

DeepBreastCAD: Deep Learning based Computer Aided Diagnosis System for Breast Mammograms

M. Arfan Jaffar

College of Computer and Information Sciences,

Al Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia

Abstract

Breast cancer is one of the most dangerous, leading and widespread cancers in the world specially in women. For breast analysis, digital mammography is the most suitable tool used to take mammograms for detection of cancer. It has been proved in the literature, that if it can be detected at early and initial stages, then there are many chances to cure timely and efficiently. Therefore, initial screening of mammograms is the most important to detect cancer at initial stages. In this paper, a solution has been proposed by using deep Convolutional Neural Network (CNN) with Support Vector Machine (SVM). Proposed method first perform preprocessing to resize the image so that it can be suitable for CNN and perform enhancement quality of the images can be enhanced. Deep Convolutional Neural Network (CNN) has been used for features extraction and classification with Support Vector Machine (SVM). Standard dataset MIAS and DDMS has been used for validation of the proposed method by generating new images from these datasets by the process of augmentation. Accuracy, Sensitivity, Specificity and area under the receiver operating curve (AUC) has been used as a quantitative measure and compared with state of the art existing methods. Results shows that proposed method has achieved accuracy 93.35% and 93% sensitivity.

Keywords:

segmentation, breast mammograms, classification, texture

1. Introduction

Cancer is the most dangerous and leading cause of death in the whole world wide. There are different types of cancers in the different organs of human. For women, breast is the most important organ. At the baby birth, mother women used his breast to feed her milk to her child. Therefore, breast is the most important organ specially for women. There is a special care required so that it can escape from any type of disease or cancer. Due to milk transfer to child's, there are chances that may be cancer also shifted to child's if it uncured or due to unaware of such type of diseases. Breast cancer is the most common cancer specially in the women. Thus there is special attention required to solve this problem. Breast cancer is one of the conspicuous

causes for deaths of women, and according to a survey and report published in 2016 statistics shows [1], that 61,000 new circumstances of breast cancer are prophesied. It has been well established and well reported in the literature that if breast cancer [2] is perceived at early stages by mammographic screening process then there are many chances to survive for these types of cancerous patients and even survival rate can be increased more than 90%. Therefore, for analysis and identification of breast cancer, the digital screening mammography [3] is broadly employed by radiologists. Mammography is a process that can be used to detect cancer in breast. Radiologists are the most expensive in the whole world wide. It is very difficult for a common person to bear too much expenses. Second this cancer is also diagnosed very carefully. Most of the time, it is recommended to take second opinion from other radiologist. Due to lack of funds or expenses, it is very difficult to take second opinion. Now a day, in this digital world, it is possible to introduce computer based solution to diagnose such type of cancers [4]. In the literature, many different Computer Aided Diagnosis (CAD) systems available to help the radiologist to take second opinion. Most of the existing system has some problems due to imaging poor quality. Some systems did not perform well in the case of noises or due to low radiation may be image has low quality or poor quality due to low contrast. There is CAD system available that guarantees the solution. Still, there is a room to improve the performance of these CAD systems [5,6]. Therefore, I have tried to propose a solution to detect cancer in the breast mammograms. Figure 1 shows some images from DDMS dataset to show benign and malignant samples. Table 1 shows a list of abbreviation used in this paper.

In the literature, it has been shown that in the previous decades, most of the researchers focus on some preprocessing and post-processing operation like segmentation, background removal, pectoral removal and some other methods for CAD. These old techniques are now old and even required more processing time as well as some expert knowledge also required to highlight that which part

is breast part and which part is not. Similarly for post-processing some features extraction methods has been performed like texture features, some geometrical features, statistical features or some other features to classify images. Again there is no rule that which features are most suitable. Classification has been performed by using some old classifiers like artificial neural network, KNN, Bayesian. But most of the times, these features extraction and classifiers are not suitable. Even best features selection and classifiers selection is also a difficult task. With the concept of new advanced concept of deep learning, all these preprocessing and post-processing steps has been removed. Deep learning used original images as input, extract features automatically and learn it from direct pixels. It has been used in may real time applications and to solve complex problem. Basically, deep learning converts the original input images to features representation at different layers. These layers work based upon some filters like edge

detection filters or line detection. Most of the studies in literature shows that the most suitable model is deep Convolutional Neural Network (CNN) that works just like Multi-Layer Perceptron (MLP) with some modifications. So in this paper, CNN has been used with SVM to extract features automatically and then classify by using deep concepts learning. Some preprocessing has been performed to resize the images. Most of the time the images used in mammography are 1024x1024 so it is very difficult to process it directly on such huge size of the image. Therefore, resizing has been applied first. Then to smooth the images while retaining the edges preserved Bilateral filter has been employed with log transformation process. It enhances the quality of the images as shown in Figure 2 and Figure 3.

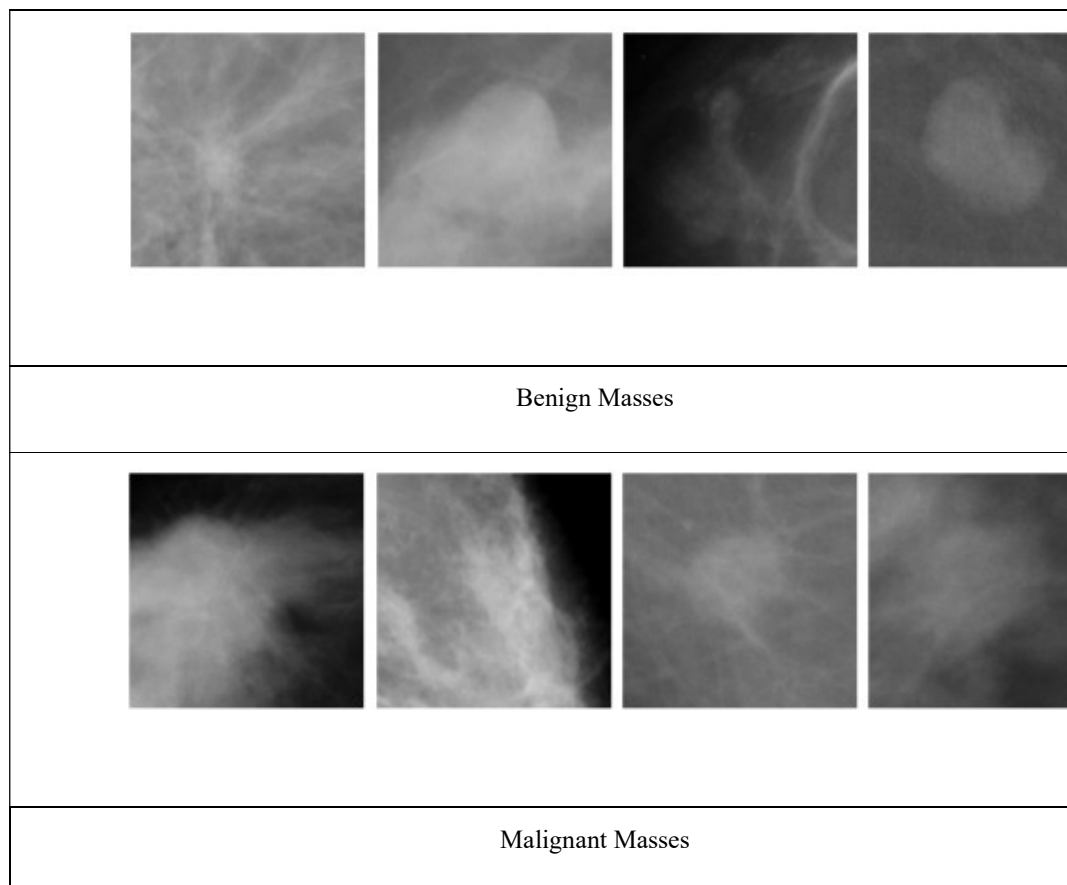


Figure 1: Benign and Malignant Masses from DDMS dataset

Table 1: List of Abbreviation used in this paper

Abbreviation	Complete Name
Log Transformation	Logarithmic Transformation
SVM	Support Vector Machine
Max Pooling	Maximum Pooling
CNN	Convolutional Neural Network
CAD	Computer Aided Diagnosis
MIAS	Mammographic image analysis society
DDMS	Digital database for screening mammography
AUC	Area under the receiver operating curve
ReLU	Rectified Liner Unit

2. Proposed Method

Proposed method consists of different phases to complete the whole process. Figure 1 shows sample image taken from DDMS dataset and it clearly shows that this mammogram images are not high contrast. There is also required to enhance the quality of the image so that it can be shows clearly visible and easily identifiable. Thus enhancement has been performed to improve the quality of mammograms. After that, features are required to classify those regions. Deep CNN has been used with SVM for features extraction and classification. Details of all these phases has been given below in detail

2.1 Dataset Description

We have used two standard datasets MIAS-mini [20] and DDMS [21] that are mostly used in the literature for breast mammograms classification and also augmented the images to generate a large dataset that is suitable for deep convolutional neural network. Augmentation process has been explained in the preprocessing step in detail. In These both datasets are supervised that mean class labels of both datasets are given with data information. From MIAS dataset 322 images and from DDMS 1800 images has been selected to use in this study. Third dataset has been created artificially by augmentation process. For data augmentation rotation has been performed on both

MIAS and DDMS images by 0, 45, 90, 135, 180, 225, 270, 315 and 360 degrees to generate huge dataset.

2.2 Preprocessing

After taking images from these datasets, it is required to enhance the visual quality of mammogram images. Therefore, enhancement has been performed to improve the quality of mammograms. There are different enhancement methods available in the literature but most suitable enhancement method for mammograms is to improve the breast part in such a way that it also preserves the edges information. First, to decrease the size of mammograms, resizing has been applied so that it can be suitable for convolution neural network. So mammogram images are resized to suitable size 96x96 using a special technique bicubic interpolation. After resizing to proper size, bilateral filter has been applied on the images. Bilateral filter enhances the images but also preserve edges. We have tried many other methods based upon histogram but bilateral filter performs well as compare to other methods. So first images have been filtered by using bilateral filter [7] so that this filter perform smoothing on the images while it also preserves the edges.

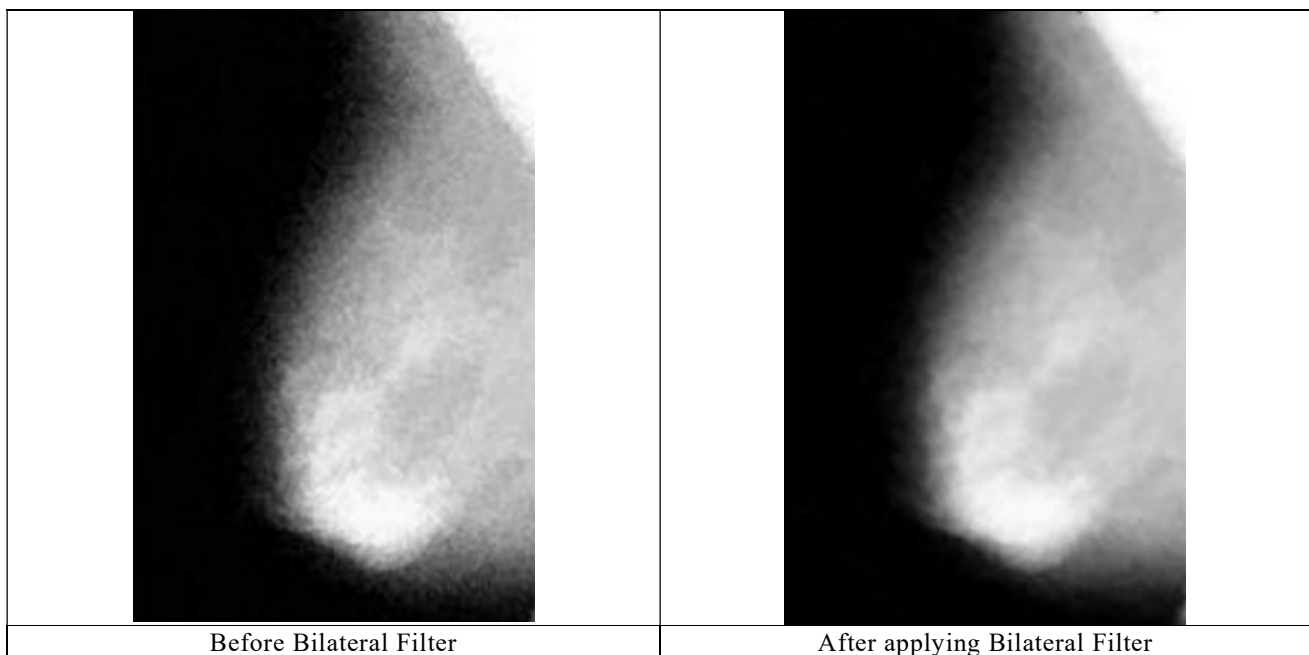
$$y(m, n) = \sum_k \sum_l h[m, n; k, l] x[k, l] \quad (1)$$

After preserving the edges and making the image smoother, logarithm transformation [8] has been applied on the image as shown in Figure 2. The basic advantage of log transformation is to increase the dynamic range specially for those areas which are dark in the mammograms as shown in Figure 3.

$$I_o = c \ln[1 + (e^\sigma - 1)I_{in}(i, j)] \quad (2)$$

Where I_o is output image, c is constant, I_{in} is input image and σ is the scaling factor that controls the input range to the logarithmic function. So this log transformation can also improve the low contrast areas

and regions available inside the images. Results of Bilateral filter and log transformation has been shown in Figure 2 and Figure 3. These images show clearly improvement by applying this process. At the end, to increase the training data, we created many images from existing data available in the datasets so that it can be suitable for convolutional neural network. For this purpose, rotation has been performed on all existing images at different angles by 0, 45, 90, 135, 180, 225, 270, 315 and 360. So in this way, a large training dataset has been created. Finally, we augmented the images to have a bigger training set. For data augmentation we rotated the images by 0, 45, 90, 135, 180, 225, 270, 315 and 360 degrees. After creating all these images, horizontal flip operation has been performed on all these artificially created images by rotation.



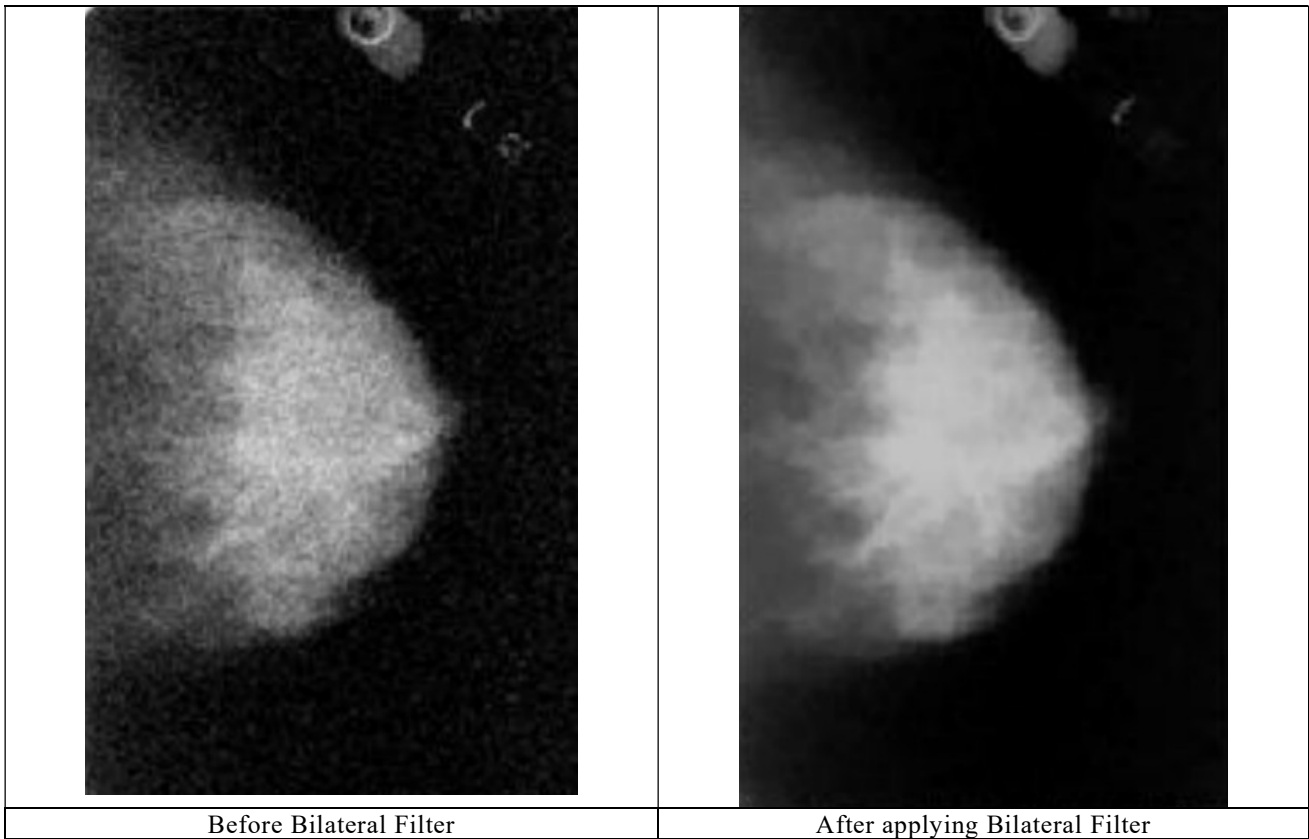


Figure 2: Results of Bilateral Filter on mammogram images

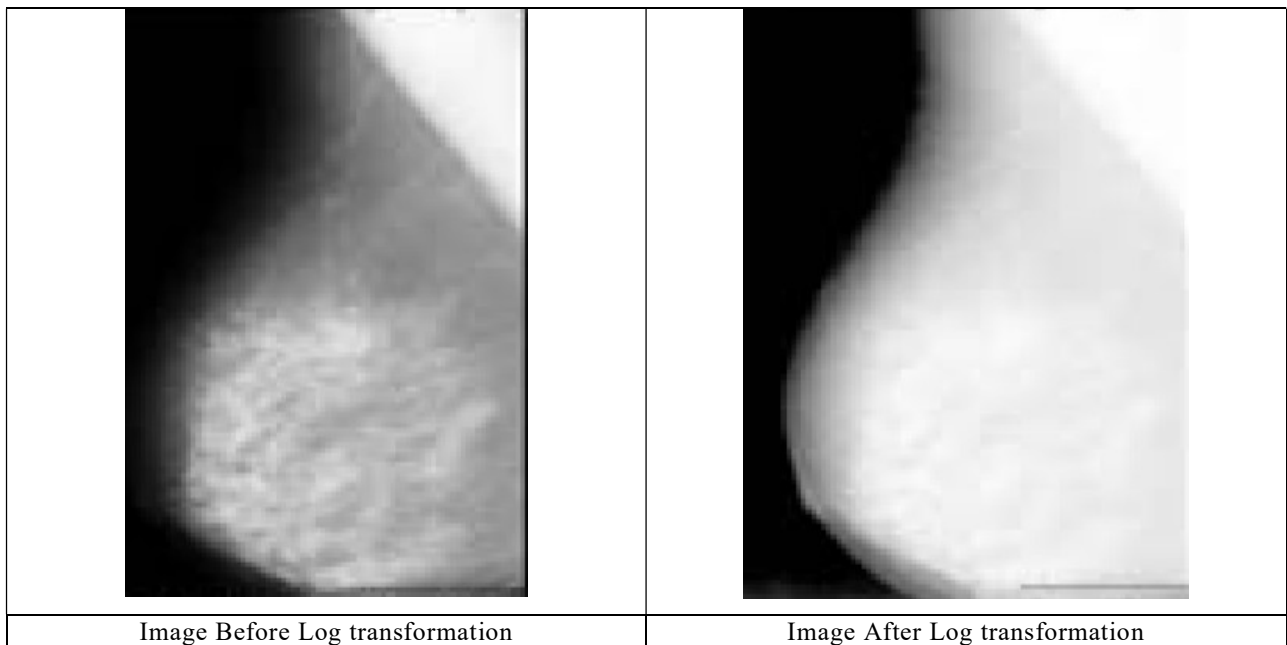


Figure 3: Results of Log Transformation on mammogram images

2.3 Deep Convolutional Neural Network

For classification, feed-forward convolutional neural network has been used in this paper to classify mammograms. The network used in this paper has been shown in Figure 4. Three convolution layers has been used for experimentation. Three layers has three filters 46, 128 and 256 respectively and each filter has a size of 5x5. Rectified Liner Unit (ReLU) has been used after these filters as activation functions. After using two convolutional layers at first two layers, max pooling has been employed at the third layer and after third layer, quadrant pooling has been applied. After that, fully connected layers has been used with 400 hidden neurons and at the end softmax layer has been used for classification. Max pooling concept has been used for filtering the process at each layer as shown in Figure 5. Biases has not been used during experimentation. The major advantage of not using biases is that it increases the speed of the network so that it can be trained very

efficiently and it also decrease the total amount of parameters that required to learn during training [10].

For training the network, stochastic gradient descent method has been used. Batch size for this stochastic gradient descent set to 64 with momentum 0.8. And decay parameter set to 1e-5. Dropout mechanism also used to the fully-connected layer by setting probability to 0.5. Training has been performed by using the model of Anna software library.

After training and features extraction along with corresponding labels of mammograms, it is used to train the Support Vector Machine (SVM) classifier. Training has been performed on different classifiers for each network. Finally, at the end, average has been used to calculate the average class score of all these different classifiers for the classification result as output.

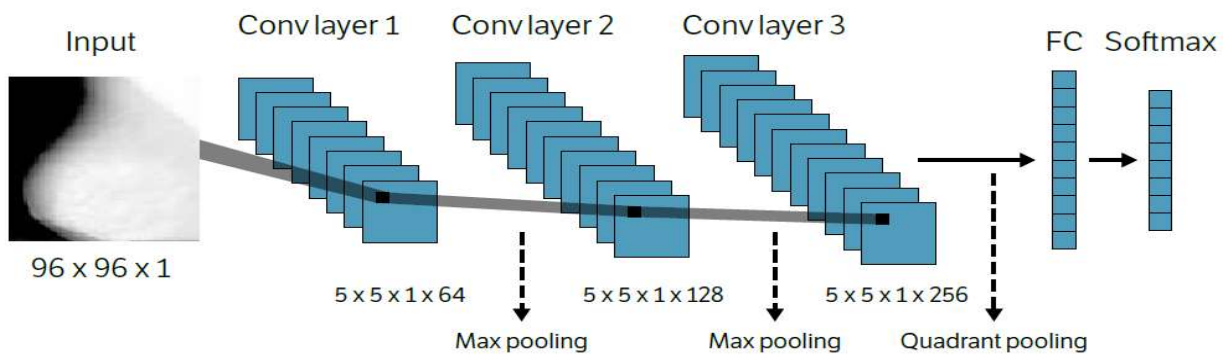


Figure 3: Deep convolutional neural network model

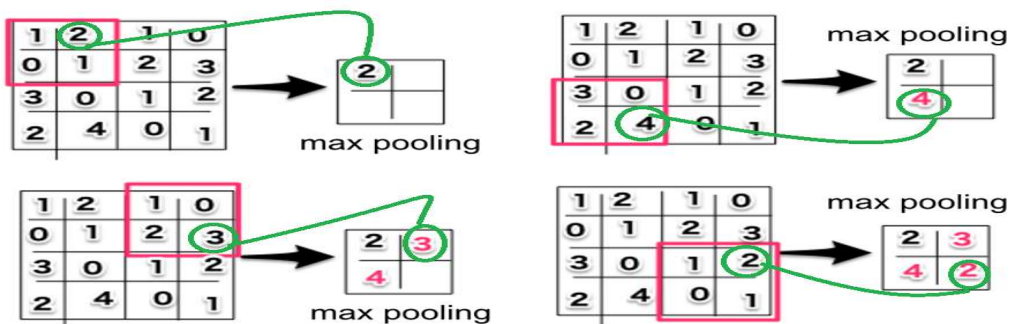


Figure 5: Process of Max Pooling applied in CNN

3. Results and Discussion

To show the performance of the proposed method, classification results has been shown in the form of accuracy in table 2. To calculate accuracy, a probability of 50% was used as threshold. For experimentation and testing, laptop with NVIDIA GTX 840M 2GB , 6 GB installed memory and Intel Core i7-4500U 1.8 GHz CPU has been used. MATLAB 2016b version has been for implementation and MatConvNet framework has been used for CNN. It takes around 120 minutes for training and for testing, it takes only 7 minutes for all test images.

Proposed method has been tested on large number of images taken from two standard datasets MIAS (322 images) and DDMS (1800 images) and also generated artificially by augmentation process explained in section 2.1 so total approximately 19000 image dataset has been generated by using this process. To generate the training and testing data, 10-fold cross validation process has been used with 50% training and 50% testing. Four different measures have been used to validate the results. Accuracy, Sensitivity, Specificity and area under the receiver operating curve

(AUC) [23] has been used to check the performance of the proposed method. These are the standard measures used for classification and specially AUC is the most common to assess the overall discrimination. Greater value of AUC shows good performance that should be close to 1. Proposed method has been compared with state-of-the-art methods that used the same concept like deep learning classification algorithms. The most common method used in this area are DeepCAD [17], CNN-Max-CAD-Qiu [15] and CNN-CAD-Jiao [16] by using same parameters and datasets.

Table 2 shows results of the proposed method on different datasets like MIAS-mini, DDMS, combination set of MIAS-mini and DDMS and combination of MIAS-mini, DDMS and artificially generated datasets. This table shows the if CNN has been trained on large dataset then it also enhances the performance of the proposed method. It clearly shows that on new artificially generated dataset the accuracy is 93.72 percent that is maximum against all other datasets. Even other measures like sensitivity is 94.19, specificity is 92.24 and AUC is 0.93. All these values shows that it performs well on big dataset that contains approximately 19000 images.

Table 2. Different Classifiers Results Without Enhancement

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
MIAS-mini	92.85	93.25	90.56	0.92
DDMS	93.02	92.84	91.35	0.91
MIAS-mini + DDMS	93.35	93.18	91.04	0.93
New Artificially generated dataset	93.72	94.19	92.24	0.93

Methodology	Dataset	Accuracy %	AUC %
Ball et al. [21]	DDMS	87	NA
Varela et al. [22]	DDMS	81	NA
DeepCAD [17]	DDMS + MIAS	91	0.91
CNN-Max-CAD-Qiu [15]	DDMS + MIAS	80	0.76
CNN-CAD-Jiao [16]	DDMS + MIAS	73	0.67
Proposed	DDMS + MIAS	93.35	0.93

Table (3) shows comparison of proposed method with other existing methods available in the literature. All these results for other methods has been taken reported in other research papers. The obtained results are reported in Table 2 and Table 3. Table 3 results has been taken from paper [17] for DeepCAD, CNN-MAX-CAD-Qiu and CNN-CAD-Jiao and for Ball and Varela has been taken from [18]. Ball shows 87% accuracy on DDMS dataset, Varela shows 81% on DDMS dataset. These methods have been not reported AUC in the paper therefore, I did not show results for these methods. DeepCAD has accuracy 91% and AUC 0.91 percent and author has tested on 600 ROI images collected from MIAS and DDMS. Author also shows results for other methods on the same collection of MIAS + DDMS dataset that CNN-Max-CAD-Qiu shows 80% and 0.76% accuracy and AUC accordingly and CNN-CAD-Jio shows 73% and 0.67% on the same collection of MIAS+DDMS dataset. Proposed method shows 93.35% accuracy and 0.93% AUC on combination dataset of MIAS-mini + DDMS. These values shows that proposed method works better as compare to all these existing methods. The proposed system does some preprocessing to enhance the performance of images like enhancement and resizing therefore it also performs well. Second, it has been trained on large set of images generated by artificially that is also helpful for good performance.

3. Conclusion and Future Work

In this paper, I have proposed a computer aided diagnosis system based upon deep convolutional neural network. There are two different tasks that has been applied. In the first task, breast preprocessing has been performed to make it suitable for CNN. During preprocessing, resizing has been performed on the images so that it can be suitable for CNN. Then dataset has been increase so that it is suitable for CNN. Different rotations have been used to generate artificially data by using the same data available in MIAS and DDMS dataset. After that, enhancement has been performed by using Bilateral filter and log transformation so that images can be smooth while preserving edges. After that CNN model has been used for features extraction classification. Results shows that it has improved the performance of CAD system by incorporating CNN. Due to these contributions, proposed system perform well. In the future, some other model of deep learning will be explored.

References

- [1] Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2016. *CA Cancer J. Clin.* 2016, 66, 7–30.
- [2] Kalager, M.; Zelen, M.; Langmark, F.; Adami, H.O. Effect of screening mammography on breast-cancer mortality in Norway. *N. Engl. J. Med.* 2010, 363, 1203–1210.
- [3] Posso, M.; Carles, M.; Rué, M.; Puig, T.; Bonfill, X. Cost-Effectiveness of double reading versus single reading of mammograms in a breast cancer screening programme. *PLoS ONE* 2016, 11, 1–13.
- [4] Bird, R.E., Professional Quality assurance for mammography screening programs, *Radiology* 177, pp.587. 1990.
- [5] Bellotti, R., De Carlo, F., Tangaro, S., Gargano, G., Maggipinto, G., Castellano, M., ... De Nunzio, G. (2006). A completely automated CAD system for mass detection in a large mammographic database. *Medical Physics*, 33(8), 3066–3075.
- [6] Varela, C., Tahoces, P. G., Méndez, A. J., Souto, M., & Vidal, J. J. (2007). Computerized detection of breast masses in digitized mammograms. *Computers in Biology and Medicine*, 37(2), 214–226.
- [7] F. Sahba, A. Venetsanopoulos, Breast mass detection using bilateral filter and mean shift based clustering, in: *Proceedings of the 2010 International Conference on Signal Processing and Multimedia Applications (SIGMAP)*, 2010, pp. 88–94.
- [8] Urvashi Manikpuri, Yojana Yadav, Image Enhancement Through Logarithmic Transformation, *International Journal of Innovative Research in Advanced Engineering (IJIRAE)* ISSN: 2349-2163 Volume 1 Issue 8, 2014
- [9] Wei, D., Chan, H. P., Helvie, M. a, Sahiner, B., Petrick, N., Adler, D. D., & Goodsitt, M. M. (1995). Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis. *Medical Physics*.
- [10] Pooya Khorrani, Tom Le Paine, Thomas S. Huang, Do Deep Neural Networks Learn Facial Action Units When Doing Expression Recognition?, *Computer Vision and Pattern Recognition*, 2017 arXiv:1510.02969
- [11] T.L. Economopoulos, P.A. Asvestas, G.K. Matsopoulos, Contrast enhancement of images using Partitioned Iterated Function Systems, *Image and Vision Computing* Volume 28, Issue 1, January 2010, Pages 45–54
- [12] Freund Y, Schapire RE, Experiments with a new boosting algorithm. In: *Machine learning: proceedings of the thirteenth international conference*, 1996, pp 325–332
- [13] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973

- [14] Lior Rokach, Ensemble-based classifiers, *Artificial Intelligence Reviews* (2010) 33:1–39
- [15] Qiu, Y.; 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 *Proceedings of the SPIE 9785, Medical Imaging 2016: Computer-Aided Diagnosis*, San Diego, CA, USA, 24 March 2016
- [16] Jiao, Z.; Gao, X.; Wang, Y.; Li, J. A deep feature based framework for breast masses classification. *Neurocomputing* 2016, 197, 221–231.
- [17] Qaisar Abbas, DeepCAD: A Computer-Aided Diagnosis System for Mammographic Masses Using Deep Invariant Features, *Computers*, 2016, 5, 28; doi:10.3390/computers5040028
- [18] Wentao Zhu, Qi Lou, Yeeleng Scott Vang, Xiaohui Xie, Deep Multi-instance Networks with Sparse Label Assignment for Whole Mammogram Classification, *arXiv preprint arXiv:1612.05968* (2016)
- [19] Suckling, J.; Parker, J.; Dance, D.; Astley, S.; Hutt, I.; Boggis, C.; Ricketts, I.; Stamatakis, E.; Cerneaz, N.; Kok, S.; et al. The mammographic image analysis society digital mammogram database. In *Excerpta Medica: International Congress Series*; Elsevier: Amsterdam, The Netherlands, 1994; Volume 1069, pp. 375–378. Available online: <http://peipa.essex.ac.uk/info/mias.html> (accessed on 28 October 2016).
- [20] Heath, M.; Bowyer, K.W.; Kopans, D. The digital database for screening mammography. In *Proceedings of the 5th International Workshop on Digital Mammography*, Toronto, ON, Canada, 11–14 June 2000.
- [21] Ball, J.E., Bruce, L.M.: Digital mammographic computer aided diagnosis (cad) using adaptive level set segmentation. In: *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. pp. 4973–4978. IEEE (2007)
- [22] Varela, C., Timp, S., Karssemeijer, N.: Use of border information in the classification of mammographic masses. *Physics in Medicine and Biology* 51(2), 425 (2006)
- [23] Metz, C.; Herman, B.; Shen, J. Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. *Stat. Med.* 1998, 17, 1033–1053