Myocardial Scar Enhancement in LGE Cardiac MRI using Localized Diffusion

Marta Hasny[1,2], Omer B. Demirel[1], Amine Amyar[1], Shahrooz Faghihroohi[2], Reza Nezafat[1]

[1] Department of Medicine (Cardiovascular Division), Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA
[2] TUM School of Computation, Information and Technology, Technical University of Munich, Munich, Germany
marta.hasny@tum.de

Abstract. Late gadolinium enhancement (LGE) imaging is considered the gold-standard technique for evaluating myocardial scar/fibrosis. In LGE, an inversion pulse is played before imaging to create a contrast between healthy and scarred regions. However, several factors can impact the contrast quality, impacting diagnostic interpretation. Furthermore, the quantification of scar burden is highly dependent on image quality. Deep learning-based automated segmentation algorithms often fail when there is no clear boundary between healthy and scarred tissue. This study sought to develop a generative model for improving the contrast of healthy-scarred myocardium in LGE. We propose a localized conditional diffusion model, in which only a region-of-interest (ROI), in this case heart, is subjected to the noising process, adapting the learning process to the local nature of our proposed enhancement. The scar-enhanced images, used as training targets, are generated via tissue-specific gamma correction. A segmentation model is trained and used to extract the heart regions. The inference speed is improved by leveraging partial diffusion, applying noise only up to an intermediate step. Furthermore, utilizing the stochastic nature of diffusion models, repeated inference leads to improved scar enhancement of ambiguous regions. The proposed algorithm was evaluated using LGE images collected in 929 patients with hypertrophic cardiomyopathy, in a multi-center, multi-vendor study. Our results show visual improvements of scar-healthy myocardium contrast. To further demonstrate the strength of our method, we evaluate our performance against various image enhancement models where the proposed approach shows higher contrast enhancement. The code is available at: https://github.com/HMS-CardiacMR/Scar_enhancement.

Keywords: Image Enhancement, Conditional Diffusion, Cardiac MR.

1 Introduction

Late gadolinium enhancement (LGE) cardiac MR (CMR) is the gold standard for evaluating myocardial scar. Using an inversion-recovery sequence, LGE images are acquired 10-15 minutes after administering gadolinium-based contrast. The delay

between inversion time and imaging is selected to null the healthy myocardium. This results in LGE images in which the scar appears bright on a dark healthy myocardium background. However, improper nulling of the myocardium [1], contrast dose/type [2], and time between contrast infusion and imaging [3] can result in suboptimal contrast between the scar and the healthy myocardium. Different imaging sequences and sampling schemes (e.g. Cartesian, radial, and spiral sequences) also impact the LGE contrast. These factors could reduce the contrast, impacting diagnostic image quality and interpretation. Furthermore, the quantification of scar burden is highly dependent on image contrast. Deep learning-based automated segmentation algorithms often fail when there is insufficient contrast between healthy and scarred tissue. Low image contrast is particularly important in the setting of non-ischemic cardiomyopathies where the scar/fibrosis is often more heterogeneous and with unclear boundaries.

Over the past two decades, significant advances have been made to improve LGE diagnostic image quality. This includes introducing phase-sensitive inversion recovery [4], more efficient acquisition sampling [5], dark-blood LGE [6], and 3D high-resolution imaging [7]. There have also been advances in analysis focusing on automating scar segmentation. With advances in deep learning models, there have been several recent advances in automating scar segmentation [8,9,10]. However, to our knowledge, no work has been performed to improve scar-healthy myocardium contrast after data acquisition. Therefore, this work sought to enhance the contrast between the scar and healthy myocardium in LGE images by developing a diffusion model to enhance the scarred region without impacting the healthy myocardium. The contrast enhancement aims to simplify image interpretation, helping to resolve ambiguities regarding the presence and extent of the scar tissue.

We introduce a generative AI framework for image-based contrast enhancement, leveraging regional diffusion models to improve scar contrast in LGE. We propose a conditional localized diffusion model where noise is applied to a predefined region-of-interest (ROI) containing the heart, rather than the whole image. We accelerate the inference speed of our model by exploring the idea of partial diffusion, starting the reverse diffusion pass from only a partially noised image. Lastly, leveraging the stochastic nature of diffusion models and adapting to the varying contrasts of the original images, we define our final enhancement as a mean of images generated by repeating the diffusion process multiple times. Thus, the enhancement level is related to the model's certainty that a scar is present in the given region, improving the enhancement of ambiguous regions.

2 Methodology

This section discusses each step of the proposed framework in detail, with each block of the method presented in Fig. 1.

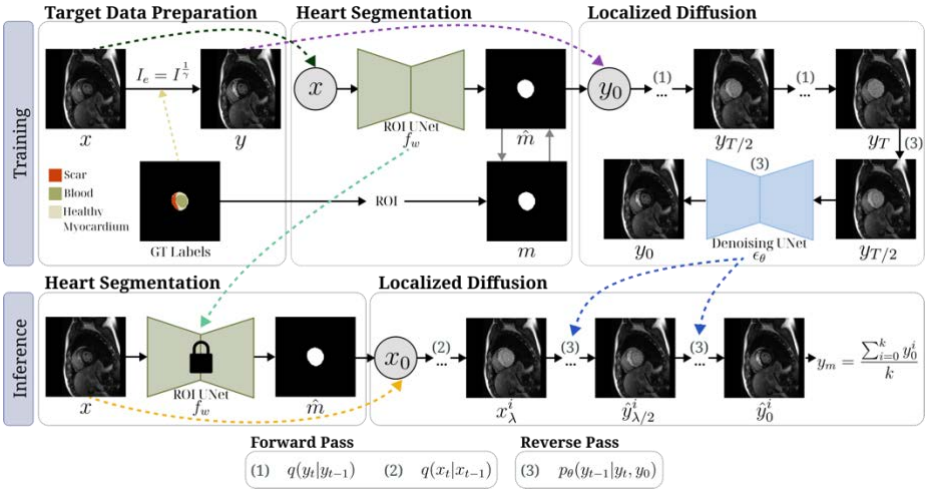

Fig. 1. Diagram representing all steps of our method. In the training process, we first generate the training data y using tissue-specific gamma correction, aided by the ground truth tissue labels. Next, a binary segmentation model f_w is trained to extract the heart region, used as the ROI. The segmentation together with the scar-enhanced images are used for the training of the diffusion model ϵ_θ. At inference, we use the trained segmentation model f_w for the ROI acquisition, which is then used for the enhancement generation using a partial diffusion scheme.

2.1 Target Data Preparation

As a first step of the framework, we generate scar-enhanced images via tissue-specific gamma correction. These images serve as targets for training purposes. First, the intensities of each image are scaled to a range of [0, 255]. For each pixel I in the original image x, gamma correction is applied using $I_e = I^{1/\gamma}$, where I_e represents the enhanced pixel and γ is the gamma correction value. Gamma values above 1 increase the contrast of the region, while values below 1 decrease it. Segmentation masks of the scar ($mask_{scar}$), healthy myocardium ($mask_{myo}$), and blood pool ($mask_{blood}$) are necessary to generate the scar enhanced target data. Masks provided with the dataset are utilized for this step. The image regions annotated as scars undergo the gamma correction using $\gamma > 1$, increasing the brightness of the scar. The strength of γ applied to the scar depends on the original contrast between the scar and the blood pool, ensuring that brightened scars are of different intensity than the blood pool. Additionally, blood pools with mean brightness above average are subjected to $\gamma < 1$. The average is calculated from the mean intensities of all the blood pools in the dataset. We denote the scar-enhanced images as $\boldsymbol{y}$. The selection of the gamma correction values was subjective to achieve a realistic looking LGE with improved scar visualization.

2.2 Localized Diffusion

Heart Segmentation. While diffusion models generally operate by gradually adding noise to the whole image, our task focuses on heart regions. To ensure the remaining regions of the image stay intact and the model's focus is defined on the area of the scar, we introduce localized diffusion. The ROI is defined as a binary mask $m = mask_{scar} + mask_{blood} + mask_{myo}$, denoting the heart. A U-Net model f_w [11] is trained using the original images and their corresponding binary labels. The ROI masks, used for both training and inference, are obtained using the trained model as $\hat{m} = f_w(x)$. The results of heart segmentation are available in Supp. Mat. Table 1.

Diffusion Process. We follow the process of the denoising diffusion probabilistic models (DDPM) [12]. DDPM is fixed to a T-step Markov chain, converting the data distribution $q(x_0)$ into normal distribution $q(x_T)$ through a series of gradually added noise ϵ under a variance schedule $\beta_t \in (0,1), t = 1, \dots, T$, where x_0 is the original image and x_T is a fully noised image. In this work, the DDPM process is modified by restricting the application of noise to ROI only rather than the entire image. Thus, we alter the forward process by introducing $\hat{m}$ as a noise generation constraint, i.e.,

$$q(x_t|x_{t-1}) := \mathcal{N}\left(x_t; \sqrt{1 - B_t}x_{t-1}, \beta_t \hat{\boldsymbol{m}}\right). \tag{1}$$

As defined by Ho *et al.* [6], the noisy image x_T can be acquired in a single step, using:

$$q(x_t|x_0) := \mathcal{N}\left(x_t; \sqrt{\overline{\alpha}_t}x_0, (1 - \overline{\alpha}_t)\hat{\boldsymbol{m}}\right), \tag{2}$$

where $\alpha_t := 1 - \beta_t$ and $\overline{\alpha}_t := \prod_{s=1}^{t} \alpha_s$. The formula can be further parameterized into

$$x_t = \sqrt{\overline{\alpha}_t}x_0 + \sqrt{1 - \overline{\alpha}_t}\hat{\boldsymbol{m}}\epsilon, \epsilon \sim \mathcal{N}(0, I). \tag{3}$$

DDPM approximates the reverse diffusion process via a Gaussian transition, conditioned on x_0, and parameterized by θ following

$$p_\theta(x_{t-1}|x_t, x_0) := \mathcal{N}\left(x_{t-1}; \mu_\theta(x_t, x_0, t), \Sigma_\theta(x_t, x_0, t)\right), \tag{4}$$

where $\Sigma_\theta(x_t, x_0, t)$ is a fixed set of scalar covariances and $\mu_\theta(x_t, x_0, t)$ is a learned posterior mean. A noise predictor network $\epsilon_\theta(x_t, x_0, t)$ is employed to predict the noise added to the data at step t. The model is trained using the scar-enhanced images, with the corresponding original images used for conditioning. The loss function of the denoising network ϵ_θ is calculated solely on the ROI and is defined as:

$$\mathcal{L} = E_{t,\epsilon,x}|(\epsilon_\theta(y_t, x_0, t) - \epsilon)\hat{m}|^2. \tag{5}$$

2.3 Inference

Partial Diffusion. Diffusion models suffer from a slow inference speed due to a high number of diffusion steps. Partial diffusion is a recent improvement on that issue. AnoDDPM [13] explored partial diffusion for the removal of large abnormalities in

brain MR, while PartDiff [14] used partial diffusion on a task of the super resolution, observing that the latent variables of high and low-resolution images are similar to each other. Given the nature of our task, the structural information of the image is expected to be preserved. Thus, it is unnecessary to remove the structural information of the image by performing the entire forward and reverse pass up to step T. We employ a partial diffusion inference scheme, where the input image x_0 is noised to an intermediate step λ, where $1 \leq \lambda \leq T$. We define x_λ as the starting point of the reverse diffusion process, i.e., $x_0 \rightarrow x_\lambda \rightarrow \hat{y}_0$. With partial diffusion, the number of sampling steps at inference is reduced, improving the speed of the framework. Through empirically tuning the λ value on the validation split, we set $\lambda = 250$, unless specified otherwise.

Mean Enhancement. Following the stochastic nature of diffusion models and ambiguities of scar borders, the framework proposes repeating the forward and reverse diffusion process k-times. The parameter θ remains the same at each repetition, while the noise initialization changes. Each of the generated k images is used to generate the final enhanced image

$$\hat{y}_m = \frac{\sum_{i=0}^{k} \hat{y}_0^{\,i}}{k}, \tag{6}$$

improving the enhancement of ambiguous regions. Through empirical tuning, we set k=5 for our experiments.

3 Experiments

3.1 Dataset

LGE CMR images from 929 patients with hypertrophic cardiomyopathy (62% male, 46±17 years) [15] were used for the training and evaluation of our framework. Participants of the study signed consent statements approved by the Institutional Review Board (IRB) of the participating centers, permitting the use of their medical information for research purposes. The data was split into training (n=630, 386 with scar), validation (n=111, 65 with scar), and testing (n=188, 117 with scar) sets. This dataset is part of a multicenter and multivendor study (GE, Philips, Siemens) at two field strengths (1.5, 3T). Masks delineating the regions of the scar, healthy myocardium, and blood pool were manually drawn for each subject. The images are centered on the left ventricle and padded to a uniform size of 256x256.

3.2 Implementation

All experiments are implemented in PyTorch and performed on a single NVIDIA Tesla V100 SXM2 32GB GPU. Diffusion models are trained using a batch size of 16 for 200 epochs. Adam optimizer with a learning rate of 10^{-4} is used and T is set to 1000. We adopt the denoising U-Net architecture of SR3 [16]. A linear variance schedule is used.

For heart segmentation, we use the U-Net architecture available through the MONAI framework [11] with a sigmoid added on the last layer. The model is trained using the ADAM optimizer with a learning rate of 10^{-3}, batch size of 64 for 100 epochs. The epoch with the best validation performance under the Dice metric is used for the generation of the heart masks.

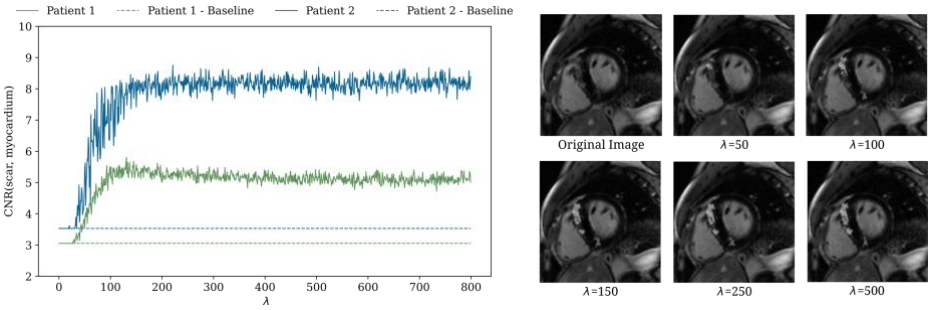

Fig. 2. The left side of the figure shows changes in the CNR along different λ values for two patients with varying scar burden, i.e., 12% for patient 1 and 3% for patient 2. Visual results for patient 1 are shown on the right side of the figure.

3.3 Evaluation Metrics

The results are evaluated based on two criteria, 1) preservation of the structural information, and 2) the contrast improvement. To assess structure preservation, we utilize the Structural Similarity Index (SSIM), Peak Signal-to-Noise Ratio (PSNR), and Learned Perceptual Image Patch Similarity (LPIPS). Due to the focus of the work on the heart, we crop the images to a size of 16x16 or larger, based on the borders of their ground truth ROIs, and calculate the metrics using these cropped images. The comparison is made between scar-enhanced images and the output images generated by the models. To evaluate contrast enhancement, we calculate the relative contrast-to-noise ratio (CNR) between the scar and the healthy myocardium (myo), as well as between the scar and the blood pool, using the following formula: $CNR(tissue1, tissue2) = |mean(tissue1) - mean(tissue2)|/std(tissue2)$.

3.4 Partial Diffusion

We perform a comprehensive analysis of the partial diffusion scheme. Fig. 2 represents the changes in the CNR between the scar and healthy myocardium along λ values from 1 to 800. The two patient cases selected for the analysis represent varying scar burden, with scar volumes of 12% and 3% for patients 1 and 2 respectively. The scar volume represents the percentage of the scar tissue in the whole myocardium. The values of the CNR increase between λ value of 1 to 200, after which they stagnate and remain at a similar level. The visual results follow that pattern. This suggests that performing a full T-step diffusion does not benefit the contrast enhancement while adding onto inference time. The optimal λ values vary between the two patients. However, while picking a

small λ can hurt the performance of the enhancement mechanism, a larger value ensures better performance while increasing the inference time. Through a grid search (Supp. Mat. Table 2), we found $\lambda = 250$ to be an empirically optimal value for our dataset.

3.5 Performance Comparison

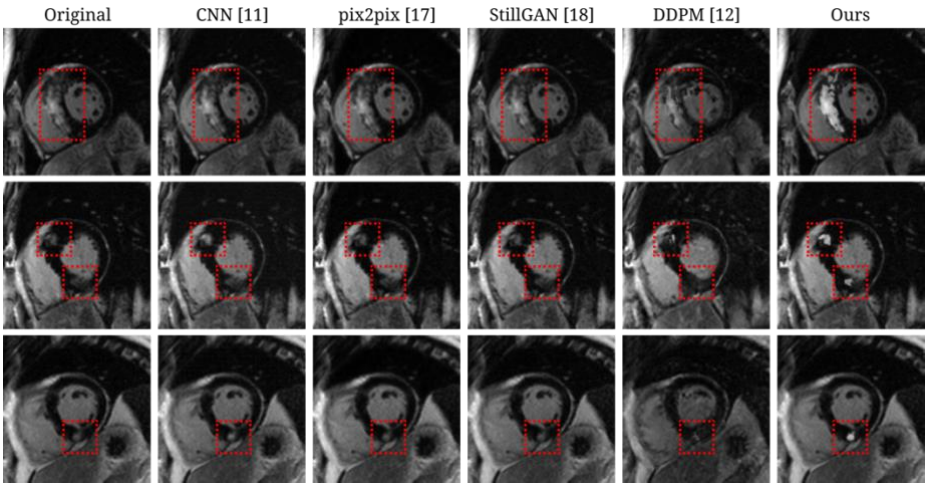

Fig. 3. Visual comparison of our framework against other methods. The red boxes highlight the areas where scars are located.

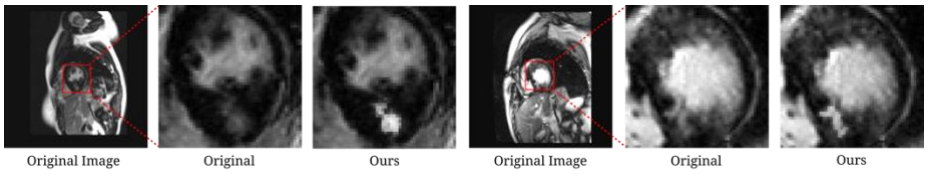

Fig. 4. A zoomed-in representation of the changes introduced by our framework.

Table 1. Quantitative results of our framework against other image enhancement methods.

Model	SSIM%↑	PSNR↑	LPIPS↓	CNR(scar,myo)↑	CNR(scar,blood)↑
CNN [11]	90.41±8.06	28.76 ±4.48	0.10±0.07	3.80±1.48	0.94±0.91
pix2pix [17]	90.05±10.2	30.04±7.12	0.10±0.09	3.59±1.52	0.85±0.81
StillGAN [18]	92.91±6.96	31.49±7.49	0.07±0.07	3.61±1.39	0.86±0.80
DDPM [12]	70.72±8.34	26.10±3.99	0.24±0.07	2.56±1.55	0.93±1.23
Ours	**96.06±6.24**	**43.75±18.04**	**0.04±0.06**	**4.97±2.42**	**1.65±1.70**
Baseline	-	-	-	3.62±1.40	0.85±0.96

We evaluate the performance of our framework against several image translation methods. Specifically, we evaluate our model against a CNN network [11], GAN-based models pix2pix [17] and StillGAN [18], and unconditional DDPM [12]. Given, that those models are not designed for our particular task, we tune each of them and present

the best results. We employ the partial diffusion scheme on the unconditional DDPM network with λ=50.

Quantitative results are presented in Table 1. The baseline represents original images. Our framework performs best in terms of all the metrics, clearly improving the contrast over the baseline values. The GAN-based models fail at improving the CNR values for both scar and myocardium and scar and blood pool. CNN achieves higher CNR than the baseline. However, the increase is not substantial. Wilcoxon signed-rank test was performed, and our method showed significant improvements compared to all other methods with $P<0.01$ on all the metrics. The global focus of those methods, with the local aim of our task, could be the reason the methods fail at enhancing the scar. Visual results, presented in Fig. 3, support this claim, showing that while the methods managed to correctly reconstruct the structure of the original image, little enhancement is applied. We present a closer view on our results in Fig. 4. The unconditional DDPM achieves the poorest performance, decreasing against the baseline CNR values.

3.6 Ablation Study

We perform an ablation study to evaluate the efficiency of our framework. The results are presented in Table 2. The study analyzes the effect of localized diffusion and mean enhancement on the performance of conditional diffusion. The addition of the localized diffusion improves on all the metrics, with the SSIM improving by 0.79 and the PSNR by 6.45dB. Both of the calculated CNR values are improved upon the inclusion of ROI-only noising, showing the efficiency of our method on local image enhancement. However, while ROI-only noising improves the results, heart segmentation missing on the scar tissue can lead to failure in the contrast enhancement. Calculation of the mean image add additional improvements on both structural and contrast enhancement levels.

Table 2. Results of the ablation study examining the strength of our methodology. Local stands for localized diffusion and mean represents inclusion of mean enhancement over k-images.

Local	Mean	SSIM%↑	PSNR↑	LPIPS↓	CNR(scar,myo)↑	CNR(scar,blood)↑
×	×	94.87±6.89	36.06±8.23	0.05±0.06	4.67±2.24	1.48±1.57
✓	×	95.66±6.54	42.51±16.89	0.04±0.06	4.86±2.43	**1.66±1.75**
✓	✓	**96.06±6.24**	**43.75±18.04**	**0.04±0.06**	**4.97±2.42**	1.65±1.70
Baseline		-	-	-	3.62±1.40	0.85±0.96

4 Conclusion

This is the first study to develop and evaluate a diffusion-based model to improve visualization of scar in LGE images. A localized diffusion model under a partial diffusion inference scheme allowed rapid LGE image enhancement. To adapt to the contrast changes in original images, we repeat the forward and reverse diffusion processes k-times to improve enhancement of the ambiguous regions of the scar. Through our framework, we show conditional diffusion models could be a promising method for scar enhancement. While localized nosing improves the results of scar enhancement, the heart segmentation network could miss the subendocardial scars and

thus exclude them from the enhancement area, constituting for a limitation of our framework. In future research this limitation can be mitigated via dilation, enlarging the segmentation region to ensure inclusion of the whole heart. In clinical practice, improper heart segmentation can be corrected by a clinician. Further research is needed to study the clinical implications of the contrast enhancement on the scar quantification.

Acknowledgments. Reza Nezafat receives grant funding from the National Institutes of Health (NIH) 1R01HL129185, 1R01HL129157, 1R01HL127015, and 1R01HL154744 (Bethesda, MD, USA).

Disclosure of Interests. The authors report no competing interest. Reza Nezafat has a research agreement with Siemens Healthineers, the manufacturer of MRI system used in a subset of patients for imaging.

References

1. Syed, I.S., Glockner, J.F., Feng, D., Araoz, P.A., Martinez, M.W., Edwards, W.D., Gertz, M.A., Dispenzieri, A., Oh, J.K., Bellavia, D., et al.: Role of cardiac magnetic resonance imaging in the detection of cardiac amyloidosis. JACC: cardiovascular imaging 3(2), 155–164 (2010)
2. Jenista, E.R., Wendell, D.C., Azevedo, C.F., Klem, I., Judd, R.M., Kim, R.J., Kim, H.W.: Revisiting how we perform late gadolinium enhancement cmr: Insights gleaned over 25 years of clinical practice. Journal of Cardiovascular Magnetic Resonance 25(1), 1–16 (2023)
3. Kellman, P., Arai, A.E.: Cardiac imaging techniques for physicians: late enhancement. Journal of magnetic resonance imaging 36(3), 529–542 (2012)
4. Kellman, P., Arai, A.E., McVeigh, E.R., Aletras, A.H.: Phase-sensitive inversion recovery for detecting myocardial infarction using gadolinium-delayed hyperenhancement. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine 47(2), 372–383 (2002)
5. Captur, G., Lobascio, I., Ye, Y., Culotta, V., Boubertakh, R., Xue, H., Kellman, P., Moon, J.C.: Motion-corrected free-breathing lge delivers high quality imaging and reduces scan time by half: an independent validation study. The International Journal of Cardiovascular Imaging 35, 1893–1901 (2019)
6. Basha, T., Roujol, S., Kissinger, K.V., Goddu, B., Manning, W.J., Nezafat, R.: Black blood late gadolinium enhancement using combined t 2 magnetization preparation and inversion recovery. Journal of Cardiovascular Magnetic Resonance 17, 1–2 (2015)
7. Rutz, T., Piccini, D., Coppo, S., Chaptinel, J., Ginami, G., Vincenti, G., Stuber, M., Schwitter, J.: Improved border sharpness of post-infarct scar by a novel self-navigated free-breathing high-resolution 3d whole-heart inversion recovery magnetic resonance approach. The international journal of cardiovascular imaging 32, 1735–1744 (2016)
8. Fahmy, A.S., Rowin, E.J., Chan, R.H., Manning, W.J., Maron, M.S., Nezafat, R.: Improved quantification of myocardium scar in late gadolinium enhancement images: deep learning based image fusion approach. Journal of Magnetic Resonance Imaging 54(1), 303–312 (2021)

9. Popescu, D.M., Abramson, H.G., Yu, R., Lai, C., Shade, J.K., Wu, K.C., Maggioni, M., Trayanova, N.A.: Anatomically informed deep learning on contrast-enhanced cardiac magnetic resonance imaging for scar segmentation and clinical feature extraction. Cardiovascular Digital Health Journal 3(1), 2–13 (2022)
10. Lau, F., Hendriks, T., Lieman-Sifry, J., Sall, S., Golden, D.: Scargan: chained generative adversarial networks to simulate pathological tissue on cardiovascular mr scans. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. pp. 343–350. Springer (2018)
11. Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., et al.: Monai: An open-source framework for deep learning in healthcare. arXiv preprint arXiv:2211.02701 (2022)
12. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in neural information processing systems 33, 6840–6851 (2020)
13. Wyatt, J., Leach, A., Schmon, S.M., Willcocks, C.G.: Anoddpm: Anomaly detection with denoising diffusion probabilistic models using simplex noise. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 650–656 (2022)
14. Zhao, K., Hung, A.L.Y., Pang, K., Zheng, H., Sung, K.: Partdiff: Image super-resolution with partial diffusion models. arXiv preprint arXiv:2307.11926 (2023)
15. Fahmy, A.S., Neisius, U., Chan, R.H., Rowin, E.J., Manning, W.J., Maron, M.S., Nezafat, R.: Three-dimensional deep convolutional neural networks for automated myocardial scar quantification in hypertrophic cardiomyopathy: a multicenter multivendor study. Radiology 294(1), 52–60 (2020)
16. Saharia, C., Ho, J., Chan, W., Salimans, T., Fleet, D.J., Norouzi, M.: Image super-resolution via iterative refinement. IEEE Transactions on Pattern Analysis and Machine Intelligence 45(4), 4713–4726 (2022)
17. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1125–1134 (2017)
18. Ma, Y., Liu, J., Liu, Y., Fu, H., Hu, Y., Cheng, J., Qi, H., Wu, Y., Zhang, J., Zhao, Y.: Structure and illumination constrained gan for medical image enhancement. IEEE Transactions on Medical Imaging 40(12), 3955–3967 (2021)