BMLDD: Breast mass lesion detection by using Deep Learning

Saeed Saffar Ardabili
PHD candidate of Tabriz University

Nasim Zolfaghari
PHD candidate of Sahand University of Technology

Afshin Ebrahimi
Professor of Electrical Engineering, Sahand University of Technology

Abstract

Early detection of breast cancer can increase treatment efficiency and therefore decline the high rate of moralities in patients with breast cancers. Computer-aided detection systems can help radiologists to detect breast mass lesions in a faster and more efficient way. This study proposed BMLDD, a new pipeline for breast mass lesion detection by a deep learning method using mam-mographic images. The proposed method uses histogram stretching, morphological operations, color-map, and a CNN with VGG16 architecture as its fundamental steps. This pipeline can efficiently improve the image condition, remove artifacts, extract breast region from the back-ground, highlight the differences in images, extracting useful features, and finally detect the breast masses. The performance evaluation on INbreast mammographic images has shown that BMLDD can yield promising results with 96% Sensitivity, 96% Precision, 96% F1-score, 96% Accuracy and 99% AUC. Moreover, the comparisons verified that BMLDD performed better than other state-of-the-art methods. Consequently, BMLDD is an efficient method for breast mass lesion detection.

Keywords: : 'Breast cancer mass, lesion detection, Deep learning, Morphological operation Mass detection, Convolutional neural network'

1. Introduction

Breast cancer is one of the fatal diseases in women, which has led to 626,700 deaths in 2018 [43]. About one-fourth of the cases with cancer have breast cancer [43]. The rate of mortality of breast cancer in developing countries is higher than the other countries [18]. A high number of cases with breast cancer imposes an enormous economic burden on societies [7, 27]. Early detection of breast cancer can significantly impact declining mortality, ineffective treatments, and imposed costs [42, 7, 22] . On the other hand, if the breast cancer is not cured in the early stage, the chance of treatment will aggravate due to the invasive metastasis of cancer to other organs [22]. Two strategies have been suggested by the World Health Organization (WHO) for cancer early detection: Distinguishing cancer in early stages, and consistent screening of the healthy populations [49]. Detecting breast cancer by specialist may result in high false-positive [46] and false-negative [31, 15]. the false diagnosis imposes the necessity of conducting further examinations [9]. Numerous computer-aided detection (CAD) methods have been proposed to decline the false detection rate. Mainly, CAD methods are attempting to pinpoint mass lesions in the mammographic images using image processing techniques. The ultimate purpose of CAD methods is to boost the accuracy of breast cancer diagnosis [21]. The proposed methods have used various machine learning classifiers such as simple logistic classifier [25, 19], k-nearest neighbor (KNN) [50, 47], decision tree [45], support vector machine (SVM) [19, 17, 41, 44], and artificial neural network (ANN) [33, 6, 32, 34, 35] to detect breast cancer lesions.

The advent of deep learning has revolutionized the machine learning field. Currently, the literature is replete with abundant deep learning methods that endeavor to address various biological issues [38, 52, 20, 12, 48]. The efficient application of deep learning has been indicated in many previous works. Maicas et al. have used deep reinforcement learning in order to detect breast lesions in dynamic contrast-enhanced magnetic resonance images (DCE-MRI) [28]. They modified the deep Q-network approach and utilized it for search policy. Their proposed method showed high speed as well as decent accuracy. De et al. have exploited a deep neural network with U-net architecture [29]. Their proposed method can both classify and segment the lesions in mammographic images. Three recent works, including the methods proposed by Ribli et al. [37], Reiazi et al. [36], and Cao et al. [10] have used Faster-RCNN network which is a regional convolutional neural network in the context of detecting breast lesions. Not only can these methods detect breast lesions, but also they can classify the lesions into malignant and benign classes. Al-Antari et al. have proposed a method for detection and classification on breast lesions from mammographic images [2]. They used YOLO classifier and boosted its performance by Inception ResNet-V2 classifier.

The previously proposed methods succeeded in obtaining promising results. However, they have high false-positive rates for proposing pretends methods with decent performance, which is still challenging and demanding. In this paper, we introduce BMLDD, a method for breast mass lesion detection using deep learning. BMLDD has several steps, including preprocessing, morphological operations, color-map, and deep neural network. This pipeline is a novel procedure in detecting mass lesions of the breast. We have implemented various types of deep neural networks, and the results showed that VGG16 is the best architecture. The performance comparisons with other state-of-the-art methods indicated that BMLDD could efficiently detect breast mass lesions.

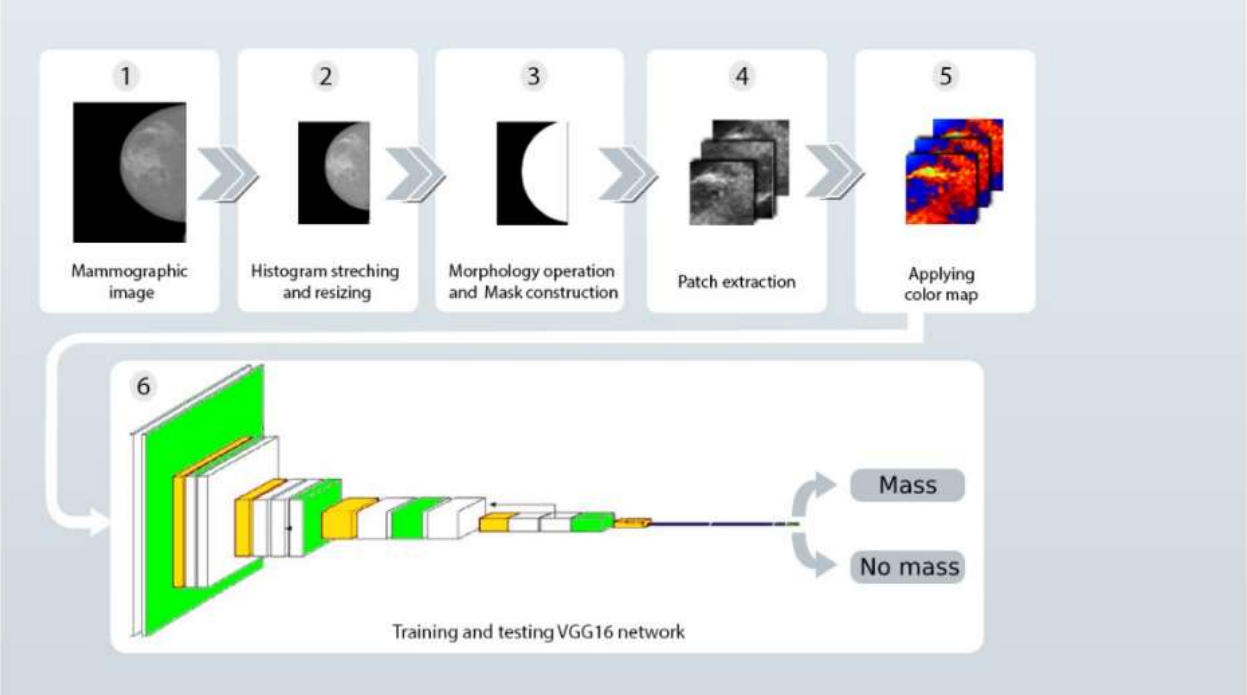

Figure 1: An overview of BMLDD steps. It takes mammographic images as the input. In the preprocessing step, it applies histogram stretching on the images and resizes them to smaller sizes. It then performs morphological operations to eliminate the noises and applies a threshold on it to specify the breast region. In the next step, the images are break down into small patches. After applying the colormap for intensifying important parts, a CNN with VGG16 architecture is applied on the images to classify them.

2. Materials and methods

2.1. Data

Mammographic images were obtained from INbreast database [30]. This database contains comprehensive cases from women and men with several types of lesions such as mass, calcification, asymmetric, and distortions. We assigned the images containing mass lesions as positive images (labeled "1") and the images without mass lesions as negative images (labeled "0"). The resulting dataset comprises 410 images that were split to 60%, 20%, and 20% for training, validation, and test sets. An overview of BMLDD steps is shown in Figure 1. These steps are described meticulously in the following subsections.

2.2. Pre-processing

Most image processing methods, apply some preprocessing steps on the images to enhance the image condition, remove noise from the image, and improve the performance of further processes. We use the histogram stretching [4] for contrast equalization besides a applying morphological operations [13].

2.2.1. Histogram stretching

We applied histogram stretching in order to improve the image conditions and have more robust results. This procedure converts the image intensities to [0, 255] interval using Eq (1).

$$p_i = \frac{p_i - min_j(p_j)}{max_j p_j - min_j p_j} \quad (1)$$

Where p_j is the intensity of the 꾉th pixel in the image. Processing large images impose much time and memory overload for the executions and may lead to occur overfitting. The converted images were resized by nearest-neighbor interpo-lation [16] to avoid this situation. This method resizes the images by a resize coefficient. In this study, we considered the resize coefficient as 1/12, which means that each image's size shrinks to the 1/12 of its original size.

2.3. Morphological operations

Morphological operations perform non-linear filters on the image, which helps eliminate noises, focus on the breasts regions, and disregard worthless parts of the image. Before applying the morphological operations, we extended the images with zero paddings. This procedure enables us to apply morphological operations on all pixels, even the edge pixels. On the other hand, this may not result in false detection since the probability of mass lesions in the edge pixels is insignificant.

Conventionally, morphological operations are done on binary images; however, they can be modified for gray-scale images. Suppose $B_{n\times m}$ be a binary image and $M_{K\times 1}$ a morphological filter. We defined the set of positive pixels by $\beta = \{(i,j)|B(i,j) = 1\}$ .Assume that (i',j') is the central pixel of morphological filter. Two basic morphological operations are dilution (D) and erosion (E), which are defined as follows:

$$D(B , M) = U_{(i,j)\epsilon\beta}\, d(i,j) \quad (2)$$

$$d(i ,j) = \{(i+h,j+g)|(B(i+h,j+g)+M(i'+h, j'+g)\geq 1; h\epsilon\{\frac{-1}{2},\cdots,\frac{1}{2}\}, g\epsilon\{\frac{-K}{2},\cdots,\frac{K}{2}\}\} \quad (3)$$

$$E(B,M) = U_{(i,j)\ \epsilon\beta}\, e(i,j) \quad (4)$$

$$E(i ,j)|if\ \forall\ h\ \epsilon\{\frac{-1}{2},\cdots,\frac{1}{2}\},\ g\ \epsilon\{\frac{-1}{2},\cdots,\frac{1}{2}\}\ d(i ,j) \quad (5)$$

$D(B,M)$ and $E(B,M)$ are positive pixels after applying dilation and erosion operations, respectively. In dilution and erosion, the morphological filter slides over the B image. Dilation operation aims to fill in the missed pixels according to the morphological filter. When the central pixel of the filter is placed under positive pixel of B, the negative surrounding pixels of B will change to positive, if the corresponding pixel in the M is positive. Therefore, dilution operation can increase the number of positive pixels.

On the contrary, erosion operation conserves solely the positive pixels that accord completely to their morphological filter; i.e., when the filter is placed on them, all positive pixels results by dilution operation positive were positive in B.

The dilation and erosion operations can be extended for gray-scale images. Let B be a gray-scale image with integer intensities in [0, 255] and M be the binary morphological structure filter. Define $\mathcal{M} = \{(p,q)|M(p,q) = 1\}$ as the positive pixels in morphological filter. The dilution and erosion are calculated as the following.

$$D(B,m)_{(i,j)} = \max_{(p,q)\epsilon M}\ \{B(i-p, j- q)\} \quad (6)$$
$$E(B,m)_{(i,j)} = \min_{(p,q)\epsilon m}\ \{B(i-p, j- q)\} \quad (7)$$

Various combinations of dilution and erosion can define numerous morphological operations. Two practical operations are opening (O) and top white hat (W) operations, which are calculated as follows.

$$O(B,m) = D(E(B,m),m) \quad (8)$$
$$W(B,m)=B-O(B,m) \quad (9)$$

The intuition of opening operation eliminates the small positive regions, whereas the top white hat removes the artifacts by eliminating all pixels that are remained from the opening operation. We applied the top white hat operation to cross-out artifacts. The considered morphological filter in this study was a Disc filter with $240r$ pixels where $r = 1/12$ denotes the resize coefficient.

2.4. Mask construction

Since the lesions are located chiefly in the breasts region, we specify the breast region and disregarded other regions. To this aim, a thresholding technique was applied to the resulted images. Specifically, each pixel was compared to a threshold and was set to zero if its intensity was lower than the threshold and set to 1, otherwise. The bulkiest object on this mask was chosen as the breast region. In this way, we could extract the breast region from the surrounding black portions. The threshold value was set to 10 in his study. After conducting the thresholding technique, a dilation operation was applied to the images to refine the resulted mask.

2.5. Patch extraction

Due to the computational burden of processing the full images, each sample image was segmented to smaller pieces called "patch." The extraction of negative and positive patches in training and validation images was done according to the following rules:

• Negative: The patches that were located completely in the breast region and contained no mass lesion. To do this, we grid sampled the breast region to patches of size 76 × 76, where every two adjacent patches had overlap.

• Positive: The patches that contain the center of the mass lesion. To extract them, we considered nine different patches: the first is a 76 × 76 patch that its central pixel is the center of the mass lesion. By moving this patch according to the following directions, we extracted eight more positives patches.

1. Five pixels to the right
2. Five pixels to left
3. Five pixels up
4. Five pixels down
5. Five pixels to the right and five pixels up
6. Five pixels to the right and five pixels down
7. Five pixels to left and five pixels up
8. Five pixels to left and five pixels down

For images in the test set, the patch extraction was done similar to the way that was used to extract negative patches in training and validation images, because we did not have any ground truth for the test images. This procedure has led to extracting 15684, 5313, and 5071 patches for training, validation, and test set.

2.6. Color-Map

We applied color-map [3] on each patch to intensify the most important parts. This procedure maps the gray-scale image to a colorful one. Each color is assigned to a number in [0, 255] interval. The color-map uses this assignment to color each gray pixel. The results are colorful patches with enhanced visualizations, which illustrated the important regions better. Each colorful patch has 76 × 76 × 3 size. The experiments have shown that applying color-map on the patches results in higher retrieval rates of breast mass lesions.

2.7. Convolutional neural network

The colorful patches were used to train VGG16 neural network [40]. This network classifies the patches into two classes: the ones containing mass lesions labeled by "1", and the ones will without mass labeled by "0". The architecture of VGG16 is represented in Table 1. This architecture comprises 13 convolutional, five max pooling, and three dense layers. All convolutional layers convoluted the images by 3 × 3 filters and then applied the rectified linear unit (ReLU) activation function [5] on them. The fundamental role of convolutional layers is to process the patches and extract their efficient features. Moreover, some pooling layers are embedded in this network to down-sample the resulted features and avoid the computational burden. Dense layers are fully connected to their previous layers and are responsible for further processing on the patches and finally classifying them accurately. The activation function of neurons in the last layer is the softmax function [23], which predicts the probability of belonging the patches to each class.

This network has 39,896,898 trainable parameters learned on the training samples with a batch size of 128, and the epoch number equals 9. The stochastic gradient descent (SGD) [8] with a learning rate of 0.01, the momentum of 0.9, and weight decay equals 0.0005 to learn the model parameters.

It is worth mentioning that all hyper-parameters was finely-tuned and set to the numbers that yielded the best results. The implementations was done by Python 3.0 and Keras library on a system with 8G RAM and Intel Corei5 CPU.

Table 1: The architecture of VGG16. Conv2D, MaxPooling2D, Dense represent the convolutional, max pooling, and fully connected layers.

Layer name	Layer type	Output shape	Number of parameters
input 3D	(InputLayer)	(76, 76, 3)	0
block1 conv1	(Conv2D)	(76, 76, 64)	1792
block1 conv2	(Conv2D)	(76, 76, 64)	36928
block1 pool	(MaxPooling2D)	(38, 38, 64)	0
block2 conv1	(Conv2D)	(38, 38, 128)	73856
block2 conv2	(Conv2D)	(38, 38, 128)	147584
block2 pool	(MaxPooling2D)	(19, 19, 128)	0
block3 conv1	(Conv2D)	(19, 19, 256)	295168
block3 conv2	(Conv2D)	(19, 19, 256)	590080
block3 conv3	(Conv2D)	(19, 19, 256)	590080
block3 pool	(MaxPooling2D)	(9, 9, 256)	0
block4 conv1	(Conv2D)	(9, 9, 512)	1180160
block4 conv2	(Conv2D)	(9, 9, 512)	2359808
block4 conv3	(Conv2D)	(9, 9, 512)	2359808
block4 pool	(MaxPooling2D)	(4, 4, 512)	0
block5 conv1	(Conv2D)	(4, 4, 512)	2359808
block5 conv2	(Conv2D)	(4, 4, 512)	2359808
block5 conv3	(Conv2D)	(4, 4, 512)	2359808
block5 pool	(MaxPooling2D)	(2, 2, 512)	0
fc1	(Dense)	(4096)	8392704
fc2	(Dense)	(4096)	16781312
predictions	(Dense)	(2)	8194

3. Results

3.1. Evaluation metrics

When the model was learned on the training set, its performance was evaluated using the following criteria that have been widely used in the classification problems.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{10}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{11}$$

$$\text{F1 - score} = \frac{2 \times sensitivity \times precision}{sensitivity + precision} \tag{12}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{13}$$

$$\text{FPR} = \frac{FP}{FP+TN} \tag{14}$$

where TP, FP, TN, FN denote true positive, false positive, true negative, and false negative, respectively. The values of these criteria depend on the threshold value applied on the predicted probabilities. The CNN used at the final stage of method, yielded the probability of lesion presence in each image. The probability was converted to a binary label by comparing to a threshold. In addition to the foregoing criteria, AUC is computed as another criterion for assessing the model performance which is invariant to the threshold value. AUC is the area under the ROC curve which plots sensivity versus FPR for all thresholds. All of these criteria lies within the [0, 1] interval,

and each of them have different point of view in evaluating the model performance. Moreover, the binary cross entropy (BCE) of the predicted probabilities is also computed as Formula 15.

$$BCE = -\frac{1}{N}\sum_{i=1}^{N} yi.\ \text{Log}\ t_i + (1 - yi)\ \log\ (1 - t_i) \quad (15)$$

where y_i is the true label of the i th patch and $t_i$is the predicted probability of method for the i th patch.

3.2. Performance of BMLDD

To find out the best number of training epoch, we executed BMLDD on multiple epochs and evaluated the BCE on the training and validation set in each epoch. Figure 2 illustrated the trend of BCE alterations during the training epochs. Both of BCE curves has declining trends, while training BCE declines steeply. After 9 epochs the training error approaches to zero and the validation error is plausible. The ROC curve of BMLDD on test data is shown in Figure 3 It shows that BMLDD performed efficiently on test data and obtained high AUC. Moreover, we performed BMLDD not only with VGG16 as the architecture of CNN in the last step, but also with some other CNN architectures. Three other CNN architecture including ALEXNET [39], ZFNET [51], and LENET [26] was considered for the last step of BMLDD. The results of BMLDD with various CNN architectures are presented in Table 2. The high values of all classification criteria as well as low BCE for all architectures verifies that BMLDD is an efficient and robust method that is not hugely dependent to the CNN architecture; nevertheless, the best results were achieved with VGG16 architecture. The sensitivity,precision, F1-score and Accuracy remained unchanged in these five models. However, VGG16 succeeded to obtain higher AUC and lower BCE.

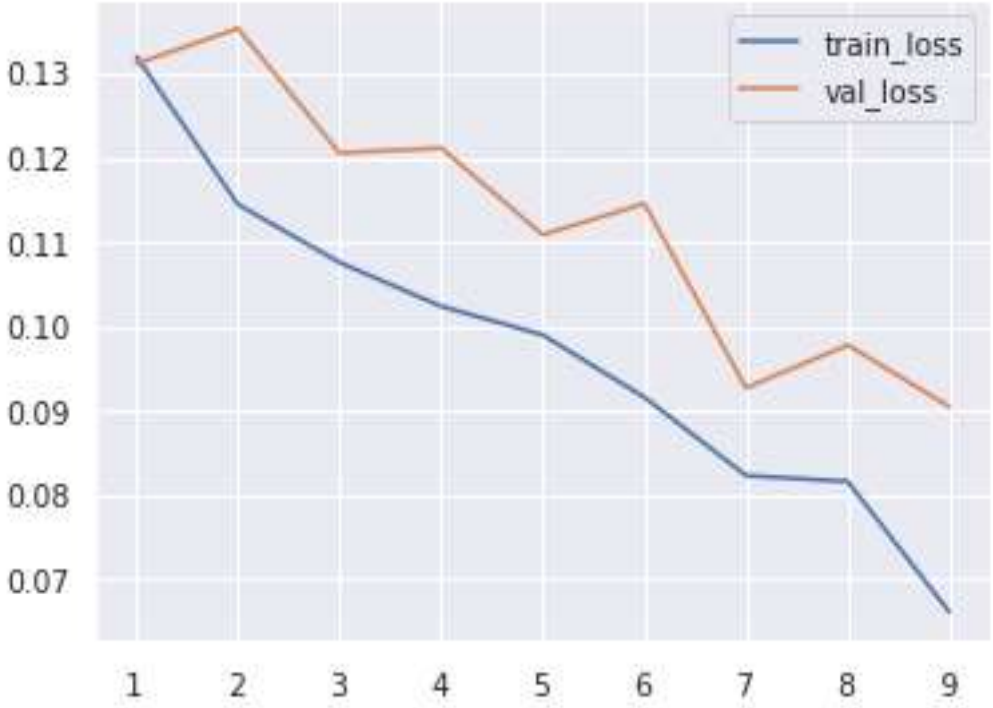

Figure 2: The changes in BCE values on training and validation set during the different epochs.

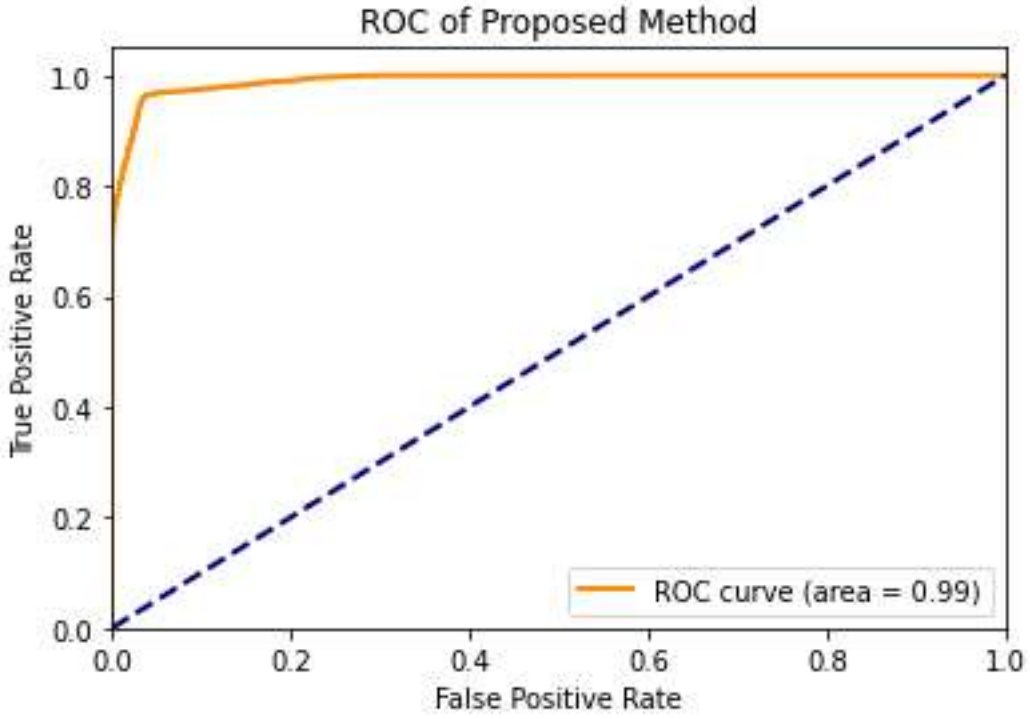

Figure 3: ROC curve of BMLDD.

3.3. Comparison with other methods

To further assess BMLDD performance, we compared its results with multiple state-of-the-art methods. Table 3 represents the method outline and performance of other methods on INbreast database. The evaluated criteria substantiated that BMLDD outperformed other methods. Therefore, BMLDD can be used in the accurate detection of breast masses.

Table 2: The performance of BMLDD with various CNN architectures.

CNN architecture	*Sensitivity*	*Precision*	*F1-score*	*Accuracy*	*AUC*	*BCE*
VGG16	0.96	0.96	0.96	0.96	0.99	0.12
ALEXNET	0.96	0.96	0.96	0.96	0.96	0.63
ZFNET	0.96	0.96	0.96	0.96	0.97	0.31
LENET	0.96	0.96	0.96	0.96	0.99	0.14

Table 3: The performance of state-of-the-art methods on INbreast database.

Reference	Method	Accuracy
BMLDD	Histogram stretching, morphological operations, colormap, and VGG16	0.96
Carneiro et al. [11]	Deep learning with CNN-F, Random forest classifers, and Boosting classifers	0.90
Dhungel et al. [14]	ROI extraction with DBN and GMM, RCNN, and Random forest	0.90
Kozegar et al. [24]	Adaptive threshold with other machine learning approaches	0.87

4. Conclusion

Breast cancer is one of the four leading cancers in the world. The high rate of breast cancer cases in women has led to a huge economic burden and high mortality in women. Since early detection of breast cancer helps significantly decrease fatalities and prosper the treatments, early detection of breast cancer is highly important. Computational methods that detect breast lesions can aid in the early detection of breast cancers and thereby increase patients' chance of survival.

This study presented a novel pipeline, called BMLDD, for breast mass lesion detection using deep learning in mam-mographic images. The proposed method uses histogram stretching and morphological operations as the preprocessing steps to improve the image conditions, remove artifacts, determine the breast region, and exclude the background portion of images. Then we divided the images into small patches to avoid the computational burden. The patches were colored using a color-map process to better illustrate the differences and nuances in the images. The executions have shown that using color-map increases the number of retrieved lesions. The modified patches were used to learn a CNN with VGG16 architecture. This network was responsible for both the feature extraction and mass lesion detection in patches.

The performance evaluations on INbreast database have shown that BMLDD achieved very accurate results. More over, by changing the final CNN architecture, we investigated the best architecture for CNN and verified that VGG16 have led to the results with less error. It has 96% Sensitivity, 96% Precision, 96% F1-score, 96% Accuracy, 99% AUC, and 0.12 binary cross-entropy. Furthermore, the comparison of BMLDD with several state-of-the-art methods verified that BMLDD outperformed other methods. Therefore, BMLDD can be efficiently used for detecting breast lesions in mammographic images.

References

[1] Agarwal, R., Diaz, O., Yap, M.H., Lladó, X., Martí, R., 2020. Deep learning for mass detection in full field digital mammograms. Computers in Biology and Medicine , 103774.
[2] Al-antari, M.A., Kim, T.S., 2020. Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital x-ray mammograms. Computer Methods and Programs in Biomedicine , 105584.
[3] Al Azzeh, J., Alhatamleh, H., Alqadi, Z.A., Abuzalata, M.K., 2016. Creating a color map to be used to convert a gray image to color image. International Journal of Computer Applications 153, 31–34.
[4] Alparslan, E., Fuatince, M., 1981. Image enhancement by local histogram stretching. ITSMC 11, 376–385.
[5] Ang-bo, J., Wei-wei, W., 2018. Research on optimization of relu activation function [j]. Transducer and Microsystem Technologies 2.
[6] Ayer, T., Chen, Q., Burnside, E.S., 2013. Artificial neural networks in mammography interpretation and diagnostic decision making. Computational and mathematical methods in medicine 2013.
[7] Blumen, H., Fitch, K., Polkus, V., 2016. Comparison of treatment costs for breast cancer, by tumor stage and type of service. American health & drug benefits 9, 23.
[8] Bottou, L., 2012. Stochastic gradient descent tricks, in: Neural networks: Tricks of the trade. Springer, pp. 421–436.
[9] Brennan, M., Houssami, N., 2016. Discussing the benefits and harms of screening mammography. Maturitas 92, 150–153.
[10] Cao, Z., Yang, Z., Liu, X., Zhang, Y., Wu, S., Lin, R.S., Huang, L., Han, M., Ma, J., 2019. Deep learning based lesion detection for mammograms, in: 2019 IEEE International Conference on Healthcare Informatics (ICHI), IEEE. pp. 1–3.
[11] Carneiro, G., Nascimento, J., Bradley, A.P., 2017. Automated analysis of unregistered multi-view mammograms with deep learning. IEEE transactions on medical imaging 36, 2355–2365.
[12] Choi, J., Park, S., Ahn, J., 2020. Refdnn: a reference drug based neural network for more accurate prediction of anticancer drug resistance. Scientific reports 10, 1–11.
[13] Comer, M.L., Delp, E.J., 1999. Morphological operations for color image processing. J. Electronic Imaging 8, 279–289.
[14] Dhungel, N., Carneiro, G., Bradley, A.P., 2017. A deep learning approach for the analysis of masses in mammograms with minimal user intervention. Medical image analysis 37, 114–128.
[15] Dromain, C., Boyer, B., Ferre, R., Canale, S., Delaloge, S., Balleyguier, C., 2013. Computed-aided diagnosis (cad) in the detection of breast cancer. European journal of radiology 82, 417–423.
[16] Dunlop, G., 1980. A rapid computational method for improvements to nearest neighbour interpolation. Computers & Mathematics with Applications 6, 349–353.
[17] Eltoukhy, M.M., Gardezi, S.J.S., Faye, I., 2014. A method to reduce curvelet coefficients for mammogram classification, in: 2014 IEEE Region 10 Symposium, IEEE. pp. 663–666.
[18] Gardezi, S.J.S., Elazab, A., Lei, B., Wang, T., 2019. Breast cancer detection and diagnosis using mammographic data: systematic review. Journal of medical Internet research 21, e14464.

[19] Gardezi, S.J.S., Faye, I., Eltoukhy, M.M., 2014. Analysis of mammogram images based on texture features of curvelet sub-bands, in: Fifth International Conference on Graphic and Image Processing (ICGIP 2013), International Society for Optics and Photonics. p. 906924.
[20] Hao, J., Kim, Y., Kim, T.K., Kang, M., 2018. Pasnet: pathway-associated sparse deep neural network for prognosis prediction from high throughput data. BMC bioinformatics 19, 1–13.
[21] Hela, B., Hela, M., Kamel, H., Sana, B., Najla, M., 2013. Breast cancer detection: A review on mammograms analysis techniques, in: 10th International Multi-Conferences on Systems, Signals & Devices 2013 (SSD13), IEEE. pp. 1–6.
[22] Killelea, B.K., Long, J.B., Chagpar, A.B., Ma, X., Wang, R., Ross, J.S., Gross, C.P., 2014. Evolution of breast cancer screening in the medicare population: clinical and economic implications. JNCI: Journal of the National Cancer Institute 106.
[23] Kouretas, I., Paliouras, V., 2019. Simplified hardware implementation of the softmax activation function, in: 2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST), IEEE. pp. 1–4.
[24] Kozegar, E., Soryani, M., Minaei, B., Domingues, I., et al., 2013. Assessment of a novel mass detection algorithm in mammograms. Journal of cancer research and therapeutics 9, 592.
[25] Landwehr, N., Hall, M., Frank, E., 2005. Logistic model trees. Machine learning 59, 161–205.
[26] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE 86, 2278–2324.
[27] Luengo-Fernandez, R., Leal, J., Gray, A., Sullivan, R., 2013. Economic burden of cancer across the european union: a population-based cost analysis. The lancet oncology 14, 1165–1174.
[28] Maicas, G., Carneiro, G., Bradley, A.P., Nascimento, J.C., Reid, I., 2017. Deep reinforcement learning for active breast lesion detection from dce-mri, in: International conference on medical image computing and computer-assisted intervention, Springer. pp. 665–673.
[29] de Moor, T., Rodriguez-Ruiz, A., Mérida, A.G., Mann, R., Teuwen, J., 2018. Automated lesion detection and segmentation in digital mam-mography using a u-net deep learning network, in: 14th International Workshop on Breast Imaging (IWBI 2018), International Society for Optics and Photonics. p. 1071805.
[30] Moreira, I.C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M.J., Cardoso, J.S., 2012. Inbreast: toward a full-field digital mammographic database. Academic radiology 19, 236–248.
[31] Nelson, H.D., O'Meara, E.S., Kerlikowske, K., Balch, S., Miglioretti, D., 2016. Factors associated with rates of false-positive and false negative results from digital mammography screening: an analysis of registry data. Annals of internal medicine 164, 226–235.
[32] Oliver, A., Lladó, X., Torrent, A., Martí, J., 2014. One-shot segmentation of breast, pectoral muscle, and background in digitised mammo grams, in: 2014 IEEE International Conference on Image Processing (ICIP), IEEE. pp. 912–916.
[33] Pal, N.R., Bhowmick, B., Patel, S.K., Pal, S., Das, J., 2008. A multi-stage neural network aided system for detection of microcalcifications in digitized mammograms. Neurocomputing 71, 2625–2634.
[34] Pratiwi, M., Harefa, J., Nanda, S., et al., 2015. Mammograms classification using gray-level co-occurrence matrix and radial basis function neural network. Procedia Computer Science 59, 83–91.
[35] Rampun, A., Morrow, P.J., Scotney, B.W., Winder, J., 2017. Fully automated breast boundary and pectoral muscle segmentation in mammograms. Artificial intelligence in medicine 79, 28–41.
[36] Reiazi, R., Paydar, R., Ardakani, A.A., Etedadialiabadi, M., 2018. Mammography lesion detection using faster r-cnn detector .
[37] Ribli, D., Horváth, A., Unger, Z., Pollner, P., Csabai, I., 2018. Detecting and classifying lesions in mammograms with deep learning. Scientific reports 8, 1–7.
[38] Rohani, N., Eslahchi, C., 2019. Drug-drug interaction predicting by neural network using integrated similarity. Scientific reports 9, 1–11.
[39] Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer. pp. 234–241.
[40] Simonyan, K., Zisserman, A., 2015. International conference on learning representations.
[41] Singh, B., Jain, V., Singh, S., 2014. Mammogram mass classification using support vector machine with texture, shape features and hierarchical centroid method. Journal of Medical Imaging and Health Informatics 4, 687–696.
[42] Society., A.C., 2016. Breast cancer survival rates, by stage. http://www.cancer.org/cancer/breastcancer/detailedguide/ breast-cancer-survival-by-stage.
[43] Society, A.C., 2018. Global cancer: Facts and figures, 4th edition. https://www.cancer.org/research/cancer-facts-statistics/ global.html.
[44] Sonar, P., Bhosle, U., Choudhury, C., 2017. Mammography classification using modified hybrid svm-knn, in: 2017 International Conference on Signal Processing and Communication (ICSPC), IEEE. pp. 305–311.
[45] Sumbaly, R., Vishnusri, N., Jeyalatha, S., 2014. Diagnosis of breast cancer using decision tree data mining technique. International Journal of Computer Applications 98.

[46] Svahn, T.M., Macaskill, P., Houssami, N., 2015. Radiologists' interpretive efficiency and variability in true-and false-positive detection when screen-reading with tomosynthesis (3d-mammography) relative to standard mammography in population screening. The Breast 24, 687–693.
[47] Tan, P.N., Steinbach, M., Kumar, V., 2016. Introduction to data mining. Pearson Education India.
[48] Wei, H., Liao, Q., Liu, B., 2020. ilncrnadis-fb: identify lncrna-disease associations by fusing biological feature blocks through deep neural network. IEEE/ACM Transactions on Computational Biology and Bioinformatics .
[49] (WHO), W.H.O., 2020. Who guide to cancer early diagnosis. http://www.who.int/iris/handle/10665/254500.
[50] Wu, X., Kumar, V., Quinlan, J.R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G.J., Ng, A., Liu, B., Philip, S.Y., et al., 2008. Top 10 algorithms in data mining. Knowledge and information systems 14, 1–37.
[51] Zeiler, M.D., Taylor, G.W., Fergus, R., 2011. Adaptive deconvolutional networks for mid and high level feature learning, in: 2011 International Conference on Computer Vision, IEEE. pp. 2018–2025.
[52] Zhang, Y., Zhan, L., Thompson, P.M., Huang, H., 2019. Biological knowledge guided deep neural network for brain genotype-phenotype association study, in: Multimodal Brain Image Analysis and Mathematical Foundations of Computational Anatomy. Springer, pp. 84–92.