Pre-trained Diffusion Models for Plug-and-Play Medical Image Enhancement

Jun Ma[1,2,3], Yuanzhi Zhu[4], Chenyu You[5], and Bo Wang[1,2,3,6,7(✉)]

[1] Peter Munk Cardiac Centre, University Health Network, Toronto, Canada
bowang@vectorinstitute.ai
[2] Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Canada
[3] Vector Institute for Artificial Intelligence, Toronto, Canada
[4] Department of Information Technology and Electrical Engineering, ETH Zürich, Zürich, Switzerland
[5] Department of Electrical Engineering, Yale University, New Haven, USA
[6] Department of Computer Science, University of Toronto, Toronto, Canada
[7] AI Hub, University Health Network, Toronto, Canada

Abstract. Deep learning-based medical image enhancement methods (e.g., denoising and super-resolution) mainly rely on paired data and correspondingly the well-trained models can only handle one type of task. In this paper, we address the limitation with a diffusion model-based framework that mitigates the requirement of paired data and can simultaneously handle multiple enhancement tasks by one pre-trained diffusion model without fine-tuning. Experiments on low-dose CT and heart MR datasets demonstrate that the proposed method is versatile and robust for image denoising and super-resolution. We believe our work constitutes a practical and versatile solution to scalable and generalizable image enhancement.

1 Introduction

Computed Tomography (CT) and Magnetic Resonance (MR) are two widely used imaging techniques in clinical practice. CT imaging uses X-rays to produce detailed, cross-sectional images of the body, which is particularly useful for imaging bones and detecting certain types of cancers with fast imaging speed. However, CT imaging has relatively high radiation doses that can pose a risk of radiation exposure to patients. Low-dose CT techniques have been developed to address this concern by using lower doses of radiation, but the image quality is degraded with increased noise, which may compromise diagnostic accuracy [9].

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43898-1_1.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14222, pp. 3–13, 2023.
https://doi.org/10.1007/978-3-031-43898-1_1

MR imaging, on the other hand, uses a strong magnetic field and radio waves to create detailed images of the body's internal structures, which can produce high-contrast images for soft tissues and does not involve ionizing radiation. This makes MR imaging safer for patients, particularly for those who require frequent or repeated scans. However, MR imaging typically has a lower resolution than CT [18], which limits its ability to visualize small structures or abnormalities.

Motivated by the aforementioned, there is a pressing need to improve the quality of low-dose CT images and low-resolution MR images to ensure that they provide the necessary diagnostic information. Numerous algorithms have been developed for CT and MR image enhancement, with deep learning-based methods emerging as a prominent trend [5,14], such as using the conditional generative adversarial network for CT image denoising [32] and convolutional neural network for MR image Super-Resolution (SR) [4].

These algorithms are capable of improving image quality, but they have two significant limitations. First, **paired images are required for training,** e.g., low-dose and full-dose CT images; low-resolution and high-resolution MR images). However, acquiring such paired data is challenging in real clinical scenarios. Although it is possible to simulate low-quality images from high-quality images, the models derived from such data may have limited generalization ability when applied to real data [9,14]. Second, **customized models are required for each task.** For example, for MR super-resolution tasks with different degradation levels (i.e., 4x and 8x downsampling), one may need to train a customized model for each degradation level and the trained model cannot generalize to other degradation levels. Addressing these limitations is crucial for widespread adoption in clinical practice.

Recently, pre-trained diffusion models [8,11,21] have shown great promise in the context of unsupervised natural image reconstruction [6,7,12,28]. However, their applicability to medical images has not been fully explored due to the absence of publicly available pre-trained diffusion models tailored for the medical imaging community. The training of diffusion models requires a significant amount of computational resources and training images. For example, openai's improved diffusion models [21] took 1600–16000 A100 hours to be trained on the ImageNet dataset with one million images, which is prohibitively expensive. Several studies have used diffusion models for low-dose CT denoising [30] and MR image reconstruction [22,31], but they still rely on paired images.

In this paper, we aim at addressing the limitations of existing image enhancement methods and the scarcity of pre-trained diffusion models for medical images. Specifically, we provide two well-trained diffusion models on full-dose CT images and high-resolution heart MR images, suitable for a range of applications including image generation, denoising, and super-resolution. Motivated by the existing plug-and-play image restoration methods [26,34,35] and denoising diffusion restoration and null-space models (DDNM) [12,28], we further introduce a paradigm for plug-and-play CT and MR image denoising and super-resolution as shown in Fig. 1. Notably, it **eliminates the need for paired data**, enabling greater scalability and wider applicability than existing paired-image dependent

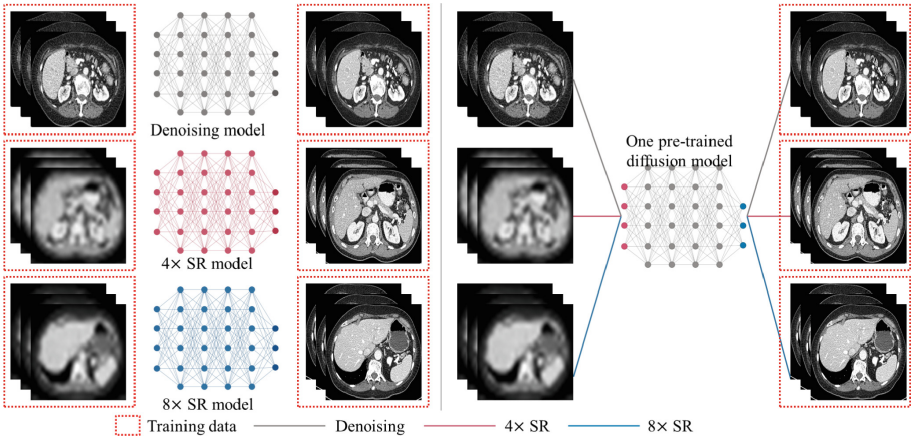

Fig. 1. Comparison of (a) the common paired-image dependent paradigm and (b) the plug-and-play paradigm for medical image enhancement. The former needs to build customized models for different tasks based on paired low/high-quality images, while the latter can share one pre-trained diffusion model for all tasks and only high-quality images are required as training data. The pre-trained model can handle unseen images as demonstrated in experiments.

methods. Moreover, it **eliminates the need to train a customized model for each task**. Our method does not need additional training on specific tasks and can directly use the single pre-trained diffusion model on multiple medical image enhancement tasks. The pre-trained diffusion models and PyTorch code of the present method are publicly available at https://github.com/bowang-lab/DPM-MedImgEnhance.

2 Method

This section begins with a brief overview of diffusion models for image generation and the mathematical model and algorithm for general image enhancement. We then introduce a plug-and-play framework that harnesses the strengths of both approaches to enable unsupervised medical image enhancement.

2.1 Denoising Diffusion Probabilistic Models (DDPM) for Unconditional Image Generation

Image generation models aim to capture the intrinsic data distribution from a set of training images and generate new images from the model itself. We use DDPM [11] for unconditional medical image generation, which contains a diffusion (or forward) process and a sampling (or reverse) process. The diffusion process gradually adds random Gaussian noise to the input image x_0, following a Markov Chain with transition kernel $q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t\mathbf{I})$,

where $t \in \{1, \cdots, T\}$ represents the current timestep, x_t and x_{t-1} are adjacent image status, and $\beta_t \in \{\beta_1, \cdots, \beta_T\}$ is a predefined noise schedule. Furthermore, we can directly obtain x_t based on x_0 at any timestep t by:

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)\mathbf{I}), \quad e.g.,\ x_t = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{(1-\bar{\alpha}_t)}\epsilon, \tag{1}$$

where $\alpha_t := 1-\beta_t$, $\bar{\alpha}_t := \Pi_{s=1}^{t}\alpha_s$, and $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. This property enables simple model training where the input is the noisy image x_t and the timestep t and the output is the predicted noise ϵ_θ (θ denotes model parameters). Intuitively, a denoising network is trained with the mean square loss $\mathbb{E}_{t,x}||\epsilon - \epsilon_\theta(x_t, t)||^2$. The sampling process aims to generate a clean image from Gaussian noise $x_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, and each reverse step is defined by:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(x_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta(x_t, t)\right) + \beta_t z; \quad z \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \tag{2}$$

2.2 Image Enhancement with Denoising Algorithm

In general, image enhancement tasks can be formulated by:

$$y = Hx + n, \tag{3}$$

where y is the degraded image, H is a degradation matrix, x is the unknown original image, and n is the independent random noise. This model can represent various image restoration tasks. For instance, in the image denoising task, H is the identity matrix, and in the image super-resolution task, H is the downsampling operator. The main objective is to recover x by solving the minimization problem:

$$x^* = \arg\min_x \|y - Hx\|^2 + R(x), \tag{4}$$

where the first data-fidelity term keeps the data consistency and the second data-regularization term $R(x)$ imposes prior knowledge constraints on the solution. This problem can be solved by the Iterative Denoising and Backward Projections (IDBP) algorithm [26], which optimizes the revised equivalent problem:

$$x^*, \hat{y}^* = \arg\min_{x,\hat{y}} \|\hat{y} - x\|_{H^TH} + R(x) \quad s.t.\ \hat{y} = H^\dagger y, \tag{5}$$

where $H^\dagger := H^T(HH^T)^{-1}$ is the pseudo inverse of the degradation matrix H and $\|f\|_{H^TH} := f^TH^THf$. Specifically, x^* and $\hat{y}^*$ can be alternatively estimated by solving $\min_x \|\hat{y} - x\|_2^2 + R(x)$ and $\min_{\hat{y}} \|\hat{y} - x\|_2^2$ $s.t.\ H\hat{y} = y$. Estimating x^* is essentially a denoising problem that can be solved by a denoising operator and $\hat{y}^*$ has a closed-form solution:

$$x^* = Denoiser(\hat{y}); \quad \hat{y}^* = H^\dagger y + (I - H^\dagger H)\hat{x}. \tag{6}$$

Intuitively, IDBP iteratively estimates the original image from the current degraded image and makes a projection by constraining it with prior knowledge. Although IDBP offers a flexible way to solve image enhancement problems, it still requires paired images to train the denoising operator [26].

2.3 Pre-Trained Diffusion Models for Plug-and-play Medical Image Enhancement

We introduce a plug-and-play framework by leveraging the benefits of the diffusion model and IDBP algorithm. Here we highlight two benefits: (1) it removes the need for paired images; and (2) it can simply apply the single pre-trained diffusion model across multiple medical image enhancement tasks.

First, we reformulate the DDPM sampling process [11] $x_{t-1} \sim p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \beta_t \mathbf{I})$ into:

$$x_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}} \left(x_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(x_t, t) \right), \tag{7}$$

and

$$x_{t-1} = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t} x_{0|t} + \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\alpha_t} x_t + \beta_t z. \tag{8}$$

Intuitively, each sampling iteration has two steps. The first step estimates the denoised image $x_{0|t}$ based on the current noisy image x_t and the trained denoising network $\epsilon_\theta(x_t, t)$. The second step generates a rectified image x_{t-1} by taking a weighted sum of $x_{0|t}$ and x_t and adding a Gaussian noise perturbation.

As mentioned in Eq. (3), our goal is to restore an unknown original image x_0 from a degraded image y. Thus, the degraded image y needs to be involved in the sampling process. Inspired by the iteration loop in IDBP, we project the estimated $x_{0|t}$ on the hyperplane $y = Hx$:

$$\hat{x}_{0|t} = H^\dagger y + (I - H^\dagger H) x_{0|t}. \tag{9}$$

It can be easily proved that $H\hat{x}_{0|t} = HH^\dagger y + Hx_{0|t} - HH^\dagger H x_{0|t} = y$. By replacing $x_{0|t}$ with $\hat{x}_{0|t}$ in Eq. (8), we have:

$$x_{t-1} = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t} \hat{x}_{0|t} + \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\alpha_t} x_t + \beta_t z. \tag{10}$$

Algorithm 1 shows the complete steps for image enhancement, which inherit the denoising operator from DDPM and the projection operator from IDBP. The former employs the strong denoising capability in the diffusion model and the latter can make sure that the generated results match the input image. Notably, the final algorithm is equivalent to DDNM [28], but it is derived from different perspectives.

Algorithm 1. Pre-trained DDPM for plug-and-play medical image enhancement

Require: Pre-trained DDPM ϵ_θ, low-quality image y, degradation operator H

1: Initialize $x_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.
2: **for** $t = T$ **to** 1 **do**
3: $\quad x_{0|t} = \frac{1}{\sqrt{\bar{\alpha}_t}}\left(x_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(x_t, t)\right)$ *// Denoise x_t with pre-trained DDPM*
4: $\quad \hat{x}_{0|t} = x_{0|t} - \Sigma_t H^\dagger (H x_{0|t} - y)$ *// Project $x_{0|t}$ on the hyperplane $y = Hx$*
5: $\quad x_{t-1} = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\hat{x}_{0|t} + \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}x_t + \beta_t z,\ z \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ *// Sampling*
6: **end for**
7: **return** Enhanced image x_0

3 Experiments

Dataset. We conducted experiments on two common image enhancement tasks: denoising and SR. To mimic the real-world setting, the diffusion models were trained on a diverse dataset, including images from different centers and scanners. The testing set (e.g., MR images) is from a new medical center that has not appeared in the training set. Experiments show that our model can generalize to these unseen images. Specifically, the denoising task is based on the AAPM Low Dose CT Grand Challenge abdominal dataset [19], which can be also used for SR [33]. The heart MR SR task is based on three datasets: ACDC [1], M&Ms1-2 [3], and CMRxMotion [27]. Notably, the presented framework eliminates the requirement of paired data. For the CT image enhancement task, we trained a diffusion model [21] based on the full-dose dataset that contains 5351 images, and the hold-out quarter-dose images were used for testing. For the MR enhancement task, we used the whole ACDC [1] and M&Ms1-2 [3] for training the diffusion model and the CMRxMotion [27] dataset for testing. The testing images were downsampled by operator H with factors of 4× and 8× to produce low-resolution images, and the original images served as the ground truth.

Evaluation Metrics. The image quality was quantitatively evaluated by the Peak Signal-to-Noise Ratio (PSNR), Structural SIMilarity index (SSIM) [29], and Visual Information Fidelity (VIF) [24], which are widely used measures in medical image enhancement tasks [9,17].

Implementation Details. We followed the standard configuration in [21] to train the diffusion model from scratch. Specifically, the diffusion step used a linear noise schedule $\beta \in [1e-4, 0.02]$ and the number of diffusion timesteps was $T = 1000$. The input image size was normalized to 256×256 and the 2D U-Net [23] was optimized by Adam [13] with a batch size of 16 and a learning rate of 10^{-4}, and an Exponential Moving Average (EMA) over model parameters with a rate of 0.9999. All the models were trained on A100 GPU and the total training time was 16 d. The implementation was based on DDNM [28]. For an efficient sampling, we used DDIM [25] with 100 diffusion steps. We followed the degradation operator settings in DDNM. Specifically, we used the identity matrix I as the degradation operator for the denoising task and scaled

the projection difference $H^{\dagger}(Hx_{0|t}-y)$ with coefficient Σ to balance the information from measurement y and denoising output $x_{0|t}$. The downsampling operator implemented with *torch.nn.AdaptiveAvgPool2d* for the super-resolution task. The pseudo-inverse operator $H^{\dagger}$ is I for the denoising task and upsampling operator for the SR task.

Comparison with Other Methods. The pseudo-inverse operator $H^{\dagger}$ was used as the baseline method, namely, $x^{*} = H^{\dagger}y$. We also compared the present method with one commonly used image enhancement method DIP [10] and two recent diffusion model-based methods: IVLR [6], which adopted low-frequency information from measurement **y** to guide the generation process towards a narrow data manifold, and DPS [7], which addressed the intractability of posterior sampling through Laplacian approximation. Notably, DPS used 1000 sampling steps while we only used 100 sampling steps.

Table 1. Performance (mean±standard deviation) on CT denoising task. The arrows indicate directions of better performance.

Methods	PSNR ↑	SSIM ↑	VIF ↑
Baseline	24.9 ± 2.4	0.778 ± 0.07	0.451 ± 0.07
DIP [10]	25.9 ± 2.4	0.783 ± 0.06	0.444 ± 0.07
IVLR [6]	25.8 ± 2.3	0.695 ± 0.11	0.432 ± 0.09
DPS [7]	26.5 ± 2.3	0.791 ± 0.08	0.475 ± 0.09
Ours	**28.3 ± 2.8**	**0.803 ± 0.11**	**0.510 ± 0.10**

4 Results and Discussion

Low-Dose CT Image Enhancement. The presented method outperformed all other methods on the denoising task in all metrics, as shown in Table 1, with average PSNR, SSIM, and VIF of 28.3, 0.803, and 0.510, respectively. Supplementary Fig. 1 (a) visually compares the denoising results, showing that the presented method effectively removes the noise and preserves the anatomical details, while other methods either fail to suppress the noise or result in loss of tissue information.

We also used the same pre-trained diffusion model for simultaneously denoising and SR by setting H as the downsampling operator. Our method still achieves better performance across all metrics as shown in Table 2. By visually comparing the enhancement results in Fig. 1 (b) and (c), our results can reconstruct more anatomical details even in the challenging noisy 8× SR task. In contrast, DIP tends to smooth tissues and ILVR and DPS fail to recover the tiny structures such as liver vessels in the zoom-in regions.

Table 2. Evaluation results of joint denoising and super-resolution for CT images.

Methods	Noisy 4× SR			Noisy 8× SR		
	PSNR ↑	SSIM ↑	VIF ↑	PSNR ↑	SSIM ↑	VIF ↑
Baseline	20.4±0.6	0.583±0.06	0.191±0.01	18.4±0.6	0.484±0.06	0.087±0.01
DIP [10]	21.8±0.7	0.642±0.06	0.243±0.02	19.6±0.7	0.560±0.07	0.146±0.02
IVLR [6]	25.1±2.5	0.715±0.10	0.417±0.09	24.6±2.8	0.702±0.12	0.395±0.11
DPS [7]	25.1±2.8	0.172±0.13	0.416±0.12	24.7±3.1	0.705±0.14	0.398±0.12
Ours	**26.2±2.6**	**0.743±0.10**	**0.446±0.10**	**25.9±3.1**	**0.731±0.13**	**0.431±0.12**

Table 3. Evaluation results of heart MR image super-resolution tasks.

Methods	4× SR			8× SR		
	PSNR ↑	SSIM ↑	VIF ↑	PSNR ↑	SSIM ↑	VIF ↑
Baseline	23.5±0.6	0.734±0.03	0.288±0.01	19.9±0.7	0.646±0.04	0.123±0.01
DIP [10]	**27.0±0.8**	0.784±0.02	0.392±0.02	**22.6±0.8**	**0.712±0.03**	0.224±0.02
IVLR [6]	25.4±0.7	0.731±0.03	0.355±0.02	21.2±0.8	0.622±0.04	0.180±0.02
DPS [7]	25.6±0.8	0.741±0.03	0.335±0.03	21.2±0.8	0.635±0.04	0.177±0.02
Ours	26.9±0.8	**0.805±0.025**	**0.416±0.03**	22.4±0.8	0.700±0.03	**0.226±0.02**

MR Image Enhancement. To demonstrate the generality of the presented method, we also applied it for the heart MR image 4× and 8× SR tasks, and the quantitative results are presented in Table 3. Our results still outperformed IVLR and DPS in all metrics. DIP obtains slightly better scores in PSNR for the 4× SR task and PSNR and SSIM for the 8× SR tasks, but visualized image quality is significantly worse than our results as shown in Supplementary Fig. 2, e.g., many anatomical structures are smoothed. This is because perceptual and distortion qualities are in opposition to each other as theoretically proven in [2]. DIP mainly prioritizes the distortion measures for the noise-free SR tasks while our results achieve a better trade-off between the perceptual and distortion quality.

5 Conclusion

In summary, we have provided two well-trained diffusion models for abdomen CT and heart MR, and introduced a plug-and-play framework for image enhancement. Our experiments have demonstrated that a single pre-trained diffusion model could address different degradation levels without customized models. However, there are still some limitations to be solved. The degradation operator and its pseudo-inverse should be explicitly given, which limits its application in tasks such as heart MR motion deblurring. Although the present method is in general applicable for 3D images, training the 3D diffusion model still remains prohibitively expensive. Moreover, the sampling process currently requires multiple network inferences, but it could be solved with recent advances in one-step

generation models [15] and faster algorithms [16]. Despite these limitations, the versatile and scalable nature of the presented paradigm has great potential to revolutionize medical image enhancement tasks. Future work could focus on developing more efficient algorithms for 3D diffusion models and expanding this paradigm to more clinical applications such as low-dose PET denoising.

Acknowledgement. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN-2020-06189 and DGECR-2020-00294), Canadian Institute for Advanced Research (CIFAR) AI Catalyst Grants, and CIFAR AI Chair programs. We thank the IDDPM [21], guided-diffusion [8], and DDNM [28] team, as their implementation served as an important basis for our work. We want to especially mention Jiwen Yu, who provided invaluable guidance and support. We also thank the organizers of AAPM Low Dose CT Grand Challenge [20], ACDC [1], M&Ms1-2 [3], and CMRxMothion [27] for making the datasets publicly available.

References

1. Bernard, O., et al.: Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE Trans. Med. Imaging **37**(11), 2514–2525 (2018)
2. Blau, Y., Michaeli, T.: The perception-distortion tradeoff. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018)
3. Campello, V.M., et al.: Multi-centre, multi-vendor and multi-disease cardiac segmentation: the m&ms challenge. IEEE Trans. Med. Imaging **40**(12), 3543–3554 (2021)
4. Chaudhari, A.S., et al.: Super-resolution musculoskeletal MRI using deep learning. Magn. Reson. Med. **80**(5), 2139–2154 (2018)
5. Chen, H., et al.: Low-dose CT denoising with convolutional neural network. In: International Symposium on Biomedical Imaging, pp. 143–146 (2017)
6. Choi, J., Kim, S., Jeong, Y., Gwon, Y., Yoon, S.: ILVR: conditioning method for denoising diffusion probabilistic models. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 14347–14356 (2021)
7. Chung, H., Kim, J., Mccann, M.T., Klasky, M.L., Ye, J.C.: Diffusion posterior sampling for general noisy inverse problems. In: International Conference on Learning Representations (2023)
8. Dhariwal, P., Nichol, A.: Diffusion models beat GANs on image synthesis. In: Advances in Neural Information Processing Systems, pp. 8780–8794 (2021)
9. Diwakar, M., Kumar, M.: A review on CT image noise and its denoising. Biomed. Sig. Process. Control **42**, 73–88 (2018)
10. Dmitry, U., Vedaldi, A., Victor, L.: Deep image prior. Int. J. Comput. Vis. **128**(7), 1867–1888 (2020)
11. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. In: Advances in Neural Information Processing Systems, vol. 33, pp. 6840–6851 (2020)
12. Kawar, B., Elad, M., Ermon, S., Song, J.: Denoising diffusion restoration models. In: Advances in Neural Information Processing Systems (2022)
13. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. In: International Conference on Learning Representations (2014)

14. Lin, D.J., Johnson, P.M., Knoll, F., Lui, Y.W.: Artificial intelligence for MR image reconstruction: an overview for clinicians. J. Magn. Reson. Imaging **53**(4), 1015–1028 (2021)
15. Liu, X., Gong, C., Liu, Q.: Flow straight and fast: learning to generate and transfer data with rectified flow. In: International Conference on Learning Representations (2023)
16. Lu, C., Zhou, Y., Bao, F., Chen, J., Li, C., Zhu, J.: DPM-solver: a fast ode solver for diffusion probabilistic model sampling in around 10 steps. In: Advances in Neural Information Processing Systems (2022)
17. Mason, A., et al.: Comparison of objective image quality metrics to expert radiologists' scoring of diagnostic quality of MR images. IEEE Trans. Med. Imaging **39**(4), 1064–1072 (2019)
18. Mazurowski, M.A., Buda, M., Saha, A., Bashir, M.R.: Deep learning in radiology: an overview of the concepts and a survey of the state of the art with focus on MRI. J. Magn. Reson. Imaging **49**(4), 939–954 (2019)
19. McCollough, C.H., et al.: Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 low dose CT grand challenge. Med. Phys. **44**(10), e339–e352 (2017)
20. Moen, T.R., et al.: Low-dose CT image and projection dataset. Med. Phys. **48**(2), 902–911 (2021)
21. Nichol, A.Q., Dhariwal, P.: Improved denoising diffusion probabilistic models. In: International Conference on Machine Learning, pp. 8162–8171 (2021)
22. Peng, C., Guo, P., Zhou, S.K., Patel, V.M., Chellappa, R.: Towards performant and reliable undersampled MR reconstruction via diffusion model sampling. In: Medical Image Computing and Computer Assisted Intervention, pp. 623–633 (2022)
23. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241 (2015)
24. Sheikh, H.R., Bovik, A.C.: Image information and visual quality. IEEE Trans. Image Process. **15**(2), 430–444 (2006)
25. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. In: International Conference on Learning Representations (2021)
26. Tirer, T., Giryes, R.: Image restoration by iterative denoising and backward projections. IEEE Trans. Image Process. **28**(3), 1220–1234 (2018)
27. Wang, S., et al.: The extreme cardiac MRI analysis challenge under respiratory motion (cmrxmotion). arXiv preprint arXiv:2210.06385 (2022)
28. Wang, Y., Yu, J., Zhang, J.: Zero-shot image restoration using denoising diffusion null-space model. In: International Conference on Learning Representations (2023)
29. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
30. Xia, W., Lyu, Q., Wang, G.: Low-dose CT using denoising diffusion probabilistic model for 20x times speedup. arXiv preprint arXiv:2209.15136 (2022)
31. Xie, Y., Li, Q.: Measurement-conditioned denoising diffusion probabilistic model for under-sampled medical image reconstruction. In: Medical Image Computing and Computer Assisted Intervention, pp. 655–664 (2022)
32. Yi, X., Babyn, P.: Sharpness-aware low-dose CT denoising using conditional generative adversarial network. J. Digit. Imaging **31**, 655–669 (2018)
33. You, C., et al.: Ct super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-circle). IEEE Trans. Med. Imaging **39**(1), 188–203 (2020)

34. Zhang, K., Li, Y., Zuo, W., Zhang, L., Van Gool, L., Timofte, R.: Plug-and-play image restoration with deep denoiser prior. IEEE Trans. Pattern Anal. Mach. Intell. **44**(10), 6360–6376 (2021)
35. Zhang, K., Zuo, W., Gu, S., Zhang, L.: Learning deep CNN denoiser prior for image restoration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3929–3938 (2017)