BI-RADS Classification of Calcification on Mammograms

Yanbo Zhang[1], Yuxing Tang[1], Zhenjie Cao[1], Mei Han[1], Jing Xiao[2], Jie Ma[3], and Peng Chang[1]

[1] PAII Inc., Palo Alto, USA
pengchang@gmail.com
[2] Ping An Insurance (Group) Company of China, Shenzhen, China
[3] Department of Radiology, Shenzhen People's Hospital, Shenzhen, China

Abstract. Calcification is one of the most common and important lesions in mammograms, and a higher BI-RADS category indicates a higher cancer risk. In this paper, we present the first deep learning-based six-class BI-RADS classification for each individual calcification in mammograms. We propose an attention ROI generation strategy to highlight calcification features. Moreover, by incorporating malignancy information, the designed new loss function effectively boosts the performance of the model. We also design a novel evaluation metric for BI-RADS classification, which considers the severity of malignancy. Experimental results have demonstrated the superior classification performance of the proposed approach to the competing methods.

Keywords: Mammogram · Deep Learning · BI-RADS · Calcification · Quadratic weighted kappa

1 Introduction

According to the World Health Organization, breast cancer became the most commonly diagnosed type of cancer worldwide, with 2.3 million cases in 2020, surpassing the number of new cases of lung cancer for the first time. Breast cancer now accounts for 11.7% of all new annual cancer cases globally [20]. Mammography is the best and widely used approach for early detection of breast cancer [4], with about 39 million mammograms performed each year in the United States.

Breast calcifications are common findings on mammograms. While most breast calcifications are benign (noncancerous), some calcification patterns, such as tight clusters with irregular shapes and fine appearance, may indicate breast cancer or precancerous changes in breast tissue. To assess the cancer risk and assure the imaging quality, the American College of Radiology has established the Breast Imaging-Reporting and Data System (BI-RADS) [1]. BI-RADS category or level reporting enables radiologists to standardize mammogram interpretation and demonstrates a close correlation with the risk of breast malignancy: 1-healthy, 2-benign, 3-probably benign, 4-suspicious abnormality, 5-highly suspicious of malignancy, 6-biopsy proven malignant.

Despite the recent advances in deep neural networks (DNN)-based approaches for computer-aided diagnosis/detection (CADx/CADe) in mammography [6,11–14,21–23], automated BI-RADS categorization of breast calcifications with robust performance remains a challenging issue, due to the fuzzy nature of the calcifications. Previous works either classified calcifications simply as benign or malignant [17], or classified them into incomplete BI-RADS categories. For example, Avalos-Rivera et al. [2] developed an artificial neural network to classify calcification ROIs (Region of Interest) into BI-RADS categories of 2, 3, and 4 (ignored 5), which does not meet the current BI-RADS standard. Some existing works have studied the BI-RADS categorization of whole mammographic images [7,10,18] instead of individual lesions or a certain lesion type (e.g., mass or calcification), which posed significant difficulties on the interpretability of the CADx/CADe system. Generally, radiologists evaluate each lesion in the breast and report the largest BI-RADS score for the breast. Hence, BI-RADS category prediction at the lesion-level is more advantageous to assist radiologists in clinical practice. Recently, there have been growing research interests on mass classification [5]. Yet hardly any work has been done for the classification of breast calcifications using the latest BI-RADS standard.

BI-RADS categorization is different from the traditional multi-class classification problem (e.g., natural object or lesion classification), in the sense that the BI-RADS levels are essentially a series of ordinal and discrete labels, which are inherently ordered according to the likelihood of malignancy. Directly applying multi-class classification models with cross-entropy loss that does not consider ordinal information is thus of inferior performance. In addition to the subjective BI-RADS labels assessed by the radiologists, binary biopsy results are considered to be the gold standard for lesion malignancy. Generally, these two sources of labels are highly correlated: biopsy-proven benign/malignant calcifications are expected to have smaller/larger BI-RADS scores.

In this work, we consider this label consistency and explicitly model it into our ordinal classification loss function, to penalize BI-RADS predictions that are inconsistent with biopsy results. More specifically, we design a malignancy adjusted, weighted BI-RADS classification loss that penalizes heavily for inconsistent predictions (e.g., predicting BI-RADS 5 for a biopsy-proven benign lesion, or BI-RADS 2 for a malignant lesion) and lightly for consistent predictions (e.g, predicting BI-RADS 2 for a benign lesion and 5 for a malignant lesion). Besides, since there might be multiple calcifications in an image patch, we design an attention mechanism to pre-process the input patch such that the network could focus on the specific calcification region to extract more discriminative features. To evaluate the performance of the proposed approach, we also propose a new metric that considers the malignancy of the calcification for a fair evaluation.

The main contribution of this work is three-fold: 1) we develop the first six-class BI-RADS classification algorithm for each calcification in mammograms, and 2) we design a specific attention-based ROIs and the malignancy adjusted loss that effectively boosts the feature learning and model optimization; 3) we introduce a novel metric, malignancy adjusted quadratic weighted

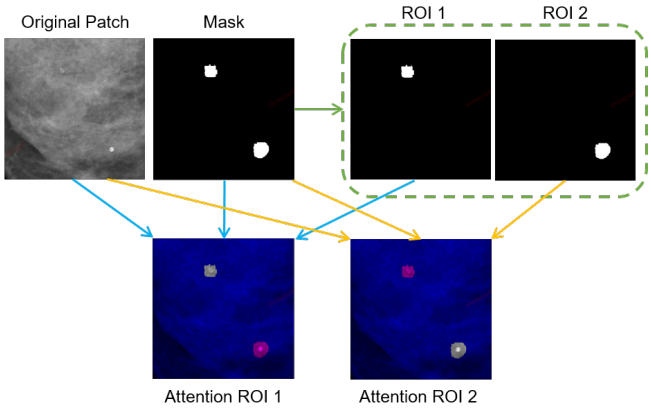

Fig. 1. The generation of calcification attention ROIs.

Kappa (MAQWK), which is a general evaluation metric for rating tasks that can be used in the medical imaging domain.

2 Methods

2.1 Method Overview

We propose a six-class (BI-RADS 2, 3, 4A, 4B, 4C, and 5) classification method for each individual calcification in mammograms. We first obtain calcification masks and BI-RADS categories annotated by radiologists, together with biopsy results. Image patches containing calcifications are then extracted, and we use masks to highlight the calcifications. In this way, the calcification patches with attention ROIs are generated as input data as described in section 2.2. With malignancy information, we introduce a weighting coefficient to the loss function as derived in section 2.3. The malignancy adjusted loss makes predicted BI-RADS more acceptable in practical applications. Finally, we design a novel evaluation metric for medical imaging rating tasks in section 2.4.

2.2 Attention-based Pre-processing

Fig. 1 shows the generation of calcification attention ROI patches. For an image patch extracted from the original image and the corresponding binary mask, we aim to predict the BI-RADS category for each individual calcification. As shown in Fig. 1, an ROI mask for each calcification is split from the mask of each patch. Each generated attention ROI consists of three channels: the first channel is the original image patch, providing all breast tissue information within this patch; the second channel is the element-wise product of the original patch and the mask, highlighting all the calcifications and their spatial relations; and the third

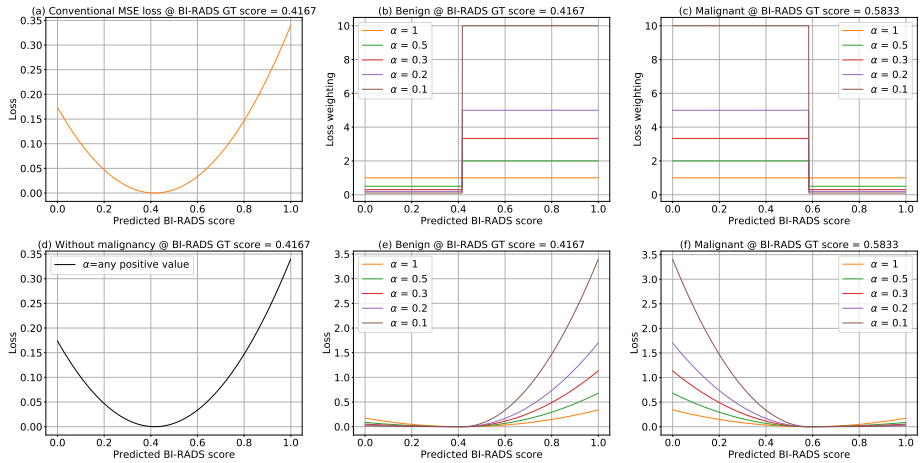

Fig. 2. Malignancy adjusted BI-RADS loss. (a) is the conventional MSE loss, (b) and (c) present the loss weighting of benign and malignant cases, respectively, and (e) and (f) are the corresponding loss curves. For the case without malignancy, the loss curve in (d) is identical to (a).

channel is the element-wise product of the original patch and one ROI mask, focusing on the calcification to be classified. The attention-based ROIs with such a design contain richer information from various perspectives, therefore, the features extracted from these patches are more meaningful spatially to enhance the classification performance.

We treat the classification for ratings as a regression task. The inputs to the classification model are the generated attention ROI patches, and the output is a scalar between [0, 1], representing the approximate risk probability. Consequently, this scalar is linearly mapped to predict a BI-RADS category.

2.3 Malignancy Adjusted Loss

According to the BI-RADS standard, score 1 means no lesions, 0 indicates incomplete information and requires follow-up, and 6 represents biopsy-proven breast cancer. We exclude the above three clearly defined BI-RADS categories and only consider the remaining six BI-RADS categories: 2, 3, 4A, 4B, 4C, and 5, for calcification classification as a rating problem. Let the number of calcification classes be N[4], and the cancer risk consistently increases from the first class to the Nth class. Thus, we convert the classification task into a linear regression task. We uniformly divide the risk range [0, 1] into N segments, thus the value range and center for the nth class range are $[(n-1)/N, n/N]$ and $(2n-1)/2N$.

[4] Previous BI-RADS standard utilizes a single category 4 instead of subcategories 4A, 4B, 4C, hence the number of calcification classes reduces to $N = 4$.

The regression loss using the conventional Mean Squared Error (MSE) is defined as:

$$Loss(b^{gt}, b^{pred}) = \left(b^{gt} - b^{pred}\right)^2, \tag{1}$$

where b^{gt} and b^{pred} denote the BI-RADS risk score of ground truth and prediction, respectively, and b^{gt} is the center of the ground truth segment. We use the following formula to measure if the bias direction of the predicted BI-RADS score is consistent with the biopsy malignancy:

$$C(b^{gt}, b^{pred}, m^{gt}) = -\left(b^{gt} - b^{pred}\right)\left(m^{gt} - m^{mid}\right), \tag{2}$$

where m^{gt} is the binary biopsy malignancy, with 0 and 1 indicate benign and malignant, respectively, m^{mid} is an auxiliary constant, which is set to be 0.5. If a mammogram does not have a biopsy result, m^{gt} is assigned with the same value as m^{mid}. A positive $C(\cdot)$ indicates consistency and a negative $C(\cdot)$ represents inconsistency between the BI-RADS prediction and the biopsy label, while $C(\cdot) = 0$ indicates unavailable biopsy malignancy or an accurate BI-RADS prediction. We give a higher penalty to the inconsistent case and a smaller one to the consistent case. Therefore, we propose a malignancy-adjusted loss weight:

$$W(b^{gt}, b^{pred}, m^{gt}) = \alpha^{\frac{C(b^{gt}, b^{pred}, m^{gt})}{|C(b^{gt}, b^{pred}, m^{gt})| + \epsilon}}, \tag{3}$$

where α is a weighting coefficient with the value range $(0, 1]$. ϵ is a very small positive constant to ensure that $W(\cdot)$ in Eq. 3 is continuous and differentiable.

The final weighted loss is denoted as:

$$Loss(b^{gt}, b^{pred}, m^{gt}) = \alpha^{\frac{C(b^{gt}, b^{pred}, m^{gt})}{|C(b^{gt}, b^{pred}, m^{gt})| + \epsilon}} \left(b^{gt} - b^{pred}\right)^2. \tag{4}$$

Fig. 2 shows some examples of the proposed malignancy-adjusted MSE loss for BI-RADS classification. We set loss weighting parameter α to 0.2 as default.

2.4 Evaluation metric

Quadratic weighted kappa (QWK) measures the agreement between two ratings [19]. This metric typically varies from 0 (random agreement between raters) to 1 (complete agreement between raters). A negative value means the classifier performs worse than random choice. The quadratic weighted kappa is calculated between the scores assigned by the human rater and the predicted scores. QWK is defined as:

$$k = 1 - \frac{\sum_{i,j} \mathbf{W}_{i,j} \mathbf{O}_{i,j}}{\sum_{i,j} \mathbf{W}_{i,j} \mathbf{E}_{i,j}}, \tag{5}$$

where $\mathbf{W}$ is a $N \times N$ matrix and its element $\mathbf{W}_{i,j} = (i - j)^2$ is the weighted cost associated with misclassifying label i as label j, and matrices $\mathbf{O}$ and $\mathbf{E}$ are the confusion matrix and the expected rating matrix [10], respectively.

Although QWK has been commonly used in medical image rating tasks, it is still not a proper metric for malignancy rating. As the examples presented

Table 1. A comparison of evaluation metrics on two classifiers. Consider an example list of six samples with various BI-RADS categories, all samples are correctly predicted except for the one with BI-RADS 4A. Both calcification example cases 1 and 2 are with the BI-RADS category 4A, and final confirmed as malignant and benign, respectively.

Cases	Classifiers	Predicted BI-RADS	QWK	MAQWK	
				MAQWK-M	MAQWK-B
Case1	Classifier 1	2	0.9024	0.8356	-
(4A, Malignant)	Classifier 2	4C	0.8919	**1.0**	-
Case 2	Classifier 1	4C	0.8919	-	0.8153
(4A, Benign)	Classifier 2	2	0.9024	-	**1.0**

in Table 1, the classifiers 1 and 2 predict case 1 (malignant, BI-RADS 4A) as BI-RADS 2 and 4C. Because these two predictions have the same distance to the ground truth, they have very close QWK values. However, in clinical settings, BI-RADS 2 is considered as benign and patients with BI-RADS 4C are required to conduct further actions to confirm its malignancy. Thus, classifier 1 may result in missed diagnosis for case 1. Likewise, classifier 1 may also lead to a false alarm for case 2. Although classifier 2 performs much better than classifier 1, they have extremely close QWK values.

For a more appropriate evaluation, we propose a malignancy-adjusted quadratic weighted kappa (MAQWK) and define the malignancy adjusted weighting matrix as:

$$\mathbf{W}^c = \mathbf{W} * \mathbf{I}^{u/l}, \tag{6}$$

where "c" is an indicator of either "m" or "b", indicating malignant/benign case, respectively. $\mathbf{W}^c$ represents the designed weighting matrix, $\mathbf{I}^{u/l}$ is an unit upper triangular matrix $\mathbf{I}^u$ for malignant or an unit lower triangular matrix $\mathbf{I}^l$ for benign. The matrices $\mathbf{I}^u$ and $\mathbf{I}^l$ find out the elements of higher/lower predicted BI-RADS cases in the confusion matrix, respectively. The operator $*$ means the element-wise product, therefore, $\mathbf{W}^c$ is either an upper or lower triangular part of the original quadratic weight matrix $\mathbf{W}$. In this way, $\mathbf{W}^m$ considers only the cases where the predicted BI-RADS is lower than the ground truth for malignant cases, and $\mathbf{W}^b$ considers only the cases where the predicted BI-RADS is higher than the ground truth for benign cases. MAQWK is defined as:

$$k^c = 1 - \frac{\sum_{i,j} \mathbf{W}^c_{i,j} \mathbf{O}^c_{i,j}}{\sum_{i,j} \mathbf{W}^c_{i,j} \mathbf{E}^c_{i,j}}, \tag{7}$$

where $\mathbf{O}^c$ represents either $\mathbf{O}^m$, which is the confusion matrix computed with only malignant cases, or $\mathbf{O}^b$, which is the counterpart of benign cases. Likewise, $\mathbf{E}^m$ and $\mathbf{E}^b$ are corresponding expected matrices for malignant and benign cases, respectively. We denote k^m and k^b as the malignancy adjusted quadratic weighted kappa for malignant (MAQWK-M) and benign (MAQWK-B), respectively. Table 1 shows that classifier 2 achieves higher values than classifier 1 in

Table 2. Number of calcification in each BI-RADS categories.

BI-RADS	2	3	4A	4B	4C	5	**Total**
Number	3180	1751	323	241	778	902	**7175**

terms of both MAQWK-M and MAQWK-B, implying a more reasonable metric for this task.

3 Experimental Setting

3.1 Datasets

We collaborated with a hospital to build a dataset for this task[5]. Mammograms were collected with two vendors' digital mammography machines, the SIEMENS Mammomat Inspiration (Germany) and the GIOTTO Image MD (Italy). Two radiologists delineated calcification regions and labeled the corresponding BI-RADS categories, before being finally checked by an experienced radiologist. The biopsy results indicating benign or malignant were confirmed by histopathology. The dataset consists of 708 patients with 1776 mammograms containing calcifications. There were 2731 malignant and 5426 benign calcifications, and the numbers of BI-RADS categories are listed in Table 2, where BI-RADS 0 and 6 are excluded. The patients were randomly split by 3:1:1 as the training, validation and test sets. A sliding window moved in mammograms with a step of 100 pixels to extract the image and mask patches, each patch was 400 $\times$ 400 pixels.

3.2 Implementation details

In this work, we use the ResNet-18 [9] as our DNN backbone (pre-trained on ImageNet). First, an attention ROI is created from original image and mask as the input of backbone, and the output of backbone is a cancer risk score. In the training stage, the obtained risk score is directly used for computing the loss. In the inference stage, if this score falls in the range $[(n-1)/6, n/6]$, then the predicted BI-RADS category is the nth category of the list [2, 3, 4A, 4B, 4C, 5].

We applied SGD with momentum as the optimizer for training. The initial learning rate was 0.001 and decreases by 0.3 every 10 epochs, with momentum as 0.99 and batch size as 32. A sampling strategy was used to balance the number of various BI-RADS classes in each mini-batch. The training stopped after 50 epochs. The method was implemented on a Linux workstation with two NVIDIA V100 GPUs (16G memory each).

Table 3. Performance comparison of various methods: standard multi-class (MC) classification, the regression-based method with MSE loss and the proposed MAMSE loss, and the impact of attention ROIs.

Method	QWK	MAQWK	
		MAQWK-M	MAQWK-B
MC	0.6583	0.1121	-0.0027
MSE	0.6877	0.1544	0.0302
MAMSE	0.7657	0.1468	0.1054
MC + attention ROI	0.8216	0.2517	0.2493
MSE + attention ROI	0.8650	0.2858	0.3242
MAMSE + attention ROI (Proposed)	**0.8870**	**0.3489**	**0.3786**

4 Experimental Results

Table 3 presents the performance comparison of standard multi-class (MC) classification, the regression-based method with standard MSE loss and with the proposed MAMSE loss. To show the impact of the proposed attention ROIs, we also compare them with the original gray image patches as the input for all models. For a fair comparison, all methods used the same backbone and training settings. It is clear that the introduced attention strategy remarkably promotes classification performance in terms of QWK and MAQWK for all methods. Moreover, the designed specific loss can further improve the performance, achieving the best results.

Fig. 3 shows the performance of the proposed method with different choices of loss weighting parameter α. The proposed approach is able to obtain relatively stable high performance when α is between 0.2 and 0.5. QWK and MAQWK fluctuate when α is approaching 0, and both decline steadily with the increase of α toward 1. The proposed method degrades to conventional MSE-based regression when α is 1.

5 Discussion and Conclusion

The core contributions of the proposed model are the designed loss and attention mechanism, which are independent of architecture. Hence, without loss of generality, we only used ResNet as the backbone to evaluate the efficacy of our method in this work. Other advanced architectures can be directly employed as a backbone, and similar advantages of the proposed modules can be expected.

The designed malignancy adjusted loss and the corresponding evaluation metrics consider the practical needs in clinical settings. They can be easily applied to other medical image rating tasks, such as for liver (LI-RADS) [15], gynecology (GI-RADS) [3], colonography (C-RADS) [16], diabetic retinopathy

[5] This retrospective case-control study was approved by the ethics review and institutional review board, which waived the requirement for individual informed consent.

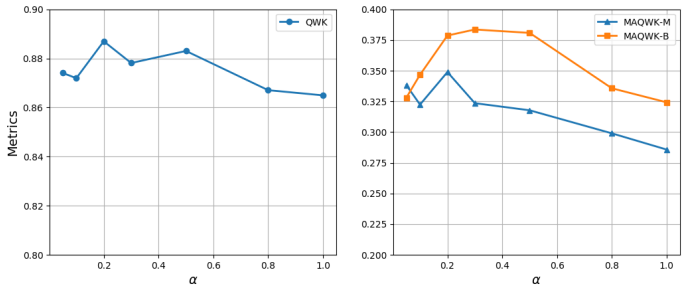

Fig. 3. Plots of metrics of the proposed method with different choices of α.

diagnosis [8], etc. The incorporation of the malignancy-based adjustment is an intuitive and interpretable way to transfer the domain knowledge and special request from the professionals to the DL framework. The loss is also flexible, accommodating the samples without malignancy information by assigning the pseudo label m^{mid}.

In this paper, we proposed a specific BI-RADS classification method and an evaluation metric for mammographic calcifications. The developed attention strategy and malignancy adjusted MSE loss effectively improve the classification performance. It shows great potential to be expanded to other tasks and domains. We have invited the radiologists from our collaborating hospitals to further evaluate its practical efficacy and other use cases.

References

1. American College of Radiology: ACR BI-RADS Atlas: Breast Imaging Reporting and Data System; Mammography, Ultrasound, Magnetic Resonance Imaging, Follow-up and Outcome Monitoring, Data Dictionary. ACR, American College of Radiology (2013) 1
2. Avalos-Rivera, E.D., Pastrana-Palma, A.: Classifying region of interests from mammograms with breast cancer into birads using artificial neural networks. Advances in Science, Technology and Engineering Systems Journal **2**(3), 233–240 (2017) 2
3. Basha, M.A.A., Refaat, R., Ibrahim, S.A., Madkour, N.M., Awad, A.M., Mohamed, E.M., El Sammak, A.A., Zaitoun, M.M., Dawoud, H.A., Khamis, M.E., et al.: Gynecology imaging reporting and data system (gi-rads): diagnostic performance and inter-reviewer agreement. European radiology **29**(11), 5981–5990 (2019) 8
4. Bleyer, A., Welch, H.G.: Effect of three decades of screening mammography on breast-cancer incidence. New England Journal of Medicine **367**(21), 1998–2005 (2012) 1
5. Boumaraf, S., Liu, X., Ferkous, C., Ma, X.: A new computer-aided diagnosis system with modified genetic feature selection for bi-rads classification of breast masses in mammograms. BioMed Research International **2020** (2020) 2
6. de Cea, M.V.S., Diedrich, K., Bakalo, R., Ness, L., Richmond, D.: Multi-task learning for detection and classification of cancer in screening mammography. In: Inter-

national Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 241–250. Springer (2020) 2
7. Geras, K.J., Wolfson, S., Shen, Y., Wu, N., Kim, S., Kim, E., Heacock, L., Parikh, U., Moy, L., Cho, K.: High-resolution breast cancer screening with multi-view deep convolutional neural networks. arXiv preprint arXiv:1703.07047 (2017) 2
8. Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., et al.: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA **316**(22), 2402–2410 (2016) 9
9. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016) 7
10. Liu, X., Fan, F., Kong, L., Diao, Z., Xie, W., Lu, J., You, J.: Unimodal regularized neuron stick-breaking for ordinal classification. Neurocomputing **388**, 34–44 (2020) 2, 5
11. Liu, Y., Zhang, F., Zhang, Q., Wang, S., Wang, Y., Yu, Y.: Cross-view correspondence reasoning based on bipartite graph convolutional network for mammogram mass detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3812–3822 (2020) 2
12. Liu, Y., Zhou, Z., Zhang, S., Luo, L., Zhang, Q., Zhang, F., Li, X., Wang, Y., Yu, Y.: From unilateral to bilateral learning: Detecting mammogram masses with contrasted bilateral network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 477–485. Springer (2019) 2
13. Lotter, W., Diab, A.R., Haslam, B., Kim, J.G., Grisot, G., Wu, E., Wu, K., Onieva, J.O., Boyer, Y., Boxerman, J.L., et al.: Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. Nature Medicine pp. 1–6 (2021) 2
14. McKinney, S., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C., King, D., Shetty, S.: International evaluation of an ai system for breast cancer screening. Nature **577**, 89–94 (01 2020) 2
15. Mitchell, D.G., Bruix, J., Sherman, M., Sirlin, C.B.: Li-rads (liver imaging reporting and data system): Summary, discussion, and consensus of the li-rads management working group and future directions. Hepatology **61**(3), 1056–1065 (2015) 8
16. Pooler, B.D., Kim, D.H., Lam, V.P., Burnside, E.S., Pickhardt, P.J.: Ct colonography reporting and data system (c-rads): benchmark values from a clinical screening program. American journal of roentgenology **202**(6), 1232–1237 (2014) 8
17. Shen, L., Rangayyan, R.M., Desautels, J.L.: Detection and classification of mammographic calcifications. International Journal of Pattern Recognition and Artificial Intelligence **7**(06), 1403–1416 (1993) 2
18. Vanderheyden, R., Xie, Y.: Mammography image bi-rads classification using ohplall. In: 2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService). pp. 120–127. IEEE (2020) 2
19. Warrens, M.J.: Some paradoxical results for the quadratically weighted kappa. Psychometrika **77**(2), 315–323 (2012) 5
20. World health Organization: Breast cancer now most common form of cancer: WHO taking action. https://www.who.int/news/item/03-02-2021-breast-cancer-now-most-common-form-of-cancer-who-taking-action 1

21. Wu, N., Phang, J., Park, J., Shen, Y., Huang, Z., Zorin, M., Jastrzebski, S., Fevry, T., Katsnelson, J., Kim, E., Wolfson, S., Parikh, U., Gaddam, S., Lin, L., Ho, K., Weinstein, J., Reig, B., Gao, Y., Pysarenko, H., Geras, K.: Deep neural networks improve radiologists' performance in breast cancer screening. IEEE transactions on medical imaging **39**(4), 1184–1194 (2019) 2
22. Yang, Z., Cao, Z., Zhang, Y., Han, M., Xiao, J., Huang, L., Wu, S., Ma, J., Chang, P.: Momminet: Mammographic multi-view mass identification networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 200–210. Springer (2020) 2
23. Zhang, F., Luo, L., Sun, X., Zhou, Z., Li, X., Yu, Y., Wang, Y.: Cascaded generative and discriminative learning for microcalcification detection in breast mammograms. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 12578–12586 (2019) 2