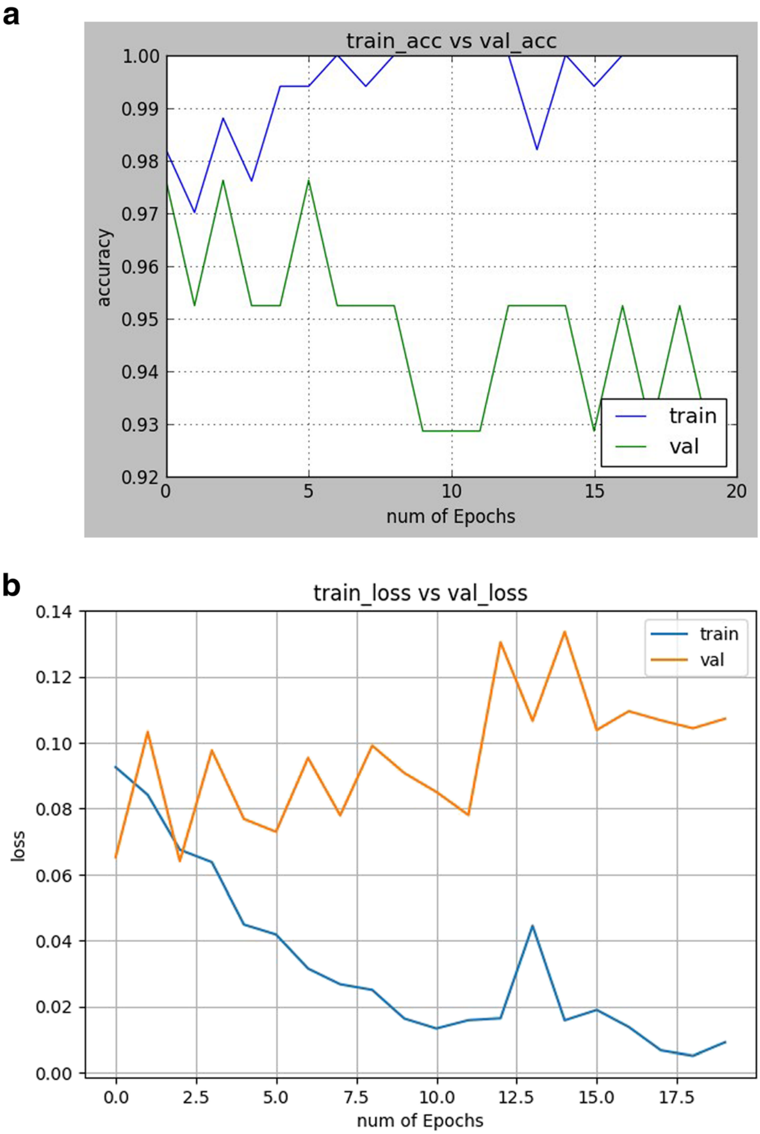


**Fig. 6** Representation of the learning rate. Smaller the learning rate slower the learning process is and the classification also degrades. (**a**) Classification accuracy with less learning rate as 0.01, Momentum coefficient = 0. (**b**) Loss rate graph in accordance with the learning rate

accuracy. To get better clarification Fig. 5. shows the one single image's classification from each classification.

This observation shows that the region outside the mask may contain the pattern that is same or like a region of interest. Such information could improve the network learning which in turn improves the prediction rate. To effectively utilize the image information obtained for

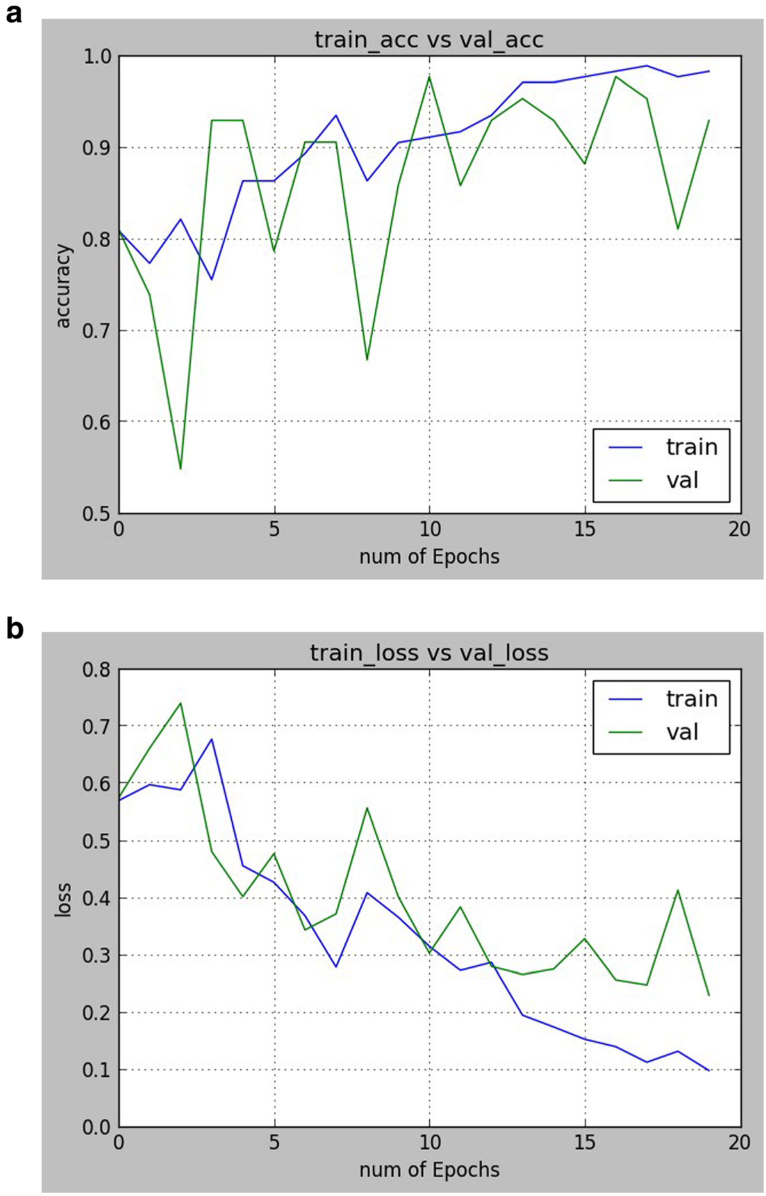


**Fig. 7** Demonstration of the impact of neurons and activation function in the fully connected layer (**a**) neurons 128 training accuracy = 88.68% (**b**) Training loss vs. Validation loss

better prediction. To avoid laborious manual labelling process the whole images are used directly without implementing masking process.

## 5 Experimental results

Various experiments are conducted to observe and find the effectiveness of the proposed framework based on the Hyperparameter. This paper evaluates mammogram images obtained from DDSM dataset. The average classification accuracy (ACA) is used as an evaluation metric [51] in this paper. It is formulated as follows:

$$\mathrm{ACA} = \frac{1}{\mathrm{x}}\sum_{\mathrm{i}=1}^{\mathrm{x}}\mathrm{CCR}_{\mathrm{i}} \tag{5}$$

The classification accuracy of class i with x number of classes denoted as group. The mean classification accuracy (MCA) gives the overall classification accuracy rate of the mammogram images.

### 5.1 Introduction of the mammogram dataset

The mammograms used for the commitment of this work were retrieved from the database of the DDSM (Digital Database for Screening Mammography) [6, 19] and the dimension of the mammograms has been fixed to 200 × 200 pixels. This database contains 262 mammograms and the intensity of every pixel varies between 0 and 255. This database also includes information about the class and the severity of abnormalities that may be present in the mammograms, as well as the coordinates of the centre of these abnormalities. It must be mentioned that only the mammogram images and the required information are needed to divide the images into three categories: patients with benign, malignant or normal are used. The dataset is grouped in the ratio of 8:2 from the total image of 262 as training (209) and testing (53). The test images are further classified and grouped based on three classifications such as benign, malignant and normal.

### 5.2 Experiments on hyper-parameters optimization

Hyperparameter is the most important factor to be considered to study about the training hyperparameter along with the model relevant hyperparameter and the network design to also obtain higher performance rate [35]. Hence various experiments are performed to analyse the same.

The input images from the DDSM database is fed randomly without any strict partition rule and is split in the ratio of 16:4:5 for training (167), validation (42) and test (53) images respectively [18, 35, 46, 59]. Section titled Network Design and Training, lists out the optimal hyperparameter extracted from the tuning process in Tables 6 and 7. The classification

**Table 9** Performance measurement of CNN classifier

| Class | Precision | Recall | F1 Score | Support |
|---|---|---|---|---|
| Class 0 (Normal) | 0.94 | 1.00 | 0.97 | 15 |
| Class 1 (Malignant) | 0.95 | 1.00 | 0.97 | 19 |
| Class 2 (Benign) | 1.00 | 0.89 | 0.94 | 19 |

**Table 10** Mean and average classification accuracy of the proposed D-CNN with DDSM image set

| METHOD | ACCURACY ON TEST SET | |
|---|---|---|
| CNN-Whole-Img (Proposed) | ACA (%) | 97.46 |
| | MCA (%) | 96.23 |
| CNN- Align | ACA (%) | 88.86 |
| | MCA (%) | 88.71 |
| CNN- FG- Mask | ACA (%) | 85.19 |
| | MCA (%) | 86.07 |
| CNN- BG- Mask | ACA (%) | 67.46 |
| | MCA (%) | 66.54 |

accuracy obtained from the network architecture on the test set corresponding to the various options available are shown in Table 8 corresponding with the Table 6.

The learning curves of MCA are demonstrated in Figs. 6 and 7, which is the impact of the hyperparameter relevant to training, validation and testing. Table 2 depicts the optimal values of the framework considering one hyperparameter at a time.

Fig. 6 (a) represents the relation between the learning rate and the learning processes in terms of MCA and epochs, to predict the training accuracy and validation accuracy. Fig. 6 (b). shows the training loss with the same learning rate, to predict the training loss and validation loss. Also, Fig. 7 (a) & (b). demonstrates the impacts of neuron and MCA [35] with the epochs along with the training loss vs validation loss. It proves hyper-parameters of CNN have a great impact on the performance of network design. To get a high-performance hyper-parameter of CNN have to be tuned properly. From Table 4 it is inferred that the accuracy of the proposed system achieves 97.46% using DDSM dataset.

A shuffled dataset with the benign, malignant and normal is considered for training, testing and validating. It is inferred that recall, precision and F1-score with the $5 \times 5$ convolutional layer and $2 \times 2$ max pooling layer and a dropout rate of 0.5, 256 neurons in fully connected layers and ReLUs are obtained best. Table 9 depicts the performance measurement of CNN classifier.

### 5.3 Cell foreground masks impact

As described in Section titled Impact of Foreground Masks, it is inferred that the Mean Classification Accuracy (MCA) in training the CNN using background and foreground, the

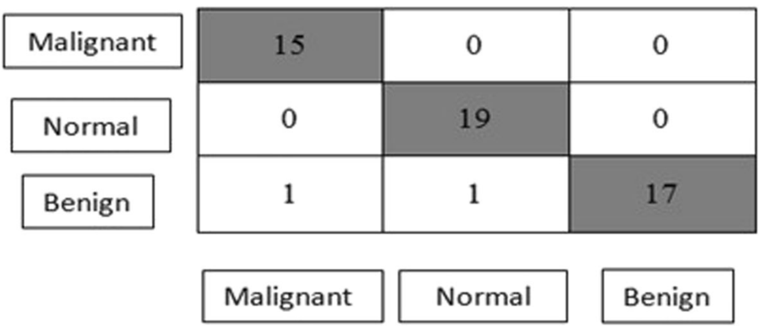


**Fig. 8** Confusion matrix with classes Benign, Malignant and Normal

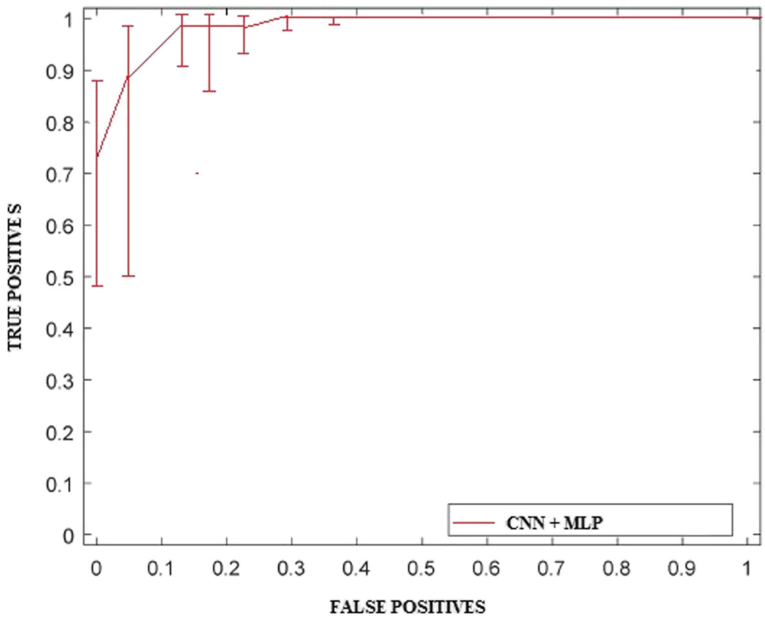


**Fig. 9** ROC plotting with the false positives and true positives for the proposed system

CNN-Whole-Img is 12.27% higher than CNN-FG-Mask. Unexpectedly, the CNN-BG-Mask achieves MCA and ACA of 66.54% and about 67.46% respectively by only using the background information of images. These results are reported in Table 10.

The overall classification performance is much appreciated. The normal and malignant classes obtain the highest classification accuracy, whereas the benign class acquires the lower accuracy rate with the error 2%. They are easily being misclassified as Normal and Malignant. This is shown as the confusion matrix of the CNN in Fig. 8.

Table 11 shows the performance analysis of deep learning convolutional network with the proposed network and other existing networks. The quantitative measurements like sensitivity, specificity, classification accuracy and receiving operating characteristic curve (AUC) based on the true positives, true negatives, false negatives and false positives. The performance result of the proposed system using CNN + MLP shows the highest classification accuracy in both

**Table 11** Performance analysis of different classification models

| Method | Test Dataset Accuracy | | | Training Dataset Accuracy | | | |
|---|---|---|---|---|---|---|---|
| | ACA(%) | SPECIFICITY | SENSITIVITY | MCA(%) | AUC | PPV(%) | NPV(%) |
| DL-CNN + SVM | 82 | 0.74 | 0.845 | 75 | 0.8832 | 84.59 | 49.50 |
| DL-CNN + NN | 88 | 0.79 | 0.9037 | 79 | 0.9131 | 88.58 | 55.42 |
| level set + DL-CNN + SVM | 91 | 0.83 | 0.93 | 81 | 0.925 | 90.22 | 94.29 |
| level set + DL-CNN + NN | 93 | 0.91 | 0.954 | 83 | 0.951 | 93.41 | 96.24 |
| DL-CNN + MLP (Proposed system) | 97.46 | 1.0 | 0.9882 | 96.23 | 0.9842 | 98.60 | 100.00 |

**Table 12** List of various classification method with accuracy %

| S.No | Method | Average Accuracy |
|---|---|---|
| 1 | Chih-Min Lin et al., (June 2014) [35] | 92% |
| 2 | John Zakos, B.V. (March 200) [25] | 88.9% |
| 3 | Mohammed Y et al., (April 2012) [37] | 67.77% |
| 4 | Kimme et al., (1977) [28] | 74% |
| 5 | S.K. Kinoshita et al., (1998) [29] | 81% |
| 6 | M. Sameti et al., (1998) [19] | 72% |
| 7 | L. Wei et al., (March 2005) [59] | 85% |
| 8 | M.A.Alolfe et al., (Nov 2009) [1] | 82.5% |
| 9 | Oliver A et al., (2006) [40] | 90% |
| 10 | Priebe et al., (July 1994) [42] | 88% |
| 11 | Proposed Method (CNN) | 96.23% |

training and testing dataset of 97.46 and 96.23% respectively. Specificity of the proposed system is 1.0 which shows that the percentage of the false positive is nearly 0.

Further from the AUC of 0.9842, shows the increase in true positives when compared to other models. The ROC curve of the proposed system is further shown in the Fig. 9. The farthest left side of the curve shows the least uncertainty and the farthest right shows the highest positivity rate. Thus, the proposed system is determined to have the high diagnostic rate.

Table 12 shows the various classification accuracy obtained from various research works proposed by authors, from which it is inferred that the proposed algorithm gives the higher accuracy rate. Currently machine learning and deep learning techniques are used in various applications [3, 5, 8, 15, 16, 22, 27, 32, 36, 39, 44, 45, 48, 49, 52–56, 58] which can be utilized for the future development in the medical image processing.

## 6 Conclusion

According to the statistics provided by American cancer society the breast cancer is termed to be the leading cause of death in women. To decrease the mortality rate and incidence rate of breast cancer in women proper treatment must be given which is followed by early detection of cancer. Thus, a novel framework for early detection of cancer has been proposed and evaluated. The network evaluated is deep convolutional neural network which utilizes both deep neural network and convolutional neural network. The database used for the research is DDSM (Digital Database for Screening Mammography) which contains chain code which promotes in fast processing of without requiring any manual labelling in contrast with the MIAS (Mammographic Image Analysis Society) database which is alternatively available. The database is first resized and cropped to $200 \times 200$ pixel and the edges are preserved using a modified bilateral filter. The features like area, smoothness, perimeter and radius is extracted and the convolutional layers are trained and classified using a multi-layer perceptron with back propagation in the fully connected layer which improved the classification rate. To improve the network training speed the grid search hyperparameter tunning was performed which increased the performance accuracy to 97.46% with dropout size as 0.5, learning rate 0.01, momentum 0.0, optimizer as adadelta and 5 hidden layers. The positive predictive values and negative predictive values were inferred as 98.60 and 100% respectively decreasing the false prediction rate which is supported by the Fig. 9. It is concluded that deep convolutional neural network with the grid search hyper parameter tuning, the overall classification accuracy and mean classification accuracy are improved to be 96.23% & 97.46% respectively. With the obtained classification rate the treatment can be initiated in the early stage. The proposed system focuses on the early detection of the breast cancer to facilitate the treatment, but due to the false positive or false negative rate the screening test is not sufficient to reduce the mortality rate. Thus, the early prediction of the breast must be performed by undergoing the diagnosis test which involves the biomarkers for evaluation.

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

## References


1. Alolfe MA, Mohamed WA, Youssef A-BM, Mohamed AS, Kadah YM (2009) Computer aided diagnosis in digital mammography using support vector machine and linear discriminant analysis classifi- cation, in Proc. 16th IEEE Int. Conf. Image Processing (ICIP). 2609–2612. doi:https://doi.org/10.1109/ICIP.2009.5413992
2. American Cancer Society, Cancer Facts and Figures (2017), American Cancer Society, Atlanta, Ga, USA from https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2012.html
3. Amin J, Sharif M, Yasmin M, Ali H et al (2017) A unique method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions. J Comput Sci 119:153–164
4. Arevalo J, González FA, Ramos-Pollán R, Oliveira JL, Lopez MAG (2015) Convolutional neural networks for mammography mass lesion classification. In Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE. 797–800. IEEE
5. Arunkumar N, Ramkumar K, Venkatraman V, Abdulhay E et al (2017) Classification of focal and non-focal EEG using entropies. Pattern Recogn Lett 94:112–117
6. Bengio Y (2012) Practical recommendations for gradient-based training of deep architectures. In neural networks: tricks of the trade. Springer, Berlin, pp 437–478
7. Bhattacherjee A, Roy S, Paul S, Roy P, Kausar N, Dey N (2016). Classification approach for breast cancer detection using back propagation neural network: a study. In Biomedical image analysis and mining techniques for improved health outcomes (pp. 210–221). IGI Global
8. Bokhari STF, Sharif M, Yasmin M et al (2017) Fundus Image Segmentation and Feature Extraction for the Detection of Glaucoma: A New Approach. Curr Med Imaging Rev 13
9. Boureau YL, Ponce J, Le Cun Y (2010). A theoretical analysis of feature pooling in visual recognition. In Proceedings of the 27th international conference on machine learning (ICML-10). 111–118
10. Carneiro G, Nascimento J, Bradley AP (2017) Automated analysis of unregistered multi-view mammograms with deep learning. IEEE Trans Med Imaging 36(11):2355–2365
11. Elmoufidi A, El Fahssi K, Jai-andaloussi S, Sekkaki A, Gwenole Q, Lamard M (2018) Anomaly classification in digital mammography based on multiple-instance learning. IET Image Process 12(3):320–328
12. Ertosun MG, Rubin DL (2015) Probabilistic visual search for masses within mammography images using deep learning. In Bioinformatics and Biomedicine (BIBM), 2015 IEEE International Conference on (pp. 1310–1315). IEEE
13. Estève J, Kricker A, Ferlay J, Parkin DM (1993) Facts and figures of cancer in the European Community. In Facts and figures of cancer in the European Community
14. Eurostat, Health statistics, Atlas on mortality in the Euro-pean Union (2002) European Commission from http://ec.europa.eu/eurostat/en/web/products-statistical-books/-/KS-30-08-357
15. Fernandes SL, Chakraborty B, Gurupur VP, Ananth PG (2016) Early skin Cancer detection using computer aided diagnosis techniques. J Integr Des Process Sci 20(1):33–43
16. Fernandes SL, Gurupur VP, Lin H, Martis RJ (2017) A novel fusion approach for early lung Cancer detection using computer aided diagnosis techniques. J Med Imaging Health Inform 7(8):1841–1850
17. Foggia P, Percannella G, Soda P, Vento M (2013) Benchmarking HEp-2 cells classification methods. IEEE Trans Med Imaging 32(10):1878–1889
18. Gao Z, Wang L, Zhou L, Zhang J (2017) Hep-2 cell image classification with deep convolutional neural networks. IEEE J Biomed Health Inform 21(2):416–428
19. Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics 249–256
20. Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS (2016) Deep learning for visual understanding: a review. Neurocomputing 187:27–48

21. Jadoon MM, Zhang Q, Haq IU, Butt S, Jadoon A (2017) Three-Class Mammogram Classification Based on Descriptive CNN Features. BioMed Res Int
22. Jain VK, Kumar S et al (2017) Extraction of emotions from multilingual text using intelligent text processing and computational linguistics. J Comput Sci 21:316–326
23. Jamieson AR, Drukker K, Giger ML (2012) Breast image feature learning with adaptive deconvolutional networks. SPIE Med Imaging, Strony 2012:831506–831506
24. Jiao Z, Gao X, Wang Y, Li J (2016) A deep feature based framework for breast masses classification. Neurocomputing 197:221–231
25. John Zakos BV (2001). A Computer-Aided Diagnosis System for Digital. IEEE
26. Kamal S, Dey N, Nimmy SF, Ripon SH, Ali NY, Ashour AS et al (2018) Evolutionary framework for coding area selection from cancer data. Neural Comput & Applic 29(4):1015–1037
27. Khan MW, Sharif M, Yasmin M et al (2016) A new approach of cup to disk ratio based Glaucoma detection using fundus images. J Integr Des Process Sci 20(1):77–94
28. Kimme C, O'Loughlin BJ, Sklansky J (1977) Automatic detection of suspicious abnormalities in breast radiographs. In: Klinger A, Fu KS, Kunii TL (eds) Data structures, computer graphics, and pattern recognition. Academic, New York, pp 427–447
29. Kinoshita SK, Marques PMA, Slates AFF, Marana HRC, Ferrari RJ, Villela RL (2016) Detection and characterization of mammographic masses by artificial neural network, in Proc. 4th Int. Workshop Digital Mammography 489–490
30. Koay J, Herry C, Frize M (2004) Analysis of breast thermography with an artificial neural network. The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, pp 1159–1162
31. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems 1097–1105
32. Kumar S, Jain N et al (2017) Rough set based effective technique of image watermarking. J Comput Sci 119:121–137
33. Lan K, Wang DT, Fong S, Liu LS, Wong KK, Dey N (2018) A survey of data mining and deep learning in bioinformatics. J Med Syst 42(8):139
34. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
35. Lin CM, Hou YL, Chen TY, Chen KH (2014) Breast nodules computer-aided diagnostic system design using fuzzy cerebellar model neural networks. IEEE Trans Fuzzy Syst 22(3):693–699
36. Martis RJ, Lin H, Gurupur VP et al (2017) Editorial: Frontiers in Development of Intelligent Applications for Medical Imaging Processing and Computer Vision. Comput Biol Med
37. Mohammed Rasoul A Al-Hadidi MY-G (2012) Solving mammography problems of breast cancer detection using. Ind J Sci Technol 5(4)
38. Mourali N, Haro on the leading cause of mortality in the world (2010) Livret Santé, https://livretsante.com/a_la_une/haro-premiere-cause-mortalite-monde/
39. Naqi SM, Sharif M, Yasmin M, et al (2017) Lung Nodule Detection Using Polygon Approximation and Hybrid Features from Lung CT Images. Curr Med Imaging Rev 13
40. Oliver A, Marti J, Marti R, Bosch A, Freixenet J (2009) A new approach to the classification of mammographic masses and normal breast tissue," in Proc. 18th Int. Conf. Pattern Recognition 4; 707–710. doi:https://doi.org/10.1109/ICPR.2006.113
41. Petersen K, Nielsen M, Diao P, Karssemeijer N, Lillholm M (2014) Breast tissue segmentation and mammographic risk scoring using deep learning. In international workshop on digital mammography. Springer, Cham, pp 88–94
42. Priebe CE, Lorey RA, Marchette DJ, Solka JL, Rogers GW (1994) Nonparametric spatio-temporal change point analysis for early detection in mammography," in Proc. 2nd Int. Workshop Digital Mammography 111–120
43. Quellec G, Lamard M, Cozic M, Coatrieux G, Cazuguel G (2016) Multiple-instance learning for anomaly detection in digital mammography. IEEE Trans Med Imaging 35(7):1604–1614
44. Rajinikanth V, Sri Madhava Raja N, Satapathy SC et al (2017) Otsu's multi-thresholding and active contour snake model to segment Dermoscopy images. J Med Imaging Health Inform 7(8):1837–1840
45. Rajinikanth V, Satapathy SC et al (2017) Entropy based segmentation of tumor from brain MR images – a study with teaching learning based optimization. Pattern Recogn Lett 94:87–95
46. Sameti M, Morgan-Parkes J, Ward RK, Palcic B (1998) Classifying image features in the last screening mammograms prior to detection of a malignant mass. Comput Imaging Vis 13:127–134
47. Selvarasu N, Nachaippan A, Nandhitha NM (2012) Image processing techniques and neural networks for automated cancer analysis from breast thermographs-a review. Ind J Comput Sci Eng (IJCSE) 3.1:133–137
48. Shabbira B, Sharifa M, Nisara W, Yasmina M et al (2016) Automatic cotton wool spots extraction in retinal images using texture segmentation and Gabor wavelet. J Integr Des Process Sci 20(1):65–76

49. Shah JH, Chen Z, Sharif M, Yasmin M et al. (2017) A Novel Biomechanics Based Approach For Person Re-identification By Generating Dense Color Sift Salience Features. J Mech Med Biol 17(7)
50. Sharif Razavian A, Azizpour H, Sullivan J, Carlsson S (2014) CNN features off-the-shelf: an astounding baseline for recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops 806–813
51. Sickles EA (1997) Breast cancer screening outcomes in women ages 40-49: clinical experience with service screening using modern mammography. JNCI Monographs 1997(22):99–104
52. Sri Madhava Raja N, Rajinikanth V et al (2017) Segmentation of breast thermal images using Kapur's entropy and hidden Markov random field. J Med Imaging Health Inform 7(8):1825–1829
53. Steven L, Fernandes G, Bala J (2016) ODROID XU4 based implementation of decision level fusion approach for matching computer-generated sketches. J Comput Sci 16:217–224
54. Steven L, Fernandes G, Bala J (2016) Fusion of sparse representation and dictionary matching for identifications of humans in uncontrolled environment. Comput Biol Med 76:215–237
55. Steven L, Fernandes G, Bala J (2017) A novel decision support for composite sketch matching using fusion of probabilistic neural network and dictionary matching. Curr Med Imag Rev 13(2):176–184
56. Steven L Fernandes G Bala J (2017) A Comparative Study on various State of the Art Face Recognition Techniques under varying Facial Expressions. Int Arab J Info Technol 14(2)
57. Tan M, Zheng B, Leader JK, Gur D (2016) Association between changes in mammographic image features and risk for near-term breast Cancer development. IEEE Trans Med Imaging 35(7):1719–1728
58. Tyagi SKS, Jain DK, Muhuri P et al (2017) Thermal-aware power-efficient deadline based task allocation in multi-Core processor. J Comput Sci 119:112–120
59. Wei L, Yang Y, Nishikawa RM, Jiang Y (2005) A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications. IEEE Trans Med Imaging 24(3):371–380
60. Zhang XS (2014) A new approach for clustered MCs classification with sparse features learning and TWSVM. Sci World J

**Saranyaraj D.** Currently pursuing Ph. D in the area medical image processing from Anna University- MIT Campus, Chennai. She received the M. Tech degree from Vel Tech University, Chennai and did her M. Tech project from C-DAC, Pune on High Performance computation and Image processing. She has 2 International Journal Publications and 3 International Conference publications. She received a funding from Vision Group of science and Technology, A Karnataka state government fund. Her Area of Interest is Medical Image Processing, High Performance computation, Pattern recognition, Machine learning.

**Manikandan M.** completed UG and PG from Madras University and his Ph.D. from Anna University in the area of Image processing. He has 27 International Journal publication and 33 conference publications. His area of specialization VLSI, embedded systems and computer architectures. He visited various universities in France, Italy, Switzerland and Malaysia for research collaboration and paper presentation. He also conducted short term courses in the area of VLSI and Image Processing.

**Maheswari S.** She pursed BE degree from Madras Institute of Technology, Chennai. She attended 3rd and 4th International Conference on Signal Processing, Communication and Networking. Her area of interest is Medical Image Processing, Machine learning and Neural networks.

## Affiliations

**Saranyaraj D.**[1] **· Manikandan M.**[1] **· Maheswari S.**[1]

Manikandan M.
maniiz@annauniv.edu

Maheswari S.
maheswariss96@gmail.com

[1] Department of Electronics Engineering, MIT Campus, Anna University, Chennai, Tamil Nadu, India