Temporal Neural Cellular Automata: Application to modeling of contrast enhancement in breast MRI

Daniel M. Lang[1,2], Richard Osuala[3], Veronika Spieker[1,2], Karim Lekadir[3,4], Rickmer Braren[5,6], and Julia A. Schnabel[1,2,7]

[1] Institute of Machine Learning in Biomedical Imaging, Helmholtz Munich, Germany
lang@helmholtz-munich.de
[2] School of Computation, Information and Technology, Technical University of Munich, Germany
[3] Departament de Matematiques i Informatica, Universitat de Barcelona, Spain
[4] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
[5] Institute for Diagnostic and Interventional Radiology, School of Medicine & Health, Klinkum Rechts der Isar, Technical University of Munich, Germany
[6] German Cancer Consortium (DKTK), Partner Site Munich, Germany
[7] School of Biomedical Engineering & Imaging Sciences, King's College London, UK

Abstract. Synthetic contrast enhancement offers fast image acquisition and eliminates the need for intravenous injection of contrast agent. This is particularly beneficial for breast imaging, where long acquisition times and high cost are significantly limiting the applicability of magnetic resonance imaging (MRI) as a widespread screening modality. Recent studies have demonstrated the feasibility of synthetic contrast generation. However, current state-of-the-art (SOTA) methods lack sufficient measures for consistent temporal evolution. Neural cellular automata (NCA) offer a robust and lightweight architecture to model evolving patterns between neighboring cells or pixels. In this work we introduce TeNCA (Temporal Neural Cellular Automata), which extends and further refines NCAs to effectively model temporally sparse, non-uniformly sampled imaging data. To achieve this, we advance the training strategy by enabling adaptive loss computation and define the iterative nature of the method to resemble a physical progression in time. This conditions the model to learn a physiologically plausible evolution of contrast enhancement. We rigorously train and test TeNCA on a diverse breast MRI dataset and demonstrate its effectiveness, surpassing the performance of existing methods in generation of images that align with ground truth post-contrast sequences. Code: https://github.com/LangDaniel/TeNCA

Keywords: NCA · Image Synthesis · Dynamic Contrast Enhancement

1 Introduction

Dynamic contrast-enhanced - magnetic resonance imaging (DCE-MRI) is the most sensitive modality for breast cancer detection, outperforming conventional

imaging with mammography, digital breast tomosynthesis and ultrasound [12]. The method images changes in tissue enhancement over time. To achieve this, multiple MRI sequences are acquired after contrast injection. While currently reserved for supplemental screening of high-risk patients, a growing body of evidence suggests that patients with lower risk profiles may also benefit from its use. However, wide adoption of DCE-MRI for breast cancer screening is hindered by its high costs and lengthy acquisition times [4]. To address this limitations, *Kuhl et al.* [11] developed an abbreviated imaging protocol that uses only one post-contrast image. Nevertheless, this approach comes at the cost of losing time-resolved contrast kinetics, which enhance specificity and enable malignancy assessment. Ideally, a breast MRI protocol should strike a balance between high spatial resolution and high temporal resolution, allowing for optimal diagnostic performance [12].

Recent studies have shown the potential of deep learning models to predict contrast uptake from unenhanced acquisitions. For example, *Schreiter et al.* [25] developed a U-Net architecture that predicts T1-weighted subtraction images from T1, T2, and diffusion weighted imaging (DWI). Additionally, *Osuala et al.* [19] explored the use of latent diffusion models (LDMs) to model contrast uptake on T1-weighted breast images, conditioning on acquisition time and supplementary imaging information using a ControlNet [27]. Furthermore, the capabilities of generative adversarial networks (GANs) have also been investigated [18,10,17].

Neural cellular automata (NCA) are a class of models that simulate the communication and progression of cells living on a grid, which can be effectively represented by convolutional neural networks (CNNs) [6]. The growing NCA variant [16] is designed to iteratively model the evolution of complex patterns. In the medical domain, *Manzanera et al.* [14] extended the architecture to simulate nodule growth in lung cancer computed tomography (CT). Additionally, *Kalkhof et al.* [8,9] developed NCAs for segmentation, while *Deutges et al.* [2] extended the architecture to classification tasks. Furthermore, NCAs have also been merged with diffusion models [3,15] and applied for image registration [21]. The ability of NCA to model temporal textures has been investigated by *Pajouheshgar et al.* [20]. They developed a model for dynamic texture synthesis on real-world temporally-dense video data. Moreover, *Richardson et al.* [22] designed a nested NCA architecture to learn spatio-temporal patterns on artificially generated datasets featuring a uniform temporal spacing.

Typically, MRI scan times are within the order of several minutes, with the exact duration dependent on the specific acquisition protocol. However, the uptake and washout of contrast agent is a dynamic process that evolves continuously over time. Therefore, a model designed to predict dynamic contrast enhancement must be capable of learning from temporally sparse data while ensuring a continuous evolution over time. NCAs are inherently suited for this task due to their iterative nature, which can be leveraged to guarantee a continuous progression. However, this feature of NCA is usually not made use of, with the iterative process being viewed solely as a means to reach a static output state after a fixed number of update steps. In this work, we introduce TeNCA

(Temporal Neural Cellular Automata), a novel approach to model temporally consistent evolution over time. TeNCA extends the capabilities of NCAs to effectively learn from temporally sparse, non-uniformly sampled imaging data and capitalize on their iterative nature. To achieve this, we advance the training strategy of our model to be able to adaptively condition intermediate states. Furthermore, we define the update step to reflect physical progression of time, enabling the model to simulate the continuous process of contrast enhancement. We evaluate the performance of TeNCA and compare it to two reference methods, surpassing current SOTA performance in generation of images that align with ground truth DCE-MRI. Furthermore, we prove superior performance of TeNCA with respect to temporal stability and sequential consistency. Our key contributions are as follows:

- We introduce TeNCA, a novel neural cellular automata based approach enabling training on temporally sparse, non-uniformly sampled imaging data.
- We adapt TeNCA to model contrast enhancement on breast MRI and rigorously train and test it on a diverse dataset, involving different subcohorts and imaging protocols, with a large variety of acquisition times.
- We evaluate TeNCA in comparison to two reference methods, improving current SOTA performance in terms of image generation that stays close to ground truth post-contrast acquisitions and prove TeNCAs superiority in learning temporal patterns.

2 Background: Neural Cellular Automata

NCAs are designed to learn update rules that allow transformation of an initial state S_0 into a final state S_{fin}, with the updates being iterative applied via [20]

$$S_{t+1} = \mathcal{F}(S_t) = S_t + \frac{\partial S}{\partial t}\Delta t. \tag{1}$$

The transition function $\mathcal{F}$ consists of a *perception* and a *update* part, that can be represented utilizing a neural network [6]. The global state $S \in \mathbb{R}^{h\times w\times d}$ represents a grid of cells $s_{ij} \in \mathbb{R}^d$ that can communicate with each other.

During the perception stage, each cell gathers information from its neighbors to form the perception vector $z_{ij} \in \mathbb{R}^{n\cdot d}$, where n represents the number of possible communication pathways. Typically, two pathways between nearest neighbors are employed, which can be modeled using learnable convolutional kernels [2]. This is combined with an identity kernel representing the cell's own state, resulting in $n = 3$. However, techniques that enable global communication can also be applied [8,20]. The global state S is divided into two parts: visible and hidden. The visible part $S_{\text{vis}} = \{s_{ijk} : i \in \{1, ..., h\}, j \in \{1, ..., w\}, k \in \{1, ..., c\}\}$ is initialized with a seed or image of dimensionality $\mathbb{R}^{h\times w\times c}$ and the remaining hidden part stores information about cell communication. The update part of the transition function $\mathcal{F}$ can be modeled by a multilayer perceptron (MLP) via

$$\frac{\partial s_{ij}}{\partial t} = \text{MLP}(z_{ij}) \odot M, \tag{2}$$

where M denotes a random binary variable, introduced for stochasticity [20].

During training, the model weights are optimized to ensure that the iterative updates converge to a visual part of the final state, which reflects a static target image. Typically, the number of iterative update steps N_{steps} is defined as a hyperparameter that is empirically selected.

3 Method: NCA for temporally sparse representations

Predicting contrast uptake requires a more flexible approach than the standard NCA training procedure, as it involves modeling a varying amount of post-contrast sequences at different time points. To address this challenge, we extend and further develop the NCA architecture to enable sequential loss computation. Additionally, we define the update stage to reflect a physical progression in time, allowing us to adaptively condition the model precisely at time points for which a ground truth post-contrast sequence is available, see Figure 1.

Let $\{x, \{\{y_1, ..., y_k\}, \{t_1, ..., t_k\}\}\}$ denote a pair of a pre-contrast image x and its corresponding post-contrast sequences y_i acquired at times t_i after contrast injection. We initialize the visible part of the state S with the pre-contrast image, i.e. $S_0^{\text{vis}} = x \in \mathbb{R}^{h \times w \times 1}$, while the hidden part is zero initialized. Our goal is to have S_{vis} gradually transition from a pre-contrast to a post-contrast state, while ensuring that intermediate states also take physiologically meaningful post-contrast states. To achieve this, we define the update step to reflect a progression in time Δt, and require S_{vis} to approximate y_i after $t_i/\Delta t$ updates. Specifically, for all update steps $\{t_i/\Delta t : i \in \{1, \dots, k\}\}$ we compute the loss between $\hat{y}_i = S^{\text{vis}}_{t_i/\Delta t}$ and y_i. The overall loss is then given by

$$\mathcal{L} = \sum_{j=0}^{m} \sum_{i=0}^{k_j} \mathcal{L}_{\text{img}} \left(y_i^j, S^{\text{vis}}_{t_i/\Delta t} \right) = \sum_{j=0}^{m} \sum_{i=0}^{k_j} \mathcal{L}_{\text{img}} \left(y_i^j, \hat{y}_i^j \right), \tag{3}$$

with k_j depicting the number of post-contrast sequences for patient j, and m the number of cases involved, i.e. given in the (mini)batch. The loss $\mathcal{L}_{\text{img}}$ depicts a standard pixel/image based loss, e.g. mean squared error (MSE). A detailed training strategy of TeNCA is given in Algorithm 1.

By training the model in this manner, we constrain it to learn a smooth and continuous transition of $S^0_{\text{vis}} = x$ into a final state S^N_{vis}, while being conditioned to physiologically meaningful intermediate states as reflected in the training data.

4 Experiments and Results

Unlike previous studies [17,19,25], we train and evaluate our model on a diverse dataset comprising images from multiple subcohorts, each with distinct imaging protocols. This diversity presents a unique challenge, as the number and timing of DCE acquisitions can vary substantially between protocols. For instance, one

Algorithm 1: Training strategy for TeNCA.

Input: $\mathcal{D}_{\text{train}}$: training set with pairs $\{x, \{y_1, ..., y_k\}, \{t_1, ..., t_k\}\}$, $\mathcal{F}$: NCA transition function, S: NCA state, N_{steps} : update steps, Δt : time-delta, m: batch size

for *number of training epochs* **do**
  for $\{x^j, \{\{y_1^j, \ldots, y_{k_j}^j\}, \{t_1^j, ..., t_{k_j}^j\}\}\}_{j=0}^{m}$ *in* $\mathcal{D}_{train}$ **do**
    for j *in* $0, \ldots, m$ **do**
      $S_0^{\text{vis}} \leftarrow x^j$
      $t \leftarrow 0$
      for l *in* $0, \ldots, N_{steps}$ **do**
        $t \leftarrow t + \Delta t$
        $S_{l+1} \leftarrow \mathcal{F}(S_l)$
        for i *in* $1, \ldots, k_j$ **do**
          if t *equals* t_i^j **then**
            $\hat{y}_i^j \leftarrow S_{l+1}^{\text{vis}}$
    $\mathcal{L} \leftarrow \sum_{j=0}^{m} \sum_{i=0}^{k_j} \mathcal{L}_{img}\left(y_i^j, \hat{y}_i^j\right)$
    perform back-propagation and optimize weights of $\mathcal{F}$

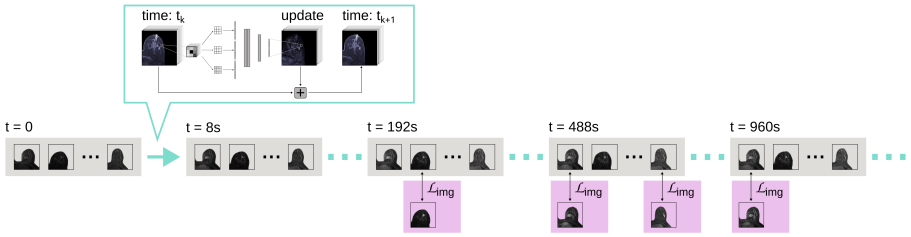

Fig. 1. Overview of TeNCA. For each step, our NCA backbone transitions the images gradually to reflect the next time point. During Training, intermediate states are conditioned at all time points with a given ground truth DCE-MRI available.

clinical center might capture only two post-contrast sequences shortly after injection, whereas another may employ five sequences to also illustrate contrast washout. To provide a comprehensive comparison, we test our approach against two SOTA methods: a U-Net model and a latent diffusion model.

4.1 Dataset

For all experiments, we utilize the public MAMA-MIA dataset [5] (License CC BY 3.0 & CC BY-NC 4.0), which comprises T1-weighted fat-saturated breast DCE-MRI scans. We adhere to the provided training-test split, which consists of 300 test cases. To augment the training cohort, we incorporate additional T1-weighted fat-saturated cases from the Duke-Breast-Cancer-MRI dataset [24] (License CC BY-NC 4.0) that are not part of MAMA-MIA. From this combined dataset, we randomly select 200 patients for validation, resulting in a training set of 1604 cases. For analysis, we consider a maximum of five post-contrast sequences and a acquisition time of up to 1024 seconds. All images are resampled to a uniform voxel spacing of 1 mm and intensity values are linearly rescaled

between zero and one based on the 0.02 and 99.98 percentiles of the respective pre-contrast image. Additionally, images are cropped to patches of size 168×168 following the basic procedure established by *Osuala et al.* [19].

4.2 Implementation

U-Net For the implementation of the U-Net, we follow the structure of MCO-Net [25], which models sequential post-contrast sequences through different output channels. The method depicts a smaller variant of the standard U-Net architecture [23]. However, since MCO-Net was trained on five input sequences, including T2 and diffusion weighted imaging, we perform an empirical grid search to optimize the hyperparameter. We find that for our case the standard U-Net structure, combined with batch normalization layers, trained on mean absolute error (MAE), yields the best results. The code for our U-Net model is available for reproducibility[1].

CC-Net We employ the latent diffusion model-based architecture proposed by *Osuala et al.* [19] and adapt it to our dataset. Specifically, we leverage the CC-Net$_{Any}$ model and retrain both the latent diffusion model and the ControlNet architecture for 100 epochs with the given hyperparameter settings. For encoding and decoding, we utilize the *2-1-base* stable diffusion autoencoder.

TeNCA In the perception state, TeNCA utilizes two learnable kernels of size 3×3 for communication between neighboring cells in combination with a kernel retrieving the cells own state. We pad the input images, featuring one color channel, to a channel size of 24, and set the temporal resolution to $\Delta t = 8$ seconds. For the update stage, a two layered MLP exhibiting a hidden size of 128 with the first layer using ReLU activation is employed. We train TeNCA with MSE as the image loss and perform empirical hyperparameter optimization. We make the code for TeNCA available for reproducibility[1].

4.3 Evaluation Metrics

Image evaluation metrics include learned perceptual image patch similarity (LPIPS) [28], the structural similarity index measure (SSIM) and multi-scale SSIM (MS-SSIM) [26], as well as peak signal-to-noise ratio (PSNR). Distribution measures involve Fréchet inception distance (FID) [7] and Fréchet radiomics distance (FRD) [19]. As models were trained to optimize different losses, i.e. MSE and MAE, and are, therefore, likely biased towards their respective training objective, we do not include those metrics in our analysis. As a lower/upper bound, we compute the difference between each post-contrast acquisition and its respective pre-contrast image, the result of which we denote as baseline.

[1] https://github.com/LangDaniel/TeNCA

4.4 Results

The overall model performance on the test set, calculated as the mean across all post-contrast phases, is presented in Table 1. Qualitative results are visualized in Figure 2. Notably, TeNCA achieves the highest overall performance, surpassing

Table 1. Comparison of Image metrics and distribution measures on the test set, alongside model parameter counts. Notably, TeNCA excels in image metrics while maintaining competitive distribution measure values. CC-Net achieves higher values for distribution measures, but its image metric performance suggest a tendency to hallucinate image parts. TeNCA requires significantly less parameters.

Method	LPIPS↓	SSIM↑	MS-SSIM↑	PSNR [dB]↑	FID↓	FRD↓	param.↓
baseline	0.13	0.86	0.88	29.24	30.20	154.20	-
U-Net	0.13	0.88	0.92	31.93	32.00	50.96	$31 \cdot 10^6$
CC-Net	0.14	0.80	0.80	30.19	**21.28**	**20.00**	$12 \cdot 10^8$
TeNCA(ours)	**0.12**	**0.89**	**0.93**	**32.26**	27.83	48.68	$\mathbf{13 \cdot 10^3}$

all other methods in terms of image-level metrics. However, CC-Net outperforms TeNCA in distribution similarity metrics, specifically FID and FRD, which assess the similarity between the set of generated images and the ground truth DCE dataset. This suggest that CC-Net is capable of producing more realistic-looking images. Nevertheless, CC-Net's performance in image-level metrics reveals a significant limitation: at pixel level, the generated images deviate substantially from the ground truth post-contrast sequences, as indicated by LPIPS and (MS-)SSIM values below the baseline. This implies that CC-Net is prone to hallucinating parts in the images, a known issue with diffusion models [1]. An example of this can be seen in the first row of Figure 2, CC-Net generates a realistic-looking example that, however, fails to reflect the actual post-contrast sequence. In contrast, the U-Net architecture performs the worst, with image metrics lower than TeNCA and the lowest results for distribution measures. The qualitative examples in Figure 2 suggest a segmentation-like behavior, which is consistent with the task for which the U-Net architecture was initially designed [23]. As depicted in Table 1, TeNCA requires significantly less parameters than other methods.

Temporal stability Figure 3 illustrates the image metric performance calculated individually for each post-contrast phase. TeNCA consistently features superior temporal ability. Other methods exhibit a decline in performance as time progresses/for later phases, with the U-Net mostly achieving its best results for the first phase. In contrast, TeNCA achieves stable temporal performance with MS-SSIM even improving for later phases.

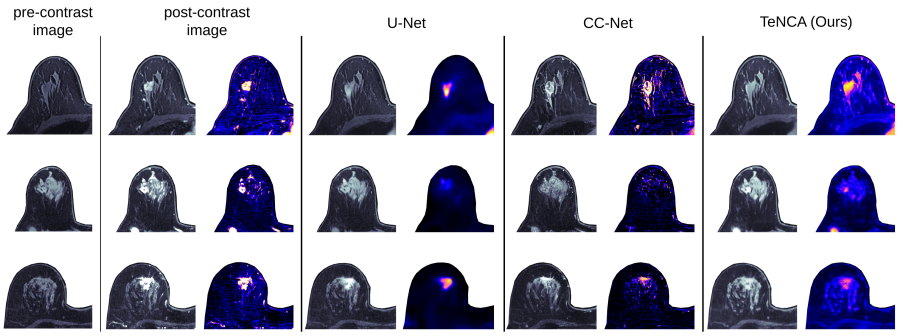

Fig. 2. Example test set results, involving a (predicted) post-contrast image and a subtraction between the pre- and the post-contrast image, highlighting contrast uptake. TeNCA successfully models detailed structures, outperforming the U-Net. Additionally, it avoids hallucination of artifacts, a limitation evident in the first example of CC-Net.

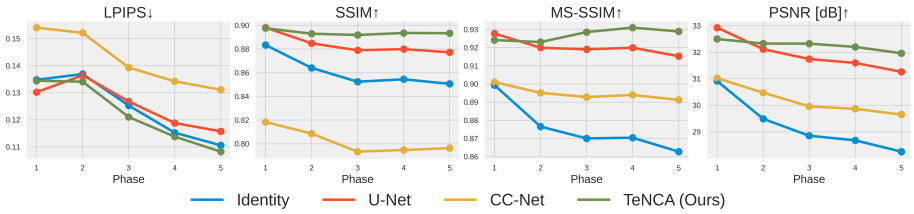

Fig. 3. Mean image metric values for the test set across all post-contrast phases. TeNCA maintains consistent performance throughout all phases, while other methods exhibit a noticeable decline in later phases.

Sequential consistency Example videos illustrating the temporal evolution of contrast enhancement are provided online[2]. Notably, TeNCA exhibits the best sequential consistency, characterized by a continuous evolution. In contrast, CC-Net's output changes significantly between consecutive frames, highlighting the inevitable need for temporal stability measures to be taken into account. Additionally, the segmentation-like behavior of the U-Net model is evident, with its output remaining relatively static.

5 Discussion and Conclusion

This paper introduces TeNCA, a novel approach that enhances the training procedure of neural cellular automata to effectively model temporally sparse, non-uniformly sampled imaging data. We train TeNCA to predict contrast enhancement on breast MRI and comprehensively evaluate its performance on a

[2] https://langdaniel.github.io/TeNCA/

challenging dataset including diverse sub-cohorts with varying imaging protocols and acquisition times. Our results demonstrate the superiority of TeNCA over two existing methods, surpassing current SOTA performance in generation of images that align with ground truth post-contrast sequences. Furthermore, we prove TeNCA's strong temporal capabilities, which performs consistent over all post-contrast phases and evolves its output continuously over time. Notably, TeNCA is less susceptible to hallucinations, an issue with diffusion based contrast prediction, which poses a significant concern in the medical domain where algorithm reliability is paramount for clinical applicability [13]. Additionally, TeNCA requires substantially fewer parameters than other methods, making in easily deployable, even in resource-constrained settings.

The strong performance of TeNCA motivates us to further enhance its capabilities to 3D modeling and evaluate its clinical applicability in future work. Its flexible training strategy also opens new opportunities for application, e.g. in cine MRI or 4D CT.

Acknowledgments. DML and JAS received funding from HELMHOLTZ IMAGING, a platform of the Helmholtz Information & Data Science Incubator. VS is partially supported by the Helmholtz Association under the joint research school "Munich School for Data Science (MUDS)". This project (RO, KL) has received funding from the EU Horizon Europe and Horizon 2020 research and innovation programme under grant agreement No 101057699 (RadioVal) and No 952103 (EuCanImage), respectively. RO acknowledges a research stay grant from the Helmholtz Information and Data Science Academy (HIDA).

Disclosure of Interests. All authors declare that they have no conflicts of interest.

References

1. Aithal, S.K., Maini, P., Lipton, Z., Kolter, J.Z.: Understanding hallucinations in diffusion models through mode interpolation. Advances in Neural Information Processing Systems **37**, 134614–134644 (2025)
2. Deutges, M., Sadafi, A., Navab, N., Marr, C.: Neural Cellular Automata for Lightweight, Robust and Explainable Classification of White Blood Cell Images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 693–702. Springer (2024)
3. Elbatel, M., Kamnitsas, K., Li, X.: An Organism Starts with a Single Pix-Cell: A Neural Cellular Diffusion for High-Resolution Image Synthesis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 656–666. Springer (2024)
4. Gao, Y., Reig, B., Heacock, L., Bennett, D.L., Heller, S.L., Moy, L.: Magnetic resonance imaging in screening of breast cancer. Radiologic Clinics of North America **59**(1), 85 (2020)
5. Garrucho, L., Reidel, C.A., Kushibar, K., Joshi, S., Osuala, R., Tsirikoglou, A., Bobowicz, M., del Riego, J., Catanese, A., Gwoździewicz, K., et al.: MAMA-MIA: A Large-Scale Multi-Center Breast Cancer DCE-MRI Benchmark Dataset with Expert Segmentations. arXiv preprint arXiv:2406.13844 (2024)

6. Gilpin, W.: Cellular automata as convolutional neural networks. Physical Review E **100**(3), 032402 (2019)
7. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: GANs trained by a two time-scale update rule converge to a local nash equilibrium. Advances in neural information processing systems **30** (2017)
8. Kalkhof, J., González, C., Mukhopadhyay, A.: Med-NCA: Robust and lightweight segmentation with neural cellular automata. In: International Conference on Information Processing in Medical Imaging. pp. 705–716. Springer (2023)
9. Kalkhof, J., Mukhopadhyay, A.: M3D-NCA: Robust 3D Segmentation with Built-In Quality Control. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 169–178. Springer (2023)
10. Kim, E., Cho, H.H., Kwon, J., Oh, Y.T., Ko, E.S., Park, H.: Tumor-attentive segmentation-guided GAN for synthesizing breast contrast-enhanced mri without contrast agents. IEEE journal of translational engineering in health and medicine **11**, 32–43 (2022)
11. Kuhl, C.K., Schrading, S., Strobel, K., Schild, H.H., Hilgers, R.D., Bieling, H.B.: Abbreviated breast magnetic resonance imaging (MRI): first postcontrast subtracted images and maximum-intensity projection—a novel approach to breast cancer screening with MRI. Journal of Clinical Oncology **32**(22), 2304–2310 (2014)
12. Leithner, D., Moy, L., Morris, E.A., Marino, M.A., Helbich, T.H., Pinker, K.: Abbreviated MRI of the breast: does it provide value? Journal of Magnetic Resonance Imaging **49**(7), e85–e100 (2019)
13. Lekadir, K., Frangi, A.F., Porras, A.R., Glocker, B., Cintas, C., Langlotz, C.P., Weicken, E., Asselbergs, F.W., Prior, F., Collins, G.S., et al.: FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare. bmj **388** (2025)
14. Manzanera, O.E.M., Ellis, S., Baltatzis, V., Nair, A., Le Folgoc, L., Desai, S., Glocker, B., Schnabel, J.A.: Patient-specific 3d cellular automata nodule growth synthesis in lung cancer without the need of external data. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). pp. 5–9. IEEE (2021)
15. Mittal, A., Kalkhof, J., Mukhopadhyay, A., Bhavsar, A.: Medsegdiffnca: Diffusion models with neural cellular automata for skin lesion segmentation. arXiv preprint arXiv:2501.02447 (2025)
16. Mordvintsev, A., Randazzo, E., Niklasson, E., Levin, M.: Growing neural cellular automata. Distill **5**(2), e23 (2020)
17. Müller-Franzes, G., Huck, L., Tayebi Arasteh, S., Khader, F., Han, T., Schulz, V., Dethlefsen, E., Kather, J.N., Nebelung, S., Nolte, T., et al.: Using machine learning to reduce the need for contrast agents in breast MRI through synthetic images. Radiology **307**(3), e222211 (2023)
18. Osuala, R., Joshi, S., Tsirikoglou, A., Garrucho, L., Pinaya, W.H., Lang, D.M., Schnabel, J.A., Diaz, O., Lekadir, K.: Simulating Dynamic Tumor Contrast Enhancement in Breast MRI using Conditional Generative Adversarial Networks. arXiv preprint arXiv:2409.18872 (2024)
19. Osuala, R., Lang, D.M., Verma, P., Joshi, S., Tsirikoglou, A., Skorupko, G., Kushibar, K., Garrucho, L., Pinaya, W.H., Diaz, O., et al.: Towards learning contrast kinetics with multi-condition latent diffusion models. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 713–723. Springer (2024)
20. Pajouheshgar, E., Xu, Y., Zhang, T., Süsstrunk, S.: DyNCA: Real-time dynamic texture synthesis using neural cellular automata. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 20742–20751 (2023)

21. Ranem, A., Kalkhof, J., Mukhopadhyay, A.: NCA-Morph: Medical Image Registration with Neural Cellular Automata. arXiv preprint arXiv:2410.22265 (2024)
22. Richardson, A.D., Antal, T., Blythe, R.A., Schumacher, L.J.: Learning spatiotemporal patterns with Neural Cellular Automata. PLOS Computational Biology **20**(4), e1011589 (2024)
23. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. pp. 234–241. Springer (2015)
24. Saha, A., Harowicz, M.R., Grimm, L.J., Kim, C.E., Ghate, S.V., Walsh, R., Mazurowski, M.A.: A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features. British journal of cancer **119**(4), 508–516 (2018)
25. Schreiter, H., Eberle, J., Kapsner, L.A., Hadler, D., Ohlmeyer, S., Erber, R., Emons, J., Laun, F.B., Uder, M., Wenkel, E., et al.: 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. Springer (2024)
26. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003. vol. 2, pp. 1398–1402. Ieee (2003)
27. Zhang, L., Rao, A., Agrawala, M.: Adding conditional control to text-to-image diffusion models. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 3836–3847 (2023)
28. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 586–595 (2018)