VMRA-MaR: An Asymmetry-Aware Temporal Framework for Longitudinal Breast Cancer Risk Prediction

Zijun Sun[1][0009−0004−1848−6966], Solveig Thrun[2][0009−0006−7349−9449], and Michael Kampffmeyer[2,3][0000−0002−7699−0405]

[1] Department of Computer Science and Engineering, University of Bologna, Bologna, Italy zijun.sun@studio.unibo.it
[2] Department of Physics and Technology, UiT The Arctic University of Norway, Tromsø, Norway
[3] Norwegian Computing Center, Oslo, Norway

Abstract. Breast cancer remains a leading cause of mortality worldwide and is typically detected via screening programs where healthy people are invited in regular intervals. Automated risk prediction approaches have the potential to improve this process by facilitating dynamically screening of high-risk groups. While most models focus solely on the most recent screening, there is growing interest in exploiting temporal information to capture evolving trends in breast tissue, as inspired by clinical practice. Early methods typically relied on two time steps, and although recent efforts have extended this to multiple time steps using Transformer architectures, challenges remain in fully harnessing the rich temporal dynamics inherent in longitudinal imaging data. In this work, we propose to instead leverage Vision Mamba RNN (VMRNN) with a state-space model (SSM) and LSTM-like memory mechanisms to effectively capture nuanced trends in breast tissue evolution. To further enhance our approach, we incorporate an asymmetry module that utilizes a Spatial Asymmetry Detector (SAD) and Longitudinal Asymmetry Tracker (LAT) to identify clinically relevant bilateral differences. This integrated framework demonstrates notable improvements in predicting cancer onset, especially for the more challenging high-density breast cases and achieves superior performance at extended time points (years four and five), highlighting its potential to advance early breast cancer recognition and enable more personalized screening strategies. Our code is available at this URL.

Keywords: Breast cancer · Longitudinal mammogram · Risk Prediction.

1 Introduction

Breast cancer remains the most prevalent malignancy worldwide, imposing a significant public health burden [1,8]. Early detection through precise risk prediction is pivotal for improving clinical outcomes and reducing mortality. Despite

debates over screening intervals, optimal starting age, and the utility of supplemental imaging, digital mammography remains the primary screening modality due to its cost-effectiveness and broad accessibility [21,6]. These ongoing controversies underscore the pressing need for enhanced risk assessment models that enable personalized screening regimens while minimizing unnecessary interventions [14].

While traditional risk models [5,19,2] have achieved only moderate predictive accuracy, deep learning–based methods have substantially advanced breast cancer risk stratification [3]. A seminal model is Mirai [23], which achieves strong performance by integrating risk factor data, managing missing information, and ensuring consistency across different mammography machines. Nonetheless, its primary focus on the most recent mammogram overlooks the abundant temporal information in sequential mammograms, an aspect that could reveal subtle indicators of disease progression and further refine early detection strategies.

There has been a growing interest in addressing this shortcoming by incorporating longitudinal imaging data. While early methods typically rely on two or three years of patient history [25,12], recent work such as LoMaR [10] expands upon this approach using a hybrid CNN–Transformer [20] architecture that leverages more extensive imaging history for more accurate risk predictions. Despite demonstrating the benefits of comprehensive longitudinal data, LoMaR and similar models often depend on positional embeddings, which may underutilize the full temporal richness available in large-scale clinical datasets.

In this study, we introduce VMRA-MaR, a novel recurrent-based framework designed to address limitations in modeling dynamic breast tissue changes over time. Drawing inspiration from AsymMirai [4], we further incorporate bilateral symmetry analysis into the risk prediction process, thereby accounting for contralateral asymmetry within an end-to-end learning pipeline. The model is validated using the CSAW-CC dataset from Karolinska University Hospital, which provides comprehensive longitudinal mammographic data. Our primary contributions are threefold: a) We propose a VMRNN-based [18] framework that leverages time-step information to track and model longitudinal breast tissue changes more effectively than models reliant on positional embeddings alone; b) We incorporate asymmetry features into risk prediction, building on AsymMirai's insight that contralateral asymmetry strongly correlates with future cancer risk, while extending its application to a longitudinal context; c) By utilizing multi-year sequential mammography data, our approach improves risk prediction accuracy and supports earlier detection of subtle disease progression, particularly in high-density cases.

2   Method

2.1   Overall Pipeline

Figure 1 provides an overview of our unified approach, which integrates and extends existing spatiotemporal and asymmetry-detection techniques to address
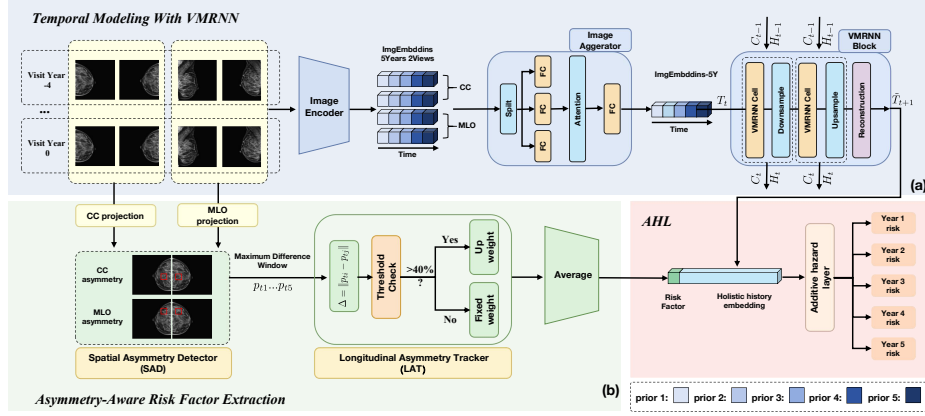
Fig. 1: Network architecture of our proposed VMRNN-Asymmetry Mammogram Risk model (VMRA-MaR).

limitations of prior models that underutilize temporal features and breast symmetry. Specifically, we fuse the temporal representations from the Vision Mamba RNN (VMRNN), with asymmetry cues within an Additive Hazard Layer to generate a five-year risk prediction. This multisource information fusion strategy more effectively captures evolving breast tissue characteristics, identifies persistent asymmetry signals, and offers a comprehensive view of each patient's risk trajectory.

Our model processes five yearly screenings, each comprising four standard mammographic views (LCC, RCC, LMLO, RMLO). At each time step t and viewpoint v, the corresponding image x_t^v is first encoded via a pre-trained image encoder; the resulting multi-view features are then fused using a Transformer module. These fused features are passed to the VMRNN, which learns longitudinal dependencies through the evolving hidden states $\{H^1, \ldots, H^T\}$. Concurrently, a Spatial Asymmetry Detector (SAD) identifies regional discrepancies between left-right views, and a Longitudinal Asymmetry Tracker (LAT) aggregates these differences over time to highlight persistent anomalies. Finally, the AHL combines both the temporal representations and asymmetry cues to produce a unified five-year risk prediction that accounts for evolving breast tissue features as well as persistent asymmetry signals.

2.2   Temporal Modeling within VMRA-MaR

The current state-of-the-art (SOTA) approach to leverage longitudinal imaging data is LoMaR [10], which leverages a Transformer backbone and positional embeddings to integrate the temporal information. However, this design relies heavily on the global attention mechanism of the transformer and its static positional embeddings, struggling to fully capture temporal coherent features. VMRA-MaR instead builds on a VMRNN block [18], that processes data recurrently updat-

ing hidden states at each time step, allowing us to address this shortcoming, resulting in improved accuracy when modeling longitudinal changes.

VMRNN Block The VMRNN Block (see Figure 1(a)) processes, at each time step t, a fused feature vector T_t (obtained from the Transformer Block shown in Figure 1(a)), along with the previous hidden state H_{t-1} and cell state C_{t-1}. The block incorporates downsampling and upsampling layers to capture multi-scale representations, and a reconstruction layer to restore spatial resolution for subsequent predictions.

Within the block, the VMRNN Cell first projects T_t and H_{t-1} into an intermediate representation using a linear projection (LP) layer that fuses and reshapes the features into a spatial format suitable for further processing by a VSS Block:

$$X_t = \text{LP}\left(T_t, H_{t-1}\right) \tag{1}$$

The VSS refines X_t into the gating signal F_t by fusing directional and direct spatial features. First, X_t is reshaped into an image-like tensor A. In the primary stream, a depth-wise 3×3 convolution followed by a SiLU activation produces.

$$A_1 = \text{SiLU}\left(\text{DWConv}_{3\times3}(A)\right) \tag{2}$$

which is expanded into four directional sequences $\{A_{1,v}\}_{v=1}^{4}$. Each sequence is processed by a directional module $S6$ [7], and the resulting directional features $\{\bar{A}_v\}$ are merged and normalized:

$$\bar{A}_v = S6\left(\text{expand}\left(A_1, v\right)\right), \quad A_3 = \text{LayerNorm}\left(\text{merge}\left(\bar{A}_1, \bar{A}_2, \bar{A}_3, \bar{A}_4\right)\right) \tag{3}$$

In parallel, a secondary stream applies a SiLU activation [9] to A. The two streams are combined by element-wise addition,

$$B_1 = \text{SiLU}(A), \ Y = A_3 + B_1 \tag{4}$$

and a sigmoid activation is then applied to produce $F_t = \sigma(Y)$. By fusing directional and direct spatial features, the VSS provides a context-aware gating signal that enhances the model's ability to capture fine-grained spatial dependencies across screening time steps.

Subsequently, Y is used to derive a gating signal F_t via a sigmoid $\sigma(Y)$ and the cell and hidden states are updated as follows:

$$C_t = F_t \odot \left(\tanh(Y) + C_{t-1}\right), \ H_t = F_t \odot \tanh\left(C_t\right) \tag{5}$$

2.3   Asymmetry-Aware Risk Factor Extraction

Building on the insight of AsymMirai [4] that asymmetry is a pivotal marker for risk prediction in single time-step mammographic screenings, we incorporate asymmetry into our end-to-end framework and extend their approach to further consider the longitudinal dynamics inherent in mammographic screenings. Our

novel VMRNN-based framework leverages a Spatial Asymmetry Detector (SAD) to capture interbreast differences and a Longitudinal Asymmetry Tracker (LAT) to integrate temporal variations, thereby enhancing overall risk evaluation.

In a single screening, following [4], SAD aligns the left and right feature maps L and R by flipping R to obtain R^F, and then computes the local feature differences D to determine the maximum asymmetry value $D_{\max}$ along with its spatial coordinates $\mathbf{p}$.

$$D_{\mathrm{norm}}(h, w) = \sqrt{\sum_{c=1}^{C} D_{c,h,w}^2}, \ \ D_{\max} = \max_{h,w} D_{\mathrm{norm}}(h, w), \tag{6}$$

where $\mathbf{p}$ denotes the coordinates at which $D_{\max}$ is attained. Building on the single time-step asymmetry, our VMRA-MaR consists of the LAT to analyze the Euclidean displacement of $\mathbf{p}$ across screenings; if the displacement is below a predefined threshold (40% of the window size), the region is considered persistently abnormal, and its risk score is upweighted via weighted fusion of historical measurements.

2.4 Additive Hazard Layer (AHL)

The additive-hazard layer (AHL) [11] integrates the entire screening history embedding, $\mathcal{H}$, with an asymmetry-aware risk factor, r_{AA}, extracted from the dedicated module. These components are concatenated to form $\widetilde{R} = \mathrm{concat}\,[\mathcal{H}, r_{AA}]$, which serves as the input for subsequent risk estimations. A baseline risk, $B(\widetilde{R})$, is computed via a $Linear_k(\widetilde{R})$. In parallel, timespecific marginal hazards are estimated through separate linear layers followed by ReLU activations, thereby ensuring non-negative hazard increments and enforcing the monotonicity of cumulative risk predictions. Formally, for each time interval K, the marginal hazard is defined as:

$$H_k(\widetilde{R}) = \mathrm{ReLU}\left(\mathrm{Linear}_k(\widetilde{R})\right) \tag{7}$$

and the cumulative risk of developing cancer by time k is given by

$$P\left(t_{\mathrm{cancer}} \leq k \mid \widetilde{R}\right) = B(\widetilde{R}) + \sum_{i=0}^{k-1} H_i(\widetilde{R}) \tag{8}$$

which facilitates end-to-end optimization via log-likelihood maximization on the observed screening follow-up data while ensuring temporal consistency in risk estimation.

3 Experiments

3.1 Dataset

Following LoMAR [10], this study uses the publicly available CSAW-CC (mammography) dataset [17] to ensure reproducibility. In total, the dataset includes

873 breast cancer cases out of 1,103 initially identified patients, and 7,850 healthy controls out of 10,000 randomly selected individuals. For each subject, all screening mammograms from multiple time points are provided, along with the dates of examination and patient age. Additionally, pixel-level tumor annotations are included for cases, indicating the precise region of the tumor. We adhere to the preprocessing protocol outlined in LoMaR [10]; detailed information on class distribution can be found therein, as the dataset utilized in the current study closely aligns with that work due to the identical preprocessing strategy.

3.2   Experimental details

We initialized the VMRNN blocks with the publicly released weights from [18] and initialized the image encoder with Mirai [23] weights. Additionally, the image embeddings used by the VMRNN and asymmetry block are obtained from Mirai's image encoder, which remains frozen during training. We employ AdamW with an initial learning rate of $1e-3$, weight decay of $1e-4$, and a cosine decay schedule. The number of training epochs is 30. To handle class imbalance and improve risk prediction accuracy, we use a weighted cross-entropy loss with class weights derived from the training set distribution, following LoMaR[10]. Following, conventional risk prediction assessment [23,10], we utilized the C-index and area under the receiver operating characteristic (ROCAUC) curve to evaluate the performance of all methods in the 1–5 year risk prediction task. All experiments were conducted using $8\times$ NVIDIA A100 GPUs with 80GB RAM.

3.3   Results

The performance of risk prediction in various models is shown in Table 1. Notably, OncoNet, Mirai, and AsymMirai do not incorporate historical data, whereas the other models leverage four years of such data to provide a richer temporal context. Our approach emphasizes two primary components: the VM-RNN module for robust temporal encoding and the asymmetry module for capturing spatial discrepancies over time. For clarity, these components are denoted as "VMRNN" (VMR) and "Asymmetry" (Asym).

   The results demonstrate that compared to the current SOTA model LoMaR, our VMRA-MaR model which integrates both the VMR and Asym modules, shows notable improvements, particularly at the 4 and 5 year follow-up intervals (C index of 0.82 with ROCAUC scores of 0.84 at both time points). Although the p-value of 0.061 suggests only a borderline significant difference from LoMaR, the trend indicates potential gains from combining temporal encoding with asymmetry features. Further analyses among VMRNN-based models suggest that while incorporating asymmetry can improve performance—especially at earlier follow-up intervals—the magnitude of these benefits may vary over longer horizons. The superior performance of VMRNN is attributed to its dynamic recurrent updating, gated mechanisms, and integrated spatial feature extraction, which allow it to effectively capture long-range dependencies and complex temporal

Table 1: Ablation results on breast cancer 5-year risk prediction with C-index and ROCAUC scores. * indicates that the model uses historical data. VMR: Vision-Mamba RNN; Asym: Asymmetry-Aware Risk Factor Extraction.

Model	Module		C-index	Follow-up year ROCAUC				
	VMR	Asym		1-year	2-year	3-year	4-year	5-year
OncoNet (Only-DL)	✗	✗	0.75 (0.72-0.78)	0.82 (0.80-0.84)	0.80 (0.78-0.82)	0.75 (0.73-0.78)	0.73 (0.71-0.75)	0.68 (0.67-0.69)
Mirai	✗	✗	0.82 (0.78-0.83)	0.92 (0.90-0.94)	0.84 (0.81-0.87)	0.80 (0.78-0.82)	0.79 (0.78-0.80)	0.77 (0.76-0.79)
AsymMirai	✗	✓	0.81 (0.78-0.83)	0.80 (0.74-0.86)	0.68 (0.64-0.71)	0.68 (0.66-0.70)	0.67 (0.65-0.69)	0.64 (0.60-0.68)
LoMaR*	✗	✗	0.82 (0.78-0.85)	0.92 (0.89-0.95)	0.83 (0.85-0.86)	0.82 (0.84-0.86)	0.80 (0.78-0.82)	0.83 (0.85-0.87)
VMR_MaR*	✓	✗	**0.84** (0.79-0.89)	0.91 (0.90-0.93)	0.82 (0.80-0.84)	**0.84** (0.82-0.86)	0.82 (0.80-0.84)	**0.86** (0.84-0.86)
VMRA_MaR*	✓	✓	0.82 (0.80-0.83)	**0.94** (0.93-0.95)	**0.88** (0.86-0.89)	0.82 (0.79-0.84)	**0.84** (0.82-0.86)	0.84 (0.82-0.86)

Table 2: Subgroup results on risk detection across low, medium, and high breast density were combined with follow-up year ROCAUC.

Model	1-year			2-year			3-year			4-year			5-year		
	Low	Med	**High**	Low	Med	**High**	Low	Med	**High**	Low	Med	**High**	Low	Med	**High**
OncoNet	0.85	0.82	0.79	0.88	0.78	0.70	0.76	0.76	0.73	0.73	0.75	0.71	0.72	0.74	0.58
Mirai	0.94	**0.95**	0.87	0.80	0.86	0.86	0.81	0.80	0.79	0.81	0.77	0.79	0.78	0.76	0.77
AsymMirai	0.84	0.86	0.70	0.72	0.66	0.66	0.73	0.68	0.63	0.68	0.65	0.66	0.70	0.62	0.60
LoMaR	**0.95**	0.93	0.88	**0.92**	0.82	0.75	**0.89**	0.82	0.75	0.84	0.81	0.75	**0.87**	0.83	0.79
VMR_MaR	0.93	0.89	0.90	0.90	0.80	0.84	0.83	**0.84**	**0.85**	0.82	0.80	**0.84**	0.80	**0.91**	**0.87**
VMRA_MaR	0.92	**0.95**	**0.97**	0.85	**0.89**	**0.90**	0.85	0.78	**0.85**	**0.86**	**0.82**	**0.84**	0.80	0.87	0.85

relationships—providing a significant advantage over Transformer models that rely on fixed positional embeddings.

Note, while our design for incorporating the asymmetry score averages these scores, we have further experimented with alternative approaches such as concatenating all asymmetry scores across time steps. Our comparisons demonstrated that the average fusion consistently yielded the best performance, reinforcing the effectiveness of this approach.

3.4   Subgroup Analysis on Risk Detection Across Breast Density

To further provide insights into the benefits of leveraging temporal information compared to the current state-of-the-art model LoMaR, we analyzed performance stratified according to breast density (see Table 2). In particular, we categorized the data into three breast density levels (low, medium, and high) based on the metric `libra_densearea` in the CSAW-CCS dataset[17]. We utilize this absolute quantity of dense breast tissue `libra_densearea` as it has been shown that absolute dense tissue tends to provide more relevant information on breast cancer risk compared to percentage density (see for instance[16]). This metric, which represents the dense area in cm^2 estimated by the Libra software, was first sorted for all patients and then trichotomized to define the low, medium, and
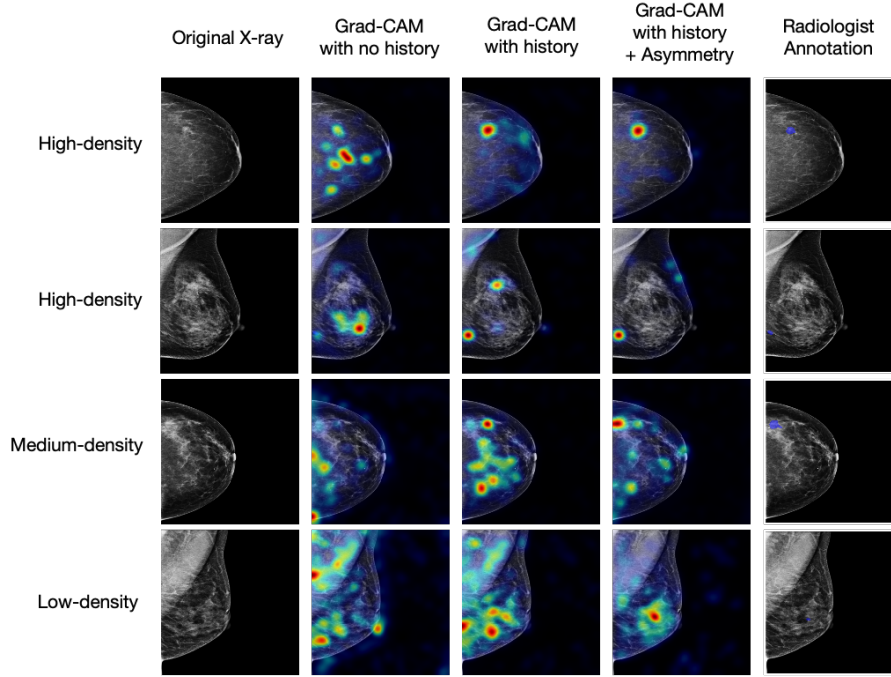
Fig. 2: Grad-CAM visualizations of VMR-MaR and VMRA-MaR models for five representative subjects (two high-, one medium-, and one low-density).

high subgroups. This stratification is especially relevant because higher breast density not only increases the intrinsic risk of developing breast cancer but also complicates the detection of lesions on mammograms.

Notably, our VMRA-MaR model demonstrates a clear advantage in the high-density subgroup: In the 1-year follow-up, it achieves a ROCAUC of 0.97 compared to 0.88 for LoMaR, 0.90 for VMR_MaR, and lower scores for OncoNet (0.79)[24,22,13] and Mirai (0.87). This superior performance persists at 2 years (ROCAUC of 0.90) and remains consistent across the 3-, 4-, and 5-year intervals, underscoring its robust ability to extract and utilize complex features from dense mammary tissues.

We hypothesize that while models lacking temporal analysis perform adequately in low density cases, where tissue characteristics are less challenging the incorporation of temporal features is particularly beneficial for detecting tumors in high-density breast tissue. Our asymmetry-aware, recurrent temporal framework effectively captures intricate tissue dynamics, thereby enhancing performance in this subgroup.

3.5   Exploration of model focus by Grad-cam

Figure 2 presents Grad-cam [15] saliency maps derived from mammographic X-ray images, with the rightmost column displaying expert annotations of cancer lesions. Comparative analysis indicates that incorporating temporal information significantly improves lesion localization, while integrating asymmetry features further refines its precision.

4   Conclusion

In this study, we introduce an Asymmetry-Aware Temporal Framework for breast cancer risk prediction, VMRA-MaR. Our unified approach combines a VMRNN architecture with asymmetry modules to dynamically capture evolving spatial cues and long-range temporal dependencies, yielding accurate risk estimates. The results indicate that our framework surpasses the previous state-of-the-art, LoMaR, by achieving higher accuracy at extended time steps and improved risk prediction for the challenging high-density breasts. Future work will focus on refining the alignment of mammography views to further enhance lesion localization and overall predictive precision.

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

References

1. Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., Jemal, A.: Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians **74**(3), 229–263 (2024)
2. Brentnall, A.R., Harkness, E.F., Astley, S.M., Donnelly, L.S., Stavrinos, P., Sampson, S., Fox, L., Sergeant, J.C., Harvie, M.N., Wilson, M., et al.: Mammographic density adds accuracy to both the tyrer-cuzick and gail breast cancer risk models in a prospective uk screening cohort. Breast Cancer Research **17**, 1–10 (2015)
3. Carriero, A., Groenhoff, L., Vologina, E., Basile, P., Albera, M.: Deep learning in breast cancer imaging: State of the art and recent advancements in early 2024. Diagnostics **14**(8), 848 (2024)
4. Donnelly, J., Moffett, L., Barnett, A.J., Trivedi, H., Schwartz, F., Lo, J., Rudin, C.: Asymmirai: Interpretable mammography-based deep learning model for 1–5-year breast cancer risk prediction. Radiology **310**(3), e232780 (2024)
5. Evans, D.G.R., Howell, A.: Breast cancer risk-assessment models. Breast cancer research **9**, 1–8 (2007)
6. Gross, C.P., Long, J.B., Ross, J.S., Abu-Khalaf, M.M., Wang, R., Killelea, B.K., Gold, H.T., Chagpar, A.B., Ma, X.: The cost of breast cancer screening in the medicare population. JAMA internal medicine **173**(3), 220–226 (2013)

7. Gu, A., Dao, T.: Mamba: Linear-time sequence modeling with selective state spaces (2024), https://arxiv.org/abs/2312.00752
8. Hakama, M., Coleman, M.P., Alexe, D.M., Auvinen, A.: Cancer screening: evidence and practice in europe 2008. European Journal of Cancer **44**(10), 1404–1413 (2008)
9. Hendrycks, D., Gimpel, K.: Gaussian error linear units (gelus) (2023), https://arxiv.org/abs/1606.08415
10. Karaman, B.K., Dodelzon, K., Akar, G.B., Sabuncu, M.R.: Longitudinal mammogram risk prediction. In: Linguraru, M.G., Dou, Q., Feragen, A., Giannarou, S., Glocker, B., Lekadir, K., Schnabel, J.A. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. pp. 437–446. Springer Nature Switzerland, Cham (2024)
11. Katzman, J.L., Shaham, U., Cloninger, A., Bates, J., Jiang, T., Kluger, Y.: Deepsurv: personalized treatment recommender system using a cox proportional hazards deep neural network. BMC medical research methodology **18**, 1–12 (2018)
12. Lee, H., Kim, J., Park, E., Kim, M., Kim, T., Kooi, T.: Enhancing breast cancer risk prediction by incorporating prior images. In: Greenspan, H., Madabhushi, A., Mousavi, P., Salcudean, S., Duncan, J., Syeda-Mahmood, T., Taylor, R. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. pp. 389–398. Springer Nature Switzerland, Cham (2023)
13. Lehman, C.D., Yala, A., Schuster, T., Dontchos, B., Bahl, M., Swanson, K., Barzilay, R.: Mammographic breast density assessment using deep learning: clinical implementation. Radiology **290**(1), 52–58 (2019)
14. Otto, P.M., Blecher, C.B.: Controversies surrounding screening mammography. Missouri medicine **111**(5), 439 (2014)
15. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Gradcam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision. pp. 618–626 (2017)
16. Stone, J., Ding, J., Warren, R.M., Duffy, S.W., Hopper, J.L.: Using mammographic density to predict breast cancer risk: dense area or percentage dense area. Breast Cancer Research **12**, 1–7 (2010)
17. Strand, F.: Csaw-cc (mammography)–a dataset for ai research to improve screening, diagnostics and prognostics of breast cancer (2022)
18. Tang, Y., Dong, P., Tang, Z., Chu, X., Liang, J.: VMRNN: Integrating Vision Mamba and LSTM for Efficient and Accurate Spatiotemporal Forecasting . In: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 5663–5673. IEEE Computer Society (2024)
19. Terry, M.B., Liao, Y., Whittemore, A.S., Leoce, N., Buchsbaum, R., Zeinomar, N., Dite, G.S., Chung, W.K., Knight, J.A., Southey, M.C., et al.: 10-year performance of four models of breast cancer risk: a validation study. The Lancet Oncology **20**(4), 504–517 (2019)
20. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need (2023), https://arxiv.org/abs/1706.03762
21. Wilkinson, A.N., Mainprize, J.G., Yaffe, M.J., Robinson, J., Cordeiro, E., Hong, N.J.L., Williams, P., Moideen, N., Renaud, J., Seely, J.M., et al.: Cost-effectiveness of breast cancer screening using digital mammography in canada. JAMA Network Open **8**(1), e2452821–e2452821 (2025)
22. 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 (2019)

23. Yala, A., Mikhael, P.G., Strand, F., Lin, G., Smith, K., Wan, Y.L., Lamb, L., Hughes, K., Lehman, C., Barzilay, R.: Toward robust mammography-based models for breast cancer risk. Science Translational Medicine **13**(578), eaba4373 (2021)
24. Yala, A., Schuster, T., Miles, R., Barzilay, R., Lehman, C.: A deep learning model to triage screening mammograms: a simulation study. Radiology **293**(1), 38–46 (2019)
25. Yeoh, H.H., Liew, A., Phan, R., Strand, F., Rahmat, K., Nguyen, T.L., Hopper, J.L., Tan, M.: Radifusion: A multi-radiomics deep learning based breast cancer risk prediction model using sequential mammographic images with image attention and bilateral asymmetry refinement (2023), `https://arxiv.org/abs/2304.00257`