Detection and Weak Segmentation of Masses in Gray-Scale Breast Mammogram Images Using Deep Learning

Young Jae Kim and Kwang Gi Kim

Department of Biomedical Engineering, Gil Medical Center, Gachon University College of Medicine, Incheon, Korea.

Purpose: In this paper, we propose deep-learning methodology with which to enhance the mass differentiation performance of convolutional neural network (CNN)-based architecture.

Materials and Methods: We differentiated breast mass lesions from gray-scale X-ray mammography images based on regions of interest (ROIs). Our dataset comprised breast mammogram images for 150 cases of malignant masses from which we extracted the mass ROI, and we composed a CNN-based deep learning model trained on this dataset to identify ROI mass lesions. The test dataset was created by shifting some of the training data images. Thus, although both datasets were different, they retained a deep structural similarity. We then applied our trained deep-learning model to detect masses on 8-bit mammogram images containing malignant masses. The input images were preprocessed by applying a scaling parameter of intensity before being used to train the CNN model for mass differentiation.

Results: The highest area under the receiver operating characteristic curve was 0.897 (Î 20).

Conclusion: Our results indicated that the proposed patch-wise detection method can be utilized as a mass detection and segmentation tool.

Key Words: Breast mammogram, detection lesion of mass, deep learning, convolutional neural network, data normalization

INTRODUCTION

Breast cancer is the fifth most common cause of death among women. In 2012, breast cancer was the most common cancer in women worldwide, with approximately 1.7 million new cases diagnosed.[1,2] Approximately 8% of women suffer from breast

cancer in modern developed countries, such as the US. Early detection of breast cancer in asymptomatic patients via the identification of microcalcifications and masses, vital indicators of the disease, can reduce the risk of mortality. Mammography screening is one of the most crucial early detection methods for breast cancer. However, some reports indicate that it is difficult for radiologists to differentiate breast carcinoma in mammography screening, with an estimated sensitivity of approximately 75%. Accordingly, early stage detection has emerged as an important issue in computer-aided diagnosis (CAD) technology for diagnostic breast cancer pathology.[3,4]

For a CAD system, detecting masses in mammography is challenging due to the ambiguity of their shape and poor contrast. In general, masses are classified as either benign or malignant to improve the biopsy yield ratio. Research indicates that regularity and radiolucency are deeply related to deciding whether or not a mass is benign.[5-7] In many studies on the detection of masses, CAD is followed by feature-based discrimination using various artificial classifiers or enhancing the contrast and morphological features: the region of mass, segmentation, and

region of interest (ROI). Several studies have been conducted on contrast enhancement and segmentation, for example, density-weighted contrast enhancement based on the object region growing technique for the extraction of masses,[8] adaptive thresholding followed by a modified Markov random field model,[9] statistics-based enhancement by multilevel thresholding,[10] and region selection.[11] Moreover, several segmentation techniques exist, including region growing, region clustering, template matching, stochastic relaxation, fuzzy techniques, bilateral image subtraction, and multiscale technique.[12-18] The features are obtained from ROI-based mass segmentation regions[19] and are classified using various methods, such as linear discriminant analysis,[20] artificial neural networks,[21] Bayesian networks,[22] and binary decision trees.[23]

As for mass detection systems using deep learning, Domingues and Cardoso[24] attempted ROI-based small patch-wise detection scanning of whole mammogram images using a deep learning model with up to three layers of trained inputs without explanation of the detailed structure, cropping ROI images to achieve a resolution of 32×32 pixels from 116 dataset images containing masses. They achieved an accuracy of 85.9% in the ROI mass differentiation task on their test dataset composed of 25% of their total mass acquisition, with no free receiver operating curve (FROC) results for the detection of masses. Dhungel, et al.[25] applied their proposed state-of-the-art mass detection algorithm using a cascade of deep learning and random forest classifiers for candidate ROI based evaluation, resulting in a true positive rate of 0.96±0.03 at 1.2 false positives per image on the INbreast data and a true positive rate of 0.75 at 4.8 false positives per image on the DDSM-BCRP data, consisting of 195 total images containing masses from the datasets.

In this article, we propose deep-learning methodology with which to enhance the mass differentiation performance of convolutional neural network (CNN)[26] based architecture, and we address weak segmentation exhibited by our method in view of the visualization of detected suspicious masses, along with numerical results, including FROC. To train the mass classifier employed in a CNN, we aimed to examine how scaling intensity (brightness) values in the input images affect the mass classifier performance. An outline of the experimental flow for the intensity scaling treatment of the input images is as follows: first, we trained mass classifying CNN models on input training images with a few steps of variation in their intensity scales in the ROI-based training modes, and then we evaluated each model against test data. Next, we employed the trained mass classifier to demonstrate the mass detection process and compared the efficiencies of the models along with each changing scale in the intensity of the grid images scanned over the entire surface of the sample mammography images. Moreover, we examined the effect of another type of intensity-scaling preprocessing wherein we assessed the detection performance of the model trained with a scaling parameter of one, that is, without scaling the original training data against only the test data of various intensity scales. The potential significance of using this preprocessing technique was supported by our numerical results in the task of mass detection. For additional issues regarding segmentation in our proposed method, we determined whether clusters of suspicious locations of the masses, inferred by the grid-scanning mass classifier, perform as segmentation in a sense by providing approximate macro shapes of the masses.
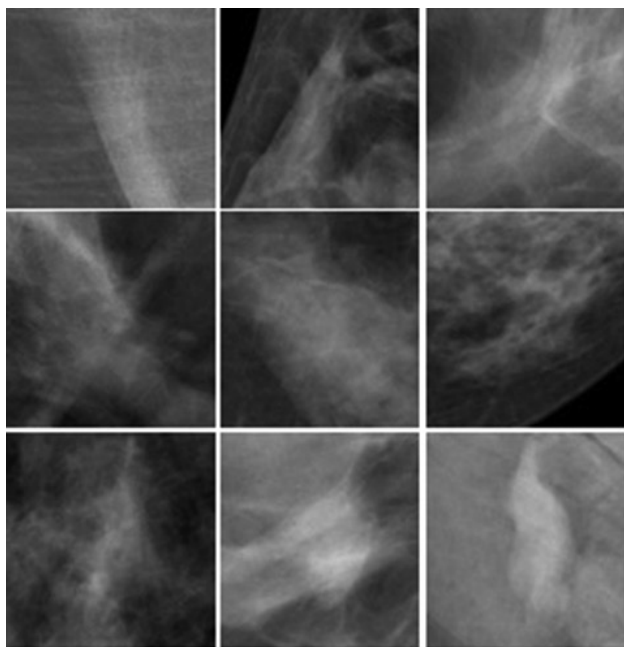
ROI of mass-like normal tissue
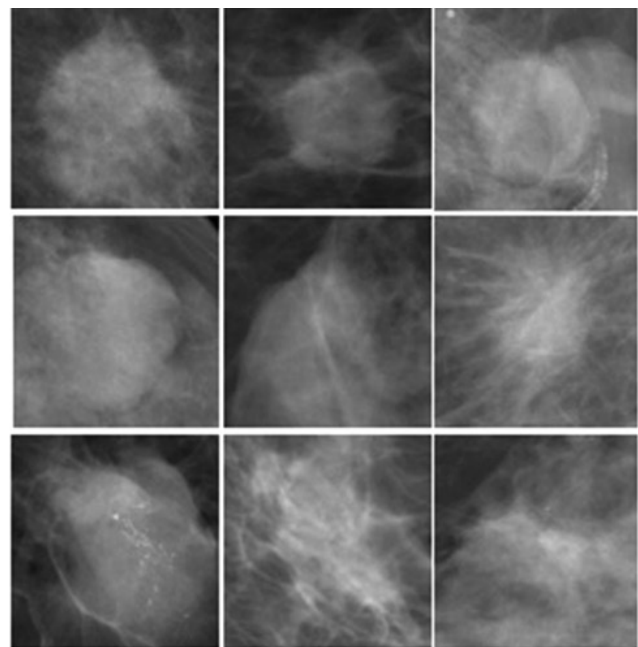
ROI of masses

Fig. 1. Representative masses and mass-like normal breast tissue used to train the models. ROI, region of interest.

In the following sections, we present the organization of our data for use in detection experiments, evaluations, and detection performance evaluation of the employed mass classifiers, discussions of our methodologies, and the conclusions of our study.

MATERIALS AND METHODS

Data acquisition

The Institutional Review Board of Gachon University Gil Medical Center approved (IRB No. GDIRB2016-088) this retrospective study. All methods were performed in accordance with the relevant guidelines and regulations. For our data, we acquired 300 breast mammography images, 281 of which contained masses. There were 146 cases of malignant masses in the mass-containing images. All images were 8-bit grayscale bitmap images. The training dataset for the CNN mass classifier model to detect masses was organized as follows: for the ROI-based CNN inference model, we utilized sample mammography images cropped into ROI patches with a resolution of 128×128, resulting in 281 mass ROI images and 285 normal ROI images of breast tissue. For the normal ROI images, we manually chose mass-like normal tissue images because of the preconception that the accuracy of mass detection might heavily depend on the detector differentiation performance between masses and mass-like normal breast tissue regions (Refer to the images of some of the masses and mass-like normal breast tissue in Fig. 1).

Augmentation in eight types of images by translation

We organized our data by cropping sample mammography images into ROI patches with a resolution of 128×128 to be used in our ROI-based CNN mass classifier. Then, we augmented each patch by shifting it slightly in eight radial directions from the center (Fig. 2). We did not employ any other augmentation methods for the translated augmentation dataset. An argument
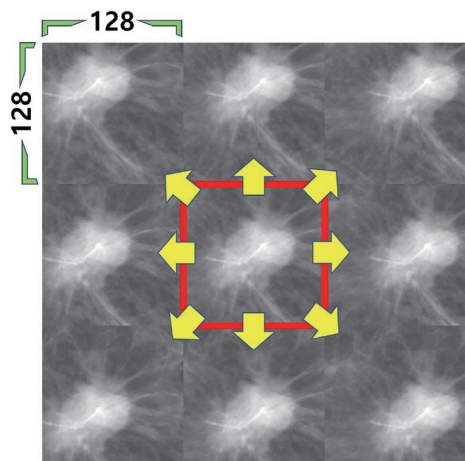
can still be made for the utility of other types of augmentation, such as flipping, being applied to the pre-augmented datasets. The art of augmentation is not the main scope of this study, and we did not compare the functional properties of different augmentation methods in this study. In a previous study[27] and the references therein, one may find issues related to data augmentation and its efficiency.

Training and test datasets

As mentioned in the previous subsection, after data augmentation, we obtained mass training data of 2529 ROI images, non-mass training data of 2565 ROI images, mass test data of 377 ROI images, and non-mass test data of 120 ROI images (Table 1). Our main interest in this study on mass differentiation was to determine how many similar mass lesion shapes are recognizable, rather than how many mass structures that are difficult to train can be classified. Hence, in our approaches employing the trained mass detection model, we tested the performance of a mass detector for test data in which, for example, each test mass image is collected by a slight shift (different location and size) from the original non-augmented mass ROI image found in the training dataset. Nevertheless, the test image (test data) was not included in the training data. In short, test images were obtained by shifting in a different direction on existing mass images in the augmented ROI images (Fig. 3). Moreover, to weigh detection performance more heavily for masses than non-mass tissue, we tested more mass images (377) than non-mass images (120).

Structure of CNN

We applied a VGG-Net[28] architecture to train the datasets, which

Table 1. Configuration of Our Sample Collection of the Mammography ROI Images

	ROI (not augmented)	ROI (augmented)	ROI (shifted + ε)
		Training data	Test data
Mass	281	2529	377
Normal tissue (mass-like)	285	2565	120
Total	566	5094	497

ROI, region of interest.

Fig. 2. Data augmentation by translation to eight radial directions.
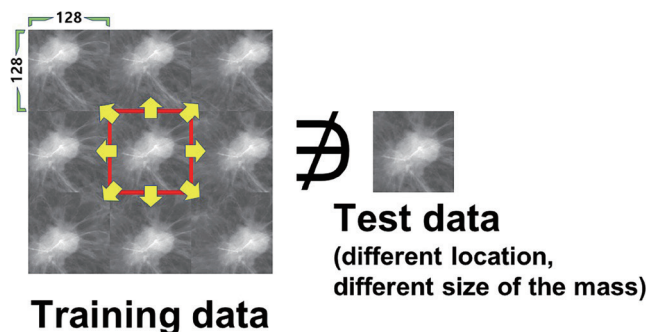

Fig. 3. Training and test data.

comprised ten convolutional layers and four pooling layers. The structure is described in detail in Table 2 and Fig. 4: in Table 2, the symbols m and n represent the size of the convolution kernel for each input channel and the number of whole kernels applied to each layer, respectively.

Training by the scaled intensity of images

To normalize the training data of the CNN mass detector, we applied mean-zero normalization. We then scaled the intensity of the input images. Let a pair of indices (i, j) represent the pixel point in the i-th position on the x-axis and j-th position on the y-axis in each input image. The corresponding pixel value is denoted by u_{ij}; by applying a scaling parameter δ, the scaled normalization v_{ij} for the training data is given as
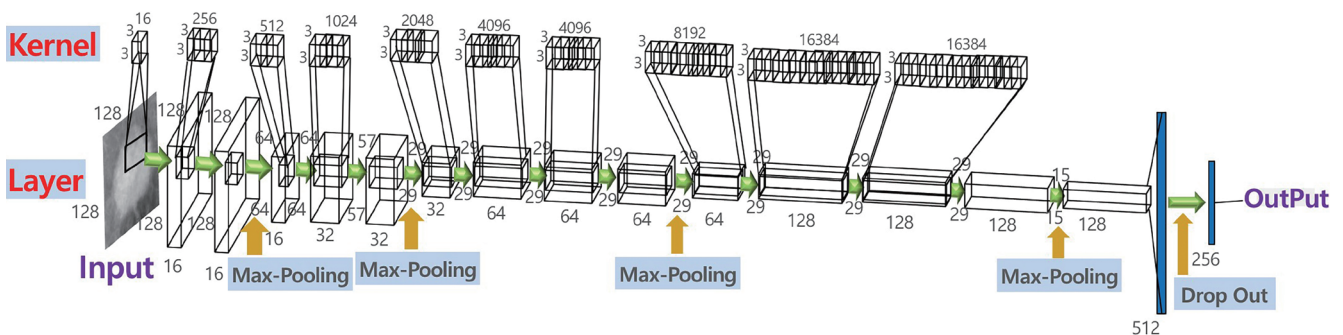
$$v_{ij} = \frac{(u_{ij} - E[u_{ij}])}{\delta}, \qquad (1)$$

where $E[u_{ij}]$ denotes the mean value of u_{ij} at position (i, j).

By defining the candidates for δ, δ∈{1, 2, 5, 10, 15, 20}, we ob-

Table 2. Training Structure of the Convolutional Neural Network (10-Conv, 4-Pool, 2-Fully-Conn Structure)

Layer	(m×m)×n	Activation	
Conv.	(3×3)×16	ReLu	
Conv.	(3×3)×256	ReLu	
Max-Pooling	kernel size: (2×2)		Strides: 2
Conv.	(3×3)×512	ReLu	
Conv.	(3×3)×1024	ReLu	
Max-Pooling	kernel size: (2×2)		Strides: 2
Conv.	(3×3)×2048	ReLu	
Conv.	(3×3)×4096	ReLu	
Conv.	(3×3)×4096	ReLu	
Max-Pooling	kernel size: (2×2)		Strides: 2
Conv.	(3×3)×8192	ReLu	
Conv.	(3×3)×16384	ReLu	
Conv.	(3×3)×16384	ReLu	
Max-Pooling	kernel size: (2×2)		Strides: 2
Fully-Conn.	512	ReLu	
Fully-Conn.	256	ReLu	(Dropout rate: 50%)
Fully-Conn.		Softmax	Output units: 2

tain the numerical results shown in Fig. 5. Here, the results of the area under receiver operating characteristic (AUROC) reveal that the CNN mass detector performance is best when δ=5, when the computational implementation is performed with 400 epochs of training, a batch size of 100, a learning rate of 0.0001, and a dropout rate of 0.5, using a standard back-propagation algorithm.[26,29,30] These parameters reflect our experience gained performing experiments to enhance the performance of the CNN architecture during this study. From these results, we attempted to evaluate the efficiency of the scaled normalization treatment using the δ parameter defined in (1).

Outline of the process of mass detection

The entire process of our mass detection examination is listed below. Fig. 6 illustrates the proposed methodology of mass detection and segmentation.

1) We subsampled the ROI patches to input into the CNN mass detector. Some arbitrary window sizing defines the cropped patches, ranging from 100 to 150 ±ε, with a constant increment of the cropping centers. Here, we set an increment of 50 pixels to move the cropping centers.

2) We resized all the subsampled patches to a resolution of 128×128, aligning for each target mammography image to determine whether there are masses and where the masses are located.

3) For the CNN models of the detector trained with δ=1, 5, we examined the detection rate and the rate of false detection per image, which resulted in the FROC for 150 types of malignant masses contained in the target mammography images.

To measure the success of correct detection, we defined a metric to approximate the diameter of each mass. If a point included in a cluster of detection points fell within the mass diameter, it counted as one instance of true detection without the allowance of multiple counts for points in the same cluster.

RESULTS

Detection results

Here, we present the FROC results of mass detection with our model. To evaluate the detection rate, we first had to establish a

Fig. 4. Convolutional neural network training architecture with 10-conv, 4-pool, and 2-fully-conn. Networks corresponding to the structure in Table 2.

standard scheme for clustering the points of detection to discriminate and identify closely detected points at the centers of the ROI patches examined by the scanning mass detector. In our experiment, we employed the well-known DBSCAN algorithm[31] to cluster some nearly detected suspicious mass points. We define a metric variable η by which the DBSCAN algorithm clusters the detection points located within a distance η from each other (refer to Fig. 7, where we employed the MATLAB code of DBSCAN given in a previous study[32]). Additionally, by parameterizing η as $[\frac{\eta}{5}, \frac{\eta}{10}, \frac{\eta}{15}, \frac{\eta}{20}, \frac{\eta}{25}, \frac{\eta}{30}]$, we estimate each FROC performance. In addition, we manually removed some falsely detected clusters of points made by DBSACN, located near the breast
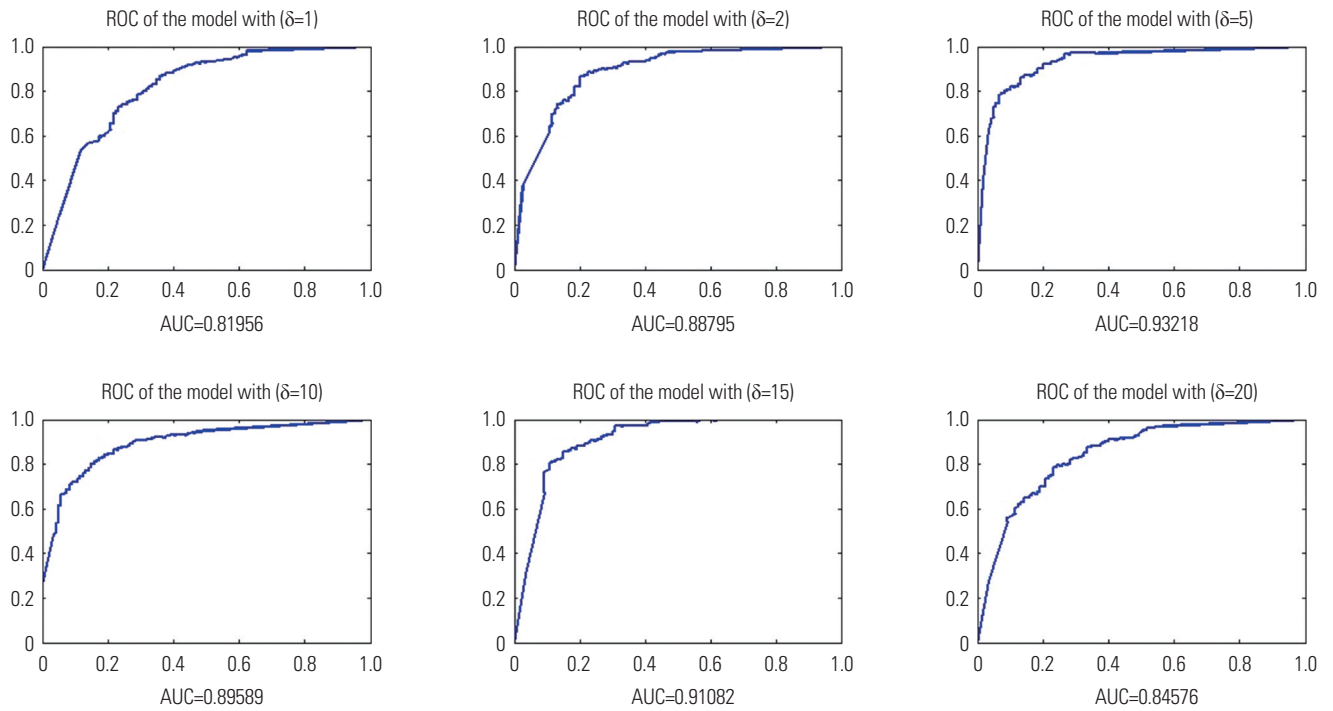

Fig. 5. ROC and AUC plots of the results, for δ∈{1, 2, 5, 10, 15, 20}. ROC, receiver operating characteristic; AUC, area under the curve.
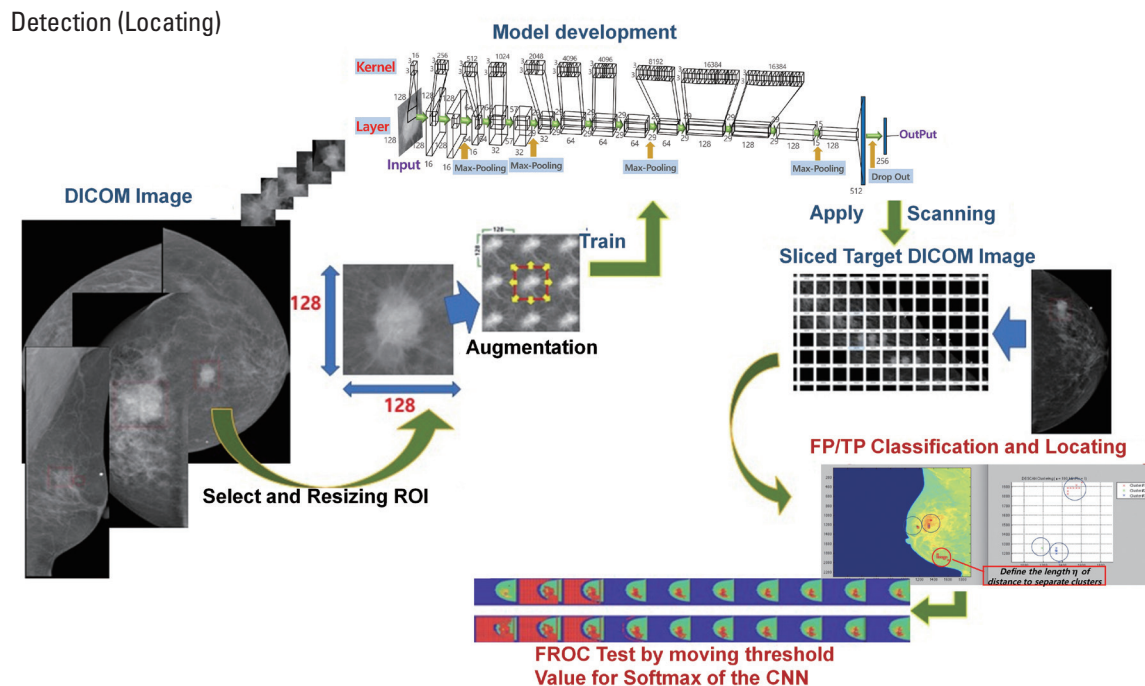

Fig. 6. Scheme of the mass detection and segmentation. ROI, region of interest; FROC, free receiver operating curve; CNN, convolutional neural network.

and nipple area boundary, which affects the total FROC results depending on η. The detection results of the mass detector trained with δ=1, 5 are shown in Fig. 8 (for δ=1) and Fig. 9 (for δ=5).

Re-estimation by pseudo scaling normalization

Along with the results shown in Figs. 8 and 9 (in the previous subsection), the FROC performance was enhanced by increasing the parameter η. However, false detection rates persisted to some degree, as it is difficult to control the dispersion of the false detection rate by the values of the thresholds splitting the Softmax classifier output values of the CNN mass detector while calculating the FROC.

We applied the so-called pseudo scaling normalization, wherein we perform scaling normalization only on the test data and not on the training dataset. This means that we trained the CNN mass detector with δ=1 and tested it for all δ∈{1, 2, 5, 10, 15,

20}. The numerical results of the primary classification performance of these models are shown in Fig. 10. As shown in Fig. 10, as the value of δ increases, the classification performance increased. In Fig. 11, we also present the FROC results of mass detection for δ=15: the results are indistinguishable from those of other values of δ, such as δ=10, 20.

DISCUSSION

In our study of mass detection in gray-scale mammograms, we categorized ROI images into those with masses and those with normal breast tissue to train a CNN model to differentiate masses, diagnosing those correctly classified images determined by medical experts in mammogram radiology to have masses as true positives. The deep-learning-based mass classifier screened
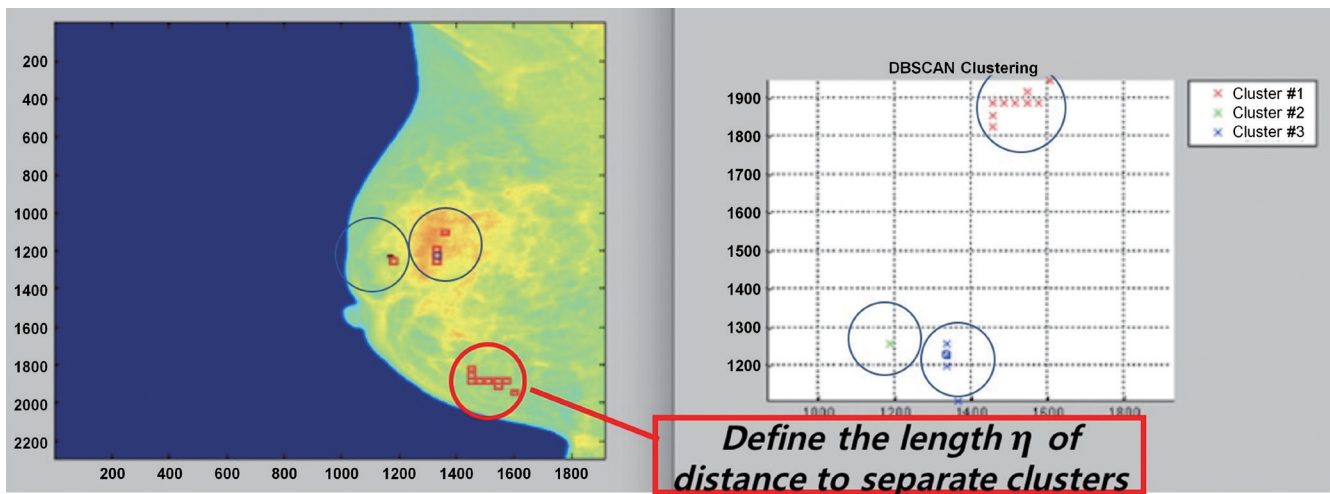

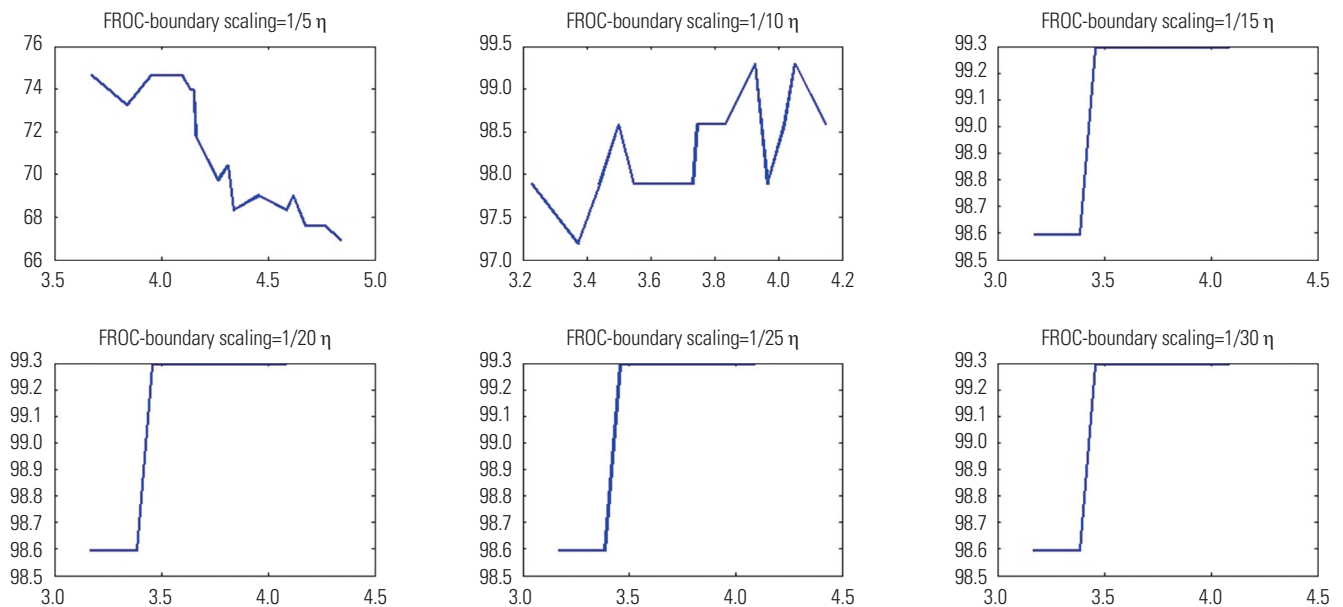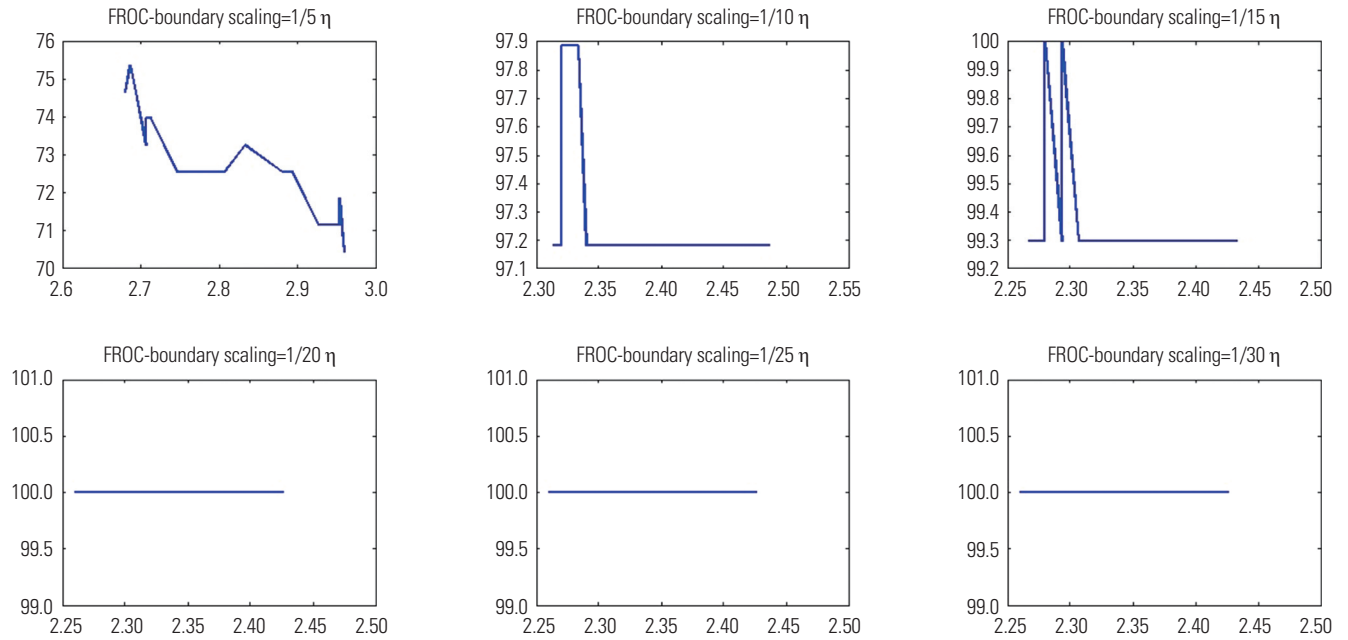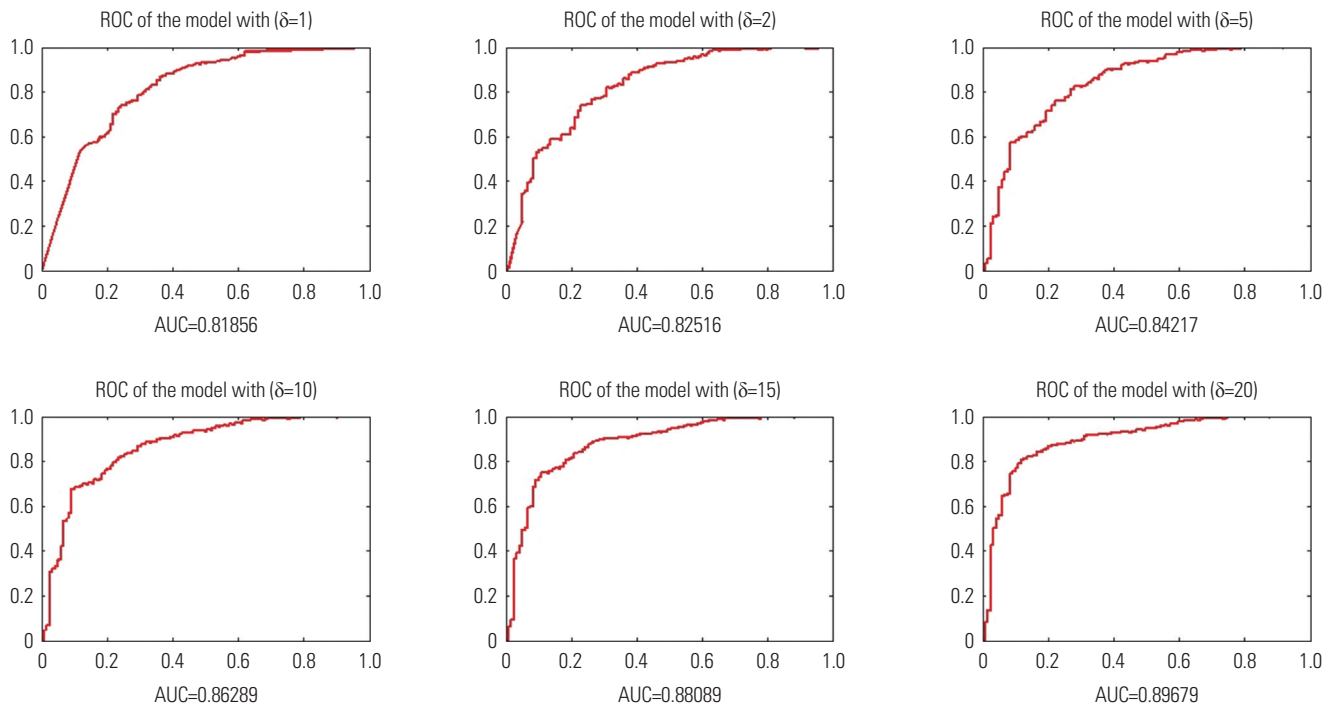**Fig. 7.** Clustering the detection points using DBSCAN to define the separating length η, with MATLAB programming.

Fig. 8. FROC plots of the results, for δ=1. FROC, free receiver operating curve.

Fig. 9. FROC plots of the results, for δ=5. FROC, free receiver operating curve.

Fig. 10. ROC and AUC plots of the results of pseudo scaling normalization, for δ∈{1, 2, 5, 10, 15, 20}. ROC, receiver operating characteristic; AUC, area under the curve.

the entirety of mammography images, and subsampled test ROI images were successively input into the mass classifier. The surviving input test ROI images were checked as suspicious mass detection points. We calculated the performance of our detection models by counting the clustering of near points. For clustering, we applied the well-known DBSCAN algorithm, which clusters points located within a predefined distance from each

nearest neighborhood point. Our mass detection experiments were conducted to identify masses with high performance in terms of the FROC, even though the actual training data for the mass classifier contained the perturbed mass ROI images found in the tested full DICOM mammography images. The reason for this experimental goal was based on the natural conjecture that, if our proposed method provides fairly clear detection or
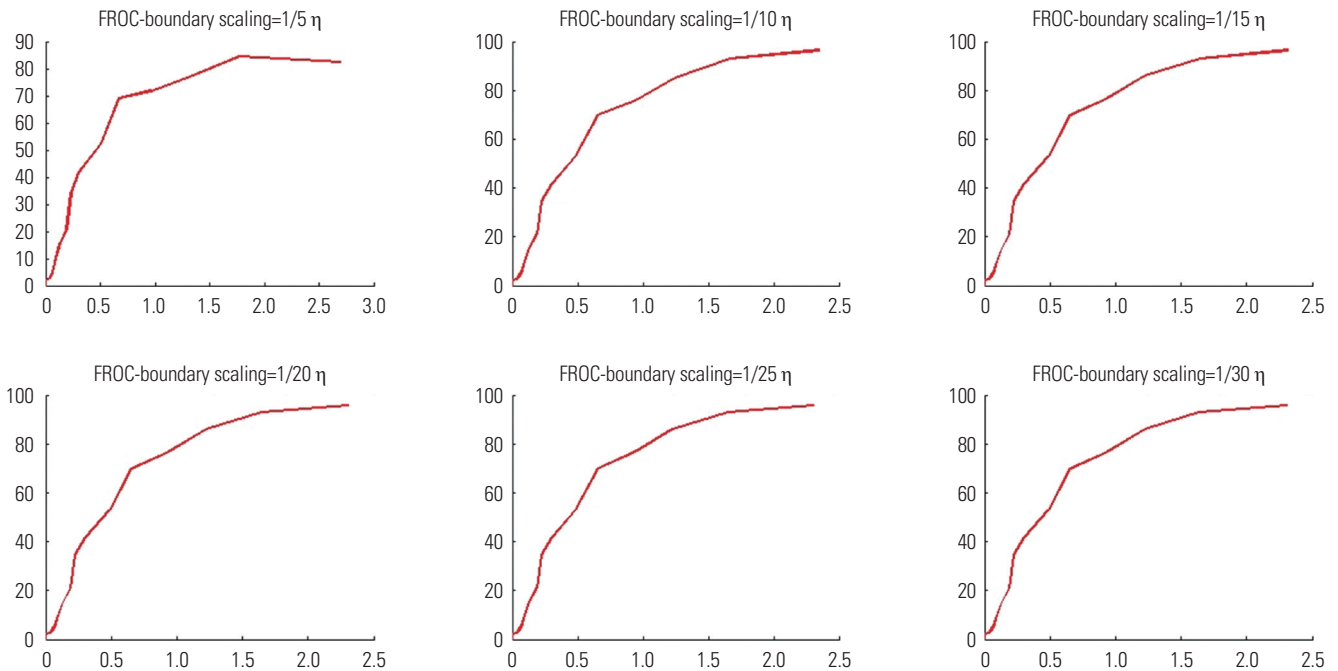
Fig. 11. FROC plots of the results of pseudo scaling normalization, for δ=15. FROC, free receiver operating curve.
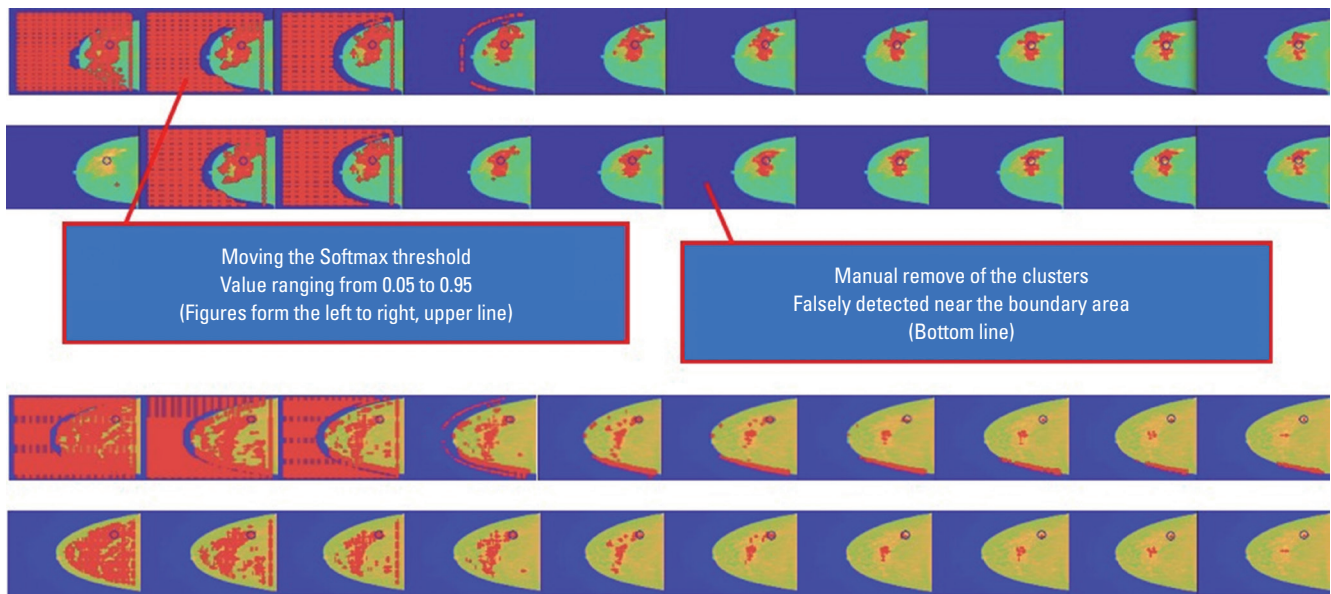

Fig. 12. Observing the segmentation property of the proposed mass detection scheme by increasing the threshold value of the output of the Softmax classifier of the convolutional neural network mass detection model.

segmentation results derived from an ROI classifier trained on small ROI images describing a specific, limited area of the original mammography, in the case of a given ideally big dataset with all characterizing properties available, such as macro shapes and image contrasts of masses, similar technology would be highly applicable with acceptable performance.

Our proposed method performed well in comparison to existing models. In the study of Ertosun and Rubin,[33] masses were detected using a deep learning-based classification engine and a localization engine, which showed an accuracy of 85% and a false positive of 0.9. Li, et al.[34] attempted to detect masses using an end-to-end network combining a Siamese-Faster-RCNN network and a region proposal network and Siamese-FC (Siamese Full Connected) networks, resulting in a true positive rate of 88% and false positives per image of 1.12. The study by Domingues and Cardoso[24] describes a similar method to our proposed method in that it used a patch-wise detection method, with an accuracy of 85.9%, compared to an area under the curve of 89.65 for our method. Furthermore, our proposed method demonstrated that even weak divisions were possible.

In regards to the ROI classifier, we employed scaling to the intensity (brightness) of the input images. The first version of the scaling method trains the classifier model using intensity scaled training data. The second version scales the input images for the test data, but not for the training data, except for δ=1 in (1). The reason we employ this parameter scaling during the process of normalization is related to the slopes of the activation function in the first layer of the CNN architecture. Let us define the convolutional operator Â, input resolution matrix X̂, output of the convolution Ŷ, and rectified linear unit (ReLu) activation function σ. Then, because the convolutional operator Â is linear, the following holds:

$$\frac{1}{\delta}\sigma(\hat{Y})=\sigma(\frac{1}{\delta}\hat{Y})=\sigma(\frac{1}{\delta}\hat{A}[\hat{X}])=\sigma(\hat{A}[\frac{1}{\delta}\hat{X}]). \qquad (2)$$

Hence, by scaling the brightness of the input images, we aimed to adjust the slope of the ReLu activation function of the first layer to perform the inference efficiently.

Both versions of the mass classifier displayed a difference in performance when evaluated for applying the efficient ROI patch discriminator to mass detection by screening the whole mammography images patch by patch. The first model (the model scaling the training input data) had better AUROC values than the other (the model scaling the test input data but not the training data). Nevertheless, the first mass classifier model failed to activate canonical shapes of the curvatures in the FROC plots. In contrast, the latter activated many typical shapes with significantly increased accuracy in the detection rate per false positive per image. These results indicated that a higher performance for an ROI-based test for differentiation does not always mean better activation in real-world detection tasks in mammogram CAD equipped with our proposed types of bare deep-learning models. In addition, such numerical results demonstrate that simply scaling input data when testing a learned model can make a remarkable difference in the performance of a given model, depending only on the types of datasets. We have still not provided a theoretical proof of why scaling the intensity of test data, but not training data, improves performance. We only suppose that the controversial issues of data normalization are deeply associated with it. Accordingly, data blurring or smoothing preprocessing methods can be used as core mechanics of scaling normalizing treatment to suit highly complex macro shapes and structures of mammogram masses. Suppose this empirical deduction is reasonable for datasets with highly complex macrostructures. In this case, one may counterintuitively blur the input datasets to be tested to obtain somewhat improved numerical results. Altogether, if two opposing datasets are given to be differentiated from each other and both possess rather obscure macro structures, both datasets can be made more distinctive from each other by blurring during preprocessing so that the trained model can even more easily differentiate between positive and negative instances with improved accuracy, compared to before blurring. Based on the idea of blur-

ring, we intended to accentuate the brightness of input images so that benign non-mass tissues in mammography look like smooth, plateaued normal tissue organizations, compared to significantly thicker and brighter mass lesions, to the trained ROI patch classifier. The second version of our employed blurring pre-process treatment, which scaled the brightness of the test input datasets without scaling the training data, displayed acceptable performance.

Several studies have addressed segmentation issues.[33,35-37] Therein, the authors implemented visualizing techniques using activation maps of the trained convolution layer of the CNN. However, none of the studies proposed a method of mass segmentation by applying detected points. In our study, we found that our proposed model of mass detection would be useful as a segmentation tool in terms of segmentation with points of detection of masses. As can be seen in Fig. 12, by increasing the threshold value of the output of the Softmax classifier in the trained CNN mass detection model, the clusters consisting of mass detection points appear to weakly describe an approximate outline of each mass shape, increasing the visual definition. We used a CNN with a basic structure; however, a variety of deeper and better-performing CNN architectures continue to be released. Therefore, it is necessary to apply our proposed method to various CNN structures and compare their performances, something that is planned through additional research in the future.

In our study of a CNN-based mass detection and segmentation approach, we determined that the characteristics of the surrounding pixel intensity of masses and their fine-scale structural similarity can significantly affect the performance of our proposed method as a segmentation-like detector. Here, for mass detection considering weak segmentation, the training data were composed of augmented mass ROI images obtained by translation in eight radial directions. Numerical results indicated that the clustering of the detection points resembles a description of the characteristics of the macro shapes of detected masses in terms of segmentation. The brightness of the input images was scaled down to some extent, and the full definition of the actual mass regions was enhanced. These results support the potential of our proposed patch-wise detection method to be utilized as a mass detection and segmentation tool.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

Conceptualization: Young Jae Kim and Kwang Gi Kim. Data curation: Young Jae Kim. Formal analysis: Young Jae Kim and Kwang Gi Kim. Funding acquisition: Young Jae Kim. Investigation: Young Jae Kim and Kwang Gi Kim. Methodology: Young Jae Kim and Kwang Gi Kim. Project administration: Kwang Gi Kim. Resources: Young Jae Kim and Kwang Gi Kim. Software: Young Jae Kim. Supervision: Kwang Gi Kim. Validation: Young Jae Kim and Kwang Gi Kim. Visualization: Young Jae Kim. Writing—original draft: Young Jae Kim. Writing—review & editing: Young Jae Kim and Kwang Gi Kim. Approval of final manuscript: all authors.

ORCID iDs

Young Jae Kim   https://orcid.org/0000-0003-0443-0051
Kwang Gi Kim   https://orcid.org/0000-0001-9714-6038

REFERENCES

1. World Health Organization. Cancer [Internet] [accessed on 2021 July 26]. Available at: http://www.who.int/mediacentre/factsheets/fs297/en/.
2. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359-86.
3. Cheng HD, Cai X, Chen X, Hu L, Lou X. Computer-aided detection and classification of microcalcifications in mammograms: a survey. Pattern Recognit 2003;36:2967-91.
4. Cheng HD, Shi XJ, Min R, Hu LM, Cai XP, Du HN. Approaches for automated detection and classification of masses in mammograms. Pattern Recognit 2006;39:646-68.
5. Zheng B, Chang YH, Wang XH, Good WF. Comparison of artificial neural network and Bayesian belief network in a computerassisted diagnosis scheme for mammography. Proceedings of the IJCNN'99. International Joint Conference on Neural Networks (Cat. No.99CH36339); 1999 Jul 10-16; Washington DC, USA: IEEE; 1999. p.4181-5.
6. Egan RL, Mosteller RC. Breast cancer mammography patterns. Cancer 1977;40:2087-90.
7. Wolfe JN. Breast patterns as an index of risk for developing breast cancer. AJR Am J Roentgenol 1976;126:1130-7.
8. Petrick N, Chan HP, Sahiner B, Wei D. An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection. IEEE Trans Med Imaging 1996;15:59-67.
9. Li HD, Kallergi M, Clarke LP, Jain VK, Clark RA. Markov random field for tumor detection in digital mammography. IEEE Trans Med Imaging 1995;14:565-76.
10. Rojas Domínguez A, Nandi AK. Detection of masses in mammograms via statistically based enhancement, multilevel-thresholding segmentation, and region selection. Comput Med Imaging Graph 2008;32:304-15.
11. Martinez VG, Gamo DM, Rios J, Vilarrasa A. Iterative method for automatic detection of masses in digital mammograms for computer-aided diagnosis. Proceedings of the Medical Imaging 1999: Image Processing; 1999 May 21; San Diego, CA, USA: SPIE;1999. p.1086-93.
12. Kallergi M, Woods K, Clarke LP, Qian W, Clark RA. Image segmentation in digital mammography: comparison of local thresholding and region growing algorithms. Comput Med Imaging Graph 1992;

16:323-31.
13. Pappas TN, Jayant NS. An adaptive clustering algorithm for image segmentation. Proc IEEE Int Conf Acoust Speech Signal Process 1989;3:1667-70.
14. Lai SM, Li X, Biscof WF. On techniques for detecting circumscribed masses in mammograms. IEEE Trans Med Imaging 1989;8:377-86.
15. Li H, Liu KR, Wang Y, Lo SCB. Nonlinear filtering enhancement and histogram modeling segmentation of masses for digital mammograms. Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 1996 Nov 3; Amsterdam, the Netherlands: IEEE; 1996. p.1045-6.
16. Sameti M, Ward RK. A fuzzy segmentation algorithm for mammogram partitioning. In: K. Doi M.L. Giger, R.M. Nishikawa, R.A. Schmidt, editors. Digital mammography. Amsterdam: Elsevier; 1996. p.471-4.
17. Zheng B, Chang YH, Gur D. Computerized detection of masses from digitized mammograms: comparison of single-image segmentation and bilateral-image subtraction. Acad Radiol 1995;2: 1056-61.
18. Brake GM, Karssemeijer N. Single and multiscale detection of masses in digital mammograms. IEEE Trans Med Imaging 1999;18: 628-39.
19. Undrill P, Gupta R, Henry S, Downing M. Texture analysis and boundary refinement to outline mammography masses. Proceedings of the IEE Colloquium on Digital Mammography; 1996 Mar 27; London, UK: IEE; 1996. p.5/1-6.
20. Petrick N, Chan HP, Wei D, Sahiner B, Helvie MA, Adler DD. Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification. Med Phys 1996;23:1685-96.
21. Bovis K, Singh S, Fieldsend J, Pinder C. Identification of masses in digital mammograms with MLP and RBF nets. Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium; 2000 Jul 27; Como, Italy: IEEE; 2000. p.342-7.
22. Viton JL, Rasigni M, Rasigni G, Llebaria A. Method for characterizing masses in digital mammograms. Opt Eng 1996;35:3453-9.
23. Brzakovic D, Luo XM, Brzakovic P. An approach to automated detection of tumors in mammograms. IEEE Trans Med Imaging 1990; 9:233-41.
24. Domingues I, Cardoso J. Mass detection on mammogram images: a first assessment of deep learning techniques. Proceedings of the 19th Portuguese Conference on Pattern Recognition (RECPAD); 2013 Nov 1; Lisbon, Portugal: APRP; 2013.
25. Dhungel N, Carneiro G, Bradley AP. Automated mass detection in mammograms using cascaded deep learning and random forests. Proceedings of the 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA); 2015 Nov 23-25; Adelaide, Australia; 2015. p.1-8.
26. LeCun YA, Bottou L, Orr GB, Müller KR. Efficient backprop. In: Montavon G, Orr G, Müller KR, editors. Neural networks: tricks of the trade. 2nd ed. Berlin: Springer; 2012. p.9-48.
27. Fawzi A, Samulowitz H, Turaga D, Frossard P. Adaptive data augmentation for image classification. Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP 2016); 2016 Sep 25-28; Phoenix, AZ, USA; IEEE; 2016. p.3688-92.
28. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv [Preprint]. 2014 [accessed on 2021 July 30]. Available at: https://arxiv.org/abs/1409.1556.
29. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE 1998; 86:2278-324.

30. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from over-fitting. J Mach Learn Res 2014;15:1929-58.

31. Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96); 1996 Aug 2-4; Portland, OR, USA: Association for the Advancement of Artificial Intelligence; 1996. p.226-31.

32. Yarpiz. DBSCAN clustering in MATLAB [Internet] [accessed on 2021 July 30]. Available at: http://yarpiz.com/255/ypml110-db-scan-clustering.

33. Ertosun MG, Rubin DL. Probabilistic visual search for masses within mammography images using deep learning. Proceedings of the 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2015 Nov 9-12; Washington DC, USA; IEEE; 2015. p.1310-5.

34. Li Y, Zhang L, Chen H, Cheng L. Mass detection in mammograms by bilateral analysis using convolution neural network. Comput Methods Programs Biomed 2020;195:105518.

35. Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, editors. Computer Vision – ECCV 2014. Cham: Springer; 2014. p.818-33.

36. Yosinski J, Clune J, Nguyen A, Fuchs T, Lipson H. Understanding neural networks through deep visualization. arXiv [Preprint]. 2015 [accessed on 2021 July 26]. Available at: https://arxiv.org/abs/1506.06579.

37. Geras KJ, Wolfson S, Shen Y, Wu N, Kim S, Kim E, et al. High-resolution breast cancer screening with multi-view deep convolutional neural networks. arXiv [Preprint]. 2017 [accessed on 2021 July 26]. Available at: https://arxiv.org/abs/1703.07047.