



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Lesion-Aware Virtual Contrast-Enhancement in Breast MRI Using CycleGAN

Kiyong Kwon*, Yeodong Yoon, and Taehwan Kim*

Polestar Healthcare Co., Ltd.

{kiyong12, david, taehwan}@polestarhc.com

* These authors contributed equally to this work.

Abstract. Dynamic contrast-enhanced breast magnetic resonance imaging plays an essential role in breast cancer detection, diagnosis, and treatment planning. However, the use of gadolinium-based contrast agents increases examination cost and time, while posing potential safety concerns for patients with contraindications. These limitations have motivated the development of image modality translation studies that transform pre-contrast MRI into post-contrast MRI (virtual contrast-enhancement). We adopted CycleGAN, which produces residuals with self-attention module to reconstruct contrast-enhanced T1-weighted MRI (virtual T1-post). This model was designed to learn the residual between pre- and post-contrast images, enabling it to focus on contrast-enhancing structures.

To further improve the lesion fidelity and clinical relevance of the virtual T1-post, we incorporated segmentation-guided Dice similarity coefficient (Dice) loss and region of interest - structural similarity (ROI-SSIM) loss. Experiments were conducted using the MAMA-Synth Challenge dataset. Compared with Rectified Flow, Our residual-prediction model showed lower performance in global image similarity (MSE: 0.71 ± 1.15 vs. 0.65 ± 1.12 ; LPIPS: 0.17 ± 0.10 vs. 0.14 ± 0.09) and FRD (12.51 vs. 9.87), while demonstrating improved lesion fidelity (ROI-SSIM: 0.41 ± 0.22 vs. 0.37 ± 0.25) and clinically relevant downstream performance (contrast AUROC: 0.86 vs. 0.77; tumor AUROC: 0.50 vs. 0.48; Dice: 0.50 ± 0.32 vs. 0.32 ± 0.37 ; HD95: 80.17 ± 114.62 vs. 198.64 ± 220.03).

Keywords: Image-to-image translation · CycleGAN · Virtual contrast-enhancement · Residual prediction

1 Introduction

Breast cancer is the most commonly diagnosed malignancy among women worldwide and remains one of the leading causes of cancer-related mortality. In 2022, approximately 2.3 million new cases of breast cancer and 660,000 breast cancer-related deaths were reported globally [1]. Accurate detection and characterization of lesion plays a critical role in treatment planning and prognostic assessment [2].

Breast magnetic resonance imaging (MRI) is an important imaging modality in breast cancer care. Dynamic contrast-enhanced MRI (DCE-MRI) can effectively visualize tumor vascularity and vascular permeability by capturing the inflow and washout patterns of contrast agents. Consequently, DCE-MRI is used as a clinical standard examination for major indications, including staging of newly diagnosed breast cancer, screening of women at high risk for breast cancer, and evaluation of response to preoperative chemotherapy [3, 4]. However, DCE-MRI requires intravenous contrast administration, increasing examination time and cost and raising safety concerns associated with gadolinium-based agents [5, 6].

To reduce the burden of acquiring medical images from different modalities, modality translation has been increasingly investigated to generate a target modality from available images [7]. Within this broader field, recent studies have explored translating pre-contrast T1-weighted MRI (T1-pre) into contrast-enhanced T1-weighted MRI (T1-post) without contrast administration [8]. Various approaches using generative models have been investigated for this task, including generative adversarial networks (GANs) [9, 10] and diffusion models [11–13]. Unlike general medical image modality translation, translating T1-pre into T1-post (virtual contrast-enhancement) emphasizes localized contrast-induced intensity changes while preserving the underlying anatomy. Accordingly, recent studies have investigated residual-prediction (RP) learning, in which the model predicts the residual between pre- and post-contrast images rather than the entire post-contrast image, demonstrating improved fidelity of virtual contrast-enhanced regions [14, 15].

We adopted CycleGAN [16] to minimize the effects of spatial misalignment between paired pre- and post-contrast breast MRI. Its cycle-consistency enables image translation with less reliance on exact pixel-wise correspondence. Building on this framework, we incorporate self-attention modules and lesion-specific supervision to guide the model toward subtle, spatially localized enhancement. To evaluate whether the proposed RP strategy is effective beyond previous approaches, we compared our method with Rectified Flow [17] and pretrained Pix2PixHD models [18].

2 Materials and Methods

2.1 Dataset

The full cohort is based on the shared dataset provided by the MAMA-Synth Challenge [19]. It contains pre-treatment breast T1-weight DCE-MRI of 1,506 patients collected from more than 25 centers throughout the United States. Detailed dataset characteristics are provided in Table 1.

For each patient, among the multiple post-contrast phases available in DCE-MRI dataset, the phase with the highest peak intensity within the tumor was selected. Then the slice with the highest lesion peak intensity was extracted to construct paired pre- and post-contrast images.

Table 1. Characteristics of the entire dataset provided by MAMA-Synth challenge

	Train (n=959)	Validation (n=241)	Test (n=306)
Age (years)	48.89 $\pm$ 10.47	48.61 $\pm$ 10.04	47.23 $\pm$ 10.40
Ethnicity (%)	Caucasian (76.2) African American (15.2) Asian (5.9) Other (2.7)	Caucasian (78.2) African American (14.6) Asian (6.7) Other (0.5)	Caucasian (72.1) African American (20.6) Asian (4.7) Other (2.6)
Field strength (%)	1.5T (72.5) 3T (27.5)	1.5T (69.3) 3T (30.7)	1.5T (73.2) 3T (26.8)
Pixel spacing (mm)	0.72 $\pm$ 0.14	0.73 $\pm$ 0.13	0.73 $\pm$ 0.15
Slice thickness (mm)	1.89 $\pm$ 0.57	1.90 $\pm$ 0.54	1.75 $\pm$ 0.68
Manufacturer (%)	GE (68.9) SIEMENS (21.8) Philips (9.3)	GE (68.0) SIEMENS (25.7) Philips (6.2)	GE (46.1) SIEMENS (45.8) Philips (8.2)
Echo time (ms)	3.04 $\pm$ 1.11	2.94 $\pm$ 1.09	2.94 $\pm$ 1.21
Repetition time (ms)	6.73 $\pm$ 3.33	6.40 $\pm$ 1.95	8.23 $\pm$ 6.36

2.2 Preprocessing

To represent each paired pre- and post-contrast image on a common intensity scale while preserving their relative contrast relationship, both images were normalized using a shared intensity range derived from the T1-pre. Intensity values were linearly normalized to $[-1, 1]$ using the range $[I_{min}, I_{min} + 2.5(I_{max} - I_{min})]$ of the T1-pre and values outside this range were clipped, where I_{min} and I_{max} are the T1-pre intensity extrema. The scaling factor of 2.5 was empirically determined from paired image statistics. Subsequently, all images were resized to 256×256 pixels for training. No additional registration was performed.

2.3 CycleGAN-based Virtual Contrast Enhancement

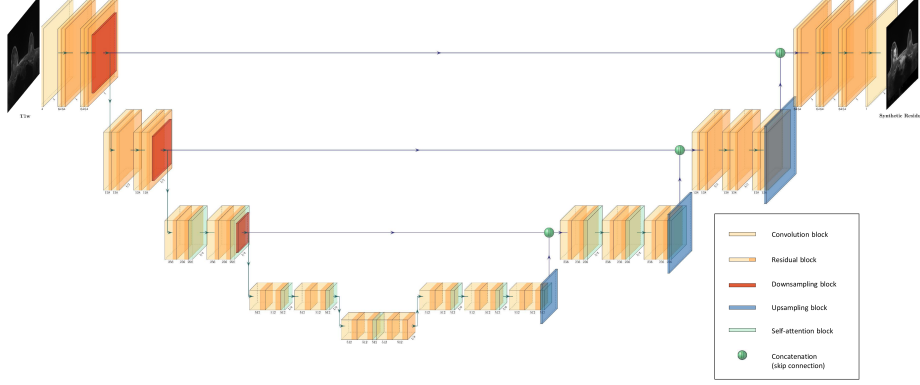
We adopted CycleGAN [16] for virtual contrast-enhancement, as its cycle-consistency reduces reliance on exact pixel-wise correspondence. The proposed framework consists of two generators and two discriminators. Each generator follows a U-Net like encoder-decoder architecture with self-attention blocks, as illustrated in Fig. 1. It predicts residual intensities for pre-to-post and post-to-pre translation, encouraging the model to focus on contrast-induced changes rather than reconstructing the entire target image.

The reconstructed virtual target image was obtained by adding the predicted residual to the source image:

$$\hat{I}_{trg} = I_{src} + \hat{I}_{res}, \quad (1)$$

where $\hat{I}_{trg}$, I_{src} , and $\hat{I}_{res}$ denote the virtual target image, source image, and predicted residual intensities, respectively.

Fig. 1. Overall architecture of generator



Each discriminator takes the channel-wise concatenation of the source images and virtual target images as input, and consists of four downsampling stages followed by a convolutional output layer. Self-attention blocks are also incorporated into the two deepest encoder layers of each discriminator.

2.4 Implementation Details

Although the generator predicts a residual intensities, reconstruction losses were computed on the virtual T1-post: the predicted residual was added to the input T1-pre to obtain virtual T1-post, which was compared with real T1-post. Training combined adversarial ($\mathcal{L}_{adv}$), cycle-consistency ($\mathcal{L}_{adv}$), pixel ($\mathcal{L}_{pixel}$), gradient ($\mathcal{L}_{grad}$), structural similarity (SSIM) ($\mathcal{L}_{SSIM}$) [20], identity ($\mathcal{L}_{iden}$), content ($\mathcal{L}_{content}$) [21], and learned perceptual image patch similarity (LPIPS) ($\mathcal{L}_{LPIPS}$) [22] losses to preserve image fidelity and structural consistency. Lesion-aware supervision was further introduced using soft Dice similarity coefficient (Dice) loss ($\mathcal{L}_{Dice}$) from a pretrained segmentation model and tumor region of interest (ROI)-SSIM loss ($\mathcal{L}_{ROI-SSIM}$), which measures SSIM between the virtual and real T1-post within the tumor ROI, to improve contrast-enhancing lesion reconstruction. Specifically, the soft Dice and ROI-SSIM losses were defined as

$$\mathcal{L}_{Dice} = 1 - \frac{2 \sum_i S_i(\hat{x}) M_i}{\sum_i S_i(\hat{x}) + \sum_i M_i}, \quad (2)$$

$$\mathcal{L}_{ROI-SSIM} = 1 - \text{SSIM}(\hat{x}|M, x|M), \quad (3)$$

where $\hat{x}$ and x are the virtual T1-post and real T1-post, respectively, $S_i(\hat{x})$ is the predicted soft tumor probability at pixel i , and M is the ground-truth lesion

mask with pixel-wise values M_i . The final loss is defined as

$$\begin{aligned}\mathcal{L}_{final} = & \lambda_{adv}\mathcal{L}_{adv} + \lambda_{cycle}\mathcal{L}_{cycle} + \lambda_{pixel}\mathcal{L}_{pixel} + \lambda_{grad}\mathcal{L}_{grad} \\ & + \lambda_{SSIM}\mathcal{L}_{SSIM} + \lambda_{iden}\mathcal{L}_{iden} + \lambda_{content}\mathcal{L}_{content} \\ & + \lambda_{LPIPS}\mathcal{L}_{LPIPS} + \lambda_{Dice}\mathcal{L}_{Dice} + \lambda_{ROI-SSIM}\mathcal{L}_{ROI-SSIM},\end{aligned}\quad (4)$$

where $\lambda_{adv} = \lambda_{content} = \lambda_{LPIPS} = \lambda_{Dice} = \lambda_{ROI-SSIM} = 1$, $\lambda_{cycle} = \lambda_{pixel} = \lambda_{grad} = \lambda_{iden} = 10$, and $\lambda_{SSIM} = 0.5$, with all weights empirically selected.

The model was trained using the Adam optimizer with an initial learning rate of 2×10^{-4} and a batch size of 1. Checkpoint selection was based on validation mean squared error (MSE), LPIPS, ROI-SSIM loss, and soft Dice loss computed using lesion masks predicted by a pretrained 2D nnU-Net segmentation model. A checkpoint was considered improved when at least one metric improved while other metrics remained within a 1% relative tolerance. The final model was selected from 19 training epochs, and no early stopping was applied.

3 Results

3.1 Quantitative Results

The proposed method was evaluated using complementary metrics that assess image fidelity, lesion-specific similarity, and downstream clinical performance. For whole image-to-image comparison, the MSE and the LPIPS were employed. Lesion-specific ROI-to-ROI similarity was evaluated using tumor-region ROI-SSIM and the fr chet radiomics distance (FRD) [23]. To assess the clinical utility of the synthesized images, we evaluated classification and segmentation performance using the official pretrained models provided by the MAMA-Synth challenge [19] without fine-tuning. Notably, pretrained 2D nnU-Net segmentation model also used for the soft Dice loss. During classification inference, the scores from multiple radiomics-based classifiers were ensembled.

For contrast classification, the classifier distinguished T1-pre from virtual T1-post within the same expert-annotated tumor mask. For tumor classification, it distinguished the tumor from its tumor mask and anatomically mirrored non-tumor contralateral ROI in virtual T1-post images. AUROC was computed for both tasks. For downstream segmentation evaluation, lesion masks generated from the virtual T1-post were compared with expert annotations using Dice and Hausdorff distance (HD95) [24].

Table. 2 summarizes the quantitative comparison with the competing methods. Rectified Flow [17] was trained from its official implementation using the same MAMA-Synth training set as our CycleGAN models. Pix2PixHD was evaluated using the pretrained model corresponding to the official MAMA-Synth baseline model [25], and achieved performance comparable to the organizers' baseline evaluation. Therefore, its results should be interpreted as a reference comparison rather than a strictly controlled benchmark. To isolate the contribution of residual prediction, we evaluated a non-RP model that directly predicted

Table 2. Quantitative comparison with other models

	MSE	LPIPS	ROI-SSIM	FRD	AUROC contrast	AUROC tumor	Dice	HD95
Pix2PixHD [18]	2.09 ± 3.35	0.24 ± 0.10	-0.07 ± 0.28	6.16	0.48	0.45	0.19 ± 0.31	287.95 ± 236.71
Rectified Flow [17]	0.65 ± 1.12	0.14 ± 0.09	<u>0.37</u> ± 0.25	9.87	0.77	<u>0.48</u>	0.32 ± 0.37	198.64 ± 220.03
Ours (non-RP)	<u>0.68</u> ± 1.10	0.20 ± 0.10	0.36 ± 0.23	<u>6.27</u>	0.86	0.46	0.53 ± 0.33	<u>87.54</u> ± 119.70
Ours (RP)	0.71 ± 1.15	<u>0.17</u> ± 0.10	0.41 ± 0.22	12.51	0.86	0.50	<u>0.50</u> ± 0.32	80.17 ± 114.62

(best in **bold**; second best underlined)

the post-contrast image while keeping the network architecture and loss functions identical to those of the proposed RP model.

Our model achieved the best ROI-SSIM, tumor AUROC, and HD95, the joint-best contrast AUROC, and the second-best LPIPS and Dice score. Rectified Flow achieved the best MSE and LPIPS, while Pix2PixHD obtained the best FRD.

3.2 Qualitative Results

Representative qualitative comparisons of the test set are shown in Fig. 2 and Fig. 3. As shown in Fig. 2, the real residual images, defined as the difference between the ground-truth T1-post and input T1-pre, highlight contrast-enhancing regions while suppressing shared anatomical information. Compared with our non-RP model, RP model shows more localized enhancement in small lesions. In Fig. 3, our model preserves subtle pathological enhancement that is less apparent in the outputs of the competing methods, particularly for small lesions. However, occasional false enhancement is also observed in non-tumor regions, particularly around the heart.

3.3 Tumor Size Subgroup Analysis

Table 3. Comparison with our non-RP model across tumor-area subgroups

Tumor size	medium (RP)	medium (non-RP)	small (RP)	small (non-RP)
AUROC contrast	0.8580	0.7101	0.7500	0.7500
AUROC tumor	0.6095	0.4586	0.6250	0.5625
Dice	0.36 $\pm$ 0.33	0.35 $\pm$ 0.32	0.12 $\pm$ 0.14	0.04 $\pm$ 0.07
HD95	90.02 $\pm$ 137.10	109.91 $\pm$ 139.28	184.81 $\pm$ 205.59	113.64 $\pm$ 94.70
ROI-SSIM	0.42 $\pm$ 0.20	0.31 $\pm$ 0.22	0.34 $\pm$ 0.20	0.26 $\pm$ 0.11
number of patients	13	13	4	4

(medium: $10 < mm^2 \leq 15$; small: $5 < mm^2 \leq 10$)

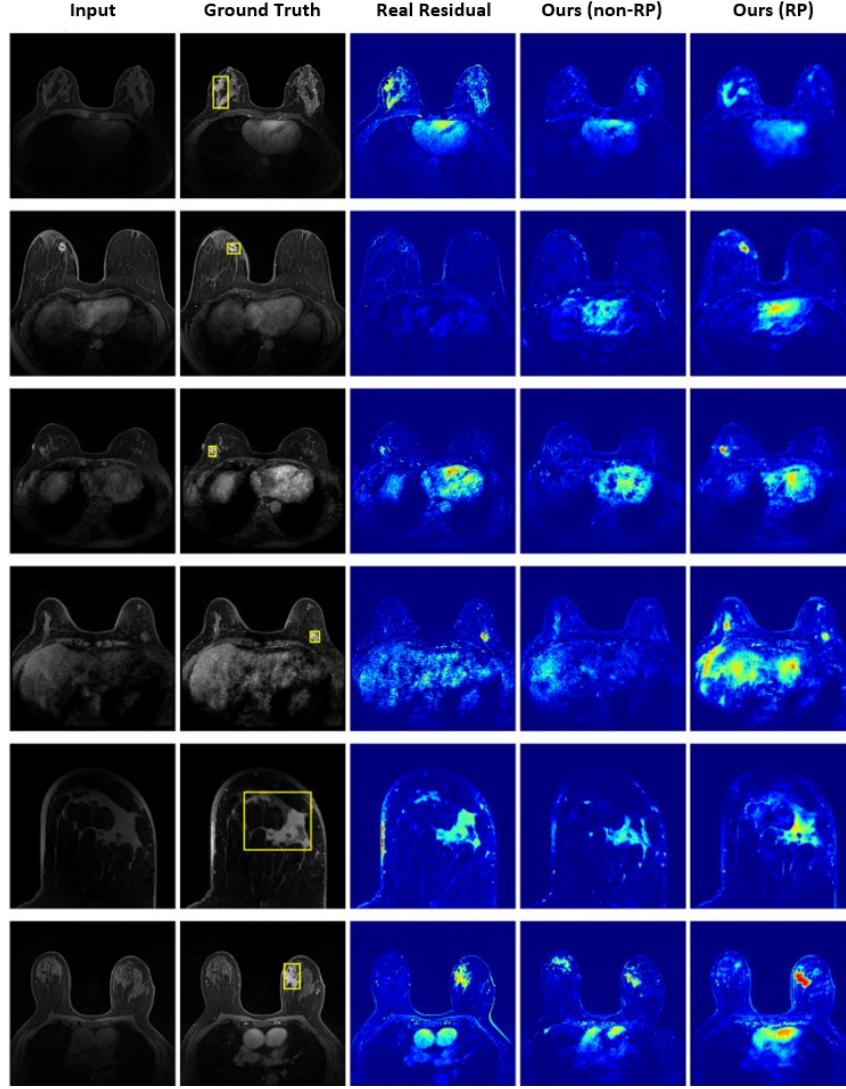
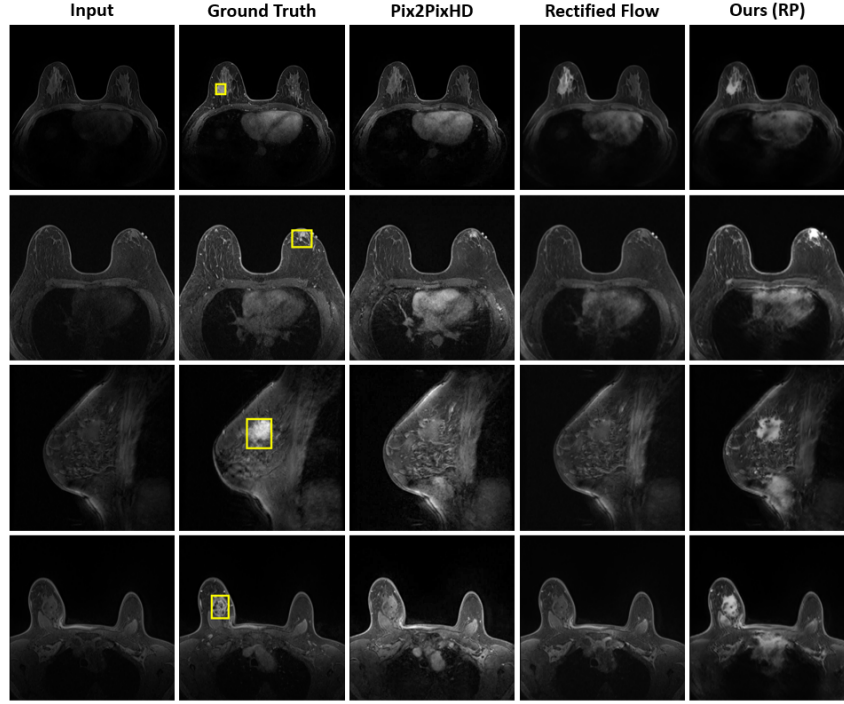
Fig. 2. Comparison of residual predictions between the non-RP and RP models

Table. 3 presents the subgroup analysis according to tumor size measured from the expert segmentation mask [26]. For medium-sized tumors, the RP model achieved higher performance in contrast AUROC, tumor AUROC, Dice, and ROI-SSIM than the non-RP model. For small tumors, the RP model achieved higher performance in tumor AUROC, Dice, and ROI-SSIM, while contrast AUROC remained identical.

Fig. 3. Qualitative comparison with other models

3.4 Ablation Study

Table 4. Ablation study with lesion-specific losses

Model	MSE	LPIPS	ROI-SSIM	FRD	AUROC contrast	AUROC tumor	Dice	HD95
RO (w/o Dice & ROI)	0.6807	0.1720	0.3699	10.8697	0.8393	0.4406	0.3732	176.7478
RO + Dice	0.6358	0.1741	0.3595	11.0550	0.8244	0.5055	0.4987	89.1414
RO + ROI	0.6897	0.1689	0.3969	13.8975	0.8157	0.5107	0.3439	149.1251
RO + Dice + ROI (Ours)	0.7195	0.1724	0.4102	12.515	0.8619	0.5079	0.5067	80.1786

(Dice: soft Dice loss; ROI: ROI-SSIM loss)

As shown in Table. 4, the full model combining both lesion-specific losses achieved the best performance in ROI-SSIM, contrast AUROC, and Dice, and HD95. The ROI-SSIM loss only model achieved the best performance in tumor AUROC, while the full model showed comparable performance. In contrast, the best performance in MSE and LPIPS were obtained by the Seg Dice loss only and ROI-SSIM loss only models, respectively.

4 Discussion and Conclusion

Quantitatively, our model showed stronger lesion-specific and downstream performance than other models despite not achieving the best global image-similarity scores. This discrepancy suggests that global metrics may not fully reflect the fidelity of localized contrast enhancement, as subtle differences within contrast-enhanced lesions may have a limited contribution to whole-image similarity.

Qualitatively, our RP model showed more localized enhancement than our non-RP model and better preserved subtle pathological enhancement compared with the competing methods. This observation is consistent with the design of the RP formulation, which focuses the reconstruction on contrast-induced intensity changes rather than the entire post-contrast image.

The tumor-size subgroup analysis further showed improvements in tumor AUROC, Dice, and ROI-SSIM for medium and small lesions with the RP model. These results suggest a potential advantage of residual learning when contrast-related changes are localized to a small spatial region, although this finding should be interpreted cautiously given the limited number of small-tumor cases.

The ablation study further demonstrated the contribution of lesion-specific supervision, which generally improved lesion-oriented and downstream clinical performance without consistent gains in global image-similarity metrics. Specifically, the soft Dice and ROI-SSIM losses were designed to encourage preservation of tumor-related features and structural similarity within the tumor region, respectively.

Several limitations remain. Variations in post-contrast acquisition phases and noise in raw MRI may contribute to false enhancement in non-tumor regions. In addition, despite the multi-institutional nature of the MAMA-Synth dataset, the proposed framework was not externally validated using independent data collected from hospitals or institutions not included in the training cohort.

Overall, the proposed residual-prediction CycleGAN framework provides a promising approach to virtual contrast-enhancement in breast MRI while preserving diagnostically important lesion characteristics. Future work will focus on improving global image similarity and robustness across external datasets.

References

1. Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). 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.
2. Harbeck, N., Penault-Llorca, F., Cortes, J., Gnant, M., Houssami, N., Poortmans, P., ... & Cardoso, F. (2019). Breast cancer *Nat Rev Dis Primers* 5: 66.
3. Mann, R. M., Cho, N., & Moy, L. (2019). Breast MRI: state of the art. *Radiology*, 292(3), 520-536.
4. Leithner, D., Wengert, G. J., Helbich, T. H., Thakur, S., Ochoa-Albiztegui, R. E., Morris, E. A., & Pinker, K. (2018). Clinical role of breast MRI now and going forward. *Clinical radiology*, 73(8), 700-714.
5. Chhor, C. M., & Mercado, C. L. (2017). Abbreviated MRI protocols: wave of the future for breast cancer screening. *American Journal of Roentgenology*, 208(2), 284-289.
6. Olchoway, C., Cebulski, K., Łasecki, M., Chaber, R., Olchoway, A., Kałwak, K., & Zaleska-Dorobisz, U. (2017). The presence of the gadolinium-based contrast agent depositions in the brain and symptoms of gadolinium neurotoxicity-A systematic review. *PloS one*, 12(2), e0171704.
7. Han, L., Zhang, T., Huang, Y., Dou, H., Wang, X., Gao, Y., ... & Mann, R. (2023, October). An explainable deep framework: towards task-specific fusion for multi-to-one MRI synthesis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 45-55). Cham: Springer Nature Switzerland.
8. Chung, M., Calabrese, E., Mongan, J., Ray, K. M., Hayward, J. H., Kelil, T., ... & Lee, A. Y. (2022). Deep learning to simulate contrast-enhanced breast MRI of invasive breast cancer. *Radiology*, 306(3), e213199.
9. Denck, J., Guehring, J., Maier, A., & Rothgang, E. (2021). Enhanced magnetic resonance image synthesis with contrast-aware generative adversarial networks. *Journal of imaging*, 7(8), 133.
10. Osuala, R., Joshi, S., Tsirikoglou, A., Garrucho, L., Pinaya, W. H., Lang, D. M., ... & Lekadir, K. (2025). Simulating dynamic tumor contrast enhancement in breast MRI using conditional generative adversarial networks. *Journal of Medical Imaging*, 12(S2), S22014-S22014.
11. Müller-Franzes, G., Huck, L., Bode, M., Nebelung, S., Kuhl, C., Truhn, D., & Lemaingue, T. (2024). Diffusion probabilistic versus generative adversarial models to reduce contrast agent dose in breast MRI. *European radiology experimental*, 8(1), 53.
12. Osuala, R., Lang, D. M., Verma, P., Joshi, S., Tsirikoglou, A., Skorupko, G., ... & Lekadir, K. (2024, October). Towards learning contrast kinetics with multi-condition latent diffusion models. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 713-723). Cham: Springer Nature Switzerland.
13. Oulmalme, C., Nakouri, H., & Jaafar, F. (2025). A systematic review of generative AI approaches for medical image enhancement: Comparing GANs, transformers, and diffusion models. *International journal of medical informatics*, 199, 105903.
14. Schreiter, H., Eberle, J., Kapsner, L. A., Hadler, D., Ohlmeyer, S., Erber, R., ... & Liebert, A. (2024, October). Virtual dynamic contrast enhanced breast MRI using 2D U-Net architectures. In *Deep Breast Workshop on AI and Imaging for Diagnostic and Treatment Challenges in Breast Care* (pp. 85-95). Cham: Springer Nature Switzerland.

15. Ibarra, S., del Riego, J., Catanese, A., Cuba, J., Cardona, J., Leon, N., ... & Osuala, R. (2025, September). Comparing Conditional Diffusion Models for Synthesizing Contrast-Enhanced Breast MRI from Pre-contrast Images. In *Deep Breast Workshop on AI and Imaging for Diagnostic and Treatment Challenges in Breast Care* (pp. 226-236). Cham: Springer Nature Switzerland.
16. Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).
17. Liu, X., Gong, C., & Liu, Q. (2022). Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*.
18. Osuala, R., Joshi, S., Tsirikoglou, A., Garrucho, L., Pinaya, W. H., Diaz, O., & Lekadir, K. (2024, April). Pre-to post-contrast breast MRI synthesis for enhanced tumour segmentation. In *Medical imaging 2024: Image processing* (Vol. 12926, pp. 226-237). SPIE.
19. Osuala, R., Joshi, S., van Dijk, J., Han, L., Cosaka, M. L., Mysler, D., ... & Diaz, O. The MAMA-SYNTH challenge: Synthesizing virtual contrast-enhancement in breast MRI (2026).
20. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4), 600-612.
21. Johnson, J., Alahi, A., & Fei-Fei, L. (2016, September). Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision* (pp. 694-711). Cham: Springer International Publishing.
22. Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018, June). The unreasonable effectiveness of deep features as a perceptual metric. In *2018 IEEE/CVF conference on computer vision and pattern recognition* (pp. 586-595). IEEE.
23. Konz, N., Osuala, R., Verma, P., Chen, Y., Gu, H., Dong, H., ... & Mazurowski, M. A. (2026). Fréchet radiomic distance (FRD): A versatile metric for comparing medical imaging datasets. *Medical image analysis*, 103943.
24. Karimi, D., & Salcudean, S. E. (2019). Reducing the hausdorff distance in medical image segmentation with convolutional neural networks. *IEEE Transactions on medical imaging*, 39(2), 499-513.
25. Osuala, R., Skorupko, G., Lazrak, N., Garrucho, L., García, E., Joshi, S., ... & Lekadir, K. (2023). medigan: a Python library of pretrained generative models for medical image synthesis. *Journal of Medical Imaging*, 10(6), 061403-061403.
26. Douglas, L., Bhattacharjee, R., Fuhrman, J., Drukker, K., Hu, Q., Edwards, A., ... & Giger, M. (2023). U-Net breast lesion segmentations for breast dynamic contrast-enhanced magnetic resonance imaging. *Journal of Medical Imaging*, 10(6), 064502-064502.