

Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification

Prabhpreet Kaur^{*}, Gurvinder Singh, Parminder Kaur

Guru Nanak Dev University, Amritsar, India



ARTICLE INFO

Keywords

Mammography
Breast cancer
Deep learning
Multiclass support vector machine
K-mean clustering
Classification learner

ABSTRACT

Breast cancer is an important cause of death in females. Early recognition of this disease with the assistance of mammography reduces the death rate. Deep learning (DL) is an approach being utilized and requested by radiologists to assist in making an accurate diagnosis, and it can help to improve outcome predictions. This paper includes a new approach, applied on the Mini-MIAS dataset of 322 images, involving a pre-processing method and inbuilt feature extraction using K-means clustering for Speed-Up Robust Features (SURF) selection. A new layer is added at the classification level, which carries out a ratio of 70% training to 30% testing of the deep neural network and Multiclass Support Vector Machine (MSVM). The outcome described herein demonstrates that the accuracy rate of the proposed automated DL method using K-means clustering with MSVM is improved as compared with a decision tree model. Experimental results show that the average accuracy (ACC) rates of the three classes, i.e., normal, benign and malignant cancer, using the proposed method, are 95%, 94% and 98%, respectively. The increased sensitivity rate is 3%, specificity is 2%, and Receiver Operating Characteristics (ROC) area is 0.99 using SVM compared to the Multi-Layer Perceptron (MLP) and J48+K-mean clustering WEKA manual approach. A 10-fold cross validation was used, and the obtained results for the Support Vector Machine (SVM), K-nearest neighbour (KNN), linear discriminant analysis (LDA) and Decision Tree were 96.9%, 93.8%, 89.7% and 88.7%, respectively.

1. Introduction

Malignancy is a serious medical ailment. Diseased cells can be irregularities or can develop into a mass known as a tumor. Breast cancer is the second most significant cause of death in females throughout the world [1]. The American Cancer Society in United States evaluated that in the year 2018, there will be approximately 266,000 instances of intrusive breast growth analysed and roughly 63,960 deaths from the disease [2]. In order to reduce the death rate, early steps in disease recognition are considered. It requires an exact basis on which to distinguish benign and malignant tumors. Most frequently, medical image reports for early recognition and diagnosis of breast cancer varieties contain mammography plus ultra-sonography data. As the number of patients increases, it becomes more difficult for radiologists to complete the diagnostic process in the limited available time. The motivation of this work is to assist radiologists in increasing the rapid and accurate detection rate of breast cancer using deep learning (DL) and to compare this method to the manual system using WEKA on single

images, which is more time consuming.

Section 2 includes a review of different computer-aided diagnosis (CAD) methods for DL, and different classification algorithms used for validation. Section 3 explains the proposed methodology, and Section 4 discusses the qualitative and quantitative analysis of the proposed workflow and validation process. Section 5 summarizes the conclusions and discusses future work.

2. Related work

The analysis and correct judgement of mammograms is an extremely large and difficult task for experienced radiologists. Thus, in this scenario, a different framework, employing CAD, feature extraction and deep learning, was hypothesized to be helpful for accurate prediction.

2.1. Computer-aided diagnosis (CAD)

CAD is distinguish from other approaches because of its highlights;

DOIs of original article: <https://doi.org/10.1016/j imu.2019.01.001>, <https://doi.org/10.1016/j imu.2019.100246>.

* Corresponding author.

E-mail address: prabhpreet.cst@gndu.ac.in (P. Kaur).

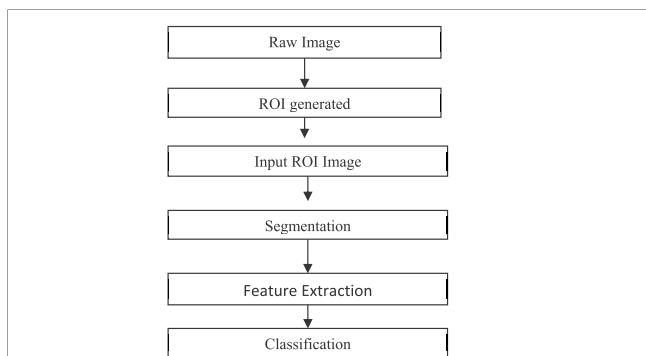


Fig. 1. Feature extraction process.

for example, the total automatization of the calculation chain (i.e., segmentation → feature extraction → classification) and the large database of mammographic images used to assess the CAD execution [3]. Recently, there has been an enthusiasm for the improvement and utilization of mammographic images, with the essential goals of expanding analytic accuracy and providing salient information to medical personnel. These methods have been utilized to create a CAD/computer-aided diagnostic (CADx) framework [9]. The CAD system is comprised of 91 breast ultrasound images, in which identifying the region of interest (ROI) is critical. Marker controlled watershed change is utilized in the division step [4]. The techniques employed are textural feature descriptors, including homogeneity, contrast, energy, correlation and entropy, along with KNN, SVM, and classification and regression trees (CART) for developing the proposed strategy. Additionally, the CAD framework [5], corporate and highlights from the Breast Imaging Reporting and Data System (BI-RADS) for breast ultrasound is addressed [28].

2.2. Feature extraction

The machine learning strategies were utilized as classifiers, and their execution was contrasted with the Student's t-test. Other arrangement

strategies can be connected [6] after texture analysis. For example, using an artificial neural network (ANN), a large portion of benign cases can be accurately distinguished from malignancies. With regard to texture features, texture analysis is connected to the assessment of textural changes in the breast tissue. In Varela et al. (2007), the conduct of an iris channel at various scales was explored to identify malignant masses on mammograms [7]. A back propagation neural system classifier was prepared to lessen the quantity of false positives. The framework steps are shown in Fig. 1 for production and assessment of two data sets. Active contour methods [8] were employed for segmentation, including an edge-based active contour method demonstrating the use of a swelling/emptying power with a damping coefficient (EM), a geometric active contour model (GAC), and an active contour without edges (ACWE). The trials utilized a test set constructed of 100 cases taken from two freely accessible databases: the Digital Database for Screening Mammography (DDSM) and the Mammographic Image Analysis Society (MIAS) database.

2.3. Deep learning techniques

In a previous study [11], the research focus was on developing a novel way to address the prediction of breast cancer risk based on deep learning technology, along with CAD. A deep learning-based CAD plot was produced for hazard appraisal, which is comprised of two modules: the adaptive feature identification module and the risk prediction module. The risk prediction module is actualized by a MLP classifier [30], which delivers a risk score to anticipate the tumor growth, and with a mammography-perceivable tumor. CAD with DL Architecture [12,13], as presented in Fig. 2, is an applicable report on DL-based CAD augmentation or CADx, for the determination of benign or malignant tumor status, by avoiding problems caused by inaccurate image processing results (e.g., boundary segmentation), as well as the classification algorithm's robust features. Hybridisation of a level set strategy with the Deep Convolution Neural Networks (DCNNs) system is supervised in the feature representation. First, multi-task transfer learning DCNN [14] was represented, followed by a single-task transfer learning technique; then a DCNN was prepared and evaluated using screen-film mammograms (SFMs). SFMs and digital mammograms (DMs) were

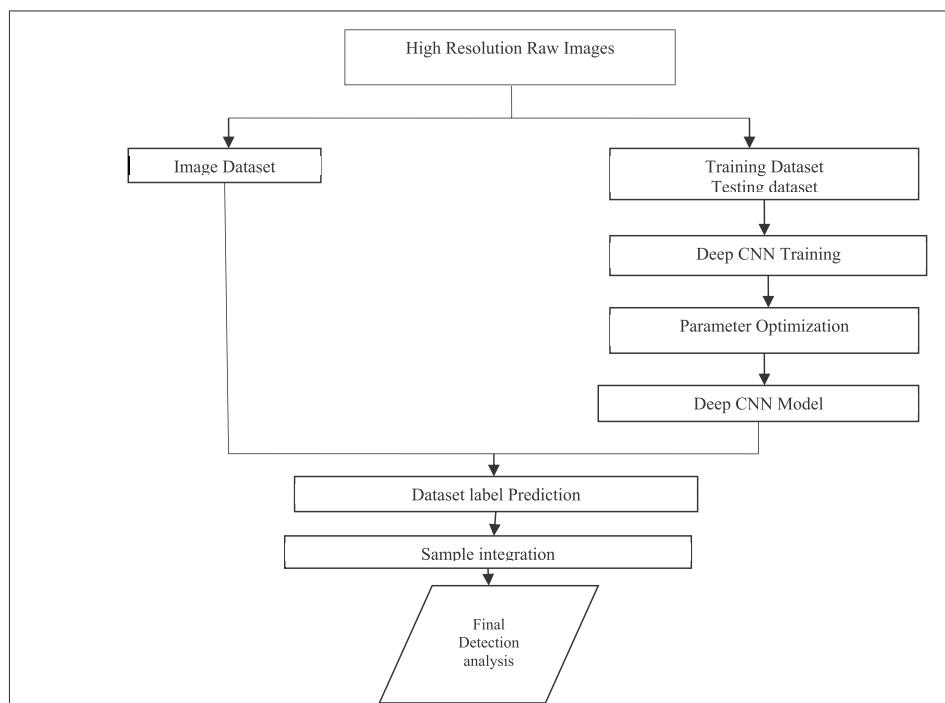


Fig. 2. Automation steps for DL approach.

Table 1
State-of-the-art classification and DL approaches for breast cancer detection in mammography images

Author	Dataset	Features Used	Classification Technique	Outcome	Parameters	Advantages	Disadvantages
Ref. No.							
H.P. (1996) [6]	86 mammograms from 54 cases (26 benign and 28 malignant) considered as case samples	Visibility rank and number of epochs	Region of Interest (ROI) and Spatial grey level dependence (SGLD) matrices	Spatial grey level dependence (SGLD) matrices	Visibility rank vary from 1 to 5 and number of epochs vary from 0 to 0000	Reduce biopsies, increase positive predictive valued function	Feature extraction and classification approaches much are refined
C.V. (2005) [7]	Mammogram images taken from the hospital at health district of Santiago de Compostable, Spain	Grey level, texture, contour, morphological Iris Filter	Adaptive threshold	Sensitivity of lesion-based and case-based evaluation were 88% and 94%, respectively	Simplicity, adaptability correctly enhances the abnormality mass of tumor	Dependency on optimal size	
R.B. (2006) [3]	Dataset taken from set of hospitals, collaborated with MAGIC-5 project	Textural features from the grey tone spatial dependence matrix GTSDM	Neural Network	ROI characterization	4.23 false positives per image are found at 80% of mass sensitivity	No need for image normalization, high capacity to learn	Unreliable, high computational time, low accurate outcome
M.K (2010) [10]	Dataset gathered from 40,075 women with breast cancer	Mortality based feature	Historical screening group & non-screening group	Mammography screening with reduction in rate of death from breast cancer	Rate ratio = 0.72 Confidence Interval(CI) = 95%	To eliminate some temporal changes, high nationwide design, large size and proportion of women in screening	Given time not enough to explore their potential, overestimation may be evaluated
M. A. (2015) [4]	91 breast UT images collected from Damietta ecology institute	Statistical & textural based features	SVM, CART, KNN	Benign & malignant were classified from focal lesion	CART classification rate = 83.75%	Efficient method, best results by features analysis	Less features to be introduced
J.S. (2015) [5]	283 pathological images obtained from database	Effective computable features	BI-RADS System	Distinguished lesions between benign and malignant cancer cells	From SVM AUC = 0.844 Accuracy = 77.7% From random forest AUC = 0.833	Individual optimum feature set, better performance	Lack of hybridized classifiers, less feature set introduced, less robust
Y.Q. (2016) [11]	270 image cases from FFDM database	Adaptive feature identification module and risk prediction module	MLP classifier, CNN, Deep learning, CAD scheme	Predicted short term risk assessment model	Accuracy = 78.5% Accuracy = 71.4% Positive predicted value = 69.2% Negative predicted value = 74.2% 10 times of 10 fold cross validation	Reliable technique, improved performance, high training skills	Less robust, limited dataset variety
J.C. (2016) [12]	Dataset approved from Taipei Veterans General Hospital and the Lung Image Database Consortium (LIDC)	Automatic noise tolerance advantage and feature exploration mechanism	SDAE-based CADx framework used to perform on mammograms images	Significant performance boost algorithm over the 2 conventional method	High noise tolerance, uncover reliable features from training dataset, high performance	Complex technique, not easily interoperated	
R. K. S. (2017) [14]	Digitized screen-film mammograms (DFMs) and digital mammograms (DMs) taken from digital database for screening mammography	Digital mammograms (DFMs) and screen-film mammograms (SFMs)	The multitask transfer learning DCNN and single-task transfer learning DCNN	The multitask transfer learning DCNN execution was essentially higher when contrasted with the single-task DCNN	Significance value, $p = 0.007$	Increased generalization capacities, high potential strategic approach	Difficult to interpret, low computation time, lack of resources for use
S. D. (2017) [13]	Standard benchmark breast cancer dataset (MIAS and BCDR)	Supervised features	Method used: SVM + DCNN	Hybridization of level set method with DCNN, used to learn the feature representation in a supervised way	Accuracy = 99% Sensitivity = 98.75% Specificity = 100% AUC = 97.37%	Increased classification accuracy, reduce false positives precisely	Highly complex valued network used, must introduce more standard datasets for future work
M. C. (2017) [8]	Digital Database for Screening Mammography (DDSM) and Mammographic Image Analysis Society (MIAS) database	Morphological features	A geometric active contour (GAC) and active contour without edges (ACWE) and damping coefficient (EM)	Formulated an adequate diagnostic hypothesis for the individual methods (malignant & benign cases together)	The average values of the EF Index amounted to: 0.07 ± 0.06 (EM), 0.07 ± 0.05 (GAC) 0.34 ± 0.32 (ACWE).	High noise reductions, accurate segmentation, lessen computational time	More contour parameters must be calculated, lack of methodology in order to define micro-califications
A.R. (2018) [15]	Hematoxylin and eosin stained breast histology microscopy image dataset delivered by a fragment of ICLAR Grand Challenge on Breast Cancer Histology Images	The image dataset is an extension of the dataset consisting of 400H&E stain images	VGG, Inception and ResNet used	Developed breast cancer histology image classification	Accuracy = 93.8% Sensitivity = 96.5% Specificity = 88% AUC = 93.37%	High diagnostic capacity, robust, simple, dynamic approach	Prone to over-fitting, low training rate

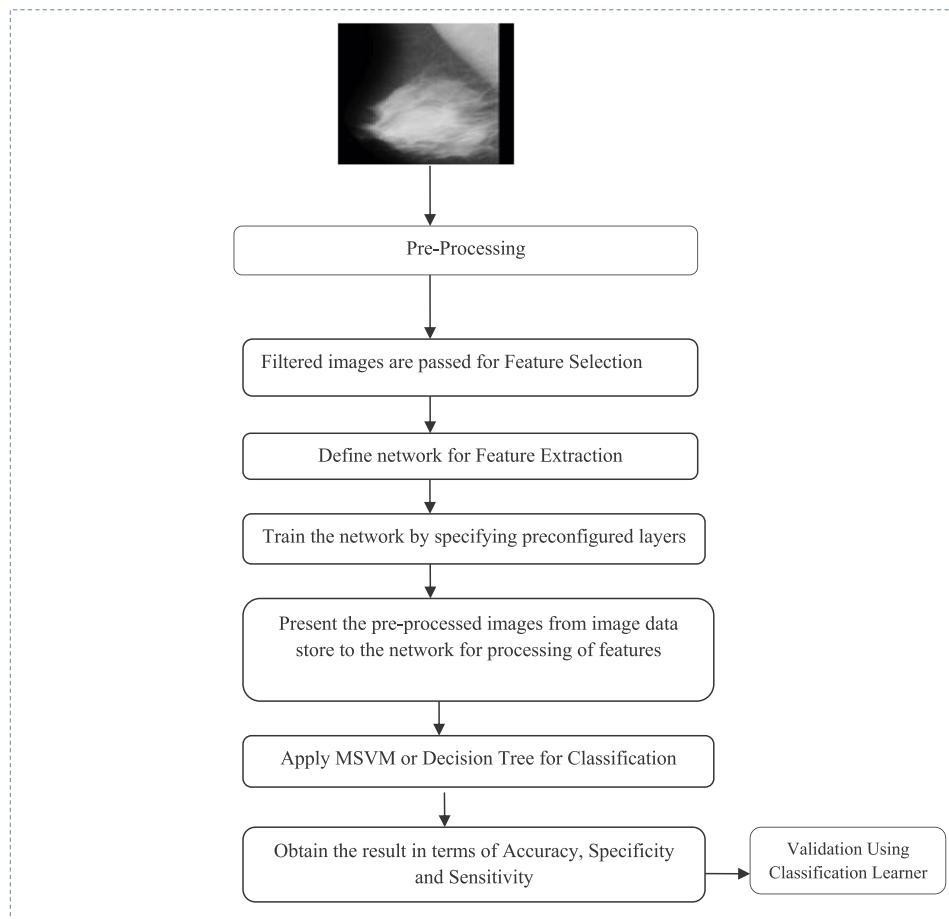


Fig. 3. Proposed methodology of breast cancer mammogram images using deep learning.

utilized to prepare the DCNN through the process of multitask transfer learning, which was then evaluated with SFMs. In the result for the autonomous test set, the multitask exchange learning DCNN was found to have extensively higher execution ($p = 0.007$) when contrasted with

the single-task transfer learning DCNN. Using DCNN for image analysis [15], a straightforward and successful strategy is proposed for the arrangement of re-coloured histological breast disease pictures with little preparation of information (using N of several hundred).

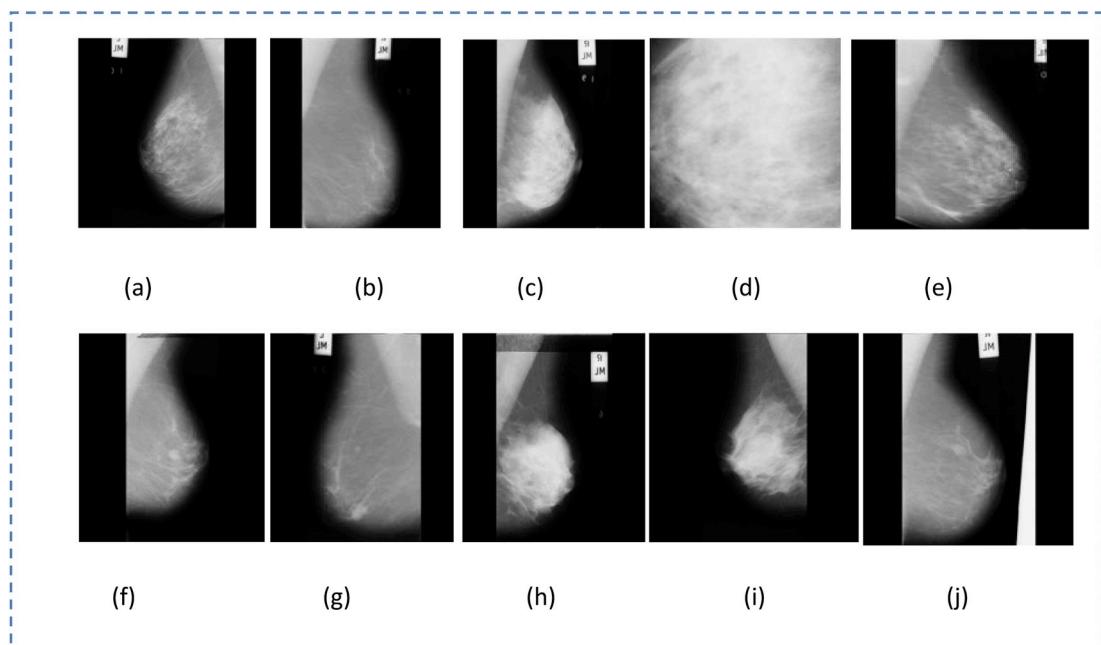


Fig. 4. Normal (a–e), benign (f–g) and malignant breast cancer (h–j) mammogram images.

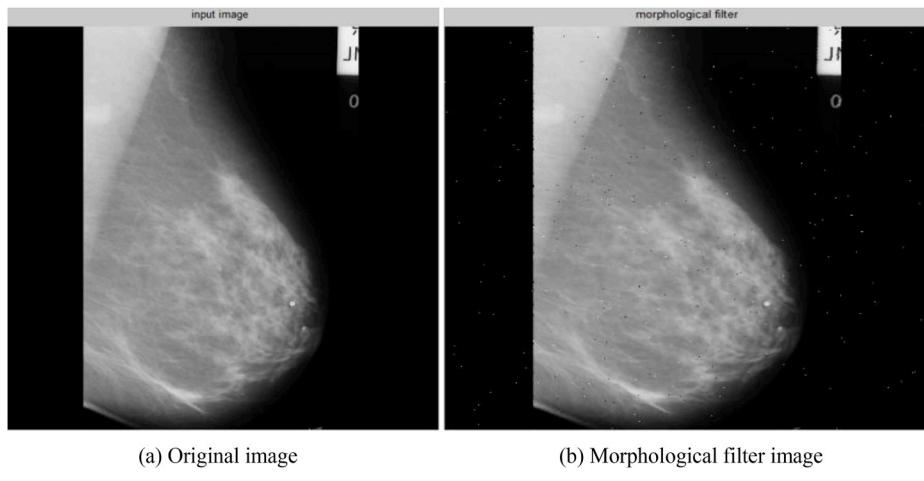


Fig. 5. Morphological operation.

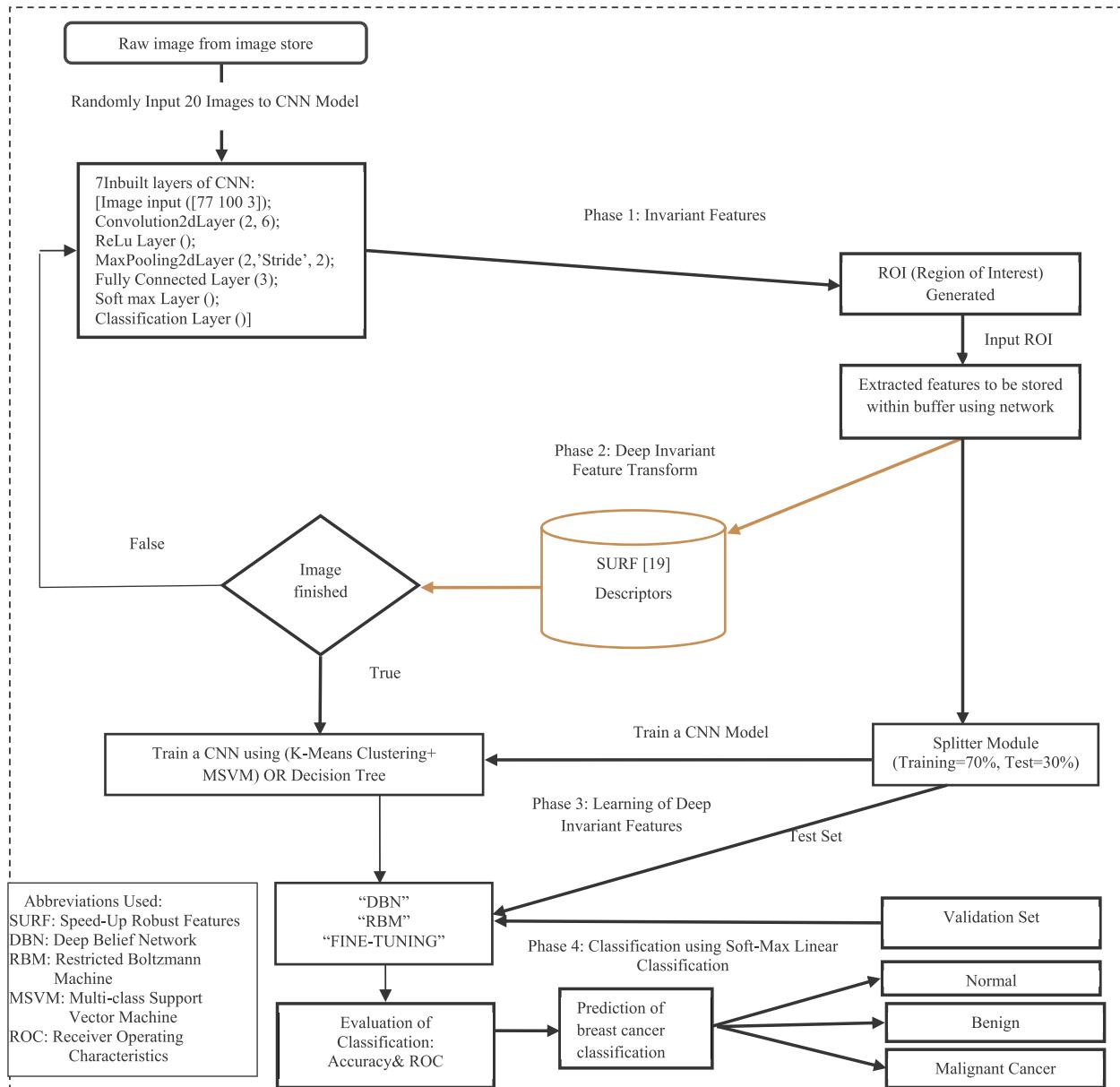


Fig. 6. The systematic internal layering process of the proposed deep CAD system for classification of normal, benign and malignant classes of mammographic masses.

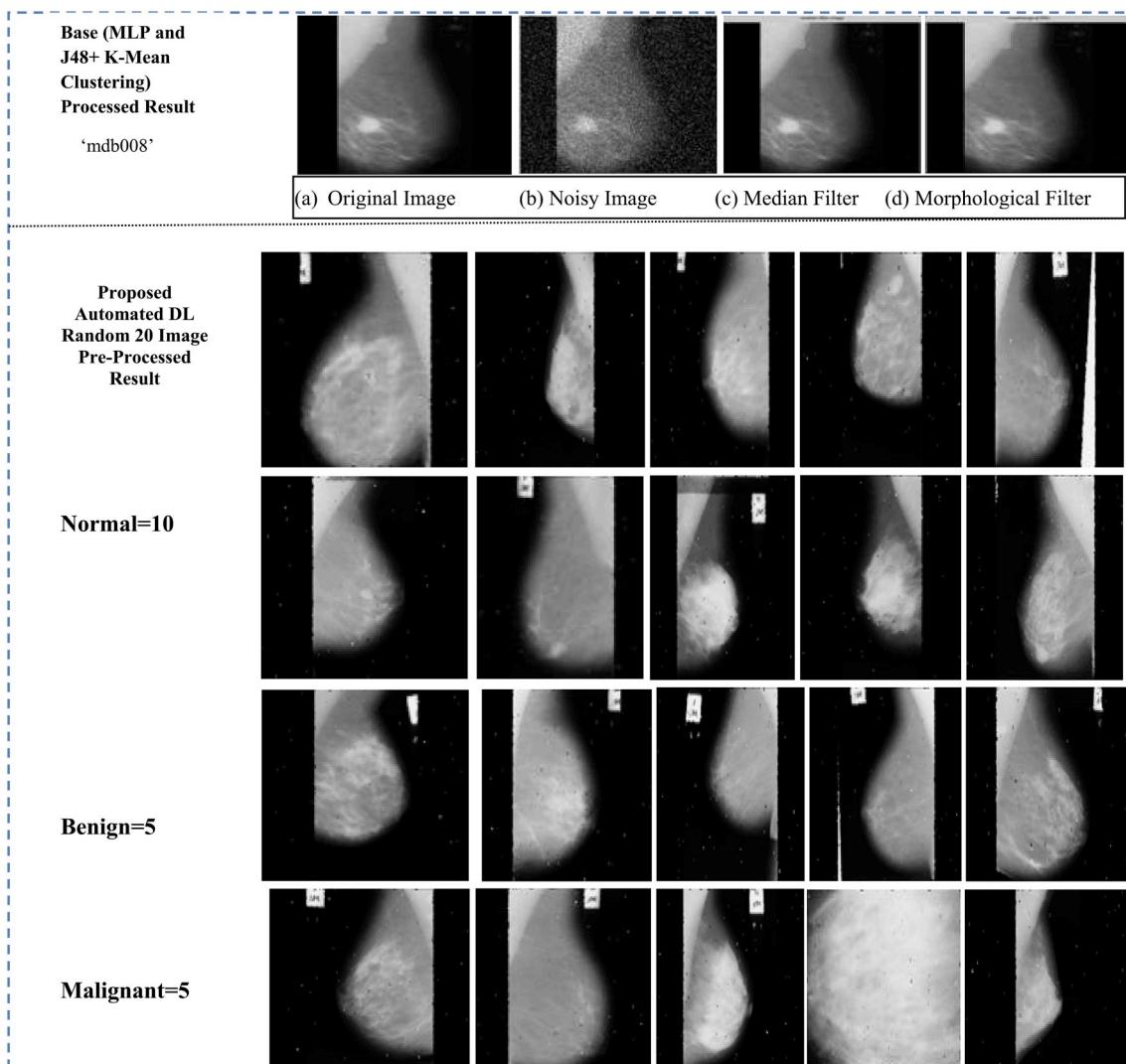


Fig. 7. Pre-processed MLP, J-48+K-mean and proposed DL method.

Table 2
Training on a single CPU

Epoch	Iteration	Time Elapsed (seconds)	Mini-batch Loss	Mini-batch Accuracy	Base Learning Rate
1	1	2.33	0.8266	54.69%	0.0005
40	40	24.36	0.5360	70.31%	0.0005

Table 3
Quantitative result of proposed DL with base

Measures	Manual/Automated		
	MLP (%)	J48+K-mean Clustering (%)	Proposed DL Technique (%)
Image Set			
NORMAL = 'mdb003. jpg'			
Accuracy	76	90	92
Specificity	74	88	90
Sensitivity	76	90	93
Precision	78	90	90
F-score	72	89	96
Recall	76	90	93

To expand the heartiness of the classifier, data augmentation was used and deep convolution features were separated at various scales with openly accessible Convolution Neural Networks (CNNs) prepared on Image-Net. At that point, over-fitting inclination and precise boosting calculation is connected. The preparation of neural systems on an immense measure of information to anticipate problematic speculation were maintained using a strategic distance measure [27]. The computer-aided design framework demonstrated the likelihood for enhancing the analytic exactness by utilizing order and DL approaches, as shown in Table 1.

2.4. Challenges & issues involved in classification of breast cancer lesion

Today, the concept of classification algorithms is emerging in the research for breast cancer detection and diagnosis purposes, but still there are gaps and challenges being faced by radiologists. Some of the issues identified after the survey of CAD and covered in the novel approach described herein are:

2.4.1. Missing training data

An automated identification approach should not only be able to match a disease to one of the known problems, but should also be able to 'reject' the ones that belong to a disease that was not the part of the training set.

Table 4

Predicted accuracy corresponding to normal (Class 1) images of Mini -MIAS dataset

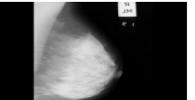
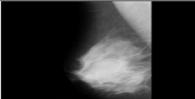
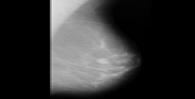
Image Set	Accuracy with Decision Tree classifiers (%)	Accuracy with MSVM (%)
	75	75.17
	78	79
	81	82
	81	82
	67	70

Table 5

Predicted accuracy corresponding to benign (Class 2) images of Mini -MIAS dataset

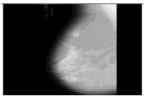
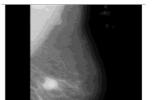
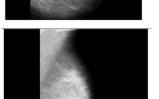
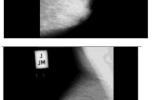
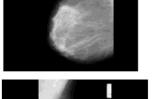
Image Set	Accuracy with Decision Tree classifiers (%)	Accuracy with MSVM (%)
	73	75
	90	91
	80	85
	70	70.63
	78	88

2.4.2. Limited live datasets

Benchmark datasets are available, but a live dataset is very difficult to collect. This is because it takes enormous effort to collect a real life scenario and then image it in a lab or in the field. If available datasets are analysed, it becomes clear that they have been created by only a few people and the specimens are from limited areas. As discussed earlier, the presence of image variation makes these datasets insufficient for wider use. Thus, for realistic classification, there is a need for developing an extensive training data that inculcates all of the characteristics and

Table 6

Predicted accuracy corresponding to malignant (Class 3) images of Mini -MIAS dataset

Image Set	Accuracy with Decision Tree classifiers (%)	Accuracy with MSVM (%)
	80	81
	82	83
	79	82
	85	89
	78	89

variations that can occur in breast cancer [29]. Also, studies should incorporate more realistic images, containing different scenarios, but yet including further examples such as multiple and overlapped. Images should be taken in a natural environment, as well as with complex backgrounds, under different lighting conditions, and during different times in the day, so that all the aspects are covered.

2.4.3. Machine learning developments

Many studies still mostly operate on small benchmark datasets. In recent times, deep learning has been extensively used for automated identification, and DCNN requires extensive training data to properly tune the large set of parameters. A limited number of studies trained CNN classifiers on large image datasets for automated breast cancer identification systems. An important work in this regard is by the research community working on the Image-Net dataset and the related benchmark. Image-Net currently contains more than 14 million images. The average number of training images per category is in the range of hundreds or thousands, and it is currently the most comprehensive image collection. Hence, more of such efforts are needed before automated identification and classification can become a reality. Until then, approaches like data augmentation can be used to address the issue of small datasets. These include simple modifications of images (such as rotation, translation, flipping and scaling). Data augmentation is being used extensively for improving the training process in computer vision.

3. Proposed methodology

This research work includes three types of breast mammography images, i.e., 208 normal images, 62 benign images and 52 malignant cancer images, from the Mini-MIAS dataset [16]. Out of these three stages, malignant cancer is the most widely recognized type of breast cancer, contributing to the death rate in women. The proposed step-by-step methodology is presented in Fig. 3.

3.1. Dataset

The image database used in this study was Mini-MIAS, which consisted of 322 mammogram breast images (available at <https://www.ka>

gle.com/kmader/mias-mammography) and a sample image set, as shown in Fig. 4. A resizing operation is applied upon the automated mechanism to fit the images into the input layer of the network.

The pre-processed images are in Portable Network Graphics (PNG) format. The deep learning approach was applied on 20 images selected randomly from the database, for which a single Centre Processing Unit (CPU) dependent on the configuration of the resources is initialized, and results were compared to MLP and J48+K-Mean Clustering WEKA manual approaches [5].

3.2. Pre-processing

Image pre-processing can be used to reduce image redundancy. For image processing, analysis of noise and the filtering process are important to achieving enhanced image quality. Filtering techniques, such as the median and mean filters, are commonly used to restore a noisy image to the original image. As mammograms require time and effort to interpret, pre-processing is helpful to increase image quality, which makes the feature extraction and segmentation stages easier and more reliable. The noise reduction in the pre-processing stage is normally used to improve images so that salient features can be detected in an easy manner. In this study, an efficient median filter and morphological operation were applied to enhance image quality.

3.2.1. Noise

A speckle is a small spot, similar to dabs of shading on the skin. In this way, 'speckle noise' alludes to the arbitrary age of various small spots in the picture. Such kinds of noise are produced in Synthetic Aperture Radar (SAR) and ultrasound pictures. For better execution and legitimate examination of the picture, a noise-free unique quality picture is required, which leads to the idea of image denoising, a fundamental advancement in picture preparation. It re-establishes a unique picture by enhancing the visual characteristics [17].

3.2.2. Filtering methods

Filtering is a technique that is used to remove undesirable information by perception, and make an image more appropriate for the next stage of image processing. Various types of filtration systems are utilized to removal speckle components of images. A mean and median filter were used in this study for de-speckling of the image [17].

3.2.3. Morphological operations

Morphological operations are related to the shape or morphology of features in an image. These operations rely only on the relative ordering of pixel values; hence, they are suitable for binary image processing [18], but can also be applied to grey scale images, as shown in Fig. 5. Consider a binary image 'B(u, v)' that can be computed using dilation, followed by subtraction:

Load a mammographic mass database

Select the initial partition of the dataset taken into K clusters such that { $c_1 \dots c_k$ }

Evaluate cluster centroids $\bar{w} = \sum_{j=1}^{K_i} w_{ij}$ where

$i = 1 \dots K$

For each w_i in dataset instances do

Re-allocate w_i to the cluster close to it, $w_i \in C_s$

And C_s moved to C_t , if

$$\|w_i - \bar{w}_t\| < \|w_i - \bar{w}_j\|, j = 1 \dots K \text{ and } j \neq t \quad \text{Equation (3)}$$

Re-evaluate centroids

Stop once cluster membership is stabilized

$$I = I \Theta H \quad (1)$$

where 'I' is the image and 'H' is the structuring element.

$$B(u, v) = \text{XOR}(I(u, v), I(u, v)) \quad (2)$$

where $B(u, v)$ = binary image, u and v are coordinates, and I = gray-scale image.

3.3. Deep learning training

The deep learning training operation begins by receiving an image set from the pre-processing phase. To create a new network, the proposed mechanism uses inbuilt layers with an input layer that accepts images of 77×100 through three channels. Training parameters are defined with a training option command using the Deep Learning Toolbox. The training network function is used to finally train the network. Fig. 6 outlines the training process, which fetches the raw image from the database, then applies seven layers of CNN. Extracted features are stored in a buffer for the next phase, where K-mean clustering + MSVM/decision tree is applied to classify normal, benign and malignant classes of mammographic masses [19].

3.4. Classification

3.4.1. K-mean clustering + MSVM

Clustering, or unsupervised classification, is useful in decision-making, machine-learning situations and pattern classification [18, 21], such as classifying patterns into groups (data items and feature vectors), to gain an insight into the data, in order to detect abnormalities and extract features. Cluster analysis is the study of methods and algorithms to group various objects, and allocate corresponding clusters to the given objects on the basis of measured characteristics or similarities [20]. Data clustering is comprised of three main purposes, including the following:

- Underlying structure to detect anomalies, locate features and gather insight into the data.
- For the classification and identification of similar features to make predictions about the identified features.
- To use as a compression technique to organize and summarize data.

K-means is the most generally used clustering algorithm for automatic classification of vector-based data [22–24]. The K-means algorithm is an iterative technique that splits an image into K clusters. This method of cluster analysis, which can divide 'n' observations into 'K' clusters, is explained as follows:

3.4.1.1. Pseudo code.

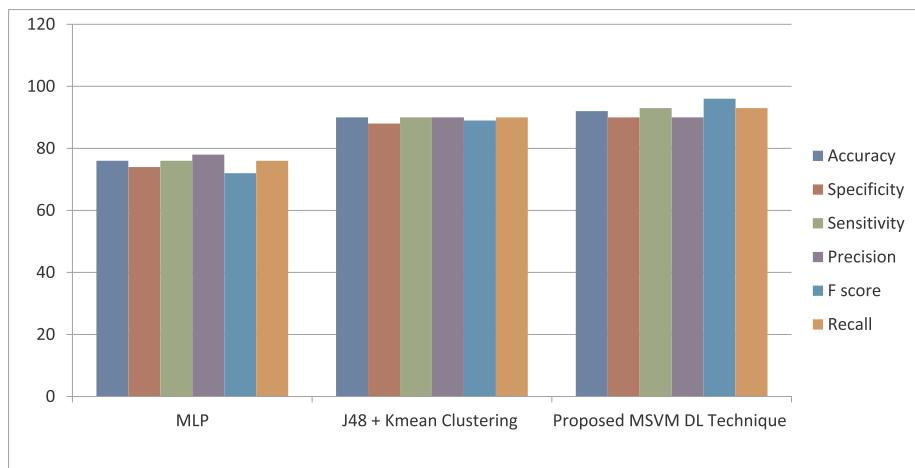


Fig. 8. Statistical analytical accuracy result of base and proposed DL approaches.

To perform the classification step, an MSVM or decision tree is used in this phase. The support vector machine uses a rules based environment to correctly reach the solution of the given problem, including outlier detection or unnecessary regions. A system was introduced to resolve the unclassified region [25,26]. In the proposed work, MSVM was utilized to realize the classification results. Optimal hyperplanes were defined to determine whether the obtained values of the membership functions satisfy the hyperplane ($D(x)$) or not.

$$\text{Satisfaction Criteria } D(X) > 1 \quad (4)$$

A one dimensional membership function $m_{ij}(x,y)$ was defined for determining optimal separating hyperplanes $D_j(x) = 0$ as follows.

1. If values of diagonal are equal ($i = j$)

$$m_{ij}(x) = \begin{cases} 1 & \text{for } D_i(X) > 1 \\ -D_i(x) & \text{for } D_i(X) < 1 \end{cases} \quad (5)$$

Rules determining correct class 1, 2 or 3

2. If values of diagonal are not equal ($i \neq j$)

$$m_{ij}(x) = \begin{cases} 1 & \text{for } D_i(X) < 1 \\ -D_i(x) & \text{for } D_i(X) > 1 \end{cases} \quad (6)$$

Rules determining correct class 1 or 2 or 3.

The procedure for classification is as follows:

3. If the pixel value x is such that $D_i(x) > 0$ and is satisfied only for that class, then it is fed into that class.
4. If $D_i(x) > 0$ and x lies between various classes, then classify the data into the class with maximum $D_i(x)$
5. If $D_i(x) \leq 0$ and x lies between various classes, then classify the data into the class with minimum $D_i(x)$

3.4.2. Decision tree

A decision tree algorithm describes the way attribute vectors behave as an instance. The algorithm maps the overall observations about an item to generate the decision making rules for the prediction of data. A pruning technique is used as a tool. In alternative algorithms, the particular category is carried out recursively until each and every leaf is definitely absolute, which is the category of data and the best possible among all selected. The aim is usually the steady generalization of a decision tree until it increases the equilibrium connected with versatility, in addition to accuracy. The steps involved in supervised classification of mammogram images by decision tree algorithm are given as:

3.4.2.1. Pseudo code.

```

Load a mammographic mass database
Apply decision tree ( $C_t$ , A, T)
If  $C_t$  includes the training data evaluation for the same class  $s_j \in S$  instances, then make decision tree T with the same
class ' $s_j$ '
Elseif  $A = \emptyset$ , then
    Make  $t_d$  into a leaf node  $s_j$ , the most frequent class in  $C_t$ 
    Else
         $\rho_0 = \text{Impurity evaluation} - 1(C_t)$ 
        For all attributes  $a_k \in a_1, a_2, \dots, a_K$  do
             $\rho_1 = \text{Impurity evaluation} - 2(a_k, C_t)$ 
        End
        Select  $a_g \in a_1, a_2, \dots, a_K$  providing biggest impurity reduction through  $\rho_0 - \rho_1$ 
        If  $\rho_0 - \rho_1 < t_{\text{threshold}}$ 
            Make T leaf node label  $s_j$ 
        Else
            Make T a decision node on  $a_g$ 
             $V_1 V_2 \dots V_n$  be the possible values of  $a_g$ 
            Divide  $C_t$  into n subsets.
            Divide to subset  $C_{t_1} C_{t_2} \dots C_{t_n}$  based on the possible values of  $a_g$ 
            For each  $C_{t_j}$  in  $C_{t_1} C_{t_2} \dots C_{t_n}$  do
                If  $C_{t_j} \neq \emptyset$ , then
                    Create edge node  $t_j$  for  $V_j$  as a child node of  $t_d$ 
                    Decision tree ( $C_{t_j}, A - \{a_g\}, t_j$ ) // here  $a_g$  removed
                End
            End
        End
    End
End

```

Table 7

Accuracy vs. classification algorithm parametric measures

Accuracy vs. Classification Algorithm	SVM (Model1)	LDA (Model 2)	Fine KNN (Model 3)	Decision Tree(Model 4)
Base Method (MLP & J48+K-Mean Clustering)	94%	94%	88.0%	90%
Proposed DL Method	96.9%	92.8%	94.8%	89.7%

4. Results & analysis

This section explains the methodology that has been followed to evaluate the proposed framework. An overview of the enhanced pre-processed result of the dataset using manual state-of-the-art techniques (MLP [5] & J48+K-mean clustering) and automated DL approach is presented in Section 4.1 [31,32]. Details of the quantitative evaluation method are explained in Section 4.2.

4.1. Manual and automated pre-processed dataset

The Mini-MIAS mammographic enhanced 20 pre-processed images that were loaded randomly and selected from three classes, i.e., normal, benign and malignant cancer [16]. The qualitative result of Proposed DL and compared result of MLP, J48+ K-Mean Clustering are presented in Fig. 7.

4.2. Quantitative analysis

To evaluate the proposed DL system, a 10-fold cross validation experiment was performed on ROI. The dataset was divided into 70% training and 30% testing using 322 ROI images. ROI is of 77×100 with three channels, and seven layers of CNN were used for feature extraction. Then, the extracted features were trained using a single CPU. The parametric result is described in Table 2.

As the number of iterations increased, the loss decreased and the accuracy rate increased by approximately 15% for Mini-Batch. The area under receiver operating curve (AUC) was calculated by evaluation criteria which are sensitivity, specificity, accuracy, precision and F-measure using the following statistical formulas:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \quad (7)$$

$$\text{Specificity} = 1 - \left(\frac{\text{FP}}{\text{FP} + \text{TN}} \right) \times 100 \quad (8)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100 \quad (9)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100 \quad (10)$$

$$\text{Fmeasure} = \frac{2 * \text{Recall} * \text{precision}}{\text{Recall} + \text{precision}} \times 100 \quad (11)$$

'TP' stands for the true positive cases in detection results and 'TN'

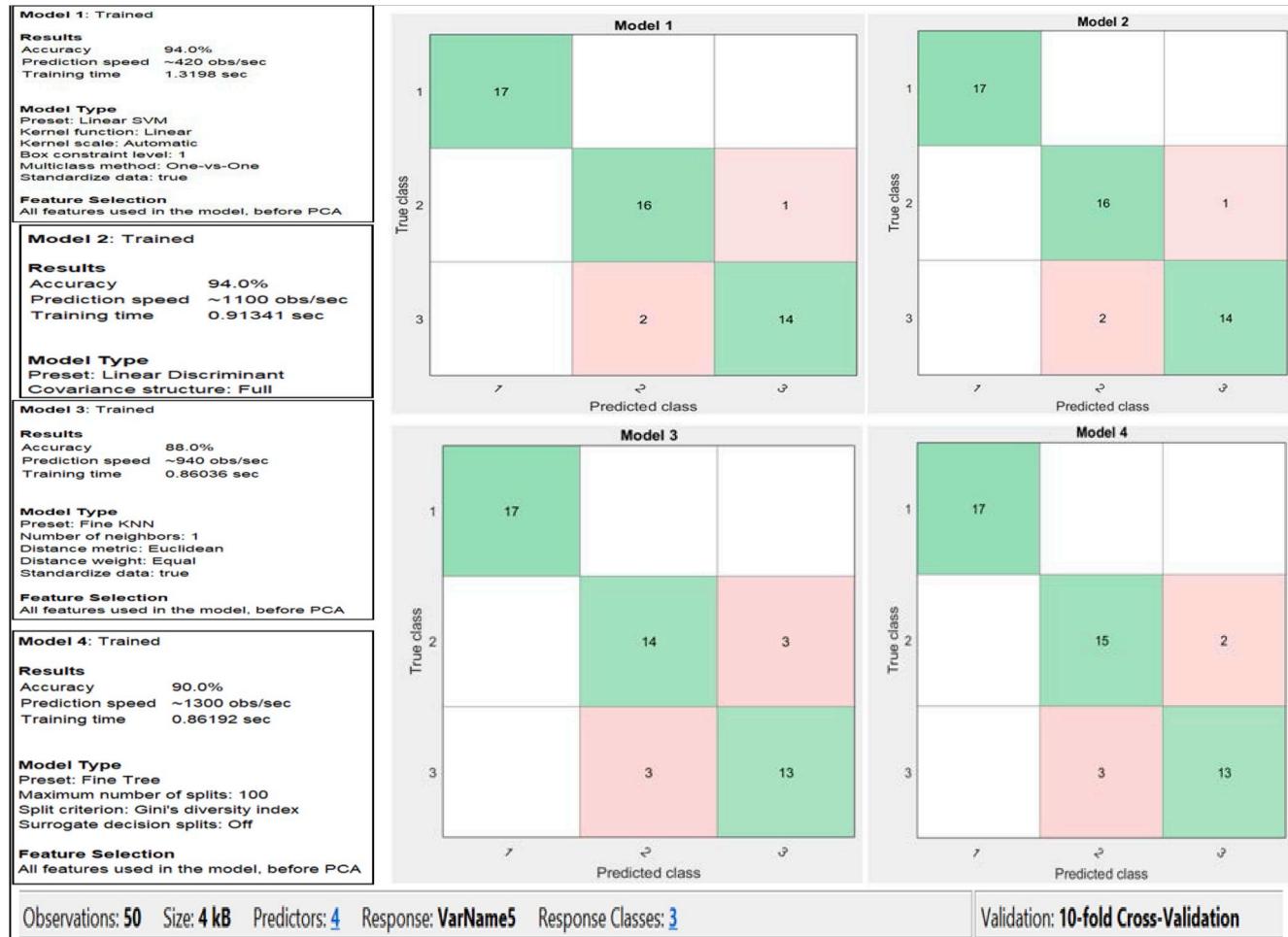


Fig. 9. Confusion matrix (number of observations) of base method.

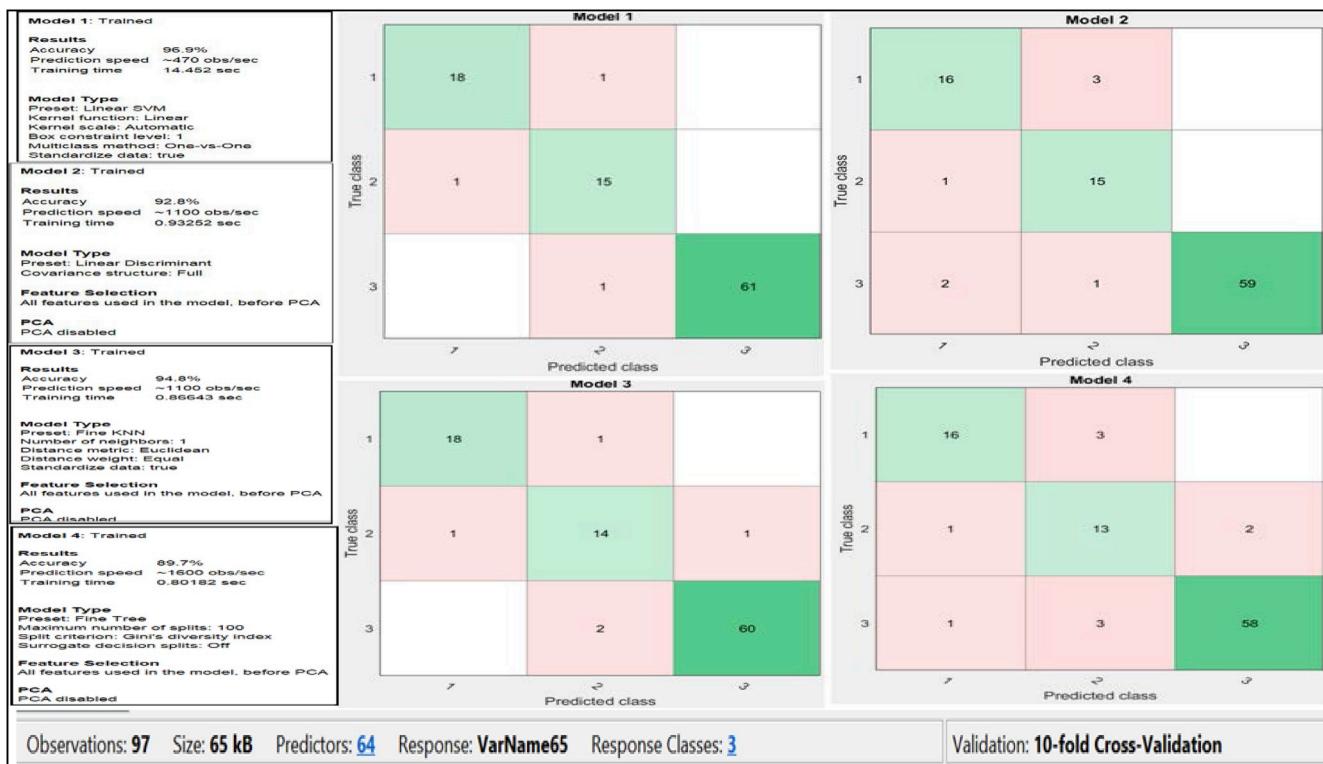


Fig. 10. Confusion matrix (number of observations) of proposed DL method.

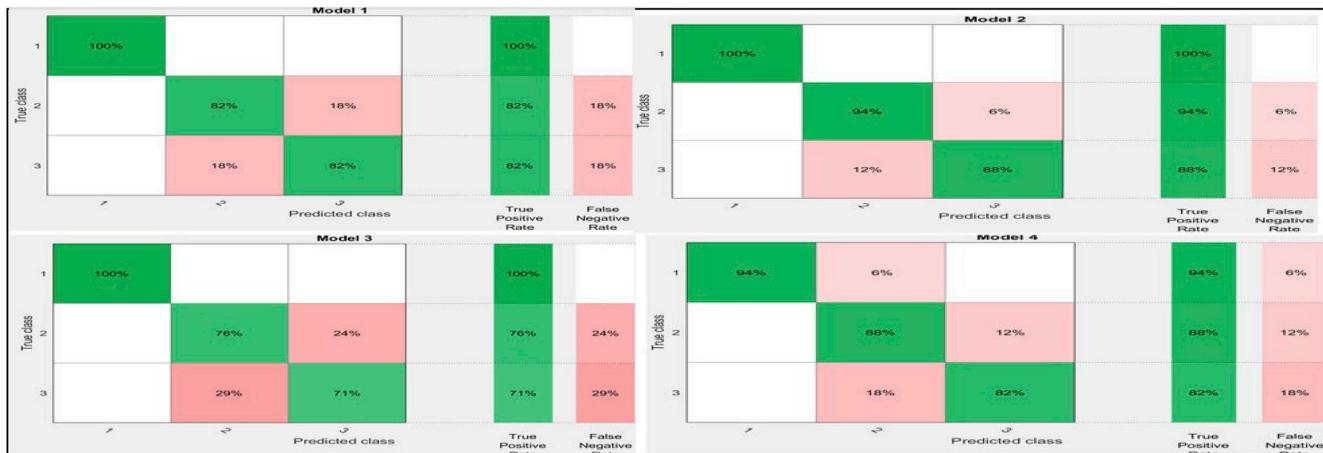


Fig. 11. Accuracy measurements (true/false positive rates) of base method.

denotes the true negative cases. In addition, 'FP' stands for the false positive cases and 'FN' equals the false negative cases, where $\text{recall} = \text{sensitivity}$. The statistical results are compared between MLP & J48+K-mean clustering using WEKA with the proposed deep learning technique (Table 3).

The interpretation of the statistical results of accuracy parameters shows that the average accuracy rate of the proposed method is increased by 2% as compared to the state-of-the-art techniques. The evaluation parameters i.e. specificity and sensitivity increased by 2% and 3% respectively, whereas precision remains constant.

4.3. Classification results

K-mean clustering was used for feature extraction. The number of features extracted was 19,968 for $K=500$. The number of features

selected equalled 97×65 . Then, the classification step included two types of algorithms, i.e., the decision tree and MSVM. The quantitative comparison result is shown in the form of accuracy. Accuracy addresses the effectiveness of the classifier that is most commonly used as an indicator that reflects precision of the results, using Equation (9). The qualitative and quantitative results applied on random images of three classes, including normal (Class 1), benign (Class 2) and malignant (Class 3) masses, are represented in Table 4, Table 5, and Table 6, respectively.

The classification algorithm analysis result proves that MSVM is better than the decision tree in all three classes of the Mini-MIAS mammographic dataset. The graphical representation of the above evaluation process was also depicted by bar chart in Fig. 8. The bar results interpreted that the statistical measures of the proposed DL method were improved as compared to state-of-the-art techniques.



Fig. 12. Accuracy measurements (true/false positive rates) of proposed DL method.

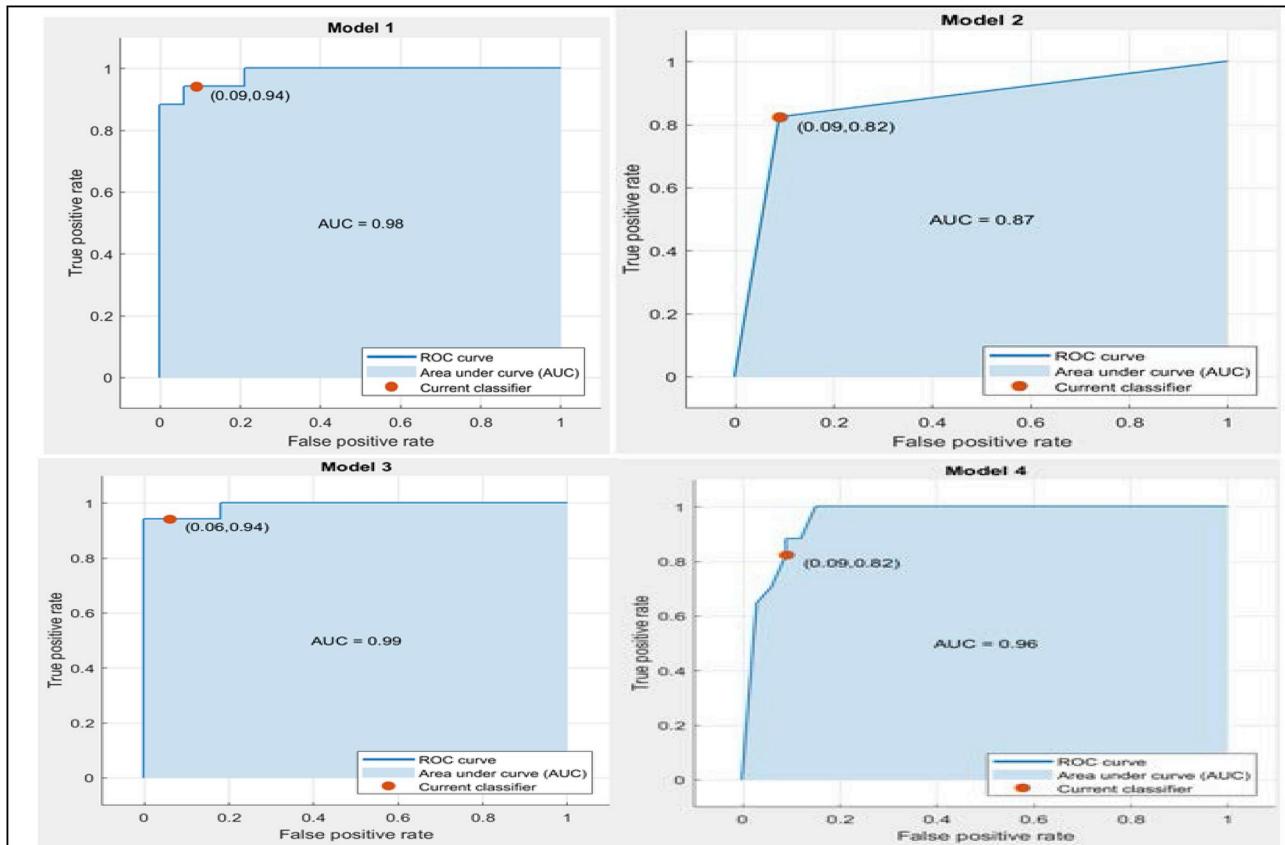


Fig. 13. ROC curve of base method.

4.4. Validation results

In addition to the classification step, the accuracy statistical measure is derived using the Classification Learner Toolbox to show the performance of the proposed Deep Learning Technique. The area used under the receiver operating characteristic (ROC) analysis was applied on the test data set. In order to investigate the confusion matrix (CM), validation was tested using inaccuracy (ACC) measurements, ROC and a

scatterplot, with the classification algorithms SVM, KNN, LDA and Decision Tree. The area under the curve (AUC) is a commonly used index to assess overall discrimination. The range of AUC lies between 0.5 and 1.0. The greater the AUC values, the higher the classification accuracy. The significance of the proposed deep learning system was compared with state-of-the-art classification algorithms, using 10-fold cross validation test on this group.

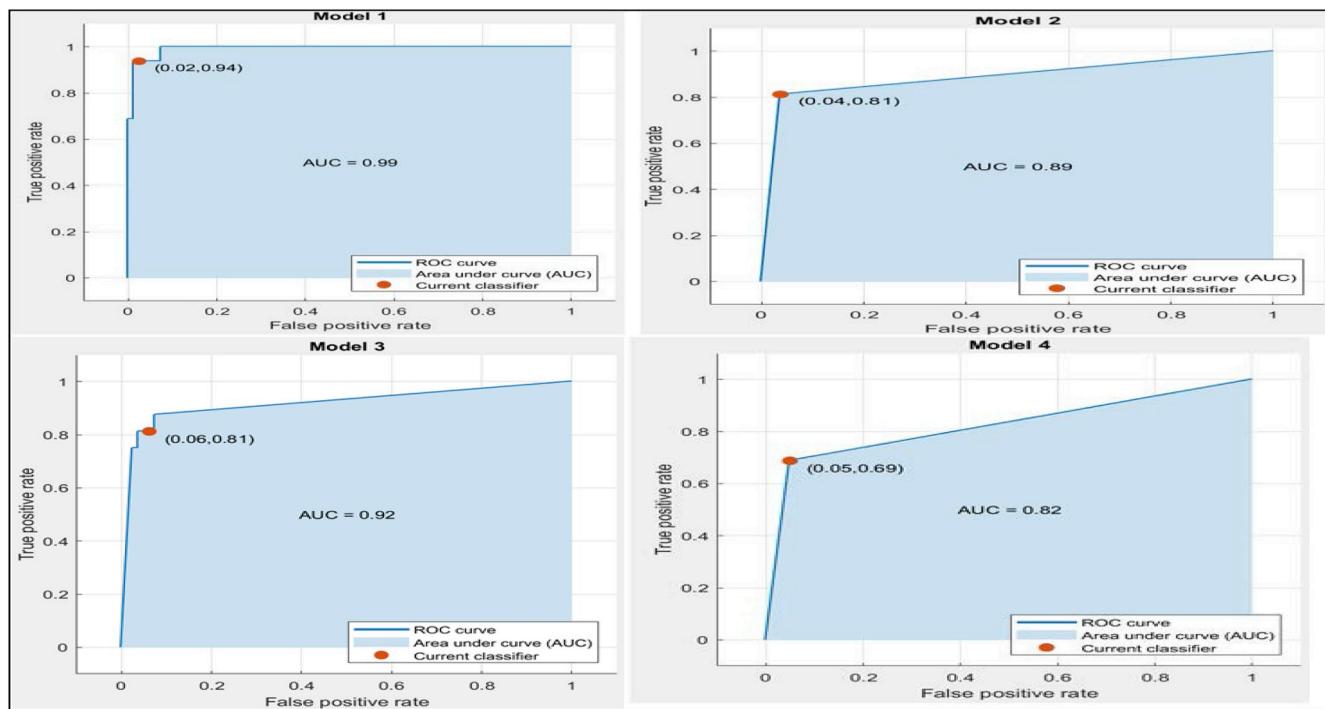


Fig. 14. ROC curve of proposed DL method.

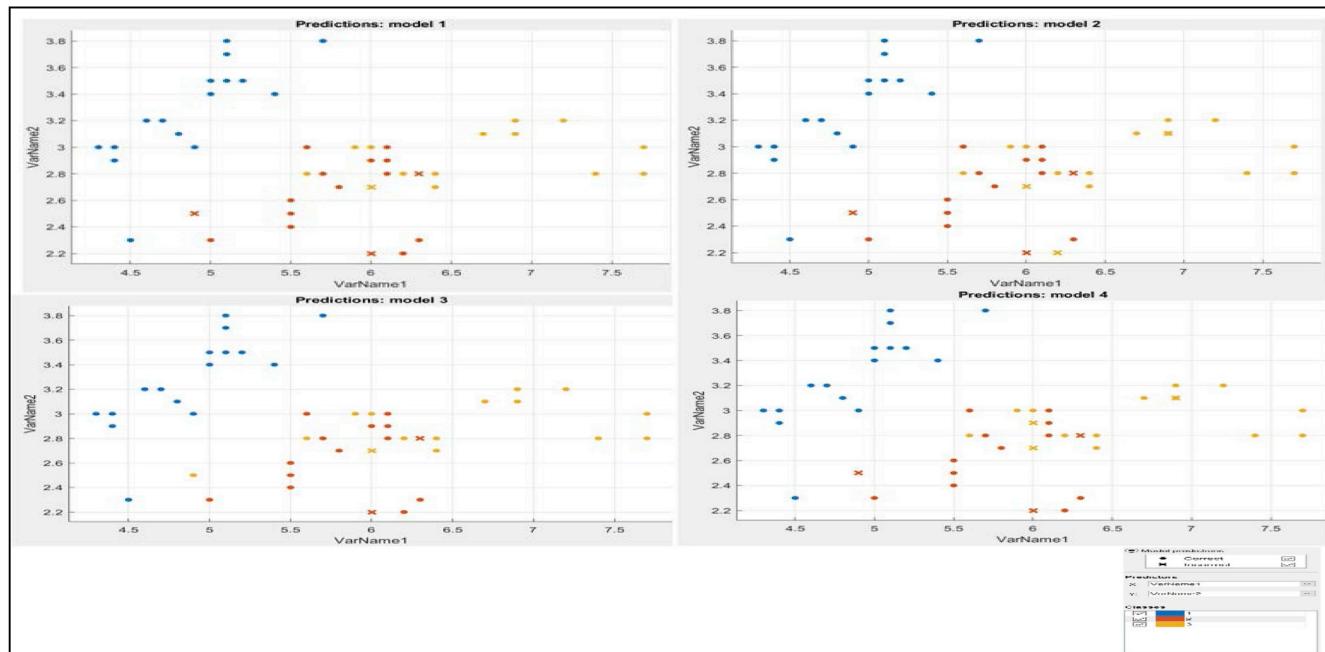


Fig. 15. Scatter plot of base method.

4.4.1. Confusion matrix (CM)

The confusion matrix was validated using different classification algorithms (Table 7) for state-of-the-art techniques (base method) and for the proposed DL method, shown in Figs. 8 and 10, respectively.

The average value of the accuracies of the base method using classification algorithms, i.e., SVM, LDA, Fine KNN and decision tree, were 92%, 94%, 88% and 88%, respectively, whereas in the proposed DL method the value of average accuracy using SVM is 96.9%, LDA is 93.8%, Fine KNN is 94.8% and decision tree is 89.7%. Through the validation process, the proposed DL method was evaluated versus the

base method. Since the ACC value was highest in SVM, it suggests that the proposed classification AUC is improved as compared to other algorithms.

4.4.2. Accuracy measurements(ACC)

Using 10-fold cross validation, with respect to the 3 response classes, ACC measurements were evaluated on the basis of the true and false positive rates for the base method and the proposed method, as shown in Figs. 11 and 12.

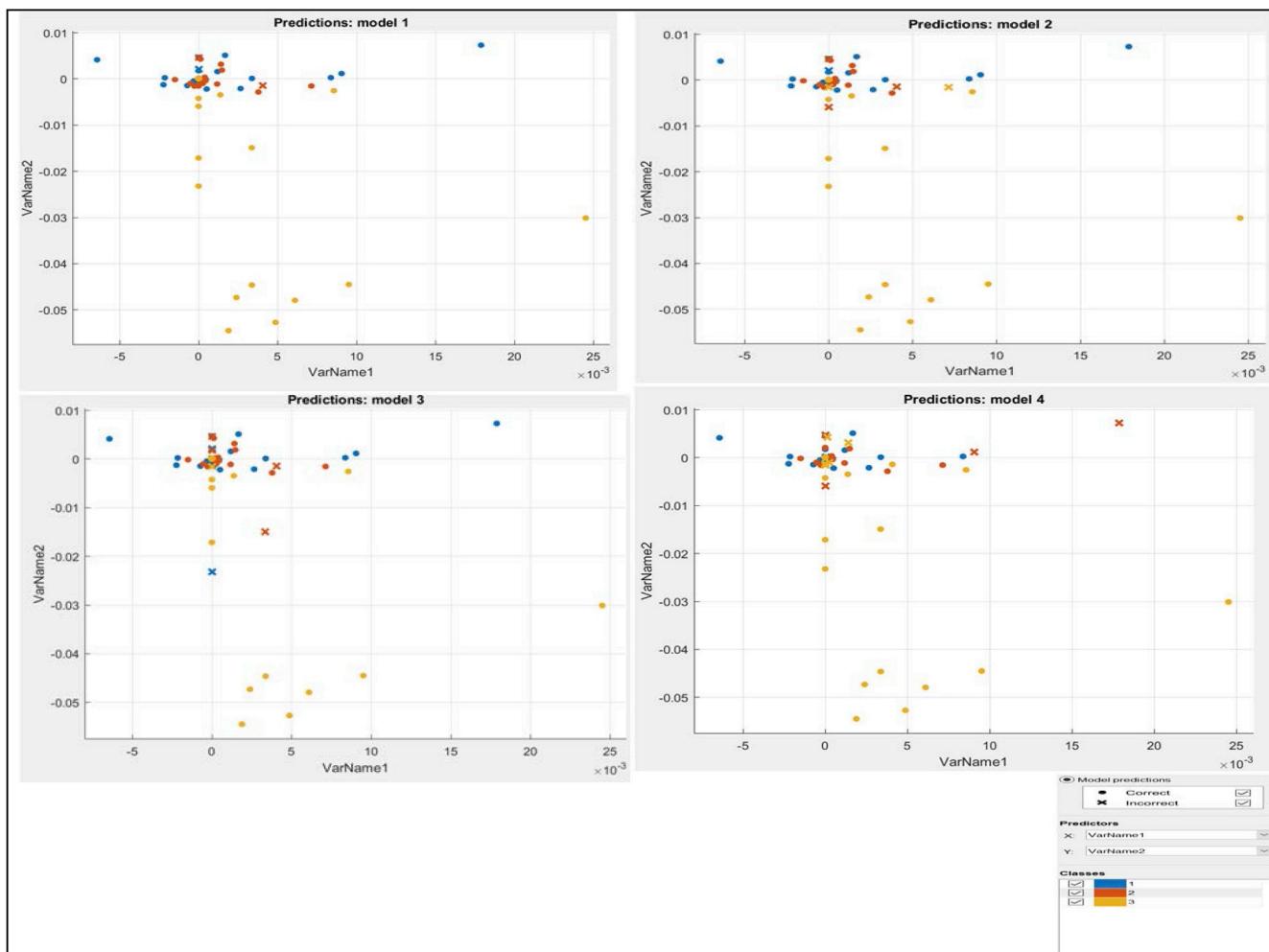


Fig. 16. Scatter plot of proposed DL method.

4.4.3. Receiver operating characteristics(ROC)/area under the curve(AUC)

The ROC curve area of the base method and the proposed DL method were validated, as shown in Figs. 13 and 14, respectively.

The ROC value was near to 1 for both the techniques. The proposed DL method was validated using SVM, whereas the base method was validated using LDA for desired accuracy.

4.4.4. Scatter plot

A scatterplot describes the correct and incorrect instances being tested by different techniques. The instances of base method and proposed DL method are shown in Figs. 15 and 16.

5. Conclusion and future scope

An automated system that utilizes a Multi-Support Vector Machine and deep learning mechanism for breast cancer mammogram images was initially proposed. The pre-processing phase is important, and is well-defined using noise handling and resizing operations. Obtained images are fed into the trained network for feature extraction, and classification is performed using MSVM. A hybrid approach of K-mean clustering and MSVM gives better results as compared to a decision tree model. The quantitative analysis and validation confirms that the proposed DL technique was better than state-of-the-art techniques, i.e., the MLP and J48+K-mean clustering WEKA approach. Overall enhancement in terms of accuracy was 2%. The main objective of this study was to determine the consistency of classification accuracy when presented with larger datasets, which were subsequently increased. The future

perspectives are to work on a large-scale network of deep learning internal layers and to help radiologists to accurately validate large datasets in less time.

References

- [1] Liu B, Cheng HD, Huang J, Tian J, Tang X, Liu J. Fully automatic and segmentation-robust classification of breast tumors based on local texture analysis of ultrasound images. *Pattern Recognit* 2010;43(1):280–98. <https://doi.org/10.1016/j.patcog.2009.06.002>.
- [2] Breast cancer facts & Figures. Atlanta: American Cancer Society; 2018.
- [3] Bellotti R. A completely automated CAD system for mass detection in a large mammographic database. *Med Phys* 2006;33(8):3066–75. <https://doi.org/10.1118/1.2214177>.
- [4] Noura M, Abdelwahed A. Computer aided system for breast cancer diagnosis in ultrasound images. *Int J Environ Health Eng* 2015;3(3):71–6. <https://doi.org/10.12785/jehc.030303>.
- [5] Shan J. Computer-aided diagnosis for breast ultrasound using computerized BI-RADS features and machine learning methods. *Ultrasound Med Biol* 2015;42(4):980–8. <https://doi.org/10.1016/j.ultrasmedbio.2015.11.016>.
- [6] Chan HP. Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network. *Physics Medical Biology* 1996;42(3):549–67. <https://doi.org/10.1088/0031-9155/42/3/008>.
- [7] Varela C, Tahoces PG, Mendez AJ, Souto M, Vidala JJ. Computerized detection of breast masses in digitized mammograms. *Comput Biol Med* 2007;37(2):214–26. <https://doi.org/10.1016/j.combiomed.2005.12.006>.
- [8] Ciechlewski M. Malignant and benign mass segmentation in mammograms using active contour methods. *Symmetry* 2017;9(277):1–22. <https://doi.org/10.3390/sym9110277>.
- [9] Posso M. Cost-effectiveness of double reading versus single reading of mammograms in a breast cancer screening programme. *PLoS One* 2016;11(7):1–13. <https://doi.org/10.1371/journal.pone.0159806>.

- [10] Kalager M. Effect of screening mammography on breast cancer mortality in Norway. *N Engl J Med* 2010;363(13):1203–10. <https://doi.org/10.1056/NEJMoa1000727>.
- [11] Yuchen Q, Wang Y, Yan S, Tan M, Cheng S, Liu H, Zheng B. An initial investigation on developing a new method to predict short term breast cancer risk based on deep learning technology. In: Proc. SPIE 9785, medical imaging: computer-aided diagnosis, vol. 978521; 2016. p. 1–6. <https://doi.org/10.1117/12.2216275>.
- [12] Cheng JZ. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. *Sci Rep* 2016;6 (24454):1–13. <https://doi.org/10.1038/srep24454>.
- [13] Duraisamy S, Emperumal S. Computer-aided mammogram diagnosis system using deep learning convolutional fully complex-valued relaxation neural network classifier. *Deep Learning in Computer Vision* 2017;11(8):656–62. <https://doi.org/10.1049/iet-cvi.2016.0425>.
- [14] Samala RK. Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms". *Physics Medical Biology* 2017;62(23):8894–908. <https://doi.org/10.1088/1361-6560/aa93d4>.
- [15] Rakhlina A, Shvets A, Iglovikov V, Kalinin AA. Deep convolutional neural networks for breast cancer histology image analysis. In: Campilho A, Karray F, ter Haar Romeny B, editors. *Image analysis and recognition. ICIAR 2018. Lecture notes in computer science*, vol. 10882. Chamdo: Springer; 2018. https://doi.org/10.1007/978-3-319-93000-8_83.
- [16] Suckling J, Parker J, Dance D. The mammographic image analysis society digital mammogram database. In: *Excerpta medica international congress series*, vol. 1069; 1994. p. 375–8. Available at: <http://peipa.essex.ac.uk/info/mias.html>.
- [17] Sundaram KM, Sasikala D, Rani PA. A study on pre-processing a mammogram image using adaptive median filter. *International Journal of Innovative Research in Science, Engineering and Technology* 2014;3(3):10333–7.
- [18] Chen CM, Chou YH, Han KC, Hung GS, Tiu CM, Chiou HJ, Chiou SY. Breast lesions on sonograms: computer-aided diagnosis with nearly setting-independent features and artificial neural networks. *Radiology* 2003;226(2):504–14. <https://doi.org/10.1148/radiol.2262011843>.
- [19] Bay H, Tuytelaars T, Gool L. SURF: speeded up robust features. *Computer Vision, Lecture Notes in Computer Science* 2006;3951:404–17.
- [20] Jain AK, Murty MN, Flynn PJ. Data clustering: a review. *ACM Comput Surv* 1999; 31(3):264–323. <https://doi.org/10.1145/331499.331504>.
- [21] Jiao Z, Gao X, Wang Y, Li J. A deep feature based framework for breast masses classification. *Neurocomputing* 2016;197:221–31. <https://doi.org/10.1016/j.neucom.2016.02.060>.
- [22] Nazeer KAA, Sebastian MP. Improving the accuracy and efficiency of the K-means clustering algorithm. In: *Proceedings of the world congress on engineering, proceedings of the world congress on engineering*, vol. 1; 2009. Available at: http://www.iaeng.org/publication/WCE2009/WCE2009_pp308-312.pdf.
- [23] Jain AK. Data clustering: 50 years beyond K-means. *Pattern Recognit Lett* 2010;31 (8):651–66. <https://doi.org/10.1016/j.patrec.2009.09.011>.
- [24] Celebi ME, Kingravi HA, Vela PA. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Syst Appl* 2013;40(1): 200–10. <https://doi.org/10.1016/j.eswa.2012.07.021>.
- [25] Wahdan P, Saad A, Shoukry A. Automated breast tumor detection in ultrasound images using support vector machine and ensemble classification. *Journal of Biomedical Engineering and Biosciences* 2016;3:4–11.
- [26] Inoue T, Abe S. Fuzzy support vector machines for pattern classification. *IEEE International Conference on Neural Network* 2001. <https://doi.org/10.1109/IJCNN.2001.939575>.
- [27] Jaffar MA. Deep learning based computer aided diagnosis system for breast mammograms. *Int J Adv Comput Sci Appl* 2017;8(7):286–90. <https://doi.org/10.14569/IJACSA.2017.080738>.
- [28] Cheng HD, Shan J, Ju W, Guo Y, Zhang L. Automated breast cancer detection and classification using ultrasound images: a survey. *Pattern Recognit* 2010;43: 299–317. <https://doi.org/10.1016/j.patcog.2009.05.012>.
- [29] Muthuselvan S, Sundaram KS, Prabashella. Prediction of breast cancer using classification rule mining techniques in blood test datasets. *International Conference on Information Communication and Embedded Systems (ICICES) 2016*: 1–5. <https://doi.org/10.1109/ICICES.2016.7518932>.
- [30] Heidari E, Sobati MA, Movahedirad S. Accurate prediction of nanofluid viscosity using a multilayer perceptron artificial neural network (MLP-ANN). *Chemometr Intell Lab Syst* 2016;155:73–85. <https://doi.org/10.1016/j.chemolab.2016.03.031>.
- [31] Shirazi F, Rashedi E. Detection of cancer tumors in mammography images using support vector machine and mixed gravitational search algorithm. In: *2016 1st conference on swarm intelligence and evolutionary computation (CSIEC); 2016*. p. 98–101. <https://doi.org/10.1109/CSIEC.2016.7482133>.
- [32] Abbas Q. Deep CAD: a computer-aided diagnosis system for mammographic mass using deep invariant features. *Computers* 2016;5:1–13. <https://doi.org/10.3390/computers5040028>.