

Calcification-Informed BI-RADS Categorization of Mammograms: A Combined Detection and Classification Approach

Debapriya Chakraborty¹[0009–0003–9159–4675], Uddalak Mukhopadhyay²[0009–0006–9643–7822], Sarbani Palit¹[0000–0002–4105–6452], and Ujjwal Bhattacharya¹[0000–0002–8546–6453]

¹ Indian Statistical Institute, Kolkata 700108, India <https://www.isical.ac.in>

² Institute of Technical Education and Research, Bhubaneswar 751030, India <https://www.soa.ac.in/iter>

Abstract. Mammogram images are classified according to the Breast Imaging Reporting and Data System (BI-RADS) scale depending on the severity of the lesions present. A higher BI-RADS class essentially indicates a greater risk of malignancy. Calcification is one of the most familiar and vital lesions found in mammograms. For the purpose of classification, it is crucial to detect the calcifications present on the mammogram. In this work, we present a novel calcification detection and classification pipeline for BI-RADS classification of the entire mammogram image while considering all the calcifications present. We propose a strategy for detection of calcification(s) in a mammogram image followed by the classification of the image itself into one of the BI-RADS category. Here, we propose an object detection strategy based on YOLOv8, since we observed that YOLOv8 is capable of highlighting the spread of calcifications in an efficient way. Further, each individual image is classified based on the BI-RADS categories depending only on the presence and spread of the calcifications spotted in the detection phase. We achieved MAP of 79.70% for the calcification detection task and classification accuracy of 81.50% and a F1 Score of 77.80% with the attention-based module aiding the residual based network (ResNet50) on a curated set of images from the combination of two publicly available datasets. The method presented here, thus tries to get closer to the analysis process of a radiologist in the real world.

Keywords: Mammogram · Calcification · BI-RADS · Attention.

1 Introduction

Among women all over the world, breast cancer continues to be the most prevalent type of cancer. Although various imaging modalities are used for breast cancer diagnosis, mammography is still the most suitable modality for breast cancer detection in early stages. Among the abnormalities found on a mammogram, mass and calcification are the most prevalent and significant ones and the presence of these lesions likely indicates the possibility of breast cancer.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026

Cite as: Chakraborty, D., et al.: Calcification-Informed BI-RADS Categorization of Mammograms: A Combined Detection and Classification Approach. In: Mahmud, M., et al. (eds.) Applied Intelligence and Informatics. AII 2026. CCIS, vol. 3206, pp. 355–369. Springer, Cham (2026).

The calcium deposits that develop in breast tissue are called breast calcification [21]. They are typically observed on mammograms as tiny or large white dots. While calcifications are usually harmless, the size, shape and scattering of the calcifications provide indications as to whether they are markers of benign condition or malignancy [3].

To evaluate the risk of cancer, the American College of Radiology proposed the Breast Imaging Reporting and Data System (BI-RADS) [8]. BI-RADS category of a mammogram sample enables radiologists to standardize the entire procedure of how to interpret a mammogram and aids in demonstrating a close correlation with risk malignancy risk. The categories are: 1-normal, 2-benign abnormality, 3-probably benign presence, 4-suspicious abnormality, 5-highly suspicious of malignancy, 6-biopsy proven cancerous lesion. Generally, a single BI-RADS label is given to a mammogram image by the radiologists based on the comparative severity of lesions present in it [14]. Therefore, only the detection of lesions present in mammograms is not sufficient, classification of the entire mammogram image into the correct BI-RADS category based on the presence of lesions is also necessary to fully replicate the analysis process of the radiologists.

Advancements made in deep learning have significantly changed the field of medical image analysis, enabling automated and highly accurate interpretation of complex imaging modalities [10, 7]. Convolutional Neural Networks (CNNs), attention mechanisms, and of late transformer-based architectures have demonstrated remarkable performance against a wide range of tasks, including disease detection, lesion localization, segmentation, and image-level classification. Among these, object detection networks such as Faster R-CNN [15], Single Shot MultiBox Detector (SSD) [9], and the YOLO family [5] have gained considerable importance owing to their ability to simultaneously localize and identify pathological findings within an image. In parallel, deep classification networks augmented with attention mechanisms have shown superior capability in learning discriminative image representations by emphasizing diagnostically relevant regions. These developments have facilitated the design of integrated computer-aided diagnosis systems that combine lesion detection with subsequent disease classification, resulting in improved interpretability and diagnostic performance.

Despite the advancement of modern machine learning algorithms, effective detection of lesions present on mammograms remain a challenging task. In addition, these lesions are often classified as either benign or malignant rather than performing the multiclass classification based on the BI-RADS scale. Some works, which have tried to classify mammograms on the BI-RADS scale, tend not to concentrate on different unique abnormalities but on general classification and we feel this is not an effective way to execute the classification task. Therefore, in this study we concentrated only on one type of abnormality: calcifications. For the categorization procedure of calcifications, in a case where more than one calcification or cluster of calcification are present, the highest BI-RADS class among the individual calcifications is given as the BI-RADS class of the whole image. The fact that calcifications, more precisely microcalcifications, appear in different shapes and clusters and can sometimes also be seen occupying a large

proportion of the mammogram, making the segregation of each calcification and labeling very difficult.

In this work, instead of classifying each individual calcification or cluster of calcification, we worked on classifying each mammogram image into one of the 4 BI-RADS labels (BI-RADS 2, BI-RADS 3, BI-RADS 4, BI-RADS 5) based solely on the morphology, size and distribution of all the calcifications present. Specifically, our objective was to study how the spread and shape of all the individual or cluster of calcifications only, determine a BI-RADS category for a single mammogram image. The following is a synopsis of our overall contributions:

- YOLOv8 [19] based framework is employed for the detection of calcifications in the first stage
- A residual based CNN architecture in unison with a block attention module [2] is proposed that achieves classification of a single mammogram image based on the morphology, size and distribution of all the calcifications present
- We focus on the image-level classification of the mammogram based on the BI-RADS categories

The remaining manuscript is structured as follows: The second section provides a brief review of the literature. Section 3 provides a detailed description of the proposed method. In the subsequent sections (4 and 5), the results and findings are discussed. Lastly, Section 6 concludes the paper.

2 Related Work

There are a diverse number of machine learning-supported approaches for the detection and classification of calcification on mammograms in the literature (Table 1). Some of the scholars proposed methods that adapt to the distinctive characteristics of calcifications. Wang et al. [22] designed a context-sensitive deep neural network that took into account local image features and the surrounding background for the detection of the calcification. Zamir et al. [25] developed a strategy that prioritized difficult pixels during training phase to tackle the issue of false-positive calcifications. Hossain et al's [4] method consisted of a number of preprocessing steps and manual selection of suspicious regions before feeding these into a UNet segmentation network. Tong et al. [20] developed an proficient methodology that integrates segmentation and quantitative analysis of breast calcification. The segmentation network was used for precise localization of the calcifications, while traditional machine learning algorithms were used for classification (benign/malignant) subsequently.

In other works, Loizidou et al.'s [11] method took advantage of a pre-processing phase and a traditional machine learning-based classification. Songsaeng et al. [18] developed a multi-scale pre-trained CNN-based approach, involving a two-branch architecture designed to extract feature maps from two distinct mammography views (Cranio-Caudal and Medio-Lateral Oblique). Moreover, Yurdusev et al. [24] used a difference filter as a pre-processing stage for the mammogram

images. Then Faster R-CNN and Yolov4 models were used for detecting microcalcifications. Likewise, Shanker et al. [17] utilized a combination of Deep Belief NN and Chimp Optimization Algorithm for classification. In [1], a new strategy for breast calcification detection was suggested by Boudouh et al. Their approach utilized pre-processing of input images and a hybrid ViT and CNN model. While all these works achieved good results over the years, none of these proposed methods concentrated on the BI-RADS classification. Some scholars have proposed methods that focused on this task. Schonenberger et al. [16] used a deep CNN method to classify microcalcifications in accordance with the BI-RADS scale by extracting region of interest (ROI) areas from mammograms. Zhang et al. [26] designed a new loss and attention mechanism for the classification of calcifications and they also trained their model on the BI-RADS scale.

In this manuscript, we focused on the simultaneous detection of calcifications on mammogram images and image-level BI-RADS classification of the entire image based on the size and scattering of the detected calcifications while utilizing a new architecture based on the swin transformer network, making our approach unique and closer to the radiologists' approach for mammogram analysis.

Table 1: Summary of the Literature Review

Study	Dataset	Image-level BI-RADS Classification	Publishing Year
Wang et al. [22]	Private	No	2018
Zamir et al. [25]	INbreast, Private	No	2021
Hossain et al. [4]	DDSM	No	2022
Tong et al. [20]	DDSM, Private	No	2024
Loizidou et al. [11]	Private	No	2020
Songsaeng et al. [18]	Private	No	2021
Yurdusev et al. [24]	DDSM	No	2022
Shanker et al. [17]	DDSM	No	2023
Boudouh et al. [1]	CBIS-DDSM	No	2024
Schonenberger et al. [16]	Private	No	2021
Zhang et al. [26]	Private	No	2021
Proposed Framework	INbreast, VinDr-Mammo, CBIS-DDSM	Yes	–

3 Materials and Method

3.1 Datasets

The framework proposed in this manuscript was validated on a combination of selected subset of two publicly available datasets: VinDr-Mammo [13] and INbreast [12] and separately on the freely available dataset CBIS-DDSM [6].

Mammogram images with only calcifications and no mass were taken into consideration for experiments in case of all the datasets, as we wanted the models to learn to classify each individual image based only on calcifications.

The INbreast dataset was published in 2012 from Portugal. It comprises of total 410 images taken from 115 studies (from which 90 have four images for each case and 25 are mastectomy patients and therefore have two images per case). The size of the original images are either 3328x4084 or 2560x3328. Among these, only 221 images, which satisfy the experiment conditions were considered for our experiments.

The VinDrMammo dataset, on the other hand, was published in 2022 from Vietnam. In this dataset, the total number of studies is 5000 and the total number of images is 20000. The size for the original images is 3518x2800. For our experiment, the number of images considered is 351. Therefore, the total number of images considered was 572. The images belonging to both these datasets are Full-Field-Digital-Mammogram (FFDM)s.

Finally, the Curated Breast Imaging Subset of the Digital Database for Screening Mammography (CBIS-DDSM) was released in 2017. The dataset comprises digitized screen-film mammogram (SFM)s collected from 1,566 patients. The dataset is divided into separate mass and calcification subsets with predefined training and testing splits. In this work, only the calcification subset was considered. After excluding mammograms that were also present in the mass subset, a total of 1,440 unique mammogram images containing only calcification abnormalities were found. Among these, only 1364 mammogram images in total belong to BI-RADS 2, 3, 4 and 5 classes and finally these 1364 images were used for the further validation of the proposed methodology.

We observed that there was no image with calcification belonging to BI-RADS 2 category in the VinDrMammo dataset. Therefore, we combined the two sets of images and split them into train and test group. We named this combined dataset as 'Dataset A' for ease in naming purposes. Dataset A consists of training set of 452 images and testing set of 120 images. We further validated our method on the CBIS-DDSM dataset. The number of images belonging to each BI-RADS category for all the datasets is given in in Table 2.

Table 2: Distribution of Images for Each BI-RADS Category

Dataset	BI-RADS 2	BI-RADS 3	BI-RADS 4	BI-RADS 5
INBreast	189	10	18	4
VinDr-Mammo	0	61	192	98
CBIS-DDSM	280	93	828	163

The images from an individual patient was either used in the training set or the test set. It was made sure that images from the same person did not get considered in both training and test sets to prevent data leakage.

3.2 Calcification Detection

The entire pipeline of the proposed method is given in Fig. 1. First part of the pipeline deals with the detection of calcifications in an input mammogram image. For this, YOLOv8 was used. The YOLOv8 network essentially consists of 3 fundamental parts: a backbone network, a neck network, and a head network. The first one, i.e. backbone network is the feature extractor, while the neck network is there for merging and refining the features collected from the backbone network. Finally, the head network is used for object detection. The loss function utilized in YOLOv8 is a combination of three losses: classification loss, localization loss, and confidence loss. While the classification loss is for measuring the difference between the predicted and ground truth class labels, the localization loss is used for the calculation of the difference between the predicted and ground truth bounding box coordinates. Lastly, the confidence loss helps to measure the difference between the predicted and ground truth object confidence scores. The total loss function is a weighted sum of these three losses. Typically, stochastic gradient descent (SGD) with momentum is employed as optimizer in YOLOv8. The accuracy of YOLOv8 is typically measured using the mean average precision (mAP) metric. The mAP metric measures the average precision of the detector across all classes. The mAP metric is calculated as follows:

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i,$$

where C is the number of classes, and AP_i is the average precision of class i . The average precision of class i is calculated as follows:

$$AP_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \frac{TP_j}{TP_j + FP_j},$$

where N_i is the number of true positives for class i , TP_j is the number of true positives for class i at IoU threshold j , and FP_j is the number of false positives for class i at IoU threshold j .

The mammogram images were converted into colored images at first. Using colored images helped in better detection of calcifications. Next, the images were divided into smaller patches, each of size 256×256 . Then these patches were fed into the YOLOv8 model to generate predictions of bounding box data. Finally, the patches with annotated bounding boxes are reconstructed into the original black and white mammogram, resulting in an image with calcifications annotated by bounding boxes.

3.3 BI-RADS Classification

The task of classification of mammogram images according to the BI-RADS norm dependent only on the presence of calcifications is dealt with here. We used a different classification network as we wanted to classify an image rather than the

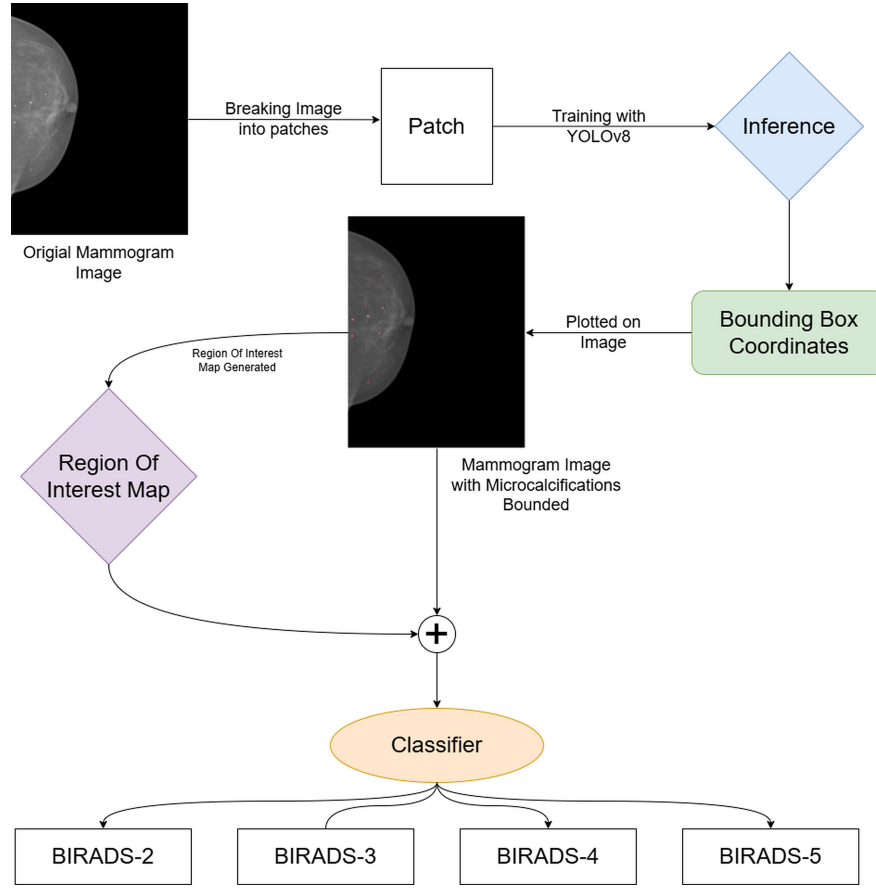


Fig. 1: Entire pipeline consisting of the detection of calcifications and classification of whole mammograms

individual calcifications found in that image. Also we wanted the classification network to concentrate mainly on the the spread and features of calcifications present in an image. Therefore, the output from the calcification detection network in the previous part of the proposed pipeline was converted into an region of interest (ROI) map and was subsequently channel-wise added to the original mammogram and supplied into the classifier network.

An attention guided residual based network was used as the classifier network. ResNets are utilized as the backbone in our network. Standard CNN architectures like ResNets are deficient when it comes to exploring global contextual information in images. Moreover, adaptive recording of information from the channel dimension appears to be essential for improving accuracy.

Hence, application of convolutional block attention modules (CBAM) [23], which comprises of both channel and spatial attention, was used by us. The overall attention function of the CBAM block is as follows:

$$\begin{aligned} F' &= M_C(F) \otimes F \\ F'' &= M_S(F') \otimes F', \end{aligned}$$

where $F \in \mathbb{R}^{C \times H \times W}$ is an intermediate feature map and taking F as the input, CBAM sequentially generates a 1D channel attention map $M_C \in \mathbb{R}^{C \times 1 \times 1}$ and a 2D spatial attention map $M_S \in \mathbb{R}^{1 \times H \times W}$. Here $\otimes$ denotes element-wise multiplication.

Correlation among channel maps is the primary goal for the channel attention sub-module. The shared network is applied to the outputs from the two pooling operations individually and the merging of resultant output feature vectors finishes the work of this sub-module.. The functioning of the channel attention sub-module can be computed as:

$$\begin{aligned} M_C(F) &= \sigma(MLP(AvgPool(F)) \\ &\quad + MLP(MaxPool(F))) \\ &= \sigma(W_1(W_0(F_{avg}^C) + W_1(W_0(F_{max}^C))), \end{aligned}$$

where σ denotes the sigmoid function while $W_0 \in \mathbb{R}^{C/r \times C}$, and $W_1 \in \mathbb{R}^{C \times C/r}$ denote the Multilayer Perceptron (MLP) weights.

On the other hand, learning relation among the pixels in the feature map is the objective of the spatial attention block. The spatial attention sub-module tries to exploit two outputs that are pooled along the channel axis and after concatenation, they are forwarded to a convolution layer which in turn produces the 2D spatial attention map. The spatial attention module can be summarized as:

$$\begin{aligned} M_S(F) &= \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \\ &= \sigma(f^{7 \times 7}([F_{avg}^S; F_{max}^S])), \end{aligned}$$

where σ denotes the sigmoid function and $f^{7 \times 7}$ represents a convolution operation with the filter size of 7×7 .

The block diagram for the classifier network encompassing the CBAM blocks is given in Fig. 2

4 Results

We used PyTorch to implement the entire proposed pipeline. The network training was implemented on a workstation equipped with 'Intel Xion' processor, an 'NVIDIA RTX A5000' GPU and 64G memory.

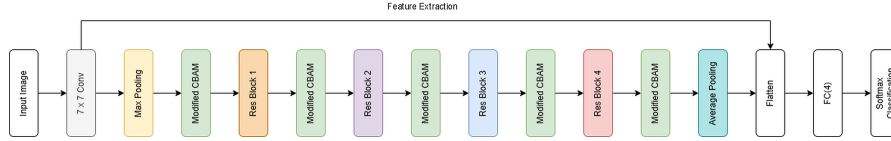


Fig. 2: Classifier network consisting of residual and CBAM blocks

Our entire process was started by training YOLOv8, setting the batch size to 90, the initial learning rate to 0.01 and momentum to 0.8. The image was colored to make calcifications easier to detect and results showed the coloring indeed made the the object detection performance better. Also, a patch based approach was taken to train the object detection network for which images were partitioned into patches of size 256x256. Precision, Recall and MAP were used as measures for object detection part of the entire procedure.

For the training of the classifier network, Adam optimizer was used along-with a batch size of 32, cross entropy as the loss function and the learning rate parameter set to 0.0001.

4.1 Calcification Detection Performance

The results of the calcification detection on both the datasets are presented in table 3 and 4

Table 3: Calcification Detection Results on Dataset A

Metric	Value (%)
Precision	75.45
Recall	76.88
mAP	79.70

The 95% confidence interval calculated by bootstrap method around the mAP is [72.51%, 86.89%].

Table 4: Calcification Detection Results on CBIS-DDSM Dataset

Metric	Value (%)
Precision	71.55
Recall	70.43
mAP	74.15

An example of calcifications detected on a patch of a sample mammogram image from Dataset A is given in Fig. 3.

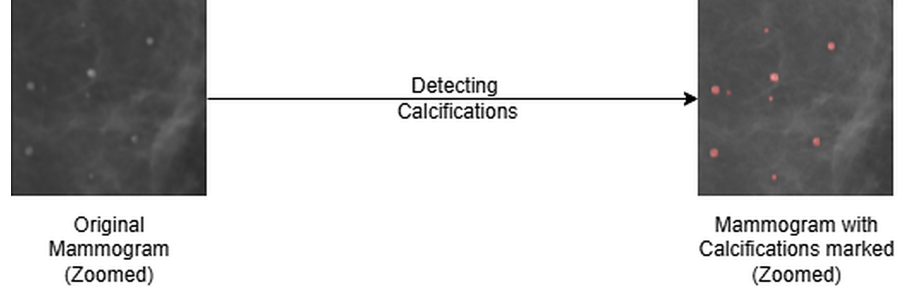


Fig. 3: Detection of calcifications on a sample mammogram image patch

4.2 Classification Performance With Different Base ResNets

At first, we tested how network depth affected the classification performance. Table 5 summarizes the classification results on four different residual network configurations. It was observed that all the networks showed very similar results overall. A small increase was observed in the evaluation metrics of ResNet50.

Table 5: Classification Results of Different ResNet Models on Dataset A

Model	Accuracy (%)	F1-Score (%)
ResNet18	65.40	62.50
ResNet34	67.75	65.00
ResNet50	70.30	67.50
ResNet101	69.40	65.30

4.3 Classification Performance Enhancement with CBAM

We evaluated the effectiveness of the addition of the CBAM module using the best baseline residual network, i.e. ResNet50 on both the datasets (shown in Table 6 and 7).

Table 6: Classification Results of the Modified ResNet50 Network on Dataset A

Network	Accuracy (%)	F1-Score (%)
ResNet50	70.30	67.50
ResNet50 + CBAM	81.50	77.80

The 95% Wilson score confidence interval around the classification accuracy for ResNet50 + CBAM model is [73.63%, 89.37%].

Table 7: Classification Results of the Modified ResNet50 Network on CBIS-DDSM

Network	Accuracy (%)	F1-Score (%)
ResNet50	65.84	64.54
ResNet50 + CBAM	76.18	75.40

The confusion matrix generated using ResNet50 with CBAM block on Dataset A is given in Table 8.

Table 8: Confusion Matrix for Dataset A

Actual / Predicted	BI-RADS 2	BI-RADS 3	BI-RADS 4	BI-RADS 5
BI-RADS 2	33	4	0	2
BI-RADS 3	2	12	2	1
BI-RADS 4	7	1	34	0
BI-RADS 5	0	2	1	19

Table 9: Class-wise Sensitivity for BI-RADS Categories for Dataset A

Category	Sensitivity (%)	Total Number of Samples
BI-RADS 2	84.61	39
BI-RADS 3	64.7	17
BI-RADS 4	80.95	42
BI-RADS 5	86.36	22

4.4 Ablation Analysis: Error Propagation

To isolate the error propagation between the upstream detection module (Stage 1) and the downstream classification module (Stage 2), we conducted an oracle ablation experiment. The classifier’s performance was evaluated under two distinct input regimes:

1. **End-to-End Pipeline (Predicted Maps):** The classifier receives region-of-interest (ROI) crops or spatial attention masks generated from the bounding boxes predicted by the Stage 1 detector (YOLOv8).
2. **Oracle Baseline (Ground-Truth Maps):** The predicted calcification maps are replaced with ground-truth (GT) spatial annotations, providing the classifier with ideal localization masks.

Table 10: Ablation Using Ground-truth Versus Predicted Calcification Maps

Pipeline Configuration	Stage 1 Input	Accuracy (%)
Oracle Baseline	Ground-Truth Maps	87.33
End-to-End Pipeline	Predicted Maps	81.50
Propagated Error Gap (Δ)	—	5.83

5 Discussion

First of all, for the object detection part, an image patch based procedure was used. Calcifications generally capture very small part of the whole image and therefore the patch based method works better than providing the whole image to the network. Next, for the classification of the entire image part, it was observed that all the different residual networks displayed similar results with a small increase in the evaluation metrics of ResNet50. This is justifiable as deeper networks essentially require larger datasets to obtain a proper training despite their larger capacities. The use of attention based CBAM came out to be effective in further finetuning the classifier's performance. The attention process essentially emphasizes on the most important features in the input (here a mammogram image) necessary for a particular task (here classification). Significant improvements of accuracy was observed compared to the naive ResNet50 architecture to support this notion. Also, the channel-wise addition of the ROI maps of the detected calcifications paved the way for the classifier network to concentrate mainly on them to classify the images.

As presented in Table 10, evaluating the classifier on ground-truth calcification maps yields an accuracy of 87.33%, compared to 81.50% when using predicted detections. This 5.83% performance degradation isolates the upper bound of error propagation introduced by Stage 1 detection inaccuracies.

From the results presented in this manuscript, it can be observed that our methodology performed poorly for the CBIS-DDSM dataset for both detection and classification tasks. This unfortunate outcome is attributed to the quality of the CBIS-DDSM dataset. As already mentioned, while INbreast and VinDr-Mammo datasets comprise of FFDM images, CBIS-DDSM is made up of SFM images. Therefore, there is significant difference in quality among the datasets and the results suggest that although the proposed framework performs effectively on FFDM images, its generalization across different mammographic domains remains a challenge. Further investigation using domain adaptation techniques and larger, heterogeneous datasets may help improve the robustness and generalization capability of the proposed approach.

6 Conclusion

In this study, we investigated the applicability of a unified approach for calcification detection and subsequent image label classification on the BI-RADS scale for mammogram depending entirely on the calcifications found in it. Here, We adopted an integrated attention mechanism with the naive deep residual network in order to enhance the classification performance. In this work, while YOLOv8 showed excellent and efficient capability in detecting even the tiniest calcifications, the attention mechanism proved beneficial for the classification task. The channel-wise addition of the ROI maps of the detected calcifications paved the way for the classifier network to concentrate mainly on them to classify the images. Although the classification accuracy obtained was good, the availability of

the limited number of images that satisfied our experiment criteria seem to hold back the accuracy a little. Given more samples for training, our method has the potential for further improvement in accuracy.

Not only our study proves to be sufficient for categorizing a mammogram when only calcifications are present, but also becomes beneficial where multiple lesions are present by taking care of the BI-RADS category for all the lesions of a particular category, i.e. calcification. Our method has the scope of seamless implementation in clinical application. After this work, We plan to handle the issue of expansion of generalization ability of the model across similar datasets. In our experiments, lower performance of our methodology on the SFM dataset (CBIS-DDSM) was observed. Also, our method is unable to detect other important lesions present in mammogram like mass and architectural distortion. We plan to work on these abnormalities and create a complete mammogram classification model in the future.

Code and Data Availability

The datasets used in this study are publicly available from their respective data providers. The source code used in this study are publicly available at <https://github.com/uddalak2005/Calcification-Informed-BI-RADS-Categorization-of-Mammograms>

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Boudouh, S.S., Bouakkaz, M.: Advancing precision in breast cancer detection: a fusion of vision transformers and cnns for calcification mammography classification: Ss boudouh and m. bouakkaz. *Applied Intelligence* **54**(17), 8170–8183 (2024)
2. Chakraborty, D., Palit, S., Bhattacharya, U.: Deep classification of mammographic breast density: Dcbarnet. In: 2023 38th International Conference on Image and Vision Computing New Zealand (IVCNZ). pp. 1–6. IEEE (2023)
3. Farshid, G., Sullivan, T., Downey, P., Gill, P., Pieterse, S.: Independent predictors of breast malignancy in screen-detected microcalcifications: biopsy results in 2545 cases. *British journal of cancer* **105**(11), 1669–1675 (2011)
4. Hossain, M.S.: Microcalcification segmentation using modified u-net segmentation network from mammogram images. *Journal of King Saud University-Computer and Information Sciences* **34**(2), 86–94 (2022)
5. Jiang, P., Ergu, D., Liu, F., Cai, Y., Ma, B.: A review of yolo algorithm developments. *Procedia computer science* **199**, 1066–1073 (2022)
6. Lee, R.S., Gimenez, F., Hoogi, A., Miyake, K.K., Gorovoy, M., Rubin, D.L.: A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific data* **4**(1), 170177 (2017)
7. Li, M., Jiang, Y., Zhang, Y., Zhu, H.: Medical image analysis using deep learning algorithms. *Frontiers in public health* **11**, 1273253 (2023)

8. Liberman, L., Menell, J.H.: Breast imaging reporting and data system (bi-rads). *Radiologic Clinics* **40**(3), 409–430 (2002)
9. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C.: Ssd: Single shot multibox detector. In: *European conference on computer vision*. pp. 21–37. Springer (2016)
10. Liu, X., Gao, K., Liu, B., Pan, C., Liang, K., Yan, L., Ma, J., He, F., Zhang, S., Pan, S., et al.: Advances in deep learning-based medical image analysis. *Health Data Science* **2021**, 8786793 (2021)
11. Loizidou, K., Skouroumouni, G., Nikolaou, C., Pitris, C.: An automated breast micro-calcification detection and classification technique using temporal subtraction of mammograms. *IEEE Access* **8**, 52785–52795 (2020)
12. Moreira, I.C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M.J., Cardoso, J.S.: Inbreast: toward a full-field digital mammographic database. *Academic radiology* **19**(2), 236–248 (2012)
13. Nguyen, H.T., Nguyen, H.Q., Pham, H.H., Lam, K., Le, L.T., Dao, M., Vu, V.: Vindr-mammo: A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography. *Scientific Data* **10**(1), 277 (2023)
14. Rawashdeh, M., Lewis, S., Zaitoun, M., Brennan, P.: Breast lesion shape and margin evaluation: Bi-rads based metrics understate radiologists’ actual levels of agreement. *Computers in Biology and Medicine* **96**, 294–298 (2018)
15. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence* **39**(6), 1137–1149 (2016)
16. Schönenberger, C., Hejduk, P., Ciritsis, A., Marcon, M., Rossi, C., Boss, A.: Classification of mammographic breast microcalcifications using a deep convolutional neural network: A bi-rads-based approach. *Investigative Radiology* **56**(4), 224–231 (2021)
17. Shanker, M., Vadivel, M.: Micro-calcification classification analysis in mammogram images with aid of hybrid technique analysis. *Wireless Personal Communications* **128**(2), 1287–1307 (2023)
18. Songsaeng, C., Woodtichartpreecha, P., Chaichulee, S.: Multi-scale convolutional neural networks for classification of digital mammograms with breast calcifications. *IEEE Access* **9**, 114741–114753 (2021)
19. Terven, J., Córdova-Esparza, D.M., Romero-González, J.A.: A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine Learning and Knowledge Extraction* **5**(4), 1680–1716 (2023)
20. Tong, Y., Jiang, J., Chen, F., Guo, G., Zhang, C., Deng, T.: A novel approach for segmentation and quantitative analysis of breast calcification in mammograms. *Frontiers in Oncology* **14**, 1281885 (2024)
21. Tse, G., Tan, P.H., Pang, A.L., Tang, A.P., Cheung, H.S.: Calcification in breast lesions: pathologists’ perspective. *Journal of clinical pathology* **61**(2), 145–151 (2008)
22. Wang, J., Yang, Y.: A context-sensitive deep learning approach for microcalcification detection in mammograms. *Pattern recognition* **78**, 12–22 (2018)
23. Woo, S., Park, J., Lee, J.Y., Kweon, I.S.: Cbam: Convolutional block attention module. In: *Proceedings of the European conference on computer vision (ECCV)*. pp. 3–19 (2018)
24. Yurdusev, A.A., Adem, K., Hekim, M.: Detection and classification of microcalcifications in mammograms images using difference filter and yolov4 deep learning model. *Biomedical Signal Processing and Control* **80**, 104360 (2023)

25. Zamir, R., Bagon, S., Samocha, D., Yagil, Y., Basri, R., Sklair-Levy, M., Galun, M.: Segmenting microcalcifications in mammograms and its applications. In: Medical Imaging 2021: Image Processing. vol. 11596, pp. 788–795. SPIE (2021)
26. Zhang, Y., Tang, Y., Cao, Z., Han, M., Xiao, J., Ma, J., Chang, P.: Bi-rads classification of calcification on mammograms. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VII 24. pp. 119–128. Springer (2021)