Semantic Feature Modulation for Mammographic Lesion Classification

Shahar Mahpod[1][0000−0001−8473−3356] and Gil Ben Artzi[1⋆][0000−0001−7449−5163]

Ariel University, Ariel, Israel {mahpods,gilba}@ariel.ac.il
https://gil-ba.com

Abstract. Most mammography AI systems predict benign versus malignant from the image alone, ignoring the structured reasoning radiologists perform through the BI-RADS descriptor lexicon. We introduce a conditional modulation framework that dynamically shifts intermediate feature statistics based on BI-RADS descriptors. Rather than considering descriptors as static labels, our method produces *sample-specific* modulation parameters: the same attribute generates different feature-space transformations depending on the image, capturing context-dependent diagnostic meaning. We show that our approach is highly effective. Across CBIS-DDSM and CMMD, our model improves AUC by up to +8.3% and F1 by up to +5.0% over the image-only baseline. By representing descriptors as conditional continuous embeddings, our framework enables exact interpolation between attributes within the same BI-RADS category; interpolation experiments confirm monotonic risk paths aligned with clinical severity, despite training only on hard categorical labels. Our model enables a post-detection decision-support mode using probabilistic outputs from emerging BI-RADS prediction models. Perturbation analysis confirms approximately 25% of joint-setting predictions change when all descriptors are removed. [1]

Keywords: Mammography · BI-RADS · Conditional normalization · Feature modulation · deep learning.

1 Introduction

Mammographic screening is a cornerstone of early breast cancer detection, yet interpretation remains challenging [15]. Deep learning has achieved strong results [18, 13, 17, 11, 14], but most lesion classifiers operate as image-only predictors, missing a key aspect of clinical reading: radiologists interpret mammograms through the BI-RADS descriptor lexicon, using structured descriptors such as breast density, mass shape and margins, and calcification morphology and distribution to organize uncertainty [1]. This creates several gaps. First, the need for *context-dependent descriptor integration.* A BI-RADS attribute does not have a fixed diagnostic effect independent of the image; its meaning is image-specific.

⋆ Corresponding author: gilba@ariel.ac.il
[1] The code is available on https://github.com/shahrma/semantic-mammo

What is missing is a mechanism that makes the descriptor's influence on feature processing depend on the image itself. Second, the need for *sub-categorical descriptor resolution*. BI-RADS forces a discrete choice onto a continuous perceptual reality: a radiologist may perceive a margin as a mixture of two adjacent attributes. The exact mixture may carry clinically meaningful information that is not utilized. Third, the need for a deployable post-detection second-reader mode. Emerging models have been reported to predict BI-RADS descriptors [19, 20, 2]. However, their outputs are probabilistic, and no current framework deploys them in a semantically transparent way that enables what-if analysis over sub-categorical ambiguity.

We address these gaps by replacing the static affine parameters of selected Layer Normalization (LN) layers with parameters predicted from BI-RADS descriptors via modulation networks [6]. A key architectural consideration is that BI-RADS descriptors are a structured multi-attribute lexicon with heterogeneous semantics operating at different spatial scales. We therefore designate a dedicated modulation module per descriptor, match descriptor types to the network's feature hierarchy where global-context descriptors influence earlier stages and lesion-level descriptors influence later stages, and allow multiple attributes to be injected jointly through interleaved layers. This yields three contributions: **(1) Context-dependent descriptors improve classification.** Each descriptor is processed by a dedicated modulation network that predicts sample-specific (γ, β) parameters. Because the backbone features they modulate are image-dependent, the same attribute produces different transformations for different images, learning context-dependent meaning rather than a fixed global association. The F1 and AUC gains over baselines that use the same descriptors in context-free ways (Concat, SE-Gate, MTL) isolate the value of this image-conditioning. **(2) Exact sub-category interpolation.** Continuous embeddings support convex combinations of attributes within a category. At inference, a radiologist who perceives margins as 60% spiculated and 40% ill-defined can input this mixture directly. Sweeping interpolation weights yields monotonic risk trajectories aligned with clinical severity, despite training exclusively on hard labels (Sec. 4.6).**(3) Second-reader candidate.** As a practical consequence, our model can accept probabilistic descriptor vectors from emerging BI-RADS prediction models, yielding a post-detection decision-support tool whose reasoning can be challenged and explored by the radiologist.

2 Related Work

Deep learning in mammography. Modern approaches use CNNs for lesion detection and classification [13, 11, 17]. ConvNeXt [9] has emerged as a competitive modernized CNN applied to medical imaging tasks [12, 21].

Clinical metadata integration. *Late fusion* concatenates metadata with pooled visual features [5] but cannot modulate intermediate representations. *Multi-task learning* (e.g., BI-RADS-Net [19, 20]) jointly predicts malignancy and descriptors from shared features. *Attention-based fusion* [2] uses cross-attention

for descriptor–image combination. *SE-style gating* [4] learns channel-wise reweighting conditioned on metadata. All these strategies operate at the representation boundary or through global reweighting, without modulating normalization statistics during intermediate processing. Critically, none address the subcategorical interpolation problem: all require hard descriptor labels and cannot represent uncertainty *within* a BI-RADS category. Our work fills both gaps.

Conditional normalization. Prior work has introduced feature-wise affine modulation conditioned on auxiliary inputs [10, 6, 16], typically applied at selected layers for style transfer or to incorporate external embeddings. In medical imaging, conditional normalization has been explored (e.g., [7]), but not for structured descriptor integration in discriminative classification. Here, we employ Layer Normalization with a dedicated modulation network that processes BI-RADS descriptors, enabling mixtures of categories within each descriptor type, and injects them in an interleaved manner across layers.

3 Method

3.1 Problem Setting

Given a mammographic ROI patch $X \in \mathbb{R}^{H \times W}$ and T radiologist-derived descriptors $\{z^{(t)}\}_{t=1}^{T}$, we learn a function $f(X, \{z^{(t)}\}) \to y$, where $y \in \{0, 1\}$ denotes the binary classification label.

3.2 Conditional Layer Normalization

We adopt ConvNeXt-Base [9] as backbone. ConvNeXt uses Layer Normalization (LN) throughout:

$$\mathrm{LN}(\mathbf{h}) = \gamma \odot \frac{\mathbf{h} - \mu(\mathbf{h})}{\sigma(\mathbf{h})} + \beta, \tag{1}$$

where $\gamma, \beta \in \mathbb{R}^{L}$ are learnable affine parameters. Unlike BatchNorm, LN operates per-sample, making its affine parameters natural control points for sample-specific conditioning.

We replace the static (γ, β) of selected LN layers with parameters predicted from clinical descriptors. For each descriptor type t, a dedicated two-layer MLP $g^{(t)}$ maps the descriptor embedding $z^{(t)}$ to modulation parameters:

$$(\gamma^{(t)}, \beta^{(t)}) = g^{(t)}(z^{(t)}), \quad \gamma^{(t)}, \beta^{(t)} \in \mathbb{R}^{L_t}. \tag{2}$$

The conditioned normalization becomes:

$$\mathrm{CLN}_l(\mathbf{h}_l, z^{(t)}) = \gamma^{(t)} \odot \hat{\mathbf{h}}_l + \beta^{(t)}, \tag{3}$$

where $\hat{\mathbf{h}}_l$ is the mean-variance normalized activation. This mechanism implements *conditional feature recalibration*: the modulation network adjusts per-channel gains and biases to amplify features that are discriminative under the given clinical context and suppress those that are not.

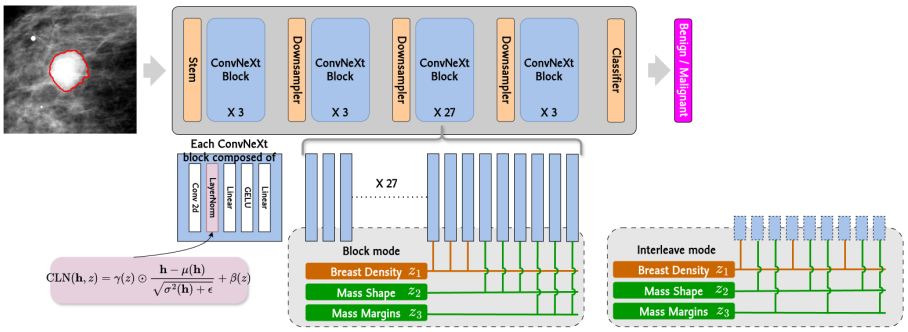

Fig. 1. Architecture overview. Dedicated MLP modulation networks predict (γ, β) for selected LN layers. **Left**: Block conditioning. **Right**: Interleaved conditioning.

3.3 Multi-Attribute Descriptor Encoding and Soft Inputs

Each descriptor is a categorical variable (e.g., mass shape $\in$ {round, oval, lobulated, irregular}). A learned embedding network (MLP) maps the descriptor $z^{(t)}$ to a dense vector in $\mathbb{R}^{d_e}$, trained jointly end-to-end. Missing descriptors default to a learned "unknown" token.

The input $z^{(t)} \in \mathbb{R}^N$ represents a descriptor with N possible attributes of this specific BI-RADS type. Each position $k \in [N]$, $z_k^{(t)}$, represents the existence of this specific attribute, where at training we only deploy hard categorical labels ($z_k^{(t)} \in \{0, 1\}$). This implies that our model supports multiple attributes and the framework naturally supports *soft descriptor inputs* at inference: a convex combination at the descriptor type, $z^{(t)} = (1-\alpha)\, z_A^{(t)} + \alpha\, z_B^{(t)}$ (with $z_A^{(t)}, z_B^{(t)}$ denoting the corresponding one-hot descriptor vectors), allows exact interpolation between adjacent BI-RADS attributes without retraining. The embedding network then maps this mixed descriptor to its dense representation. This enables what-if analysis where a radiologist can explore how sub-categorical ambiguity affects the malignancy prediction, and it also allows probabilistic descriptors from automated BI-RADS prediction models to be consumed directly.

3.4 Hierarchical Conditioning

Different descriptors carry information at different abstraction levels. We assign each descriptor type to a specific depth range: global descriptors (e.g., density) modulate earlier stages, lesion-level descriptors (e.g., margins, morphology) modulate later stages. We explore **block conditioning** (each descriptor modulates consecutive LN layers within one stage) and **interleaved conditioning** (descriptors alternate across layers within the same stage). Separate modulation networks per descriptor type preserve each descriptor's semantic identity (Fig. 1).

The conditioned backbone produces features via global average pooling, passed to a classification head trained with binary cross-entropy. Modulation outputs are initialized near identity ($\gamma \approx 1$, $\beta \approx 0$) to prevent early instability.

4 Experiments

4.1 Datasets

CBIS-DDSM [8]: 3,567 ROIs (2,924 train / 643 test), 58.5% benign, patient-level split. **CMMD** [3]: 1,642 ROIs (1,360 train / 282 test), 89% malignant, patient-level split. Both use stratified 5-fold cross-validation for hyperparameter selection and report metrics over 5 independent runs. ROIs are extracted at 1152×1152 pixels, resized to 384×384, with standard augmentations (rotation, flips, jittering, noise, random erasing). Final inputs are $3 \times 224 \times 224$.

4.2 Implementation Details

All models use ConvNeXt-Base (ImageNet-pretrained). Training: Adam, lr 1×10^{-5}, weight decay 5×10^{-4}, batch size 8, 200 epochs, 10-epoch warmup, Multi-Step LR decay at epochs 60/120/160 (factor 0.2). Mass classification uses density, shape, and margins; calcification uses density, morphology, and distribution; the joint setting uses all six plus lesion type. Each descriptor feeds a dedicated two-layer MLP. Conditioning targets Stages 3 (27 blocks) and 4 (3 blocks) of ConvNeXt.

4.3 Baselines

All baselines share the same ConvNeXt-Base backbone and training protocol: **Base**: image-only. **Concat**: descriptor embeddings concatenated with pooled features. **SE-Gate**: SE-style [4] channel gating conditioned on descriptors. **MTL**: multi-task with auxiliary descriptor prediction heads (0.3× loss weight). **Cond-LN Block / Inter**: our method with block and interleaved conditioning.

4.4 Main Classification Results

Tables 1–2 present results across mass, calcification, and joint settings. Cond-LN Inter achieves the highest (or tied-highest) AUC in 5 of 6 dataset×setting combinations (strictly highest in 4/6). On CBIS-DDSM, it improves AUC over Base by +2.2%, +4.5%, and +3.4% for mass, calcification, and joint settings, respectively, and exceeds MTL by +3.1% on mass. On CMMD calcifications, it outperforms SE-Gate by +6.0% AUC. The CMMD-mass case illustrates a regime where descriptor headroom is limited, the subset's imbalanced malignant prior may further constrain how much headroom descriptor signal can add.

Table 1. Results on CBIS-DDSM. All values: mean±std (%) over 5 runs.

	Model	Acc	AUC	F1	Rec
Mass	Base	80.3±1.1	88.7±0.8	78.5±1.4	79.1±2.2
	Concat	79.5±2.1	88.3±0.7	77.5±2.6	77.4±3.5
	SE-Gate	80.7±1.4	87.8±0.5	78.4±1.7	77.1±2.2
	MTL	81.2±1.9	87.8±1.2	79.4±2.2	79.6±2.4
	Cond-LN Block	83.4±1.5	89.9±0.4	82.1±1.6	**84.1±3.3**
	Cond-LN Inter	**84.7±0.9**	**90.9±1.3**	**83.0±0.9**	81.8±2.7
Calc	Base	80.0±1.2	86.4±0.8	69.1±2.2	72.9±3.9
	Concat	79.1±1.8	86.5±1.1	67.7±2.5	71.4±2.6
	SE-Gate	79.5±0.8	86.6±0.7	68.9±1.0	74.0±1.6
	MTL	79.4±1.4	86.7±0.9	68.5±2.2	73.1±2.8
	Cond-LN Block	80.6±1.6	88.8±0.7	70.9±2.3	77.1±2.5
	Cond-LN Inter	**82.8±1.1**	**90.9±0.4**	**74.1±0.7**	**80.2±4.0**
Both	Base	78.8±0.6	86.9±0.1	72.3±1.1	73.9±2.0
	Concat	78.9±1.1	87.1±0.7	72.8±1.1	75.2±1.1
	SE-Gate	79.4±1.4	87.0±0.7	73.2±1.6	75.2±1.0
	MTL	80.2±0.8	88.2±0.4	74.7±0.9	**78.2±1.5**
	Cond-LN Block	**81.2±1.5**	**90.3±0.3**	74.2±2.4	72.2±3.8
	Cond-LN Inter	**81.2±0.7**	**90.3±0.4**	**75.0±1.7**	75.4±4.3

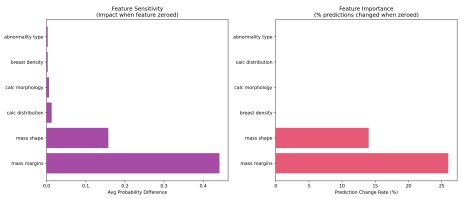

Fig. 2. Descriptor importance via perturbation on CBIS-DDSM joint. **Left:** Probability shift when each descriptor is zeroed. **Right:** Prediction flip rate. Mass margins and shape have the strongest influence.

4.5 Descriptor Sensitivity

Zeroing individual descriptors (Fig. 2) reveals that lesion-specific attributes, mass margins and shape, produce the largest probability shifts and flip rates, consistent with their diagnostic importance. When all descriptors are removed, accuracy drops from 81.0% to 68.0% and ∼25% of predictions change, confirming that the model substantively relies on descriptor conditioning for prediction.

4.6 Descriptor Interpolation and Monotonicity

To validate that the learned embedding space supports clinically coherent interpolation, we sweep descriptor embeddings as convex combinations $z^{(t)} = (1-\alpha)\, z_A^{(t)} + \alpha\, z_B^{(t)}$ for $\alpha \in \{0, 0.1, \ldots, 1.0\}$ across all 643 CBIS-DDSM test images, regardless of original descriptor, pathology, or lesion type. We evaluate mass shape (ROUND → OVAL → LOBULATED → IRREGULAR) and mass margins (CIRCUMSCRIBED → MICROLOBULATED → SPICULATED).

Table 2. Results on CMMD. All values: mean±std (%) over 5 runs.

	Model	Acc	AUC	F1	Rec
Mass	Base	85.9±0.7	**86.9±0.3**	91.8±0.5	94.5±1.7
	Concat	86.8±0.8	85.4±3.7	**92.3±0.5**	94.8±1.1
	SE-Gate	86.8±0.8	86.8±1.4	**92.3±0.5**	**95.1±0.6**
	MTL	**86.9±1.2**	85.4±1.6	**92.3±0.7**	94.4±1.4
	Cond-LN Block	86.3±1.4	85.9±1.5	92.0±0.8	94.8±1.1
	Cond-LN Inter	86.6±0.7	86.3±0.6	92.1±0.4	94.3±0.7
Calc	Base	86.5±2.0	83.3±1.7	92.5±1.1	95.3±1.0
	Concat	86.5±1.5	83.2±2.0	92.5±0.8	94.6±1.2
	SE-Gate	87.6±1.0	85.6±1.1	93.1±0.6	95.7±0.7
	MTL	88.0±0.9	82.8±1.5	93.4±0.5	**96.2±1.0**
	Cond-LN Block	87.3±2.4	86.4±2.7	92.8±1.4	94.2±3.0
	Cond-LN Inter	**89.7±1.7**	**91.6±1.1**	**94.1±1.0**	93.4±3.2
Both	Base	85.8±0.8	87.3±1.3	91.8±0.5	94.2±1.0
	Concat	85.5±0.6	85.8±1.2	91.7±0.3	93.9±0.9
	SE-Gate	86.5±0.7	87.3±0.7	92.2±0.4	94.0±0.8
	MTL	86.5±0.3	87.6±0.8	92.2±0.2	93.7±0.3
	Cond-LN Block	**88.7±0.5**	90.2±0.4	**93.6±0.3**	**97.2±0.6**
	Cond-LN Inter	88.1±0.5	**91.7±0.8**	93.2±0.3	96.6±0.4

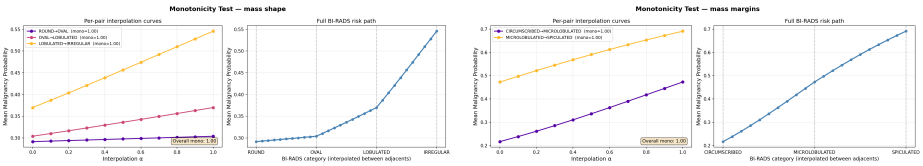

Fig. 3. Descriptor interpolation on CBIS-DDSM (joint model). **Left pair:** Mass shape. Risk rises from 0.29 (ROUND) to 0.55 (IRREGULAR). **Right pair:** Mass margins. Risk spans 0.22 (CIRCUMSCRIBED) to 0.69 (SPICULATED), a 47pp range. All adjacent pairs exhibit monotonic increase (mean score = 1.00) despite training only on hard categorical labels.

All interpolation curves are monotonically increasing (score = 1.00 for every adjacent pair). Mean malignancy probability rises from 0.29 to 0.55 for shape and from 0.22 to 0.69 for margins (Fig. 3). Margins exhibit the larger dynamic range, consistent with their higher perturbation sensitivity (Fig. 2). The ordering aligns with clinical knowledge: irregular shapes and spiculated margins are the strongest malignancy indicators.

These results are important because the model was trained exclusively on hard one-hot labels and never saw soft embeddings. The monotonic response reflects *emergent* structure in the embedding space. Furthermore, monotonicity holds across the entire test set, including images whose ground-truth descriptor differs from the one being swept, indicating a robust generalizable semantic-to-risk mapping rather than a descriptor-specific lookup.

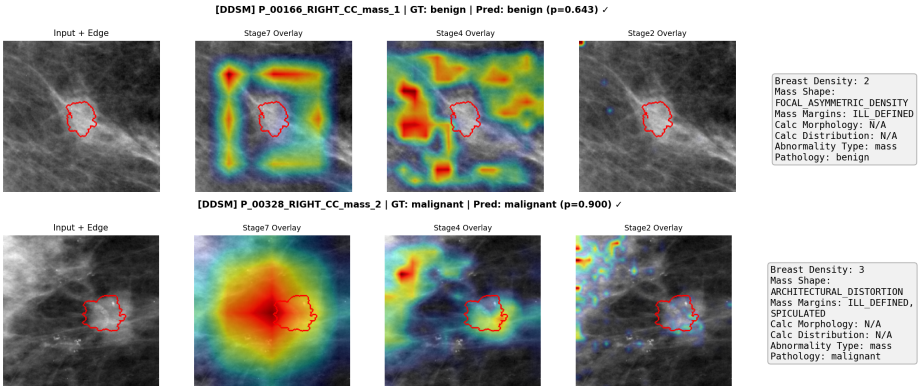

Fig. 4. Grad-CAM++ at multiple ConvNeXt depths for CBIS-DDSM mass cases. Benign cases (top row) exhibit diffuse patterns; malignant cases (bottom row) show increasingly localized activations.

4.7 Qualitative Analysis

Grad-CAM++ visualizations (Fig. 4) confirm that the conditioned model progressively focuses on clinically relevant regions, with deep-stage activations aligning with lesion boundaries in malignant cases.

5 Discussion

Emergent clinical coherence in embedding space. The interpolation experiment reveals that the model has learned an embedding geometry that faithfully mirrors BI-RADS risk ordering without explicit ordinal supervision. This has practical implications: it validates that soft descriptor inputs produce clinically meaningful outputs, supporting both the what-if analysis use case (contribution 2) and post-detection decision-support mode (contribution 3) where predicted descriptors are inherently probabilistic. An end-to-end coupling with an upstream automated BI-RADS predictor is a natural next step.

What descriptor integration adds. The additional gain from *how* descriptors are integrated, conditional LN versus concatenation, SE-gating, or multi-task learning, is consistent and most pronounced in settings where subtle morphological distinctions between benign and malignant patterns can be difficult (e.g., +5.2 F1 over SE-Gate for CBIS-DDSM calcifications). This suggests that conditional processing tied to the specific image can be the differentiating factor in more challenging cases.

Interleaved vs. block conditioning. Interleaved conditioning provides equal or better performance with lower variance. Alternating descriptors within a stage encourages complementary interactions, with the strongest effect on calcification classification where subtle morphological patterns benefit from distributed modulation.

6 Conclusion

We presented a semantic feature modulation framework that significantly improves mammographic lesion classification, by integrating structured BI-RADS descriptors into a deep network via sample-specific processing. The framework supports practical cases of descriptor uncertainty as an input: adjacent-category ambiguity and probabilistic, missing, and unknown inputs. Systematic comparison shows that improvements arise from the conditioning mechanism itself. The approach operates in a post-detection setting suitable for integration into existing pipelines.

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

Acknowledgment We acknowledge the Ariel HPC Center at Ariel University for providing computing resources.

References

1. American College of Radiology: BI-RADS atlas: Breast imaging reporting and data system (2013)
2. Ben-Artzi, G., Daragma, F., Mahpod, S.: Deep bi-rads network for improved cancer detection from mammograms. In: Pattern Recognition. pp. 17–30. Springer Nature Switzerland, Cham (2025)
3. Cui, C., Li, L., Cai, H., Fan, Z., Zhang, L., Dan, T., Li, J., Wang, J.: The Chinese mammography database (CMMD): An online mammography database with biopsy confirmed types for machine diagnosis of breast. The Cancer Imaging Archive (2021)
4. Hu, J., Shen, L., Albanie, S., Sun, G., Wu, E.: Squeeze-and-excitation networks. IEEE Trans. Pattern Anal. Mach. Intell. **42**(8), 2011–2023 (Aug 2020)
5. Huang, S.C., Pareek, A., Seyyedi, S., Banerjee, I., Lungren, M.P.: Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. NPJ Digit Med **3**, 136 (Oct 2020)
6. Huang, X., Belongie, S.: Arbitrary style transfer in real-time with adaptive instance normalization. In: IEEE International Conference on Computer Vision (ICCV). pp. 1501–1510 (2017)
7. Karani, N., Erdil, E., Chaitanya, K., Konukoglu, E.: Test-time adaptable neural networks for robust medical image segmentation. Medical Image Analysis **68**, 101907 (2021)
8. 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**, 170177 (2017)
9. Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S.: A ConvNet for the 2020s. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 11976–11986 (2022)

10. Perez, E., Strub, F., de Vries, H., Dumoulin, V., Courville, A.C.: FiLM: Visual reasoning with a general conditioning layer. In: Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI). pp. 3942–3951 (2018)
11. Ribli, D., Horváth, A., Unger, Z., Pollner, P., Csabai, I.: Detecting and classifying lesions in mammograms with deep learning. Scientific Reports **8**(1), 4165 (2018)
12. Roy, S., Koehler, G., Ulrich, C., Baumgartner, M., Petersen, J., Isensee, F., Jäger, P.F., Maier-Hein, K.H.: MedNeXt: Transformer-driven scaling of ConvNets for medical image segmentation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. pp. 405–415. Springer Nature Switzerland (2023)
13. Shen, L., Margolies, L.R., Rothstein, J.H., Fluder, E., McBride, R., Sieh, W.: Deep learning to improve breast cancer detection on screening mammography. Scientific Reports **9**(1), 12495 (2019)
14. Shen, Y., Wu, N., Phang, J., Park, J., Liu, K., Tyagi, S., Heacock, L., Kim, S.G., Moy, L., Cho, K., Geras, K.J.: An interpretable classifier for high-resolution breast cancer screening images utilizing weakly supervised localization. Medical Image Analysis **68**, 101908 (2021)
15. Siegel, R.L., Giaquinto, A.N., Jemal, A.: Cancer statistics, 2024. CA Cancer J Clin **74**(1), 12–49 (Jan 2024)
16. de Vries, H., Strub, F., Mary, J., Larochelle, H., Pietquin, O., Courville, A.C.: Modulating early visual processing by language. In: Advances in Neural Information Processing Systems (NeurIPS) (2017)
17. Wu, N., Phang, J., Park, J., Shen, Y., Huang, Z., Zorin, M., Jastrzębski, S., Févry, T., Katsnelson, J., Kim, E., Wolfson, S., Parikh, U., Gaddam, S., Lin, L.L.Y., Ho, K., Weinstein, J.D., Reig, B., Gao, Y., Toth, H., Pysarenko, K., Lewin, A., Lee, J., Airola, K., Mema, E., Chung, S., Hwang, E., Samreen, N., Kim, S.G., Heacock, L., Moy, L., Cho, K., Geras, K.J.: Deep neural networks improve radiologists' performance in breast cancer screening. IEEE Transactions on Medical Imaging **39**(4), 1184–1194 (2020)
18. Yala, A., Lehman, C., Schuster, T., Portnoi, T., Barzilay, R.: A deep learning mammography-based model for improved breast cancer risk prediction. Radiology **292**(1), 60–66 (Jul 2019)
19. Zhang, B., Vakanski, A., Xian, M.: BI-RADS-Net: An explainable multitask learning approach for cancer diagnosis in breast ultrasound images. In: IEEE International Workshop on Machine Learning for Signal Processing (MLSP) (2021)
20. Zhang, B., Vakanski, A., Xian, M.: BI-RADS-Net-V2: A composite multi-task neural network for computer-aided diagnosis of breast cancer in ultrasound images with semantic and quantitative explanations. IEEE Access **11**, 79480–79494 (2023)
21. Zhang, H., Zhong, X., Li, G., Liu, W., Liu, J., Ji, D., Li, X., Wu, J.: BCU-Net: Bridging ConvNeXt and U-Net for medical image segmentation. Computers in Biology and Medicine **159**, 106960 (2023)