Synthesis of Contrast-Enhanced Breast MRI Using T1- and Multi-b-Value DWI-Based Hierarchical Fusion Network with Attention Mechanism

Tianyu Zhang[1,2,3], Luyi Han[1,3], Anna D'Angelo[4], Xin Wang[1,2,3], Yuan Gao[1,2,3], Chunyao Lu[1,3], Jonas Teuwen[5], Regina Beets-Tan[1,2], Tao Tan[1,6(✉)], and Ritse Mann[1,3]

[1] Department of Radiology, Netherlands Cancer Institute (NKI), Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands
[2] GROW School for Oncology and Development Biology, Maastricht University, P. O. Box 616, 6200 MD Maastricht, The Netherlands
[3] Department of Diagnostic Imaging, Radboud University Medical Center, Geert Grooteplein 10, 6525 GA Nijmegen, The Netherlands
[4] Dipartimento di diagnostica per immagini, Radioterapia, Oncologia ed ematologia, Fondazione Universitaria A. Gemelli, IRCCS Roma, Roma, Italy
[5] Department of Radiation Oncology, Netherlands Cancer Institute (NKI), Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands
[6] Faculty of Applied Sciences, Macao Polytechnic University, Macao 999078, China
taotanjs@gmail.com

Abstract. Magnetic resonance imaging (MRI) is the most sensitive technique for breast cancer detection among current clinical imaging modalities. Contrast-enhanced MRI (CE-MRI) provides superior differentiation between tumors and invaded healthy tissue, and has become an indispensable technique in the detection and evaluation of cancer. However, the use of gadolinium-based contrast agents (GBCA) to obtain CE-MRI may be associated with nephrogenic systemic fibrosis and may lead to bioaccumulation in the brain, posing a potential risk to human health. Moreover, and likely more important, the use of gadolinium-based contrast agents requires the cannulation of a vein, and the injection of the contrast media which is cumbersome and places a burden on the patient. To reduce the use of contrast agents, diffusion-weighted imaging (DWI) is emerging as a key imaging technique, although currently usually complementing breast CE-MRI. In this study, we develop a multi-sequence fusion network to synthesize CE-MRI based on T1-weighted MRI and DWIs. DWIs with different b-values are fused to efficiently utilize the difference features of DWIs. Rather than proposing a pure data-driven approach, we invent a multi-sequence attention module to obtain

---

T. Zhang and L. Han—Equal contribution.

refined feature maps, and leverage hierarchical representation information fused at different scales while utilizing the contributions from different sequences from a model-driven approach by introducing the weighted difference module. The results show that the multi-b-value DWI-based fusion model can potentially be used to synthesize CE-MRI, thus theoretically reducing or avoiding the use of GBCA, thereby minimizing the burden to patients. Our code is available at https://github.com/Netherlands-Cancer-Institute/CE-MRI.

Keywords: Contrast-enhanced MRI · Diffusion-weighted imaging · Deep learning · Multi-sequence fusion · Breast cancer

1   Introduction

Breast cancer is the most common cancer and the leading cause of cancer death in women [18]. Early detection of breast cancer allows patients to receive timely treatment, which may have less burden and a higher probability of survival [6]. Among current clinical imaging modalities, magnetic resonance imaging (MRI) has the highest sensitivity for breast cancer detection [12]. Especially, contrast-enhanced MRI (CE-MRI) can identify tumors well and has become an indispensable technique for detecting and defining cancer [13]. However, the use of gadolinium-based contrast agents (GBCA) requires iv-cannulation, which is a burden to patients, time consuming and cumbersome in a screening situation. Moreover, contrast administration can lead to allergic reactions and finaly CE-MRI may be associated with nephrogenic systemic fibrosis and lead to bioaccumulation in the brain, posing a potential risk to human health [4,9,14–16]. In 2017, the European Medicines Agency concluded its review of GBCA, confirming recommendations to restrict the use of certain linear GBCA used in MRI body scans and to suspend the authorization of other contrast agents, albeit macrocyclic agents can still be freely used [10].

With the development of computer technology, artificial intelligence-based methods have shown potential in image generation and have received extensive attention. Some studies have shown that some generative models can effectively perform mutual synthesis between MR, CT, and PET [19]. Among them, synthesis of CE-MRI is very important as mentioned above, but few studies have been done by researchers in this area due to its challenging nature. Li et al. analyzed and studied the feasibility of using T1-weighted MRI and T2-weighted MRI to synthesize CE-MRI based on deep learning model [11]. Their results showed that the model they developed could potentially synthesize CE-MRI and outperform other cohort models. However, MRI source data of too few sequences (only T1 and T2) may not provide enough valuable informative to effectively synthesize CE-MRI. In another study, Chung et al. investigated the feasibility of using deep learning (a simple U-Net structure) to simulate contrast-enhanced breast MRI of invasive breast cancer, using source data including T1-weighted non-fat-suppressed MRI, T1-weighted fat-suppressed MRI, T2-weighted fat-suppressed

MRI, DWI, and apparent diffusion coefficient [5]. However, obtaining a complete MRI sequence makes the examination costly and time-consuming. On the other hand, the information provided by multi-sequences may be redundant and may not contain the relevant information of CE-MRI. Therefore, it is necessary to focus on the most promising sequences to synthesize CE-MRI.

Diffusion-weighted imaging (DWI) is emerging as a key imaging technique to complement breast CE-MRI [3]. DWI can provide information on cell density and tissue microstructure based on the diffusion of tissue water. Studies have shown that DWI could be used to detect lesions, distinguish malignant from benign breast lesions, predict patient prognosis, etc [1,3,7,8,17]. In particular, DWI can capture the dynamic diffusion state of water molecules to estimate the vascular distribution in tissues, which is closely related to the contrast-enhanced regions in CE-MRI. DWI may be a valuable alternative in breast cancer detection in patients with contraindications to GBCA [3]. Inspired by this, we develop a multi-sequence fusion network based on T1-weighted MRI and multi-b-value DWI to synthesize CE-MRI. Our contributions are as follows:

i From the perspective of method, we innovatively proposed a multi-sequence fusion model, designed for combining T1-weighted imaging and multi-b-value DWI to synthesize CE-MRI for the first time.

ii We invented hierarchical fusion module, weighted difference module and multi-sequence attention module to enhance the fusion at different scale, to control the contribution of different sequence and maximising the usage of the information within and across sequences.

iii From the perspective of clinical application, our proposed model can be used to synthesize CE-MRI, which is expected to reduce the use of GBCA.

2 Methods

2.1 Patient Collection and Pre-processing

This study was approved by Institutional Review Board of our cancer institute with a waiver of informed consent. We retrospectively collected 765 patients with breast cancer presenting at our cancer institute from January 2015 to November 2020, all patients had biopsy-proven breast cancers (all cancers included in this study were invasive breast cancers, and ductal carcinoma in situ had been excluded). The MRIs were acquired with Philips Ingenia 3.0-T scanners, and overall, three sequences were present in the in-house dataset, including T1-weighted fat-suppressed MRI, contrast-enhanced T1-weighted MRI and DWI. DWI consists of 4 different b-values (b = 0, b = 150, b = 800 and b = 1500). All MRIs were resampled to 1 mm isotropic voxels and uniformly sized, resulting in volumes of 352×352 pixel images with 176 slices per MRI, and subsequent registration was performed based on Advanced Normalization Tools (ANTs) [2].

2.2   Model

Figure 1 illustrates the structure of the proposed model. First, the reconstruction module is used to automatically encode and decode each input MRI sequence information to obtain the latent representation of different MRI sequences at multi-scale levels. Then, the hierarchical fusion module is used to extract the hierarchical representation information and fuse them at different scales.
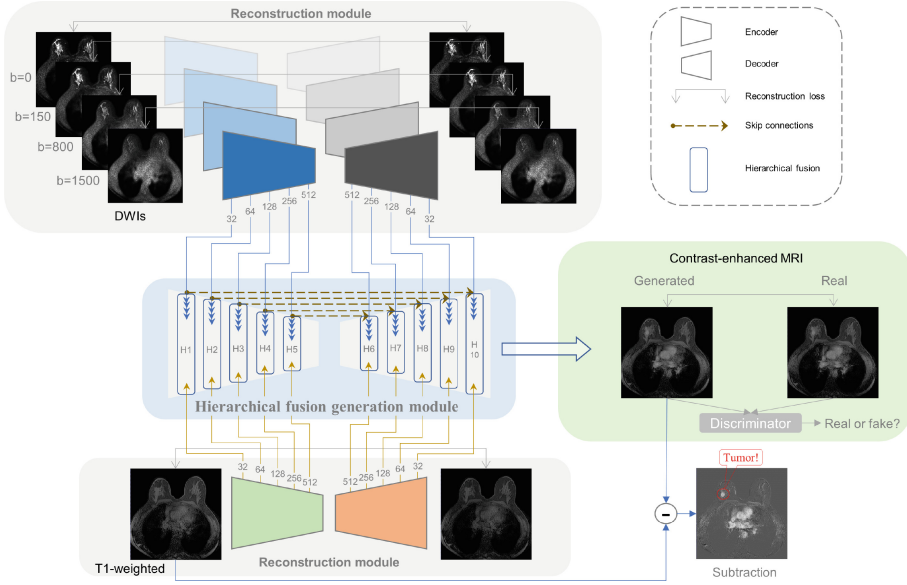

Fig. 1. Model details and flowchart for this study.

In each convolutional layer group of the reconstruction module, we use two 3×3 filters (same padding) with strides 1 and 2, respectively. The filters are followed by batch normalization, and after batch normalization, the activation functions *LeakyReLU* (with a slope of 0.2) and *ReLU* are used in the encoder and decoder, respectively. The l_1-norm is used as a reconstruction loss to measure the difference between the reconstructed image and the ground truth.

Figure 2 shows the detailed structure of the hierarchical fusion module, which includes two sub-modules, a weighted difference module and a multi-sequence attention module. The calculation of the apparent diffusion coefficient (ADC) map is shown in Eq. 1, which provides a quantitative measure of observed diffusion restriction in DWIs. Inspired by ADC, a weighted difference module is designed, in which the neural network is used to simulate the dynamic analysis of the ln function, and the element-wise subtraction algorithm is used to extract the differentiation features between DWIs with different b-values, and finally the features are weighted to obtain weighted feature maps (F_{DWI}, Eq. 2).

$$\text{ADC} = -\ln(S_h/S_l)/(b_h - b_l) = [\ln(S_l) - \ln(S_h)]/(b_h - b_l) \tag{1}$$

$$F_{\text{DWI}} = [f_{\theta^l}(S_l) - f_{\theta^h}(S_h)]/(b_h - b_l) \tag{2}$$

where S_l and S_h represent the image signals obtained from lower b value b_l and higher b_h, f_{θ^l} and f_{θ^h} represent the corresponding neural networks for DWI with a lower and higher b value.

In the multi-sequence attention module, a channel-based attention mechanism is designed to automatically apply weights (A_s) to feature maps (F_{concat}) from different sequences to obtain a refined feature map (F'_{concat}), as shown in Eq. 3. The input feature maps (F_{concat}) go through the maximum pooling layer and the average pooling layer respectively, and then are added element-wise after passing through the shared fully connected neural network, and finally the weight map A_s is generated after passing through the activation function, as shown in Eq. 4.

$$F'_{concat} \in \mathbb{R}^{C \times H \times W} = F_{concat} \otimes A_s \tag{3}$$

$$A_s = \sigma(f_{\theta^{fc}}(AvgPool(F_{concat})) \oplus f_{\theta^{fc}}(MaxPool(F_{concat}))) \tag{4}$$

where $\otimes$ represents element-wise multiplication, $\oplus$ represents element-wise summation, σ represents the sigmoid function, θ^{fc} represents the corresponding network parameters of the shared fully-connected neural network, and $AvgPool$ and $MaxPool$ represent average pooling and maximum pooling operations, respectively.

In the synthesis process, the generator $\mathcal{G}$ tries to generate an image according to the input multi-sequence MRI $(d_1, d_2, d_3, d_4, t_1)$, and the discriminator $\mathcal{D}$ tries to distinguish the generated image $G(d_1, d_2, d_3, d_4, t_1)$ from the real image y, and at the same time, the generator tries to generate a realistic image to mislead the discriminator. The generator's objective function is as follows:

$$\begin{aligned}\mathcal{L}_G(G, D) =& \mathbb{E}_{d_1,d_2,d_3,d_4,t_1 \sim pro_{data}(d_1,d_2,d_3,d_4,t_1)} \\ & [\log(1 - D(d_1, d_2, d_3, d_4, t_1, (G(d_1, d_2, d_3, d_4, t_1))))] \\ & + \lambda_1 \mathbb{E}_{d_1,d_2,d_3,d_4,t_1,y}[\|y - G(d_1, d_2, d_3, d_4, t_1)\|_1]\end{aligned} \tag{5}$$

and the discriminator's objective function is as follows:

$$\begin{aligned}\mathcal{L}_D(G, D) =& \mathbb{E}_{y \sim pro_{data}(y)}[\log D(y)] \\ & + \mathbb{E}_{d_1,d_2,d_3,d_4,t_1 \sim pro_{data}(d_1,d_2,d_3,d_4,d_1)} \\ & [\log(1 - D(G(d_1, d_2, d_3, d_4, t_1)))]\end{aligned} \tag{6}$$

where $pro_{data}(d_1, d_2, d_3, d_4, t_1)$ represents the empirical joint distribution of inputs d_1 (DWI_{b0}), d_2 (DWI_{b150}), d_3 (DWI_{b800}), d_4 (DWI_{b1500}) and t_1 (T1-weighted MRI), λ_1 is a non-negative trade-off parameter, and l_1-norm is used to measure the difference between the generated image and the corresponding ground truth. The architecture of the discriminator includes five convolutional

layers, and in each convolutional layer, 3×3 filters with stride 2 are used. Each filter is followed by batch normalization, and after batch normalization, the activation function *LeakyReLU* (with a slope of 0.2) is used. The numbers of filters are 32, 64, 128, 256 and 512, respectively.

2.3   Visualization

The T1-weighted images and the contrast-enhanced images were subtracted to obtain a difference MRI to clearly reveal the enhanced regions in the CE-MRI. If the CE-MRI was successfully synthesized, the enhanced region would be highlighted in the difference MRI, otherwise it would not.
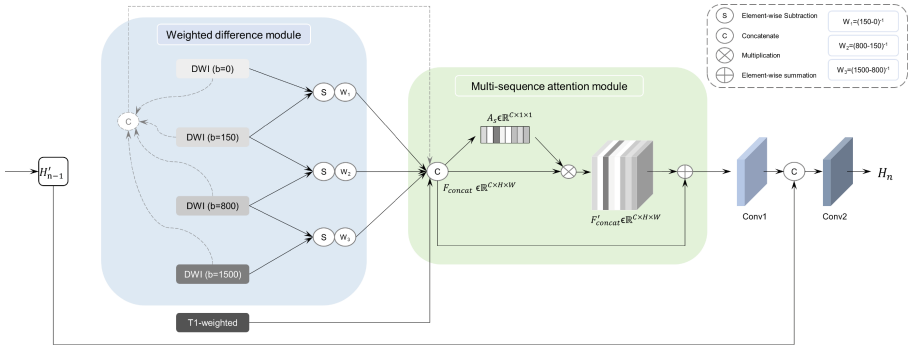

Fig. 2. Detailed structure of the hierarchical fusion module.

2.4   Experiment Settings

Based on the ratio of 8:2, the training set and independent test set of the in-house dataset have 612 and 153 cases, respectively. The trade-off parameter λ_1 was set to 100 during training, and the trade-off parameter of the reconstruction loss in the reconstruction module is set to 5. Masks for all breasts were used (weighted by a factor of 100 during the calculation of the loss between generated and real CE-MRI) to reduce the influence of signals in the thoracic area. The batch was set to 8 for 100 epochs, the initial learning rate was 1e−3 with a decay factor of 0.8 every 5 epochs (total run time is about 60 h). *Adam* optimizer was applied to update the model parameters. MMgSN-Net [11] and the method of Chung et al. [5] were used as cohort models, and all models were trained on NVIDIA RTX A6000 48 GB GPU.

2.5   Evaluation Metrics

Results analysis was performed by Python 3.7. Structural Similarity Index Measurement (SSIM), Peak Signal-to-Noise Ratio (PSNR) and Normalized Mean

Squared Error (NMSE) were used as metrics, all formulas as follows:

$$SSIM = \frac{(2\mu_{y(x)}\mu_{G(x)} + c_1)(2\sigma_{y(x)G(x)} + c_2)}{(\mu_{y(x)}^2 + \mu_{G(x)}^2 + c_1)(\sigma_{y(x)}^2 + \sigma_{G(x)}^2 + c_2)} \tag{7}$$

$$PSNR = 10\log_{10}\frac{\max^2(y(x), G(x))}{\frac{1}{N}\|y(x) - G(x)\|_2^2} \tag{8}$$

$$NMSE = \frac{y(x) - G(x)}{\|y(x)\|_2^2} \tag{9}$$

**Table 1.** Results for synthesizing breast ceT1 MRI for different models.

Method	Sequence	SSIM↑	PSNR↑	NMSE↓
MMgSN-Net [11]	T1+DWI(0)	86.61 ± 2.52	26.39 ± 1.38	0.0982 ± 0.038
Chung et al. [5]	T1+DWIs	87.58 ± 2.68	27.80 ± 1.56	0.0692 ± 0.035
Proposed	**T1+DWIs**	**89.93 ± 2.91**	**28.92 ± 1.63**	**0.0585 ± 0.026**

where $G(x)$ represents a generated image, $y(x)$ represents a ground-truth image, $\mu_{y(x)}$ and $\mu_{G(x)}$ represent the mean of $y(x)$ and $G(x)$, respectively, $\sigma_{y(x)}$ and $\sigma_{G(x)}$ represent the variance of $y(x)$ and $G(x)$, respectively, $\sigma_{y(x)G(x)}$ represents the covariance of $y(x)$ and $G(x)$, and c_1 and c_2 represent positive constants used to avoid null denominators.

3   Results

First, we compare the performance of different existing methods on synthetic CE-MRI using our source data, The quantitative indicators used include PSNR, SSIM and NMSE. As shown in Table 1, the SSIM of MMgSN-Net [11] and the method of Chung et al. [5] in synthesizing ceT1 MRI is 86.61 ± 2.52 and 87.58 ± 2.68, respectively, the PSNR is 26.39 ± 1.38 and 27.80 ± 1.56, respectively, and the NMSE is 0.0982 ± 0.038 and 0.0692 ± 0.035, respectively. In contrast, our proposed multi-sequence fusion model achieves better SSIM of 89.93 ± 2.91, better PSNR of 28.92 ± 1.63 and better NMSE of 0.0585 ± 0.026 in synthesizing ceT1 MRI, outperforming existing cohort models.

MMgSN-Net [11] combined T1-weighted and T2-weighted MRI in their work to synthesize CE-MRI. Here we combined T1-weighted MRI and DWI with a b-value of 0 according to their method, but the model did not perform well. It may be because their model can only combine bi-modality and cannot integrate the features of all sequences, so it cannot mine the difference features between multiple b-values, which limits the performance of the model. In addition, although the method of Chung et al. [5] used full-sequence MRI to synthesize CE-MRI, it would be advantageous to obtain synthetic CE-MRI images using as little data as possible, taking advantage of the most contributing sequences. They did not take advantage of multi-b-value DWI, nor did they use the hierarchical fusion module to fully fuse the hierarchical features of multi-sequence MRI.
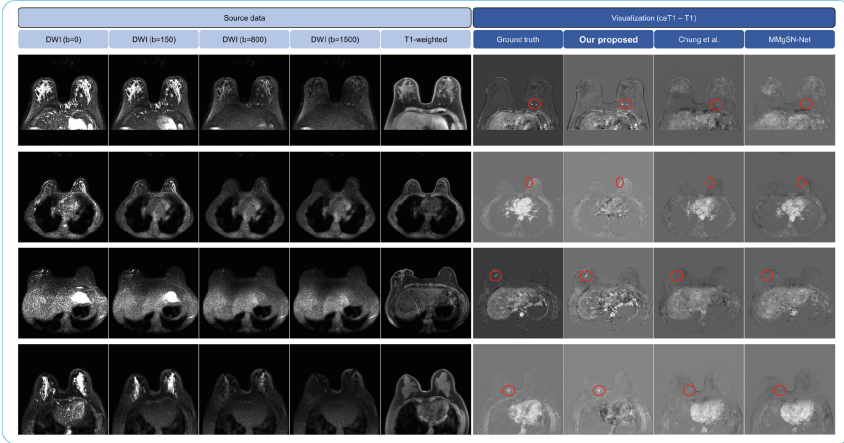
Fig. 3. Some examples of source data and visualization results

As described in Methods, the proposed model consists of several key components, including a hierarchical fusion generation module, a weighted difference module, and a multi-sequence attention module. Therefore, ablation studies were performed to demonstrate the importance and effectiveness of our three key components. Several network structures were selected for comparison, as follows: (1) Input-level fusion network without other modules (called IF-Net), (2) Hierarchical fusion generation network combined with reconstruction module, without weighted difference module and multi-sequence attention module (called HF-Net), (3) Hierarchical fusion generation network with weighted difference module (called HFWD-net), (4) Hierarchical fusion generation network with weighted difference module and multi-sequence attention module (proposed model). As shown in Table 2, IF-Net achieves SSIM of 87.25 ± 2.62 and PSNR of 26.51 ± 1.52 in synthesizing ceT1 MRI. HF-Net achieves SSIM of 88.32 ± 2.70 and PSNR of 27.95 ± 1.59 in synthesizing ceT1 MRI. After adding the weighted difference module, SSIM and PSNR were improved to 89.18 ± 2.73 and 28.45 ± 1.61, respectively. Finally, the addition of the multi-sequence attention module further improved the performance of the model, with SSIM of 89.93 ± 2.91, PSNR of 28.92 ± 1.63, and NMSE of 0.0585 ± 0.026.

Table 2. Ablation results for synthesizing breast ceT1 MRI.

Method	Sequence	SSIM↑	PSNR↑	NMSE↓
IF-Net	T1+DWIs	87.25 ± 2.62	26.51 ± 1.52	0.0722 ± 0.032
HF-Net	T1+DWIs	88.32 ± 2.70	27.95 ± 1.59	0.0671 ± 0.033
HFWD-Net	T1+DWIs	89.18 ± 2.73	28.45 ± 1.61	0.0622 ± 0.030
Proposed	**T1+DWIs**	**89.93 ± 2.91**	**28.92 ± 1.63**	**0.0585 ± 0.026**

The visualization results of random samples are shown in Fig. 3. It can be seen from the visualization results that after the difference between the generated CE-MRI and the original T1-weighted MRI, the lesion position of the breast is highlighted, the red circle represents the highlighted area, which proves that our method can effectively synthesize contrast-enhanced images, highlighting the same parts as the real enhanced position. See Supplementary Material for more visualization results, including visualizations of breast CE-MRI synthesized in axial, coronal, and sagittal planes.

4   Conclusion

We have developed a multi-sequence fusion network based on multi-b-value DWI to synthesize CE-MRI, using source data including DWIs and T1-weighted fat-suppressed MRI. Compared to existing methods, we avoid the challenges of using full-sequence MRI and aim to be selective on valuable source data DWI. Hierarchical fusion generation module, weighted difference module, and multi-sequence attention module have all been shown to improve the performance of synthesizing target images by addressing the problems of synthesis at different scales, leveraging differentiable information within and across sequences. Given that current research on synthetic CE-MRI is relatively sparse and challenging, our study provides a novel approach that may be instructive for future research based on DWIs. Our further work will be to conduct reader studies to verify the clinical value of our research in downstream applications, such as helping radiologists on detecting tumors. In addition, synthesizing dynamic contrast-enhanced MRI at multiple time points will also be our future research direction. Our proposed model can potentially be used to synthesize CE-MRI, which is expected to reduce or avoid the use of GBCA, thereby optimizing logistics and minimizing potential risks to patients.

References

1. Amornsiripanitch, N., Bickelhaupt, S., Shin, H.J., Dang, M., Rahbar, H., Pinker, K., Partridge, S.C.: Diffusion-weighted MRI for unenhanced breast cancer screening. Radiology **293**(3), 504–520 (2019)
2. Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C.: A reproducible evaluation of ants similarity metric performance in brain image registration. Neuroimage **54**(3), 2033–2044 (2011)
3. Baltzer, P., et al.: Diffusion-weighted imaging of the breast-a consensus and mission statement from the EUSOBI international breast diffusion-weighted imaging working group. Eur. Radiol. **30**, 1436–1450 (2020)
4. Broome, D.R., Girguis, M.S., Baron, P.W., Cottrell, A.C., Kjellin, I., Kirk, G.A.: Gadodiamide-associated nephrogenic systemic fibrosis: why radiologists should be concerned. Am. J. Roentgenol. **188**(2), 586–592 (2007)

5. Chung, M., et al.: Deep learning to simulate contrast-enhanced breast MRI of invasive breast cancer. Radiology **306**, 213199 (2022)

6. Goldhirsch, A., et al.: Personalizing the treatment of women with early breast cancer: highlights of the St Gallen international expert consensus on the primary therapy of early breast cancer 2013. Ann. Oncol. **24**(9), 2206–2223 (2013)

7. van der Hoogt, K.J.J., et al.: Factors affecting the value of diffusion-weighted imaging for identifying breast cancer patients with pathological complete response on neoadjuvant systemic therapy: a systematic review. Insights Imaging **12**(1), 1–22 (2021). https://doi.org/10.1186/s13244-021-01123-1

8. Iima, M., Honda, M., Sigmund, E.E., Ohno Kishimoto, A., Kataoka, M., Togashi, K.: Diffusion MRI of the breast: current status and future directions. J. Magn. Reson. Imaging **52**(1), 70–90 (2020)

9. Kanda, T., Ishii, K., Kawaguchi, H., Kitajima, K., Takenaka, D.: High signal intensity in the dentate nucleus and globus pallidus on unenhanced t1-weighted MR images: relationship with increasing cumulative dose of a gadolinium-based contrast material. Radiology **270**(3), 834–841 (2014)

10. Kleesiek, J., et al.: Can virtual contrast enhancement in brain MRI replace gadolinium?: a feasibility study. Invest. Radiol. **54**(10), 653–660 (2019)

11. Li, W., et al.: Virtual contrast-enhanced magnetic resonance images synthesis for patients with nasopharyngeal carcinoma using multimodality-guided synergistic neural network. Int. J. Radiat. Oncol.* Biol.* Phys. **112**(4), 1033–1044 (2022)

12. Mann, R.M., Cho, N., Moy, L.: Breast MRI: state of the art. Radiology **292**(3), 520–536 (2019)

13. Mann, R.M., Kuhl, C.K., Moy, L.: Contrast-enhanced MRI for breast cancer screening. J. Magn. Reson. Imaging **50**(2), 377–390 (2019)

14. Marckmann, P., et al.: Nephrogenic systemic fibrosis: suspected causative role of gadodiamide used for contrast-enhanced magnetic resonance imaging. J. Am. Soc. Nephrol. **17**(9), 2359–2362 (2006)

15. Nguyen, N.C., Molnar, T.T., Cummin, L.G., Kanal, E.: Dentate nucleus signal intensity increases following repeated gadobenate dimeglumine administrations: a retrospective analysis. Radiology **296**(1), 122–130 (2020)

16. Olchowy, C., et al.: 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 (2017)

17. Partridge, S.C., Newitt, D.C., Chenevert, T.L., Rosen, M.A., Hylton, N.M., Team, A.T., Investigators, I.S.T.: Diffusion-weighted MRI in multicenter trials of breast cancer. Radiology **291**(2), 546–546 (2019)

18. Sung, H., et al.: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: Cancer J. Clin. **71**(3), 209–249 (2021)

19. Yi, X., Walia, E., Babyn, P.: Generative adversarial network in medical imaging: a review. Med. Image Anal. **58**, 101552 (2019)