IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 9 February 2023, accepted 14 April 2023, date of publication 26 April 2023, date of current version 24 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3270562

RESEARCH ARTICLE

Regression to Classification: Ordinal Prediction of Calcified Vessels Using Customized ResNet50

HOSNA ASMA-ULL[1], IL DONG YUN[1], (Member, IEEE), AND BO LA YUN[2,3]

[1]Department of Computer and Electronic Systems Engineering, Hankuk University of Foreign Studies, Yongin 17035, South Korea
[2]Department of Radiology, Seoul National University Bundang Hospital, Seongnam 13620, South Korea
[3]Department of Radiology, Seoul National University College of Medicine, Seoul 08826, Republic of Korea

Corresponding author: Il Dong Yun (yun@hufs.ac.kr)

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science, Technology under Grant 2019R1A2C1085113.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Seoul National University Bundang Hospital Institutional Review Board under Approval No.: B-1905-541-104.

ABSTRACT A substantial percentage of women die from cardiovascular disease (CVD). Computed tomography (CT) scan helps predict/monitor CVD-related diseases. However, previous studies found relation to assessing the risk of CVD by estimating severity of breast arterial calcification (BACs). In many countries, women undergo breast cancer screening using mammograms regularly, so it is easier to access vast datasets compared to CT-scan without additional cost and absorbing radiation. Earlier, measuring BACs severity was largely depends on the performance of segmentation of calcified vessel. Nevertheless, dataset maintenance for segmentation is expensive and tedious. As pixel-wise manual ground-truth (GT) marking can only be performed by multiple experienced experts. First time ever, to the best of our knowledge, a direct ordinal measurement of BACs severity is proposed without prior segmentation. Therefore, due to the complexity of calcified vessels, cascade classification performs better than end-to-end classification. In this proposed method, first, we imply Frangi filter to deduct background automatically without any prior ground-truth marking. Then, unlike other methods, interpolation is implied to shrink cropped images rather than splitting images. Also, left view images have flipped to reduce the complexity of different directions of breasts. In addition, among various residual network (ResNet) architecture, proposed method customized ResNet50 to adjust large image size and vessel complexity. The proposed method achieve, best AUC-score compare to state-of-art, above 90 and 85 percent for Normal vs Abnormal and Minor vs Severe patients, respectively. The proposed method reduces cost, time, and tedious job of researchers as this does not require segmenting calcified vessels and provides efficient prediction of BACs severity which is helpful to assess the risk of CVD in long run.

INDEX TERMS Deep learning, medical image, mammography, breast arterial calcification, coronary artery disease, cascade, ResNet.

I. INTRODUCTION

Atherosclerotic cardiovascular disease (ASCVD) is the leading cause of mortality rate of males and females with 16% which followed by Stroke and chronic obstructive pulmonary disease are the $2^{(nd)}$ and $3^{(rd)}$ leading causes of death, responsible for approximately 11% and 6% of total deaths respectively. Moreover, the largest increase for death rising from 2 million to 8.9 million for atherosclerotic cardiovascular disease, since 2000 [1]. Although atherosclerotic cardiovascular disease (ASCVD) risk prediction algorithms, such as the pooled cohort equation (PCE), play an important role in identifying high-risk patients, many cardiovascular events also occur in women at low ASCVD risk [19], [49].

The associate editor coordinating the review of this manuscript and approving it for publication was Ziyan Wu.

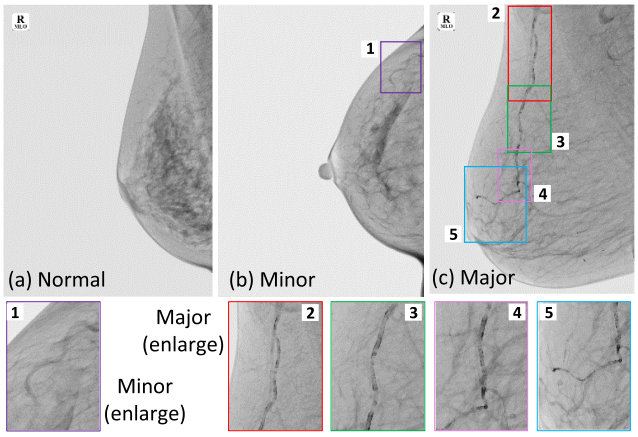


FIGURE 1. Different severity levels of patients. Illustration of 3 different levels of severity of patient using craniocaudal (CC) view of images. Fig. 1 (a) shows the mammogram of normal patients, and Fig. 1 (b) and 1 (c) show the abnormal patients of minor and severe classes. Fig. 1 (c) shows the severe patient who needs immediate attention. However, the enlarged calcified vessels of severe and minor classes are shown in last row of figure, respectively.

Though radiologists have been aware of the BACs phenomenon for a long time, the clinical utility was unknown until Baum et a. [6] first published clinical research on the significance of BACs in 1980. Since then, it has been discovered that BACs on mammograms can serve as valuable indicators of diseases like CVD [10], [31]. The majority of these studies used multivariate analysis to identify a significant relationship between BACs and both coronary artery calcium and coronary artery plaque. Breast arterial calcifications (BACs) develops along the artery wall in a straight line [8]. The American Cancer Society advises women who are 40 years of age or older to get annual mammography for the screening of breast cancer. [44]. Around 40 million mammograms are performed yearly in the USA alone [17]. Mammography is a well-accepted breast cancer screening method. Determining the severity level of breast arterial calcification (BACs) for risk assessment of cardiovascular disease, therefore, is an effective and widely applicable way to gather a large number of cases without any additional cost and excessive radiation.

BACs detection is considered as being challenging, despite the fact that deep learning models have demonstrated promise in a number of different domains. Due to the fluctuating shape, location, and contrast of calcified vessels. In addition, marking and maintaining the ground truth (GT) for segmentation is particularly expensive. Only highly experienced experts can mark the pixel-wise GT which is very tedious job. Due to the complexity of the dataset, the marking process requires participation of different expert opinions. We believe that using a multistage procedure is less effective than one-stage assessing the BAC severity levels. In this study, we proposed a direct ordinal prediction to assess the level of BAC severity using cascade with a customized ResNet50. To predict the severity of patients, this approach does not segment calcified vessels. To the best of our knowledge, this proposed method is the first of its type to categorize BACs scores without segmenting the calcified vessels. Additionally, this technique abolishes the necessity for prior pixel-to-pixel ground-truth labeling. The severity score of BACs defines by range from $0 - 12$. The proposed approach turns the traditional regression-based problem into classification by categorizing this range into well-known three categories (normal, minor, and severe) for ordinal prediction without reducing the importance of outcome of patients health. Since vessel changes widely, it is particularly difficult to stabilize depth networks for larger images with high accuracy. This suggested technique, therefore, use Frangi filter to lessen the background and imply Bicubic interpolation to shrink image size. Flipping the left view images for mediolateral oblique (MLO) and cranial-caudal (CC) breast directions is also introduced to reduce the number of possible breast orientations. Afterward, the ResNet50 network is modified to adjust the larger image size and intricate structure of the calcified vessel using cascade approach.

- We propose, first ever, classification of the severity of calcified vessel without segmentation and abolish the necessity of pixel-to-pixel ground-truth marking. Unlike previous 2-steps method, proposed method use entire image for direct ordinal prediction.
- The proposed method analyzes various ResNet and customizes ResNet50 to yield the best performance for prediction.

The remainder of the paper is organized as follows. Section II presents the related work of deep learning in medical arena, previous engineering and clinical method for estimating severity of BAC, and cascade approach in detail. Section III goes into detail on the preprocessing with the Frangi filter to remove background, proposed customized ResNet50 architecture, objective function, and ROC curve & AUC score calculation for cascade. In section IV, we describe the results of our experimental analysis, a comparison with previous approaches. This section also included the findings (advantages and shortcomings) of the proposed method. The conclusion is outlined in Section V, which also incorporates new concepts and ideas that will be used going forward.

II. RELATED WORK

The related work section subdivided into multiple parts. The relation between CAD and BAC, previous method of BAC is elaborated on separate subsections. Additionally, 2-tables summaries the key features of engineering and clinical-based paper. Previous engineering-based (Table 1) papers focus on findings based on experiment, and clinical-based (Table 2) focus on findings based on observation.

A. RELATION OF BAC AND CAD

Radiologists have long been aware of the BACs phenomena, but it wasn't until Baum et al.'s 1980 [6] publication of clinical study on the relevance of BACs that the clinical utility of the phenomenon was made known. Since then, it has been found that BACs on mammograms can be useful

TABLE 1. Summarize Deep Learning (DL) related works, methods, dataset, image type, 2-step outcome to estimate severity of breast calcification (engineering field based journal & transaction).

Year	Authors	Dataset (GT)	Method	Class	Type	Test	Size	Outcome
2017	**Wang** [47]	Private (GT:Private)	CNN	2	Split	33%	95 × 95	Segmentation†& estimate severity
2019	**Wang** [48]	Private (GT:Private)	Multiple‡	1 (BAC)	Crop	20%	Varies	Segmentation†& estimate severity
2020	**ALGhamdi** [4]	Public (GT:Private)	DU-Net	2	Shrink	20%	640 × 640	Segmentation†& estimate severity

Segmentation†= Calcified vessel segmentation. Multiple‡= YOLO, U-net and DeepLabV3+, Hessian-based multiscale filter.

Findings : All experiments require 2-step method to estimate severity of BAC & performance depends on segmentation's outcome. Maintenance of pixel-to-pixel GT marking is costly & tedious. Marking require involvement of multiple experts.

TABLE 2. Summary Clinical based paper's related work from various perspective (i.e. class, Breast arterial calcification (BAC) score range, observation) shown in below table.

Year	Authors	Dataset (GT)	Class	BAC score	Observation
2015	**Karm** [24]	Private (GT:Private)	2	0, 1∼12	BAC†Severity
2015	**Abou** [2]	Private (GT:Private)	2	0, 1∼12	BAC†Severity
2016	**Chadashvili** [10]	Private (GT:Private)	2	0, 1∼12	BAC†Severity
2016	**Csafak** [40]	Private (GT:Private)	2	0, 1∼12	BAC†Severity
2014	**Moradi** [32]	Private (GT:Private)	3	–	Cardiac Disease risk
2018	**Suh** [45]	Private (GT:Private)	3	0, 1∼6, 7∼12	BAC†Severity
2018	**Chun** [14]	Private (GT:Private)	3	0, 1∼6, 7∼12	BC†Severity
2015	**Newallo** [36]	Private (GT:Private)	4	–	Cardiac Disease risk
2015	**Hendriks** [23]	Private (GT:Private)	4	–	Cardiac Disease risk
2016	**Margolies** [28]	Private (GT:Private)	4	0, 1, 2∼3, 4∼12	BAC†Severity
2018	**Kelly** [25]	Private (GT:Private)	4	–	Cardiac Disease risk

BAC†= Breast arterial calcification.

diagnostic tools for conditions like CVD [10], [31]. To see how successfully can use the severity of BACs to determine CAD, some earlier studies attempted to demonstrate the relationship between BACs and CAD risk variables. Authors also revealed that there is affirmed relationship between BACs and Coronary Artery Calcification [29]. Most studies back up the idea of relation of CAD and existence of BACs. Though some cases it could be severe and other cases may be minor. Mostafavi et al. [34] made an effort to assess the effectiveness of logistic regression for patient CAD prediction. According to the findings, existence of BAC detection of 12 patients, among 10 patients also determined as CAD patients. Identical to this research, another author [23] performed a meta-analysis and comprehensive assessment of the literature found a link between BAC levels and cardiovascular events. Table 2 shows the summary of various clinical journals regarding relation of BACs and CAD of recent years. Authors are focus on the relation of BACs and CAD with their observation.

B. DEEP LEARNING BASED WORK IN MEDICAL ARENA

However, medical image analysis uses deep learning techniques, which have grown in popularity recently [27], [35], [41]. Examples include the segmentation of brain CT images [21], fundus images [3], low-resolution breast mammograms [47], segmentation in electron microscopy images for membrane [15], metastatic breast cancer's identification automatically [46] and identify lung nodule using CT [42], segmentation in retina images of blood vessels [43], cancer in skin and lung [5]. Inspiring by the success of deep learning in medical area, researchers shows interest to detect the BACs using deep learning. Table 1 shows the summary of recent engineering based paper's various perspective of BACs detection. Result of these methods compared with proposed method at experiment section.

C. PREVIOUS BAC AND CASCADE WORK

Breast Arterial Calcifications might appear as extended routes or short, erratic courses, but they can also vary greatly in size and shape [18]. The effectiveness of BACs identification has been the subject of recent studies. Ge et al. [18] employed k-segments to identify BACs using an image filtering clustering approach. Cheng et al. suggested using a two-step method to find the calcification to locate vessels [11], [12], [13]. According to Moordang et al. [33], BACs have been found in mammography using a multi-stage method to diagnose breast cancer. Wang et al. [47] have discovered BACs tried by 12-layer CNN by using patch-based model vessel segmentation. The tissue region identify based on the largest linked area with higher value than mean value of entire image. In another work, [4], the U-Net model that have dense connections, employed to identify BACs automatically in mammograms. Performance of other segmentation including sementics using U-Net and Dense-Net, served as the inspiration for this investigation. This approach integrates connections from both short and long skips. At the end of the long skip connections, it employed a summing process. The automatic identification of BACs in mammography remains a difficult task despite earlier successful findings in numerous domains. The BAC's variable size, shape, and contrast are

the main cause of the problem. Various severity levels are displayed with the view of craniocaudal (CC) images in Fig. 1. The variety of calcified vessels of severe and mild levels can be seen in the final row of enlarged figures.

Due to the complexity of the dataset, binary classification may also outperform end-to-end classification. Numerous disciplines, specifically medical diagnostics, commonly make use of the cascade classification method. This categorization procedure is a decision-making method that employs several binary steps. According to research, cascade classification improves accuracy more than traditional cascade classification, which classifies cascades one at a time using just single-stage citations [20], [37]. Cascade classifiers are frequently used in a range of applications, including handwriting recognition, facial identification, and medical diagnosis [39], [50]. Cascade classifiers have been employed in several research projects in medical diagnosis. As an illustration, cascade categorization was used to assess the level of fibrosis in people with chronic hepatitis C infections [20].

The precise detection of BACs from mammograms remains a challenge despite all previous studies. The algorithms excelled at performance but had a few shortcomings. It is significantly more difficult because to BACs' thin and erratic look. Some models offered had extremely complicated architectures [26], some used massive data sets, and some used the pixel-wise technique [47], which limited the data used. Additionally, when BACs are non-continuous, some may yield incorrect findings [4]. The processing is quite challenging because there aren't any large, annotated datasets of BACs is available. Compared to earlier cited studies, the research provided in this publication uses screening program mammography without the use of additional input parameters to enhance performance. This study uses deep learning with a customized approach based on dataset characteristics, which cuts down on training time and costs while creating a model that is considerably lighter and more precise. Furthermore, it discusses the issue of the severity of calcifications that may lead the way to to automatically create an index relating to the severity of BACs [30].

III. PROPOSED APPROACH

In this section, we present detail idea of first ever direct ordinal prediction without segmenting the calcified vessels to estimate the severity of BACs that is helpful to assess the cardiovascular disease.

A. PRE-PROCESSING

To train the deep networks effectively, large scale data are required. However, collecting labeled medical image is difficult. It become more challenging to mark ground-truth for segmentation for calcified vessels. Other method estimate the severity of BACs based on prior segmentation result of calcified vessel and the estimated level's of accuracy largely depends on the performance of segmentation. Additionally, segmentation requires pixel-to-pixel ground-truth marking from highly experience radiologists. Therefore, this increase the cost of maintaining the dataset and also requires tedious job from experts. Unlike other previous method, proposed method utilize entire image as input for experiment for direct ordinal prediction of BACs level without prior segmentation of calcified vessels. However, to stabilize the architecture with large image size is quite challenging due to complexity of vessels. Since, the optimizer need to get idea of all calcified vessels of individual patient at a time, it is beneficial to utilize the entire image. Proposed method, imply pre-processing using Frangi filter [16] to trimmed the background. Then imply Bicubic [9] interpolation to shrink image so the proposed method can predict class based on entire image. The whole idea breaked down in Algorithm 1.

Algorithm 1 Pre-Processing With Frangi Filter Without Ground-Truth for Ordinal Prediction

Input:
Initial training dataset without ground-truth marking: D;

Output:
Cropped & flipped images with foreground patient-wise, $I_{(i_c)}$;

Pre-processing :

for each $I_{(c_f)} \in I$ **do**

Step-1 : Apply Frangi filter patient-wise, $I_f = FF(I)$

Step-2 : Cropped foreground with different position, $I_t = (I_1, I_2, I_3, I_4, I_5, I_6)$;

Step-3 : Compare images, I_t to find image with maximum foreground, $compare(I_1, I_2, \ldots .I_6)$;

Step-4 : Cropped images for all view, $I_C = (I_{(LCC)}, I_{(RCC)}, I_{(LMLO)}, I_{(RMLO)})$;

Step-5 : Flipped all left images of CC and MLO view, $I_F = (I_{(LCC)}, I_{(LMLO)}$;

Step-6 : Combine all view images of each patients after implying Bicubic interpolation, $I_{(i_c)} = (I_C \cup I_F)$;

end for

Finalize the image dataset after preprocessing for cascade classifiers;

The preprocessing proceed dataset, D without prior ground-truth marking to reduce the background of DICOM image. First, apply patients-wise, Frangi filter $FF(I)$. After the filter, patient's image denoted as I_f. To deal with left and right view images of CC and MLO, filtered image I_f has been cropped in equal size from different position, $I_t = (I_1, I_2, I_3, I_4, I_5, I_6)$. To find cropped image with maximum foreground portion, all cropped images are compared based on the summation of the pixel of foreground object. Afterwards, using same coordinate of cropped image with maximum foreground object, cropped the original image of CC and MLO view of each patients, $I_C = (I_{(LCC)}, I_{(RCC)}, I_{(LMLO)}, I_{(RMLO)})$. To reduce the complexity with different direction of breast, flipped is imposed to all left view images of CC and MLO, $I_F = (I_{(LCC)}, I_{(LMLO)}$. Finally, cropped and filpped images of each patient is combined after

shrink the image using Bicubic interpolation, $I_{(i,C,F)} = (I_C \cup I_F)$ with image size (512 × 512). This process continues for all patients to prepare the initial dataest.

B. PROPOSED ARCHITECTURE AND OBJECTIVE FUNCTION

The implicit hierarchical feature learning method used by CNN, one of the deep learning models, makes it a popular choice for image classification. This has demonstrated a significant improvement in the general image classification problem with many scenarios of its use. Additionally, the structures are becoming more complicated to model more complex decisions. Despite having a variety of internal structures, the existing deep classification models typically represent a similar system. One of the newest additions in the field of deep learning for classification is the ResNet architecture. ResNet has various version depends on number of layers and parameter. This technique is widely used as accomplice for medical image, also. The basic model is proposed in [22] in 2016. However, unlike ResNet initial input (224 × 224 × 3), we need require architecture with bigger initial input (e.g. 512×512×4) with 4-channel as each patients have 4-images. As the size mammogram is quite large compare to calcified vessel and in various shape, the initial input image size (224 × 224) is quite small to achieve better performance of classification.

Proposed residual model (Fig. 2 (b), middle) is inspired by ResNet50 including the shortcut connection (Fig. 2 (c)) which is similar as [22]. However, this shortcut connection turns the network into its counterpart residual version. The identity shortcuts can be directly used when the input and output are of the same dimensions. When the dimensions increase we used to match dimensions adding (1×1) convolutions. The shortcuts go across feature maps of two sizes, they are performed with a stride of 2. Additionally, we increase the initial input size of image from (224 × 224 × 3) to (512 × 512 × 4). In the same time, we reduce the parameters (Table 1) of different convoluitonal layers. As a consequence, the FLOPS (the number of floating point operations that can be performed by a computing entity in one second) varies for different residual network (e.g. ResNet18, ResNet34, ResNet50, ResNet101, ResNet152) that shows in Table 2. Additionally, we perform downsampling directly by convolutional layers that have a stride of 2, same as [22]. The network ends with a global average pooling layer that is followed by fully-connected layer with softmax. The total number of weighted layers for proposed network is 62 (Table 3), that range is in between ResNet50 and ResNet101 network. However, proposed method use cascade approach, unlike end-to-end method of residual network. Fig. 2 (a), shows that at first cascade classifier, normal vs abnormal patients are predicted using proposed customized network. Afterwards, second cascade classifier is introduced for ordinal prediction between minor (abnormal) and severe (abnormal) with another customized ResNet50 network. Both of the networks structure is same and the output of first classifier used as input of second classifier. Therefore, the performance of earlier classifier has impact on later one as this has dependency on earlier network.

Like [22], in this paper, we also explicitly let shortcut connection to fit a residual mapping (Fig. 2 (c)). Connections that skip one or more layers are called shortcut connections [7], [38]. Formally, we allow the stacked nonlinear layers to fit another mapping of $F(\boldsymbol{x}) := H(\boldsymbol{x}) - \boldsymbol{x}$ identifying the desired underlying mapping as $H(x)$. Since it is simpler to optimize the residual mapping than improve the original, unreferenced mapping, the original mapping recasts into $F(\boldsymbol{x}) + \boldsymbol{x}$. In the extreme, fitting an identity mapping by a stack of nonlinear layers would be more difficult than pushing the residual to zero if an identity mapping were best. Feedforward neural networks are capable of realizing the formulation of $F(\boldsymbol{x})+\boldsymbol{x}$. In this study, identity mapping is performed by the shortcut connections, and the results are combined with those of the stacked layers (Fig. 2 (c)). Identity shortcut links don't increase computing complexity or introduce new parameters.

We apply residual learning to every few stacked levels with the following definition of building blocks. Here, the input and output vectors for the layers under consideration are $\boldsymbol{x}$ and $\boldsymbol{y}$. The residual mapping to be learned is represented by the function $F(\boldsymbol{x}, \boldsymbol{W_i})$. For the example in (Fig. 2 (c)), which has two layers, the biases have been left out to make the notations simpler. An element-wise addition and a shortcut connection carry out the operation $F + x$. The dimensions of $\boldsymbol{x}$ and F must, however, be equal. The shortcut connections can be used to conduct a linear projection $\boldsymbol{W_s}$ by shortcut connections to match the dimensions if this is not the case (for instance, while changing the input/output channels). The square matrix $\boldsymbol{W_s}$ can similarly be used to match dimensions. The residual function F can take on several shapes. The function F used in the experiments in this paper has numerous layers. Also to be noted is the fact that convolutional layers can benefit from the aforementioned notations, which, for the sake of simplicity, refer to fully-connected layers.

$$\boldsymbol{y} \;=\; F(\boldsymbol{x}, \boldsymbol{W_i}) + \boldsymbol{W_s}\boldsymbol{x}. \tag{1}$$

However, to avoid ambiguity, we use upper and lower case to represent function and variable, respectively, in this section. Unlike, [22] this experiment implemented cascade classification, so the earlier equation require to adjust as per experiment. At first classifier, normal and abnormal patients are classified. Therefore, $\boldsymbol{y_0}$ and $\boldsymbol{x_0}$ consider as output and input for the classifier. Here, i^0 representing the classifier order (e.g. classifier to identify normal vs abnormal patients). Similarly, the second classifier, minor and severe patients are classified. Therefore, $\boldsymbol{y_1}$ and $\boldsymbol{x_1}$ consider as output and input for the classifier. Here, i^1 representing the classifier order (e.g. minor vs severe). Moreover, as this classifiers are not independent, so the performance of first classifier has impact in later one. The output of first classifier is used as input for second classifier and the total output can be summarize as follows. Here, $\boldsymbol{y_n}, \boldsymbol{y_0}, \boldsymbol{y_1}$ represents the combined output of both classifiers, first classifier and second classifier, respectively.

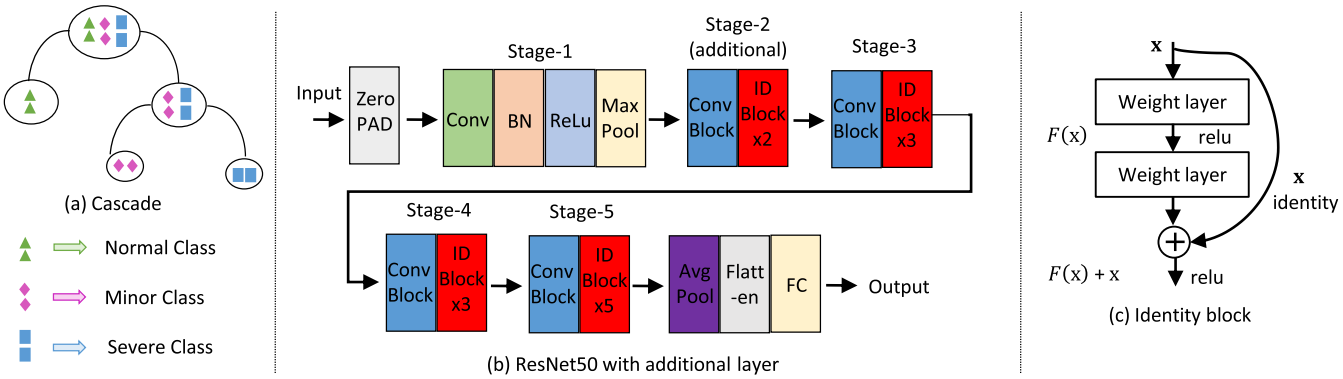


FIGURE 2. Figures illustrates the cascade and customized ResNet50 with skip connection. 2 (a), shows that proposed method use cascade approach, unlike end-to-end method of residual network where Performance of earlier classifier has impact on later one. Proposed residual model (2 (b), middle) is inspired by ResNet50 including the shortcut connection (2 (c)) which is similar as [22]. However, the initial input size is bigger compare to [22] and have fewer parameters at different convolutional layers.

TABLE 3. To adjust the network with bigger image size, proposed method customize the number of layers and also reduce the parameters of different convolutional layers. Table shows the breakdown of layer-wise parameter of ResNet50 and proposed (customized ResNet50) network.

	ResNet50	Customized ResNet50
conv2_x	$64, 64, 256$	$16, 16, 64$
conv3_x	$128, 128, 512$	$32, 32, 128$
conv4_x	$256, 256, 1024$	$64, 64, 256$
conv5_x	$512, 512, 2048$	$128, 128, 512$

Finally, the equation can be expand using earlier equation.

$$\begin{aligned} \boldsymbol{y_0} &= F(\boldsymbol{x^0}, \boldsymbol{W_{(i^0)}}) + \boldsymbol{W_{(s^0)}}\boldsymbol{x^0}, \\ \boldsymbol{y_1} &= F(\boldsymbol{x^1}, \boldsymbol{W_{(i^1)}}) + \boldsymbol{W_{(s^1)}}\boldsymbol{x^1}, \\ \boldsymbol{y_2} &= \boldsymbol{y_0} \cup \boldsymbol{y_1}, \\ \boldsymbol{y_2} &= (F(\boldsymbol{x^0}, \boldsymbol{W_{(i^0)}}) + \boldsymbol{W_{(s^0)}}\boldsymbol{x^0}) \\ &\quad \cup (F(\boldsymbol{x^1}, \boldsymbol{W_{(i^1)}}) + \boldsymbol{W_{(s^1)}}\boldsymbol{x^1}). \end{aligned} \tag{2}$$

C. ROC CURVE AND AUC SCORE CALCULATION FOR CASCADE

Proposed method can not calculate the ROC (receiver operating characteristic) curve and AUC (area under curve) score due to cascade classification. So we calculate the overall curve using probability of 2-classifiers. The equation shows calculation that defines the likelihood of the happening of an event of normal patients for each and every patinets. Here, probability of normal patients, $P(n_n)$, depends on favorable outcome of normal patients, $O(n_n)$ and Total number of favorable outcome of normal and abnormal patients, $O(n_a)$. However, the probability of not normal (or abnormal) patients, $P(n_a)$ is the opposite of probability of normal patients, $P(n_n)$, that shows in first line of the equations. So the probability of being abnormal patients can be represented using the earlier formula (in reverse) for probability of normal patients without separately calculate the probability of abnormal patients shows in second line of equation. However, from the second classifier, probability of minor patients can be calculated (shows in third line of equations). The probability of minor patients, $P(m)$, depends on favorable outcome of minor patients, $O(m)$ and Total number of favorable outcome of abnormal patients, $O(a)$. Also, from the second classifier, probability of severe patients can be calculated. The probability of severe patients, $P(s)$, depends on favorable outcome of severe patients, $O(s)$ and Total number of favorable outcome of abnormal patients, $O(a)$. Since, output of normal and abnormal patients from first classifiers, used as input for second cascade classifier, the probability of minor and severe patients require to multiply with probability of not normal patients of first classifier. So, the probability of total minor, $P(t_m)$ can be find with multiplying of probability of abnormal $P(n_a)$ and probability of minor class $P(m)$. To find the probability of total severe, $P(t_s)$ patients also calculated using the similar formula as minor patients. To find the RUC curve and AUC score, the probability of 3-classes calculated step-by-step as follows.

$$\begin{aligned} P(n_n) &= \frac{O(n_n)}{O(n_a)}, \\ P(n_a) &= 1 - P(n_n), \\ P(n_a) &= 1 - \frac{O(n_n)}{O(n_a)}, \\ P(m) &= \frac{O(m)}{O(a)}, \\ P(s) &= \frac{O(s)}{O(a)}, \\ P(t_m) &= P(n_a) \times P(m), \\ P(t_s) &= P(n_a) \times P(s). \end{aligned} \tag{3}$$

IV. EXPERIMENT

This section, we analysis the dataset of 3-different classes. Besides, it describes the preprocessing of dataset without

TABLE 4. Comparison of various ResNet and proposed (customized ResNet50) network, in terms of layer and FLOPS (the number of floating point operations that can be performed by a computing entity in one second).

	Layer	FLOPS
ResNet18	18	1.8×10^9
ResNet34	34	3.6×10^9
ResNet50	50	3.8×10^9
Customized ResNet50	62	4.6×10^9
ResNet101	101	7.6×10^9
ResNet152	152	11.3×10^9

ground-truth marking. Afterwards, present the comparison of various method after the experiment.

A. DATASET AND PRE-PROCESSING

The Seoul National University Bundang Hospital is the place where screening Mammography database is collected and maintained. The radiologist has more than ten years of expertise gathering and maintaining DICOM-format mammograms, even though this in-house dataset is not publicly accessible. The total number of patients' is 868, and their BAC scores range from 0∼12. However, parameters including length, density, and degree of the calcified vessel are utilized to estimate the severity level of vessels. The summation these 3-metrics measurements is known as the BACs score. Since there is insufficient data for each class individually to conduct a regression-based experiment, we split these into three groups (0 (normal), 1∼6 (minor), and 7∼12 (severe)) for ordinal prediction. Patients with anomalies are divided into classes as minor and severe cases (Table 2). Each instance includes information about the patient and the imaging, including mediolateral oblique (MLO) and cranial-caudal views for each breast (CC). Four images are available for each patient (CC and MLO view of each breast). There are a total of 3472 (868 × 4) images consist, in this in-house dataset. The image's dimensions are (2816, 2016) in height and width. (0.0815, 0.0815) is the pixel spacing of all images. Three separate classes have differing numbers of cases, thus different augmentation techniques (such as flipping, rotating, and adding noise) are used to balance the cases and also used to expand the dataset in order to mitigate the overfitting issue. The age range various for middle age women among the various classes (Table 5, Col. 4). It is important to note that the proposed method use 4-images (left-right breast of CC and MLO view) for each patient for all trials. However, the results obtained with just MLO (left and right) can match those obtained with four images. We use four images for each patient since unknown testing datasets might require information for CC view images. To prevent bias, we also keep patient-specific training and testing datasets, unlike [4]. This experiment never mixed up images of a single patient for multiple subgroups, always assigning 4-images for each patient to the training or testing subset. However, original images (2816 × 2016) contain lot of background (Fig. 3 (a)). We trimmed the images automatically based-on position of breast (left/right side) to to remove the backdrop implying Frangi filter (Fig. 3 (b)). After implying filter the image cropped automatically in different (left and right side) positions (Fig. 3 (c - h)) of image (upper, middle and lower). Each cropped image's height and width is (2560 × 1536). Then compare these images and find the image with maximum foreground portion (Fig. 3 (i)). Using the coordinate of selected cropped image, the original image is cropped (Fig. 3 (j)). Finally, Bicubic interpolation is used to shrink the image to 512 × 512 (Fig. 3 (k)).

B. EXPERIMENTAL ANALYSIS

In the begging of the experiment, first we conduct the experiment for Bicubic and Bilinear interpolation to compare which interpolation perform well with mammogram. Comparison of proposed customized ResNet50 with Bilinear and Bicubic interpolation shown in Table 6. Additionally, Bicubic interpolation apply with flipped images. All the experiment conducted using end-to-end approach using 4-fold cross-validation, though the final experiment conducted with cascade approach. The performance of various metrics shows that Bicubic interpolation perform better than Bilinear interpolation (i.e. accuracy 0.74390 & 0.72358). Additionally, we experimented the proposed method with Bicubic interpolation and flipped images. The last column shows the metrics performance improve for flipped image a bit as the direction of left and right view breast align to similar direction and reduce the complexity of dataset.

To compare the performance of various approaches for measuring the severity of calcified vessels without segmentation is depicted in this section. We additionally analyze regression, in addition to existing classification methods, since this is a regression-based topic. Patients' severity increases as they progress from normal to abnormal (minor and severe) class. The rest of other traditional (experimented) methods are based on classification.

In Table 7, a comparison of conventional and proposed customized ResNet50 is presented. The performance of the traditional RestNet50 cascade is shown in the 1st column, where end-to-end classifier is performed. The 2nd column, shows the performance of classification based on split images. Finally, the proposed customized ResNet50's performance shows in last column. The accuracy result shows the efficiency of proposed cascade method (0.8105) as it consider the overall idea of entire image with bigger image size (512 × 512) compare to traditional ResNet50 (224 × 224). However, traditional ResNet50 does not perform well (accuracy 0.6845) as this has smaller image size and use end-to-end classification. The split image with (512) shows a bit improvement (0.7362) due to bigger image size but that is also not efficient enough for real time performance. From this table it shows that the

TABLE 5. Dataset analysis: 3-different class information regarding age, cases of normal and abnormal patients along with ground truth of radiologist. The classification GT is also shown in this table.

Severity	Score†	GT‡	Age	Cases	Image size	Cropped size	Shrink size	Pixel spacing
Normal	0	0	40∼65	567	(2816, 2016)	(2560, 1536)	(512, 512)	(0.0815, 0.0815)
Minor	1∼6	1	40∼80	153	(2816, 2016)	(2560, 1536)	(512, 512)	(0.0815, 0.0815)
Severe	7∼12	2	50∼85	148	(2816, 2016)	(2560, 1536)	(512, 512)	(0.0815, 0.0815)
Total Cases								868
Total Images								3472

Score†= Breast artery calcification's score. To measure the calcification of vessels, metrics like length, density, and degree, are used usually. BACs score is the summation of these three matrics (actual ground-truth from doctors); GT‡= Ground-Truth (GT) refers the classification label to conduct experiment of proposed method to find severity of BAC.

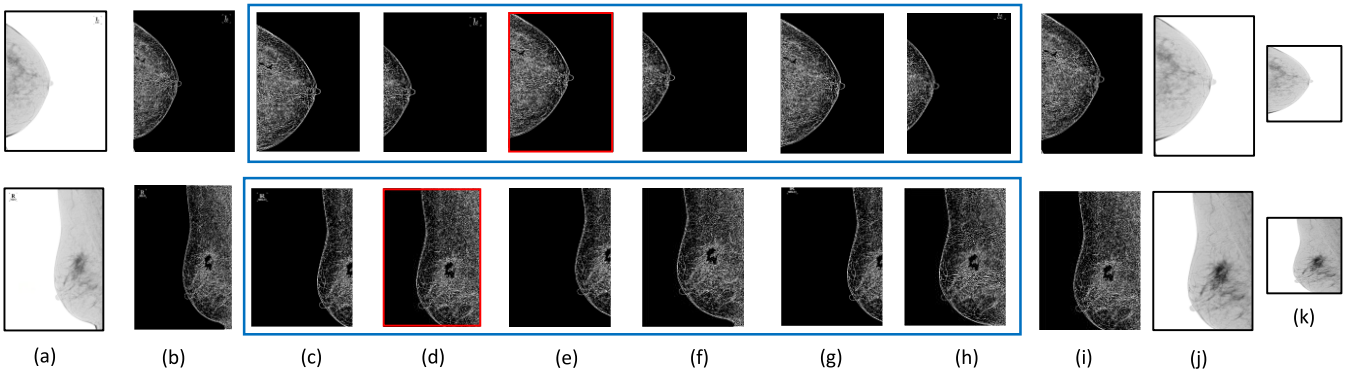


FIGURE 3. Automatic pre-processing of dataset shows in the figures. The height and width of all the mammograms are 2816 and 2016. However, (3 (a)) demonstrates that the images include a lot of backgrounds. Pre-processing trimmer the background of image by implying Frangi filter (3 (b)). After applying the filter, proposed method cropped image from different position (3 (c - h)) and compare images to find the image with most foreground object (3 (i)). Finally, the original image is cropped based on filtered image and shrink using interpolation (3 (j)-(k)) to adjust with customized ResNet50 network.

TABLE 6. Comparison of proposed customized ResNet50 with Bilinear and Bicubic interpolation. Additionally, Bicubic interpolation apply with flipped images.

	Bilinear	Bicubic	Bicubic + Flip
Accuracy	0.72358	0.74390	0.76016
Recall	0.72358	0.74390	0.76016
Precision	0.72358	0.74390	0.76016
F1-score	0.72358	0.74390	0.76016
AUC	0.72358	0.74390	0.76016

CResnet50†= Proposed method (customized Resnet50)

performance increase gradually for ResNet34 to proposed Customize ResNet50 and afterwards the performance decrease a bit for ResNet101 and ResNet152. It seems the dataset is not large enough to adopt with deeper network with increased number of parameter. We believe, the larger dataset may increase the performance with deeper network.

The proposed method, unlike previous methods, doesn't segment the calcified vessel to predict the severity level of BACs. So it is very important to find, on what basis the proposed network classifies the patients. Fig. 4 shows the activation map for some abnormal patients shown for ResNet50 and proposed customized ResNet50. This figure, shows the activation map of the mediolateral oblique (MLO) view of each patient. For profound investigation, this figure visualizes some abnormal patients' predictions using activation maps. 3-images shown for each patients (original image, activation map of ResNet50 and (proposed) customized ResNet50). Fig. 4 (b) and (e), show that ResNet50 predicts patients-1 and patients-2 as normal as this focus on the edge and little left side of the breast, respectively where no calcified vessel is found clearly. On the contrary, proposed method Fig. 4 (c) and (f), focuses on calcified vessel area and predicted patient-1 and patient-2 as abnormal. However, both of the networks predict patient-3 and patient-4 as abnormal but ResNet50 doesn't focus on the actual calcified vessel area Fig. 4 (h) and (k), when (proposed) customized ResNet50 makes prediction based on focusing the calcified vessels Fig. 4 (i) and (l). Moreover, the example case of last row shows that both ResNet50 and proposed customized ResNet50 failed to predict the patients severity level accurately.

TABLE 7. Comparison of proposed customized ResNet50 (CResnet50†) and conventional methods. Proposed cascade performs better compare to ResNet50 and with split images.

	ResNet34	ResNet50	Split image	CResnet50†	ResNet101	ResNet152
Accuracy	0.6423	0.6845	0.7362	0.8105	0.7673	0.7424
Recall	0.6754	0.7258	0.7721	0.8613	0.9046	0.8758
Precision	0.6431	0.6827	0.7568	0.8441	0.5614	0.5432
F1-score	0.6676	0.7035	0.7661	0.8216	0.6989	0.6631
AUC	0.6637	0.7122	0.7682	0.8465	0.8123	0.7857

CResnet50†= Proposed method (customized Resnet50)

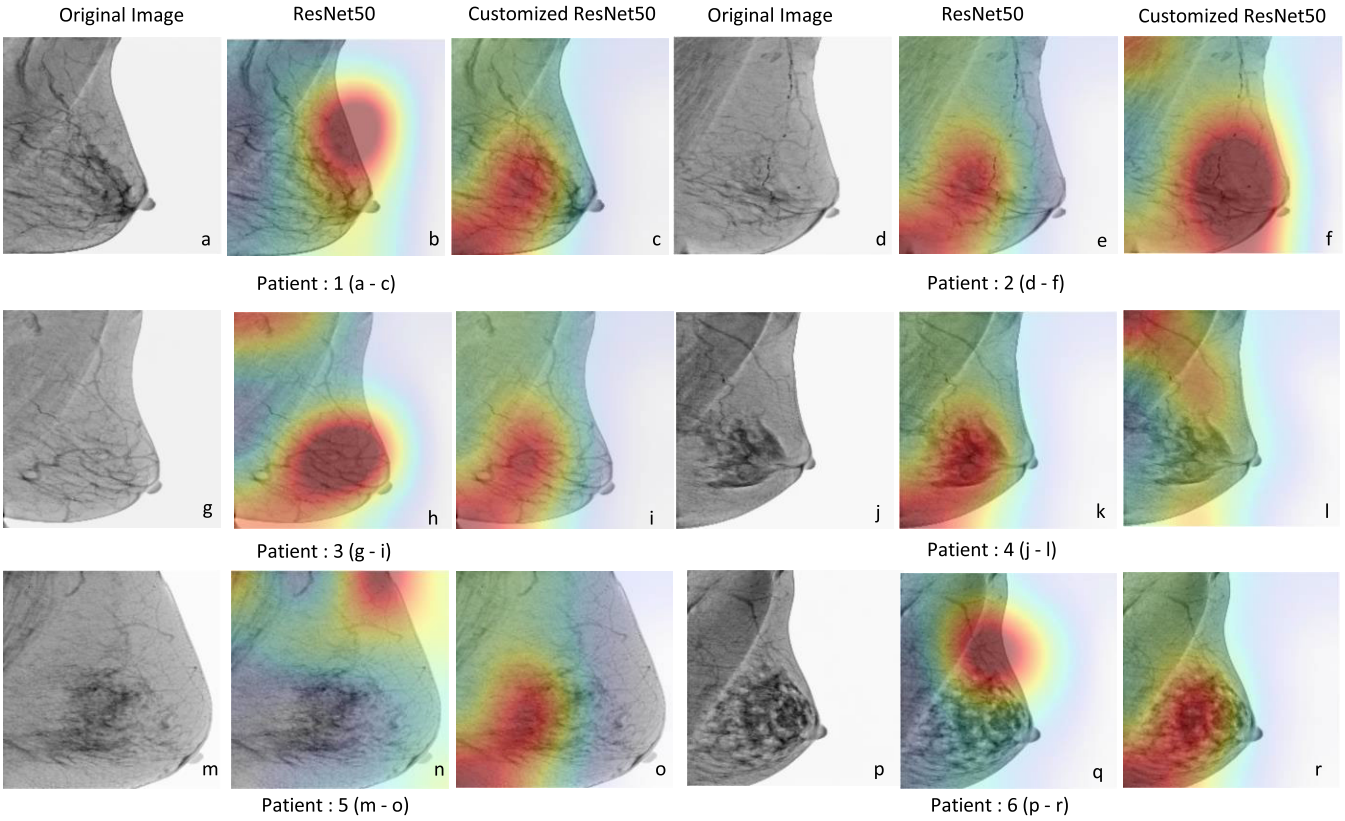


FIGURE 4. Activation map for some abnormal patients shown for ResNet50 and (proposed) customized ResNet50. This figure, shows the activation map of the mediolateral oblique (MLO) view of each patient. The proposed method, unlike previous methods, doesn't segment the calcified vessel to predict the severity level of BACs. So it is very important to find, on what basis the proposed network classifies the patients. For profound investigation, this figure visualizes some abnormal patients' predictions using activation maps. 3-images shown for each patients (original image, activation map of ResNet50 and (proposed) customized ResNet50). 4 (b) and (e), show that ResNet50 predicts patients-1 and patients-2 as normal as this focus on the edge and little left side of the breast, respectively where no calcified vessel is found clearly. On the contrary, proposed method 4 (c) and (f), focuses on calcified vessel area and predicted patient-1 and patient-2 as abnormal. However, both of the networks predict patient-3 and patient-4 as abnormal but ResNet50 doesn't focus on the actual calcified vessel area 4 (h) and (k), when (proposed) customized ResNet50 makes prediction based on focusing the calcified vessels 4 (i) and (l). Moreover, the example case of last row shows that both ResNet50 and proposed customized ResNet50 failed to predict the patients severity level accurately.

Even within one class, mammography characteristics are extremely variable. The shapes, location, and contrast of calcified vessel varies a lot (from patient to patient). For a straightforward classification problem, it is impossible to earn the progressive diversity of mammograms using single network. Thus, we suggested an cascade technique for learning and distinguishing the calcified vessel characteristic gradually. The comparison with other traditional approaches is depicted in (Fig. 5). The base-line network ResNet50 and traditional cascade is showed in 1st and 2nd bars. Later, unlike using end-to-end classification, cascade classification with progressive and split image is shown with (bars 3,

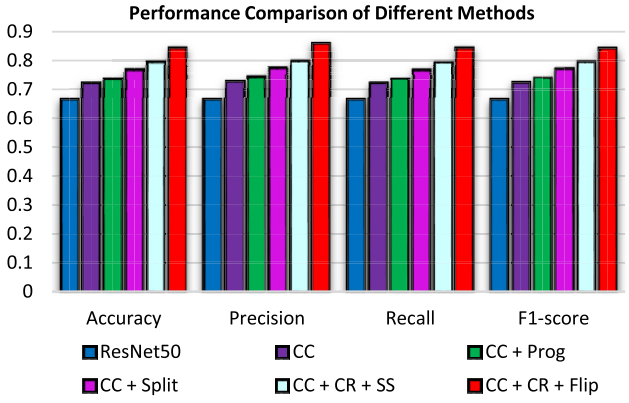


FIGURE 5. The comparison with other traditional approaches is depicted in this Figure. The base-line network ResNet50 and traditional cascade is showed in 1st and 2nd bars. Later, unlike using end-to-end classification, cascade classification with progressive and split image is shown with (bars 3, and 4, respectively). The progressive resizing approach is introduced with conventional cascade which increases the accuracy above 0.73. Afterward, modified ResNet50 is used with sub-sample and entire image that also increases the performance a bit (bar 5∼6). The last bar shows the performance goes up to 0.85 after implementing the proposed customized ResNet50 with entire shrink image and flipping the direction of mammogram.

and 4, respectively). The progressive resizing approach is introduced with conventional cascade which increases the accuracy above 0.73. Afterward, modified ResNet50 is used with sub-sample and entire image that also increases the performance a bit (bar 5∼6). The last bar shows the performance goes up to 0.85 after implementing the proposed customized ResNet50 with entire shrink image and flipping the direction of mammogram.

Table 8 compares the proposed method to standard models on a class-by-class basis using 4-fold cross-validation. This also shows how performance matches the cascading tendency. Because we categorize the normal class first, it gets the best results. Precision receives a higher score for the proposed method. However, because the severe class uses the classifier-2, it obtains the second-best performance (also for the proposed technique). Other metrics (Recall, F1-score, accuracy) are also compared in this table to assess the suggested method's performance. The loss function in regression is the mean-square-error (MSE). Although this is a regression problem, ResNet50 performs better than regression.

Comparison of ROC curve and AUC score with traditional methods shows in (Fig. 6). Left figure shows the comparison of normal class and the right figure shows the comparison of minor class. The base-line network ResNet50 and traditional cascade is showed with blue and orange color (for both figure). Later, cascade is introduced unlike using end-to-end classification, with progressive and split image is shown with (green and red color, respectively). The progressive resizing approach is introduced with conventional cascade which increases the AUC score to 0.83. Afterward, modified ResNet50 is used with sub-sample and entire image that also increases the performance a bit (purple and brown). The last purple color shows the performance goes up to 0.91 after implementing the proposed customized ResNet50 with entire shrink image and flipping the direction of mammogram.

C. COMPARE WITH PREVIOUS METHODS

The proposed customized ResNet50 perform better compare to various state-of-art ResNet. However, this is also important to compare with previous method based on mammogram. In this sub-section we comparatively analysis other methods (Table 1) and proposed customized ResNet50 in detail. However, as it shows in the table that none of the dataset / GT-marking is available publicly, so we are unable to compare previous and proposed methods on same dataset. Also, the metric to measure the outcome and number of BACs class varies for different papers. So we conducted experiment following the similar pattern of each paper and compare gradually. Reference [47] examine the viability of automated and precise BAC detection in mammography for assessing the risk CAD. To distinguish between BAC and non-BAC, they create a 12-layer convolutional neural network and implement a pixel-wise, patch-based method for BAC identification. The combination of convolutional layer, batch-normalization, non - linearity, max-pooling, and fully connected (FC) layers makes up the deep CNN architecture that is the subject of this research. This results in a neural network of 12 layers, 10 convolutional layers, and 2 FC layers. While the FC layers are employed for the final classification, the convolutional layers can be thought of as feature extractors. To assess the performance of the system, they seek help from in-house experts radiologists to conduct a reader study for ground-truth marking. However, they divide the patients into 2-groups (BAC and Non-BAC patients). Afterwards, evaluate the performance using a set of 840 full-field digital mammograms from 210 cases, using both free-response receiver operating characteristic(FROC) and calcium mass quantification analysis.

Table (1 and 9), shows the various perspective of 12-layer CNN-based method and proposed customized ResNet50. This method split mammogram with 95 × 95 pixel for segmentation where proposed customized Resnet50 trimmed the background and imply Bicubic interpolation to shrink image into 512 × 512. To compare with this method we also extract 33% of dataset for testing to use 3-folding cross-validation. The true-positive shows proposed method performance is better compare to this CNN-based 2-stage method. It reason for performance variation [47] is that the final outcome of BAC and Non-BAC patients depends on outcome of segmentation result which is quite challenging. On the other hand proposed customized ResNet50 directly estimate the severity of calcified vessels which helps to out perform the performance of this method.

Next, we conduct experiment using customized ResNet50 with this paper [48]. Authors also conducted this experiment using in-house dataset which is marked with the help of expert radiologists. This paper evaluates the performance of three deep learning architectures, YOLO, Unet and DeepLabv3+, on detecting BACs in digital mammography.

TABLE 8. Class-wise comparison with traditional methods (4-fold cross-validation). We compare regression beside ResNet and the proposed classification as this is a regression-based problem. As Patients' severity increases gradually from normal to abnormal (minor and severe) classes.

	Cases‡(Train,Test)	Regression (MSE)			ResNet50 [22]			CResNet50†		
		P	R	F1	P	R	F1	P	R	F1
Normal	(343,115)	0.56	0.53	0.54	0.81	0.78	0.78	0.85	0.81	0.82
Minor	(125,41)	0.47	0.48	0.47	0.52	0.53	0.54	0.70	0.72	0.72
Severe	(113,38)	0.51	0.52	0.52	0.66	0.65	0.65	0.78	0.79	0.78
Average	(581,194)	0.51	0.51	0.51	0.66	0.65	0.65	0.78	0.77	0.77
Accuracy	(581,194)		0.51			0.64			0.81	

CResNet50†= proposed method (customized ResNet50); ‡Though the dataset is fix, number of training and testing cases change a bit for different folding, cause of random shuffling. Here, R = Recall, P = Precision, F1 = f1-score

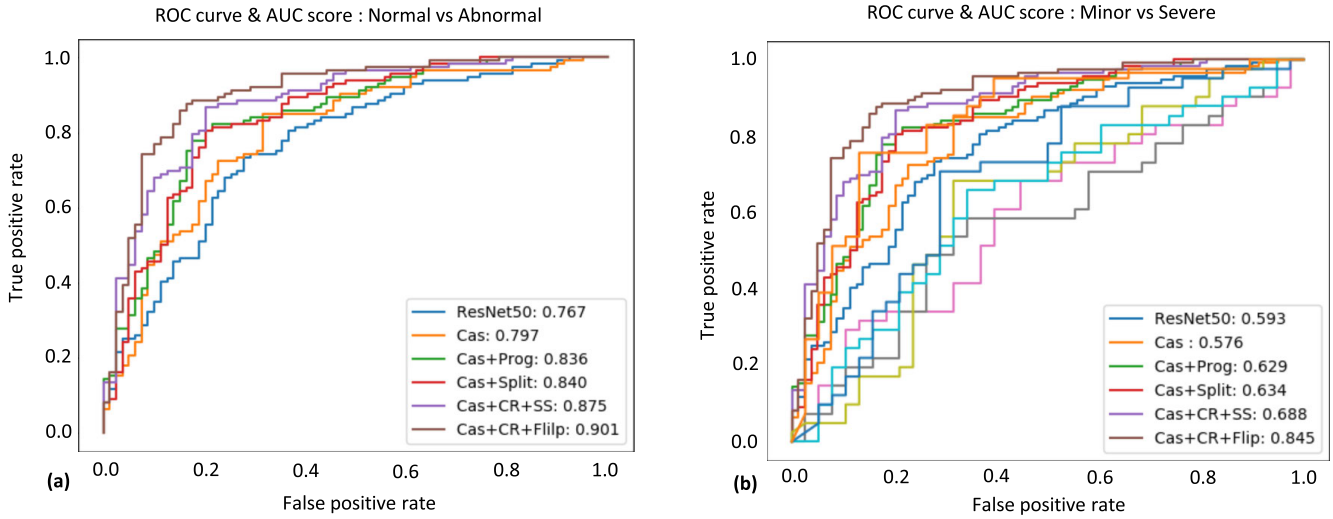


FIGURE 6. Comparison of ROC curve and AUC score with traditional methods. The comparison with other traditional approaches is depicted in this Figure. Left figure shows the comparison of normal class and the right figure shows the comparison of minor class. 6-different methods are compared, same as (Fig. 5).

TABLE 9. Performance comparison of [47] & proposed customizedResNet50 is shown in this table.

	FROC†	Calcium Mass	CResNet50‡
Image type	Split	Split	Crop & Bicubic
Pixel	95 × 95	95 × 95	512 × 512
Test Ratio	0.33	0.33	0.33
Cross-validation	3-folding	3-folding	3-folding
Threshold	0.99	0.99	0.50
Case / Image	334-image	64-case	209-case
Prediction	256-image	35-case	223-case
True-positive	0.76	0.54	0.94

FROC†= Free-Response Receiver Operating Characteristic
CResnet50‡= Proposed method (customized Resnet50)

TABLE 10. Performance comparison of [48] & proposed customizedResNet50 is shown in this table.

	DeepLabV3+	Hessian†	CResNet50‡
Image type	Crop	Crop	Crop & Bicubic
Image size	Varies	Varies	512 × 512
Test Ratio	0.20	0.20	0.20
Cross-validation	Once	Once	5-folding
Training case	108-case	108-case	790-case
Testing case	27-case	27-case	196-case
True-positive	0.7407	0.7778	0.9048

Hessain†= Hessain-based filter
CResnet50‡= Proposed method (customized Resnet50)

Instead, a straightforward filter based on Hessian is designed to improve the structure of BACs. Afterwards a thresholding technique is used to produce mask (binary) to assess the BACs. The authors designed new metric to more accurately assess the small item segmentation because BACs has diversity in shapes specially tiny in size. Authors use 135 mammograms with labeled BACs, of which 20% were used for

TABLE 11. Performance comparison of [4] & proposed customizedResNet50 is shown in this table.

	BAC	Non-BAC	Non-BAC (random)	Accuracy	Sensitivity	Specificity	f1-score	Jaccard index
DU-Net	413-case	2217-case	413-case	0.9147	0.9122	0.9201	0.9219	0.8517
CResNet50†	301-case	567-case	301-case	0.9206	0.9206	0.9344	0.9219	0.8550

CResnet50†= Proposed method (customized Resnet50)

validation and 80% were used to train deep learning networks. The cropped image varies for different patients. Table 10, shows the various perspective of this method (DeepLabv3+, Hessian-based multiscale filter which perform best) and proposed customizes ResNet50 method. This shows that the true-positive of proposed method perform better than this method.

Lastly, we compare proposed method with [4]. The authors of this research created a deep learning method using mammograms to identify BACs. They expand dense connection based U-Net to detect BACs in mammograms, motivated by the encouraging results obtained utilizing the U-Net model in several biological segmentation tasks and the DenseNet in semantic segmentation. The given model aids in reusing computation and enhances gradient flow, resulting in greater accuracy and simpler model training. A mammograms dataset were generated and gathered for this assignment to test the performance. However, it shows from Table 1 that, though the dataset are publicly available but the GT-marking is conducted by experts radiologist (in-house). Also the GT marking is not made publicly available yet. The original image size shrink to 640×640 to perform segmentation first then use various metrics (Table 11) to analyse the experiment result. To compare with this method, we also prepare the dataset following the similar patterns. As dataset has more Non-BAC patients compare to BAC patients, this method randomly chose same number of Non-BAC patients. From this table, we can see that despite we have way small number of dataset compare to this method, the performance of propose customized ResNet50 is comparatively better for various metrics (accuracy, specificity, f1-score and Jaccard index) except sensitivity.

D. FINDINGS

Given that the calcification area varies greatly, the mammography dataset for vascular segmentation is rather complicated (e.g. shape, position, size). The suggested method, in contrast to past methods, measures the risk of severity degree of patients without segmenting the calcified vessel. Numerous experiment with proposed method have backup the idea of direct measurement of severity of BAC. Additionally, it revealed, based on proposed method, that the cascade with the customized ResNet50 alpng with entire image is more effective than other traditional techniques. The proposed customized ResNet50 approach, like other deep learning-based networks, require large dataset. If this procedure only needed a tiny dataset to do the experiment, it would be of immense use to doctors in real time. Nevertheless, because calcified vessels have so many different characteristics, it can be rather challenging. Also, there is a possibility to explore the suggested method for multi-modal datasets.

V. CONCLUSION

One of the main causes of death in the human race is cardiovascular disease. However, maintaining a dataset with computed tomography (CT) is quite expensive and exposes individuals to a lot of radiation exposure when being scanned. Women over 40 get routine mammograms to screen for breast cancer in several countries. Mammograms are a practical and comprehensive method for compiling a large number of cases. To estimate the severity of patients based on segmented vessels, researchers have historically segmented calcified vessels using pixel-wise ground truth. The cost of this pixel-by-pixel ground truth marking technique is rather significant because only multiple experienced professionals can mark those accurately. We considered the notion of developing an automated system for direct ordinal prediction of BACs scores without segmenting the calcified vessels for the first time, as per best of our knowledge that outperforms classification results in terms of performance. The suggested method incorporates a Frangi filter to eliminate background and shrink image size with Bicubic interpolation since the extreme vascular diversity and huge image size make it difficult to discern the severity level of patients. Additionally, unlike other previous methods, this one enables direct ordinal prediction utilizing cascade and gradually produces the best results; it does not require pixel-to-pixel segmentation or split the image. In terms of accuracy, proposed method performs better than the traditional and equivalent models. It is obvious that the suggested approach will save researchers cost, time, and effort because it does not require segmenting the calcified vessel and offers a reliable risk assessment for BAC, even though more research is required to completely address this issue. In order to lengthen the time frame for determining the risk of CVD, we will expand our experiment to employ small and multi-modal datasets in the future.

REFERENCES

[1] *The Top 10 Causes of Death*. (2020). [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

[2] N. Abou-Hassan, E. Tantisattamo, E. T. D'Orsi, and W. C. O'Neill, "The clinical significance of medial arterial calcification in end-stage renal disease in women," *Kidney Int.*, vol. 87, no. 1, pp. 195–199, Jan. 2015.

[3] M. Al Ghamdi, M. Li, M. Abdel-Mottaleb, and M. A. Shousha, "Semi-supervised transfer learning for convolutional neural networks for glaucoma detection," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 3812–3816.

[4] M. AlGhamdi, M. Abdel-Mottaleb, and F. Collado-Mesa, "DU-Net: Convolutional network for the detection of arterial calcifications in mammograms," *IEEE Trans. Med. Imag.*, vol. 39, no. 10, pp. 3240–3249, Oct. 2020.

[5] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation," 2018, *arXiv:1802.06955*.

[6] J. K. Baum, C. H. Comstock, and L. Joseph, "Intramammary arterial calcifications associated with diabetes," *Radiology*, vol. 136, no. 1, pp. 61–62, Jul. 1980.

[7] C. M. Bishop, *Neural Networks for Pattern Recognition*. New York, NY, USA: Oxford Univ. Press, 1995.

[8] Q. M. Bui and L. B. Daniels, "A review of the role of breast arterial calcification for cardiovascular risk stratification in women," *Circulation*, vol. 139, no. 8, pp. 1094–1101, Feb. 2019.

[9] R. E. Carlson and F. N. Fritsch, "Monotone piecewise bicubic interpolation," *SIAM J. Numer. Anal.*, vol. 22, no. 2, pp. 386–400, Apr. 1985.

[10] T. Chadashvili, D. Litmanovich, F. Hall, and P. J. Slanetz, "Do breast arterial calcifications on mammography predict elevated risk of coronary artery disease?" *Eur. J. Radiol.*, vol. 85, no. 6, pp. 1121–1124, Jun. 2016.

[11] J.-Z. Cheng, C.-M. Chen, E. B. Cole, E. D. Pisano, and D. Shen, "Automated delineation of calcified vessels in mammography by tracking with uncertainty and graphical linking techniques," *IEEE Trans. Med. Imag.*, vol. 31, no. 11, pp. 2143–2155, Nov. 2012.

[12] J.-Z. Cheng, C.-M. Chen, and D. Shen, "Identification of breast vascular calcium deposition in digital mammography by linear structure analysis," in *Proc. 9th IEEE Int. Symp. Biomed. Imag. (ISBI)*, May 2012, pp. 126–129.

[13] J.-Z. Cheng, E. B. Cole, E. D. Pisano, and D. Shen, "Detection of arterial calcification in mammograms by random walks," in *Proc. Int. Conf. Inf. Process. Med. Imag.* Cham, Switzerland: Springer, 2009, pp. 713–724.

[14] E. J. Chun, S. Ahn, S. M. Kim, S. I. Choi, M. Bo La Yun, and J.-W. Suh, "Prediction of subclinical coronary artery disease with breast arterial calcification and low bone mass in asymptomatic women," Tech. Rep., 2018.

[15] D. Ciresan, A. Giusti, L. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, 2012, pp. 2843–2851.

[16] K. B. Khan, A. A. Khaliq, A. Jalil, and M. Shahid, "A robust technique based on VLM and Frangi filter for retinal vessel extraction and denoising," *PloS ONE*, vol. 13, no. 2, 2018, Art. no. e0192203.

[17] *FDA Approves First 3-D Mammography Imaging System*, US Food and Drug Administration, 2016.

[18] J. Ge, H.-P. Chan, B. Sahiner, C. Zhou, M. A. Helvie, J. Wei, L. M. Hadjiiski, Y. Zhang, Y.-T. Wu, and J. Shi, "Automated detection of breast vascular calcification on full-field digital mammograms," in *Proc. SPIE*, vol. 6915, pp. 379–385, Mar. 2008.

[19] D. C. Goff Jr., D. M. Lloyd-Jones, G. Bennett, S. Coady, R. B. D'agostino, R. Gibbons, P. Greenland, D. T. Lackland, D. Levy, C. J. O'donnell, and J. G. Robinson, "2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American College of Cardiology/American heart association task force on practice guidelines," *Circulation*, vol. 129, pp. 49–73, 2014.

[20] A. M. Hashem, M. E. M. Rasmy, K. M. Wahba, and O. G. Shaker, "Single stage and multistage classification models for the prediction of liver fibrosis degree in patients with chronic hepatitis C infection," *Comput. Methods Programs Biomed.*, vol. 105, no. 3, pp. 194–209, 2012.

[21] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle, "Brain tumor segmentation with deep neural networks," *Med. Image Anal.*, vol. 35, pp. 18–31, Jan. 2017.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[23] E. J. E. Hendriks, P. A. de Jong, Y. van der Graaf, W. P. T. Mali, Y. T. van der Schouw, and J. W. J. Beulens, "Breast arterial calcifications: A systematic review and meta-analysis of their determinants and their association with cardiovascular events," *Atherosclerosis*, vol. 239, no. 1, pp. 11–20, Mar. 2015.

[24] D. Karm, D. S. Marks, M. Wein, and A. L. Kong, "Benign arterial calcification on screening mammogram: A marker for coronary artery disease?" *J. Women's Health*, vol. 24, no. 10, pp. 795–800, Oct. 2015.

[25] B. S. Kelly, E. ScanlON, H. Heneghan, C. E. Redmond, G. M. Healy, E. Mc Dermott, E. J. Heffernan, R. Prichard, and S. Mc Nally, "Breast arterial calcification on screening mammography can predict significant coronary artery disease in women," *Clin. Imag.*, vol. 49, pp. 48–53, May 2018.

[26] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes," *IEEE Trans. Med. Imag.*, vol. 37, no. 12, pp. 2663–2674, Dec. 2017.

[27] D. C. Luvizon, D. Picard, and H. Tabia, "2D/3D pose estimation and action recognition using multitask deep learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 5137–5146.

[28] L. Margolies, M. Salvatore, H. S. Hecht, S. Kotkin, R. Yip, U. Baber, V. Bishay, J. Narula, D. Yankelevitz, and C. Henschke, "Digital mammography and screening for coronary artery disease," *JACC, Cardiovascular Imag.*, vol. 9, no. 4, pp. 350–360, Apr. 2016.

[29] M. E. Matsumura, C. Maksimik, M. W. Martinez, M. Weiss, J. Newcomb, K. Harris, and M. A. Rossi, "Breast artery calcium noted on screening mammography is predictive of high risk coronary calcium in asymptomatic women: A case control study," *Vasa*, vol. 42, no. 6, pp. 429–433, Nov. 2013.

[30] N. Mazidi, C. Roobottom, and G. Masala, "Automatic quantification of breast arterial calcification on mammographic images," in *Innovation in Medicine and Healthcare Systems, and Multimedia*. Berlin, Germany: Springer, 2019, pp. 283–292.

[31] S. Y. Molloi and C. Iribarren, "Breast arterial calcifications are correlated with subsequent development of coronary artery calcifications, but their aetiology is predominantly different," *Breast Diseases, Year Book Quart.*, vol. 19, no. 3, pp. 240–241, Jan. 2008.

[32] M. Abedi, M. Moradi, and A. Adibi, "Relationship between breast arterial calcification on mammography with CT calcium scoring and coronary CT angiography results," *Adv. Biomed. Res.*, vol. 3, no. 1, p. 79, 2014.

[33] J.-J. Mordang and N. Karssemeijer, "Vessel segmentation in screening mammograms," in *Proc. SPIE*, vol. 9414, Mar. 2015, Art. no. 94140J.

[34] L. Mostafavi, W. Marfori, C. Arellano, A. Tognolini, W. Speier, A. Adibi, and S. G. Ruehm, "Prevalence of coronary artery disease evaluated by coronary CT angiography in women with mammographically detected breast arterial calcifications," *PLoS ONE*, vol. 10, no. 4, Apr. 2015, Art. no. e0122289.

[35] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, "Speech recognition using deep neural networks: A systematic review," *IEEE Access*, vol. 7, pp. 19143–19165, 2019.

[36] D. Newallo, F. G. Meinel, U. J. Schoepf, S. Baumann, C. N. De Cecco, R. J. Leddy, R. Vliegenthart, H. Möllmann, C. W. Hamm, P. B. Morris, and M. Renker, "Mammographic detection of breast arterial calcification as an independent predictor of coronary atherosclerotic disease in a single ethnic cohort of African American women," *Atherosclerosis*, vol. 242, no. 1, pp. 218–221, Sep. 2015.

[37] K. Polat, S. Güneş, and A. Arslan, "A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square support vector machine," *Exp. Syst. Appl.*, vol. 34, no. 1, pp. 482–487, 2008.

[38] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge, U.K.: Cambridge, Univ. Press, 2007.

[39] Y. Saatci and C. Town, "Cascaded classification of gender and facial expression using active appearance models," in *Proc. 7th Int. Conf. Autom. Face Gesture Recognit.*, 2006, pp. 393–398.

[40] K. Y. Safak, A. Eratalay, E. D. Ilis, N. U. Tanju, B. D. Sencan, and T. Baysal, "The relationship of breast arterial calcification detected in mammographic examinations with cardiovascular diseases, cardiovascular risk factors, parity, and breastfeeding," *TURKISH J. Med. Sci.*, vol. 46, pp. 641–645, Jan. 2016.

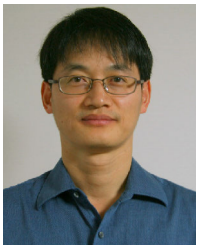
[41] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, May 2015.

[42] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, "Multi-scale convolutional neural networks for lung nodule classification," in *Proc. Int. Conf. Inf. Process. Med. Imag.* Cham, Switzerland: Springer, 2015, pp. 588–599.

[43] S. Y. Shin, S. Lee, I. D. Yun, and K. M. Lee, "Deep vessel segmentation by learning graphical connectivity," *Med. Image Anal.*, vol. 58, Dec. 2019, Art. no. 101556.

[44] R. A. Smith, V. Cokkinides, D. Brooks, D. Saslow, and O. W. Brawley, "Cancer screening in the United States, 2010: A review of current American cancer society guidelines and issues in cancer screening," *CA, A Cancer J. Clinicians*, vol. 60, no. 2, pp. 99–119, Mar. 2010.
[45] J.-W. Suh and B. La Yun, "Breast arterial calcification: A potential surrogate marker for cardiovascular disease," *J. Cardiovascular Imag.*, vol. 26, no. 3, pp. 125–134, 2018.
[46] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck, "Deep learning for identifying metastatic breast cancer," 2016, *arXiv:1606.05718*.
[47] J. Wang, H. Ding, F. A. Bidgoli, B. Zhou, C. Iribarren, S. Molloi, and P. Baldi, "Detecting cardiovascular disease from mammograms with deep learning," *IEEE Trans. Med. Imag.*, vol. 36, no. 5, pp. 1172–1181, May 2017.
[48] K. Wang, N. Khan, and R. Highnam, "Automated segmentation of breast arterial calcifications from digital mammography," in *Proc. Int. Conf. Image Vis. Comput. New Zealand (IVCNZ)*, Dec. 2019, pp. 1–6.
[49] J. Yeboah, T. S. Polonsky, R. Young, R. L. McClelland, J. C. Delaney, F. Dawood, M. J. Blaha, M. D. Miedema, C. T. Sibley, J. J. Carr, G. L. Burke, D. C. Goff, B. M. Psaty, P. Greenland, and D. M. Herrington, "Utility of nontraditional risk markers in individuals ineligible for statin therapy according to the 2013 American College of Cardiology/American heart association cholesterol guidelines," *Circulation*, vol. 132, no. 10, pp. 916–922, Sep. 2015.
[50] P. Zhang, T. D. Bui, and C. Y. Suen, "A novel cascade ensemble classifier system with a high recognition performance on handwritten digits," *Pattern Recognit.*, vol. 40, no. 12, pp. 3415–3429, Dec. 2007.

HOSNA ASMA-ULL received the B.Sc. degree from Sookmyung Womens University, Seoul, in 2011, and the M.Sc. degree from the Hankuk University of Foreign Studies, Yongin, in 2017, South Korea, where she is currently pursuing the Ph.D. degree with the Department of Computer and Electronic Systems Engineering. She was a Faculty Member, from 2012 to 2019, with the Department of Computer Science and Engineering, Stamford University Bangladesh, Dhaka, Bangladesh, (an Assistant Professor, currently on study leave). Her current research interests include deep learning, machine learning, medical and real image analysis, artificial intelligence, image processing, and computer vision.



IL DONG YUN (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in control and instrumentation engineering from Seoul National University, Seoul, South Korea, in 1989, 1991, and 1997, respectively. From 1996 to 1997, he was with Daewoo Electronics, Seoul, as a Senior Engineer. Since 1997, he has been with the Hankuk University of Foreign Studies (HUFS), Yongin, South Korea. He has been the Dean of Computer Engineering with HUFS and the Vice Chairperson of the Korean Computer Vision Society (KCVS), since 2022 March. His research interests include medical image analysis and computer vision.

BO LA YUN received the degree from the Seoul National University College of Medicine, in 2004, and the M.S. and Ph.D. degrees in medicine from Seoul National University, Seoul, South Korea, in 2011, and 2014, respectively. She completed an Internship with Seoul National University College of Medicine. From 2006 to 2010, she completed Resident Course with Seoul National University Hospital, Seoul. Since 2012, she has been with Seoul National University Bundang Hospital, Seongnam, South Korea. She is currently a Medical Doctor (M.D.) majoring in breast radiology. She is a Clinical Associate Professor in radiology with Seoul National University Bundang Hospital.