BIR
British Institute
of Radiology

OXFORD

Diffusion models for medical image reconstruction

George Webber [iD], MMathCompSci* and **Andrew J. Reader** [iD], PhD

School of Biomedical Engineering and Imaging Sciences, King's College London, London SE1 7EU, United Kingdom
*Corresponding author: George Webber, School of Biomedical Engineering and Imaging Sciences, King's College London, 5th Floor, Becket House,
1 Lambeth Palace Rd, London SE1 7EU, United Kingdom (george.webber@kcl.ac.uk)

Abstract

Better algorithms for medical image reconstruction can improve image quality and enable reductions in acquisition time and radiation dose. A prior understanding of the distribution of plausible images is key to realising these benefits. Recently, research into deep-learning image reconstruction has started to look into using unsupervised diffusion models, trained only on high-quality medical images (ie, without needing paired scanner measurement data), for modelling this prior understanding. Image reconstruction algorithms incorporating unsupervised diffusion models have already attained state-of-the-art accuracy for reconstruction tasks ranging from highly accelerated MRI to ultra-sparse-view CT and low-dose PET. Key advantages of diffusion model approach over previous deep learning approaches for reconstruction include state-of-the-art image distribution modelling, improved robustness to domain shift, and principled quantification of reconstruction uncertainty. If hallucination concerns can be alleviated, their key advantages and impressive performance could mean these algorithms are better suited to clinical use than previous deep-learning approaches. In this review, we provide an accessible introduction to image reconstruction and diffusion models, outline guidance for using diffusion-model-based reconstruction methodology, summarise modality-specific challenges, and identify key research themes. We conclude with a discussion of the opportunities and challenges of using diffusion models for medical image reconstruction.

Keywords: image reconstruction algorithms; diffusion models; score-based generative models; deep learning; MRI; CT; PET; ultrasound.

Introduction

Reconstructing medical images from scanner measurements is a necessary step for many imaging techniques, enabling medical diagnosis and research. Advances in image reconstruction research enable us to accelerate acquisition, improve image quality, or reduce necessary radiation doses.[1,2] Sophisticated reconstruction algorithms achieve this by integrating current measurement data, our understanding of imaging physics, and our prior knowledge of plausible reconstructed images.[3]

One recent method for capturing such prior knowledge is the diffusion model,[4,5] a state-of-the-art generative deep learning framework for modelling images. Crucially, diffusion models can utilise prior knowledge either with (supervised) or without (unsupervised) knowledge of a specific reconstruction task. By decoupling learning of the prior knowledge from the reconstruction task, diffusion models can overcome existing issues of costly training and poor robustness to varied scan parameters.[6] As a result, image reconstruction with diffusion models, particularly in the unsupervised setting, is pushing the frontiers of medical image reconstruction.

This review seeks to explain, explore, and review the use of diffusion models for medical image reconstruction. We focus on the inverse problem of image reconstruction, omitting approaches that just perform denoising as a post-processing step.

We begin with an accessible overview of image reconstruction and diffusion model theory, as well as how diffusion models are incorporated into image reconstruction. We offer some practical considerations for working with the algorithms discussed. We then discuss recent advances, including modality-specific research into MRI, CT, PET, and ultrasound imaging. We conclude by considering the challenges and opportunities faced for using these algorithms in a clinical setting.

Image reconstruction

Image reconstruction is an inverse problem, where we seek an image $\hat{x}$ to explain our noisy scan measurements m (eg, sinograms for CT or PET, or k-space data for MRI).[1,2] To solve this inverse problem, we first define A as the forward model that maps images x to their corresponding noise-free measurements $y = Ax$. We then find the image estimate with the greatest likelihood $L(x;m)$ of explaining m, assuming a model of the noise (ie, $L(x;m) = \mathbb{P}(m \sim \text{Noise}(y) \,|\, y = Ax) = p(m|x)$).

Incorporating prior knowledge

As measurements become noisier (eg, as scan time or radiotracer dose in PET is reduced) or less complete (eg, to accelerate MRI or CT), the resulting image reconstruction problem becomes highly *ill-posed*, meaning it has no stable, unique solution. We can compensate for this reduction in measurement information by incorporating information about the distribution of probable images x, so-called *prior* knowledge. The best-reconstructed image $\hat{x}$ then balances maximising both the likelihood $L(x;m)$, that is, the likelihood that $\hat{x}$ explains the measurements m, and the prior $p(x)$, that is, the probability that $\hat{x}$ is a valid medical image. This image $\hat{x}$ is called the *maximum a posteriori* (MAP) estimate.

Reconstruction paradigms

As Table 1 summarises, different ways to acquire and use such prior knowledge lead to image reconstruction algorithms with varying strengths and weaknesses.

Table 1. A comparison of methods for integrating prior information into image reconstruction.

Method	Training data required	Training difficulty	Reconstruction speed	Reconstruction quality	Resilience to domain shift
Hand-crafted priors	None	N/A (no training required)	Fast	Good	Excellent (no dependency on prior scanner data)
Unsupervised diffusion models	Images	Easy (just training image denoisers)	Slow (complex iterative process with 100+ iterations)	Excellent	Good (at most, weak dependency on scanner)
Supervised learning	Paired measurements and images	Hard (requires learning A or calculating gradients through A)	Fast (varies, but methods are not inherently slow)	Excellent (best possible with abundant training data)	Poor (strong dependency on training data and its associated scanner(s))

Resilience to domain shift refers to performance on data that is dissimilar to the training data, so-called out-of-distribution data.

Conventionally, using prior knowledge meant minimising hand-crafted functional. For example, we may penalise large variations between neighbouring voxel intensities, resulting in a smoother and less noisy reconstructed image. However, hand-crafted priors are limited by our ability to describe mathematically what complex medical images should look like.[2]

Viewing reconstruction as a supervised deep learning problem offers a data-driven way to incorporate more complex prior information, by learning a prior from measurements and corresponding high-quality images.[1] For example, we may train a neural network to learn a mapping from low-quality measurement data to high-quality images, resulting in a network that can perform reconstruction tasks.[7] However, the large volumes of paired training data required are difficult to acquire. Furthermore, supervised training results in reconstruction algorithms that are not robust to domain shift, performing poorly on data that is different to their training data (eg, data acquired with different scanning parameters).

An alternative paradigm is unsupervised reconstruction, which uses an image prior to learning without knowledge of the reconstruction task.[1] For example, we may incorporate an untrained deep image prior (DIP) or generative image model to regularise our reconstruction process.[8] As measurement data is not needed for training, neither is the forward model A; this helps to decouple the learnt prior from particular scanning parameters. This can lead to faster training, lower training data requirements, and greater robustness to domain shift than supervised reconstruction methods. However, the learnt prior is necessarily less specific to the reconstruction task than in supervised reconstruction, so it risks sacrificing accuracy for flexibility.

Why diffusion models?
Comparison of diffusion models to other generative models
Diffusion models belong to the machine learning paradigm of deep generative models. Generative models use training images to learn a prior probability distribution of images (eg, brain PET or cardiac MRI images), from which they then sample to generate new images.

In recent years, *diffusion models*[4,5] have become the state-of-the-art generative model for learning distributions of images. Compared to autoencoder and normalising flow approaches, diffusion models generate higher fidelity samples.[9] Compared to adversarial learning approaches such as generative adversarial networks (GANs), they offer superior mode coverage, that is, samples match the diversity of the relevant distribution of medical images more closely.[9,10] While diffusion models typically take longer to generate images than other popular approaches, this is usually an acceptable trade-off for higher-quality image reconstruction. See Table 2 for a more detailed comparison.

Diffusion models are therefore a natural choice of deep learning framework to integrate with inverse problem solving and more specifically medical image reconstruction.

Comparison of diffusion model reconstruction to state-of-the-art reconstruction
Diffusion models are inherently iterative and may be integrated with existing model-based iterative methodology to yield unsupervised reconstruction algorithms that leverage the image modelling abilities of diffusion models. As previously discussed in Reconstruction paradigms, such algorithms benefit from lower training data requirements and greater robustness to domain shift than supervised reconstruction methods. In particular, the same trained model may be used at differing acceleration factors or coil configurations for MRI, or dose levels for PET.

Furthermore, once a diffusion model has been trained and conditioned on measured data, it implicitly models the full posterior distribution of possible reconstructed images (whereas conventional methods just provide a point estimate such as the MAP estimate). In particular, by generating multiple plausible reconstructions, we may quantify reconstruction uncertainty at the voxel level.

In summary, the potential advantages of using unsupervised diffusion models for medical image reconstruction include:

- State-of-the-art image generation, with superior image quality and diversity to other generative modelling methodologies.
- Improved training and handling of domain shift, as a result of decoupling the image prior from the scanner parameters.
- The ability to sample from the posterior distribution of possible reconstructions, and thereby quantify reconstruction uncertainty.
- Natural integrations with existing model-based iterative reconstruction algorithms.

However, in use cases with many high-quality measurement datasets and unchanging scanner parameters, state-of-the-art supervised reconstruction retains an advantage over

Table 2. Comparison of 3 popular generative image modelling frameworks.

Method	Training stability	Image quality	Generation time	Mode coverage
Autoencoder (eg, VAEs, WAEs)	Good (regularised reconstruction objective is usually simple to balance)	Medium (typically lower than state-of-the-art)	Fast (varies with methods, but typically faster than diffusion models)	Good (variational inference techniques help approximate the whole distribution)
Adversarial learning (eg, GANs)	Poor (difficulty balancing the competing actors in adversarial training)	Good (state-of-the-art)	Fast (varies with methods, but typically faster than diffusion models)	Poor (adversarial training objectives are prone to mode collapse)
Diffusion model (eg, DDIM, DDPM)	Good (just requires training denoisers at multiple noise scales)	Good (state-of-the-art)	Slow (computationally heavy iterative process)	Good (inherent randomness helps to generate diverse samples)

Abbreviations: DDIM = denoising diffusion implicit model, DDPM = denoising diffusion probabilistic model, GAN = generative adversarial network, VAE = variational autoencoder, WAE = Wasserstein autoencoder.

unsupervised diffusion due to the additional information in measurement data.

Diffusion model theory

Diffusion models as a noising and denoising process

Diffusion models consist of 2 iterative processes, both shown in Figure 1: a random *forward* process and a generative *backward* process.

The forward process is simple: we take an image x_0 as input and corrupt it repeatedly by adding artificial random Gaussian noise for T steps (indexed by time $t = 0, \frac{1}{T}, \frac{2}{T}, \ldots, 1$). This process yields a pure noise image x_1 where all original information has been lost.

The backward process is the reverse of the forward process. We generate a pure noise image x_1 and iteratively denoise it for a fixed number of timesteps (eg, yielding a sequence of less-noisy images $x_{0.99}, x_{0.98}, x_{0.97}, \ldots, x_t, \ldots$). After T timesteps, the output x_0 is a high-quality image belonging to the same distribution of images as the training data.

The denoising step is learnt by training a neural network to remove different levels of artificial noise from images, a process shown in Figure 2. This is typically done with a single noise-level-dependent neural network $s_\theta(x_t \,|\, t)$, which is given the current timestep t and image x_t as inputs.

Once the denoising step has been learnt, a diffusion model can generate realistic medical images from noise. The iterative denoising can be implemented by directly predicting x_{t-1} from x_t using s_θ, in the case of DDPMs (denoising diffusion probabilistic models).[5] An important modification to DDPMs is DDIMs (denoising diffusion implicit models).[11] A DDIM uses the same neural network s_θ as a DDPM but removes noise from x_t by estimating the fully denoised image x_0 and adding back noise to get $x_{t-\varepsilon}$; this can lead to 10–50× faster image generation by skipping timesteps.

Diffusion models as score-based generative models

In the image reconstruction literature, diffusion models are more often formulated from the score-based generative model (SGM) perspective[12] (see Figure 3). The central concept in the SGM formulation is the score $\nabla_x \log p(x)$, a vector that points from image x towards highly probable images.

The score can hence guide a generative process, but generating images with the score alone is computationally infeasible for medical images.

To make this process tractable, we learn approximate versions of the score, by learning to calculate the score for noisy

images at decreasing noise levels $\sigma_1 > \ldots > \sigma_{\frac{2}{T}} > \sigma_{\frac{1}{T}} > 0$. Then, to generate samples, we start with pure noise and use decreasingly noisy score vector approximations to guide an iterative algorithm towards a highly probable image. Figure 3A visualises how score functions learnt on artificially noisy data guide the generation of high-quality images.

We may learn the approximate score vectors $\nabla_x \log p_t(x_t)$ with a noise-level-dependent neural network $s_\theta(x \,|\, t)$ (using the Denoising Score Matching algorithm), see Figure 4 for a visualisation.

From the above description, it is hopefully not surprising that our previous formulation for "denoising" is broadly equivalent to the SGM formulation.[13] We may therefore view training and sampling from diffusion models in terms of learnt denoising or guiding image generation with the score function. The SGM perspective, while more mathematically involved, has the advantage of showing how to model the probability distribution $p(x)$ of images explicitly.

Diffusion models as continuous random processes

The natural generalisation of the SGM framework is to consider the SGM as a discrete version of a continuous noising process (see Figure 3B). This leads to the stochastic differential equation (SDE) formulation of SGMs.[14]

In practice, this allows us to forgo specifying a fixed set of T noise levels in training, instead specifying a continuous noise schedule. When we train our model we can then sample random noise levels from the schedule. This is beneficial for image generation, as it allows us to vary the number of timesteps used (and hence the image quality/speed of generation) without retraining the model.

## Integrating diffusion models with image reconstruction

Unsupervised reconstruction

The prior distribution of images learnt by a diffusion model may be exploited to solve inverse problems in medical imaging.[6,15,16] Most simply, this is achieved by interleaving the diffusion model's generative denoising steps with additional steps to encourage consistency with measured data (see Figure 5).

More formally, at timestep t, instead of the noisy score $\nabla_x \log p_t(x_t)$, we use (an approximation of) the *conditional* noisy score $\nabla_x \log p_t(x_t|m)$ to guide our image generation process. We can decompose this term using Bayes' Law:
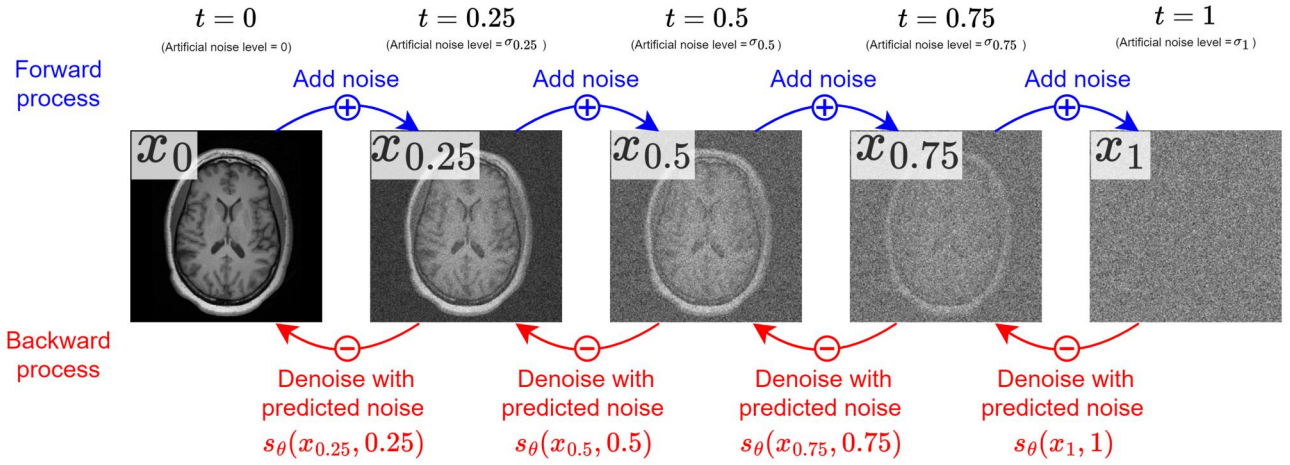
Figure 1. A diffusion model's iterative noising (additive) and denoising (subtractive) process, shown for $T=4$ timesteps with a single example image.
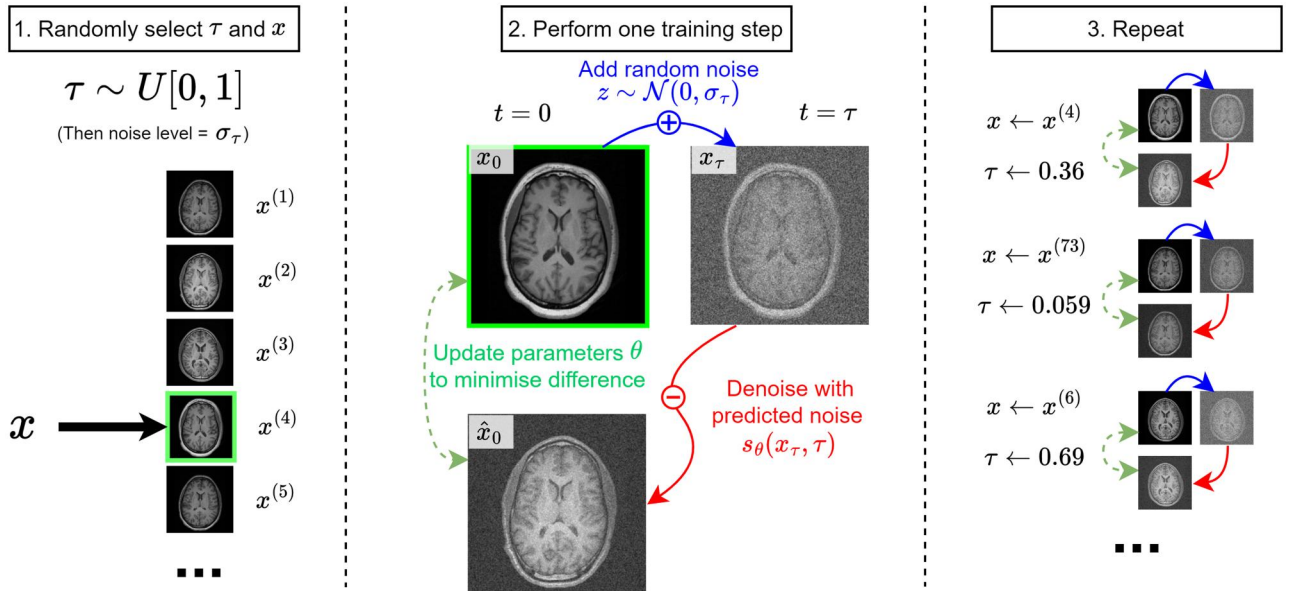
Figure 2. One possible training algorithm for a score-based generative model (SGM) with noise-level-dependent network s_θ. Using this training algorithm s_θ is trained to map artificially noisy versions of training images (with any randomly selected level of artificial noise) to their denoised counterparts. Alternative formulations learn instead to predict x_τ from $x_{\tau+\varepsilon}$, or select τ from a predefined discretisation of [0,1]. Side information can be incorporated by conditioning s_θ, for example, providing a guidance MRI image for a diffusion model trained on PET images.

$$\nabla_x \log p_t(x_t|m) = \nabla_x \log p_t(x_t) + \nabla_x \log p_t(m|x_t)$$

The first term $\nabla_x \log p_t(x_t)$ is just the unconditional noisy score so can be learnt as we've seen already. However, the second term $\nabla_x \log p_t(m|x_t)$, that is, the noisy score of the likelihood of the noisy image, is difficult to calculate. This is because for $t \neq 0$ there is a mismatch between our likelihood function L for noise-free (ie, $t=0$) data and the noisy iterate x_t we would like to apply it to.

The likelihood term $\nabla_x \log p_t(m|x_t)$ can be approximated by naïvely scaling the usual log-likelihood gradient $\nabla_x \log p(m|x)$, which is fast but inaccurate.[15,17] Diffusion posterior sampling (DPS)[18] instead proposes calculating the gradient of a combined "denoise-and-apply-likelihood" operation, which is slow but highly accurate. Other approaches such as decomposed diffusion sampling (DDS)[19] forgo direct approximation and instead apply the likelihood to the estimate of x_0 in the DDIM generation process.[20]

The choice of approximation method depends on the speed of reconstruction required, the size of the image output, the noise in the image, and the particular forward model in use.

Supervised reconstruction

While not the focus of this article, diffusion models may be used for familiar supervised learning, by inputting the measurement data to s_θ in the training process. Of course, such an approach no longer has the aforementioned benefits of alleviating the key issues of training data requirements and domain shift. Figure 5 illustrates the key paradigms for using diffusion models for reconstruction.

Practicalities

For s_θ, most researchers use a U-Net architecture with at least 3 down and upsampling stages with[6] or without[14,17,21] attention components, with a sinusoidal embedding for the timestep t.
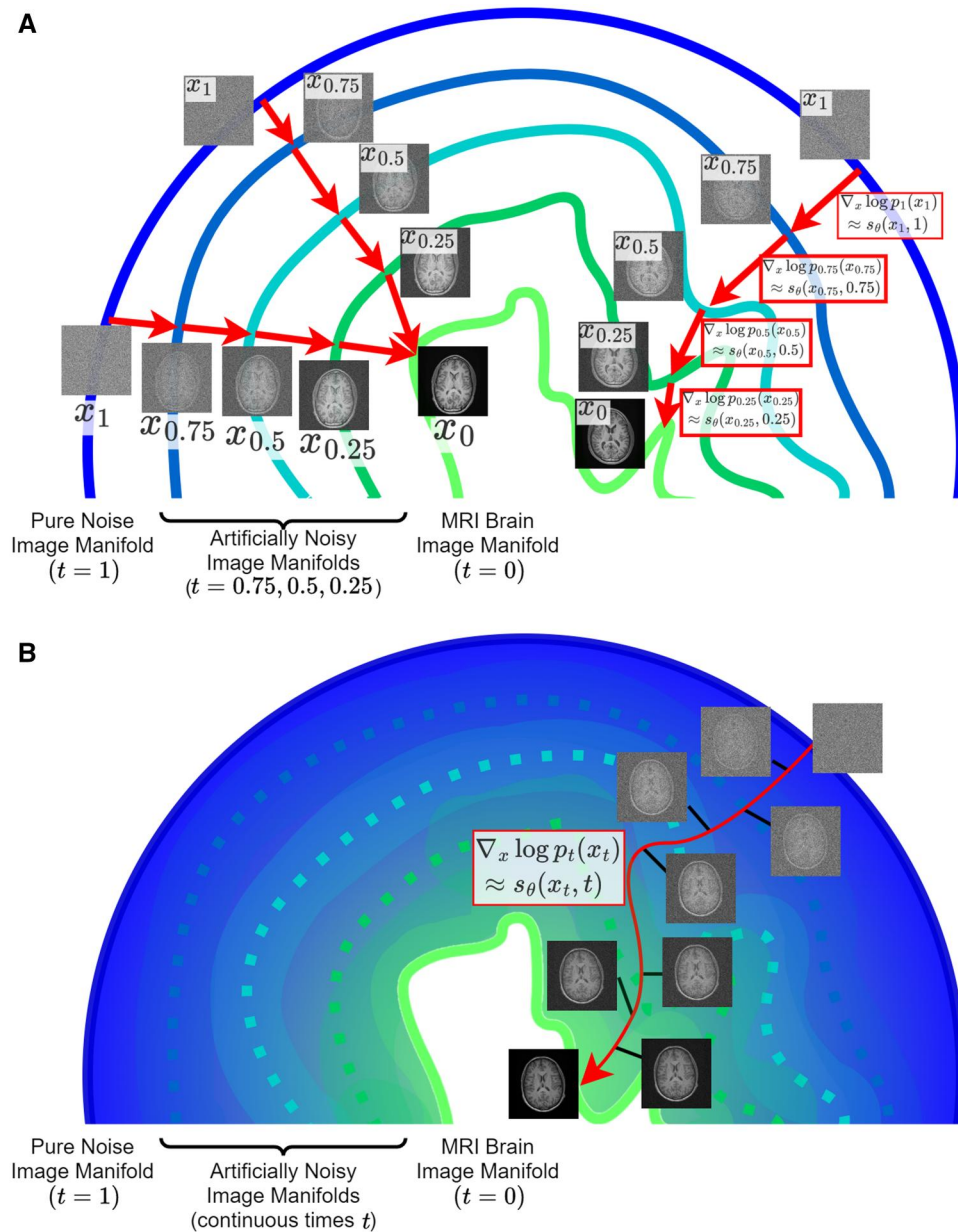
Figure 3. A 2D visualisation of score-based generative models (SGMs) from the perspective of artificially noisy image manifolds. (**A**) The coloured line segments each represent a manifold containing all possible artificially-noised medical images from our distribution of interest, at a given discrete noise level σ_t. The 9 brain images on the left visualise the diffusion model process of converting noise images into true medical images (note that this mapping may not be a one-to-one function). The path on the right shows the same process, with additional labels showing how the score vector is approximated by s_θ to guide image generation towards the medical image manifold without artificial noise. (**B**) Discrete noise levels are replaced by a continuous (monotone) noise schedule; in this setting, we see how the score function charts a continuous path from a pure noise distribution to the desired image manifold. We may discretise this process into arbitrarily small timesteps to generate high-quality images.

Some authors instead use variants of the vision transformer architecture.[22-24] Most papers report training for 100-50 000 epochs (with most < 5000).[6,24-29] Ideally, training should use as many images as possible and should continue until the loss on a validation set has been minimised (assessed with fixed timesteps and fixed random seed for stability). Adam with learning rate $\sim 10^{-4}$ is a common choice of optimiser.[30-32] Image generation (inference) is typically performed for 100-4000 iterations,[21,33] depending on image dimensions and noise schedule choice. Longer inference usually results in better samples, with diminishing returns expected beyond 1000 iterations.

Simple noise schedule choices include linear or cosine schedules.[34] The final variance should be set such that noisy images at this noise level approximate pure Gaussian noise (this should be checked visually and statistically).

Images should be cropped to remove empty space, and normalised before input for improved training. See the section on 3D reconstruction for handling of 3D data.

Most authors implement their work in Python, using either PyTorch (more flexible) or TensorFlow (shallower learning curve). Open-sourced code is an extremely valuable starting point; see Chung and Ye[6] (CT/MRI) or Singh et al[21] (PET),
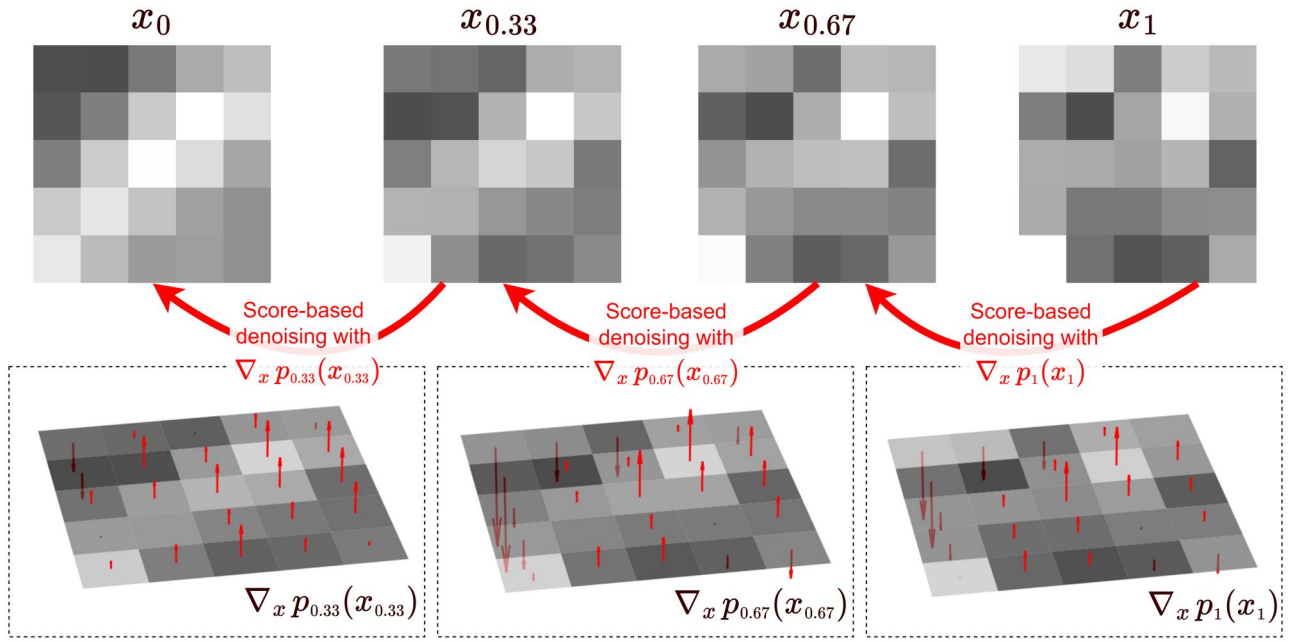
Figure 4. Diffusion model image generation from the perspective of the score vector $\nabla_x p_t(x_t)$. Here we show $T = 3$ equal timesteps of the backwards diffusion process for a small 5×5 patch, and for each timestep, we visualise the score vector as an array of arrows in red. It may be observed (particularly from the top left pixels) that the score function at each timestep points in the direction that reduces the noise in the image (where an "up" arrow represents an increase in a pixel value, ie, lighter, and a "down" arrow represents a decrease in a pixel value, ie, darker).
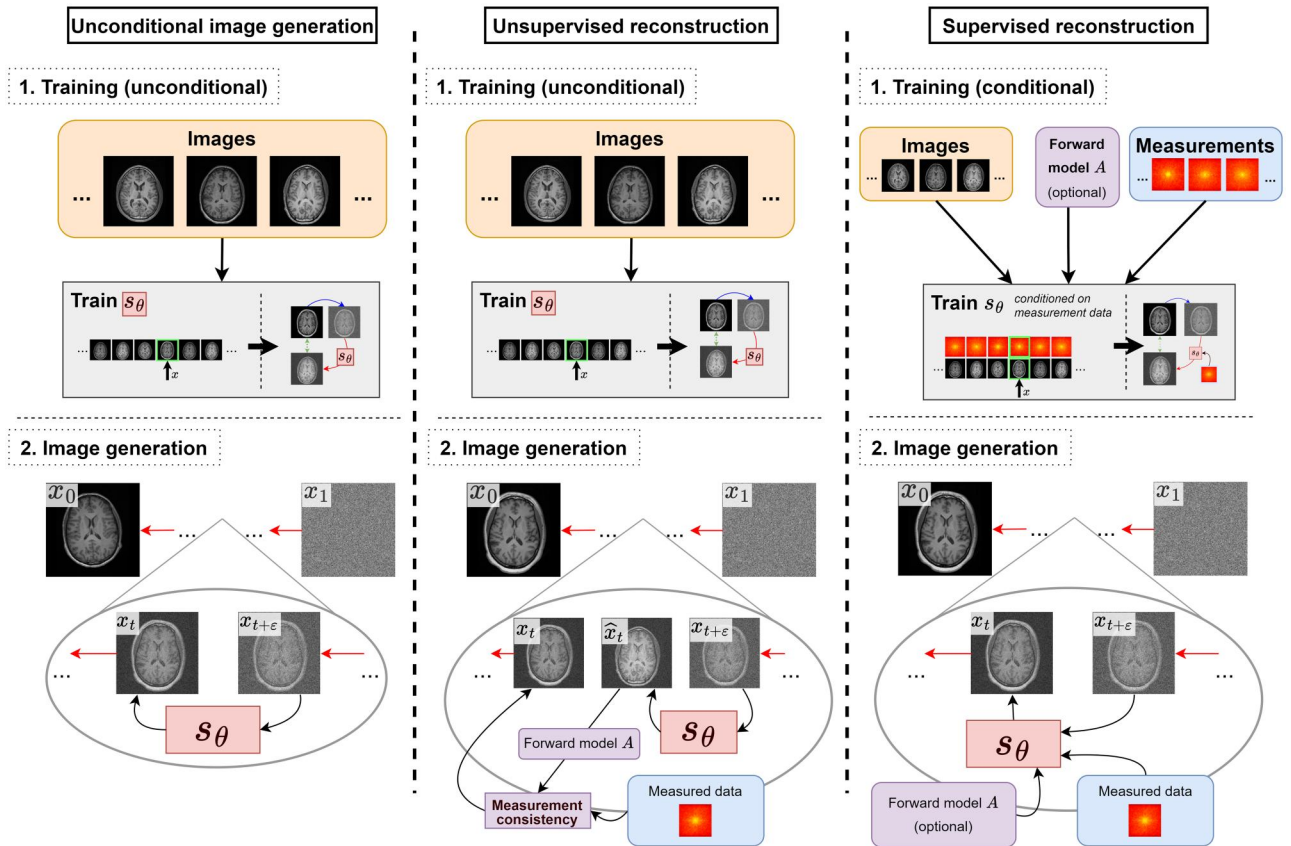
Figure 5. Varying paradigms for using diffusion models for image reconstruction; this review focuses on unsupervised reconstruction. Left: To generate images from the training distribution, the score network s_θ is first trained on example images (see Figure 2). Then, to generate a new image, pure noise x_1 is given as input, and iteratively denoised using s_θ until a new high-quality sample x_0 remains. Middle: To reconstruct in the absence of paired training data, unsupervised reconstruction can be performed. As in the leftmost panel, the score network s_θ is first trained on example images. Then, to perform reconstruction, the iterative denoising process is interleaved with steps to promote consistency with measured data. Right: If we have paired training data, we may condition our score network s_θ on the measurement data. Optionally, the forward model A can be incorporated into this process; otherwise, the scanner physics must be learnt by the model. Training in this case learns a conditional score network s_θ. To perform reconstruction, the learnt network is conditioned on the measured data and used directly to generate a reconstructed image x_0 from a pure noise image x_1. (This case is representative of conditional generation.)

Algorithm 1 Training a score-based generative model

Input: Untrained neural network s_θ with parameters θ
Input: Noising process $\nu(t)$
Input: Image data $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(M)}$
Input: Epoch number N
Output: Trained neural network s_θ

  Initialise optimiser opt
  for $i = 1, 2, \ldots, N$ **do**   ▷ Epoch for loop
    Shuffle images $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(M)}$
    for $j = 1, 2, \ldots, M$ **do**   ▷ Image for loop
      $t \sim U[0, 1]$ ▷ Get random timestep
      $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I})$ ▷ Sample noise
      $\mu_t, \sigma_t \leftarrow \nu(t)$ ▷ Get noising parameters
      $\mathbf{x}_t^{(j)} \leftarrow \mu_t \cdot \mathbf{x}^{(j)} + \sigma_t \cdot \mathbf{z}$ ▷ Define noisy image
      $\hat{\varepsilon} \leftarrow s_\theta(\mathbf{x}_t^{(j)}, t)$ ▷ Predict score (negative noise)
      $\mathcal{L} \leftarrow \|\hat{\varepsilon} + \mathbf{z}\|_2^2$ ▷ Calculate loss
      $\theta \leftarrow opt(s_\theta, \mathcal{L})$ ▷ Improve θ by minimising $\mathcal{L}$
    end for
  end for
  return s_θ

Algorithm 2 Image generation (or reconstruction) with a pre-trained score-based generative model

Input: Trained neural network s_θ
Input: Time discretisation $0 = t_{k_1} \leq \ldots \leq t_{k_T} = 1$
Input: Noising process $\nu(t)$
  Optional: measurement data $\mathbf{y}$, forward model A
Output: Generated sample $\mathbf{x}$

  $\mathbf{x}_1 \sim \mathcal{N}(0, \mathbf{I})$ ▷ Get initial noisy image
  for $k = T-1, T-2, \ldots, 1$ **do**
    $\hat{\varepsilon}_{t_{k+1}} \leftarrow s_\theta(\mathbf{x}_{t_{k+1}}, t_{k+1})$ ▷ Estimate score
    $\mu_t, \sigma_t \leftarrow \nu(t)$ ▷ Get noising parameters
    $\mathbf{x}_{t_k} \leftarrow$ REMOVE-NOISE$(\hat{\varepsilon}_{t_{k+1}}, \mu_t, \sigma_t)$
    if $\mathbf{y}, A$ provided **then**
      $\mathbf{x}_{t_k} \leftarrow$ CONDITION-ON-MEASUREMENTS$(\mathbf{x}_{t_k}, \mathbf{y}, A)$
    end if
  end for
  return $\mathbf{x}_0$

---

for example, implementations, and Algorithms 1 and 2 for pseudocode.

Modality-specific challenges

Different modalities in medical imaging come with their own challenges. For example, PET data is usually fully sampled with high-variance Poisson noise, while MRI data is usually undersampled with lower-variance Gaussian noise. As a result, the reconstruction problems associated with each modality are related but different.

In addition, modality-specific use cases inform different research focuses (eg, motion correction for cardiac MRI[35]) In this section, we survey some of the attempts to solve modality-specific challenges to image reconstruction with diffusion model methodology.

Magnetic resonance imaging

A major focus for MRI reconstruction is accelerating acquisition. Much successful work combines modern diffusion models with classical techniques for MRI acceleration such as compressed sensing,[15,36] parallel imaging,[28,37-41] and the use of non-Cartesian sampling trajectories.[42] The resulting algorithms generally report improved image quality at higher acceleration factors, and improved robustness to domain shift.

Many sub-problems for MRI reconstruction have also been addressed, for example, motion correction for high-resolution foetal[43] and adult[35,44] brains. Safari et al[35] note that these methods are particularly beneficial for dealing with diverse involuntary motions in elderly patients.

Other avenues explored include utilising multi-contrast MRI information,[45,46] dynamic MRI,[47-49] quantitative MRI,[27] and generating high-field images from low-field data.[29]

MRI case study

Cao et al[50] propose HFS-SDE (HFS = high-frequency space), a diffusion-based approach tailored to MRI that is focused on reconstructing high-frequency image details from undersampled data—see example results in Figure 6. On $\sim 6\times$ accelerated data, their algorithm can decrease the reconstruction error by $10\times$ relative to conventional parallel imaging reconstruction (from 14% to 1.4%) and performs competitively with supervised deep learning methods that require retraining for each different undersampling mask.

Computed tomography

Most diffusion-model-based reconstruction work on CT is for sparse-view[33,52-58] or limited-angle CT,[57,59-61] where the learnt prior is used to compensate for incomplete data.

Low-dose CT has also been addressed,[16,62-64] where the learnt prior compensates for a worse signal-to-noise ratio. A key distinction between efforts is whether the diffusion model acts on image space only (eg, Chung et al[19]), projection space only (eg, Guan et al[54]), or both synergistically (eg, Pan et al,[55] Xia et al[56] or Li et al[65]).

Some CT geometries are inherently 3D, so are especially exposed to challenges with 3D reconstruction[55,56,66] (see 3D reconstruction). Additionally, spectral CT has been successfully integrated with diffusion reconstructions,[67-69] while Zhou et al[70] have considered the non-standard geometry of robotic CT. In both of these cases, authors found improved image resolution relative to standard model-based iterative reconstruction algorithms.

CT case study

Vazia et al[68] propose spectral diffusion posterior sampling (SDPS), an adaptation of unsupervised diffusion reconstruction to tackle synergistic reconstruction for spectral CT. They show empirically on real data that their method is efficient, robust to different levels of noise, and outperforms state-of-the-art iterative regularised reconstruction. Figure 7 shows example reconstructions for the 80 keV energy bin. For the example shown, SDPS achieves a structural similarity index measure of 0.87, compared to 0.79 for the previous state-of-the-art and just 0.49 for the unregularised method.

Positron emission tomography

Variations in injected dose and differences between patients can lead to wide differences in dynamic range between PET images, potentially hindering the training of a diffusion model. Singh et al[21] address this with a novel normalisation scheme for images before training and during reconstruction.
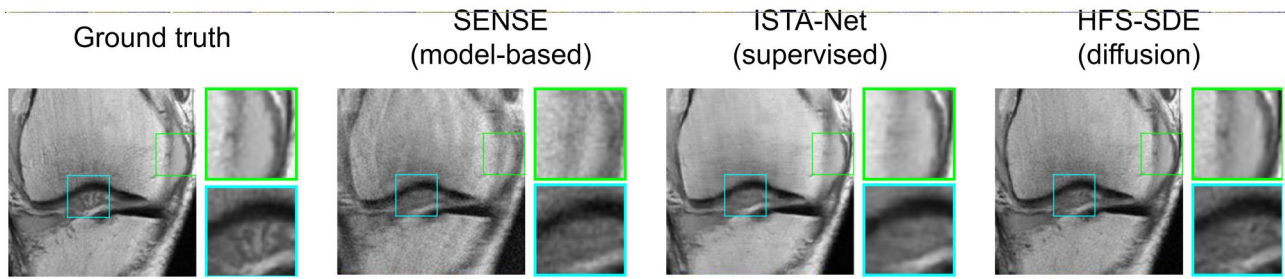
Figure 6. Example diffusion-based reconstruction of 12-fold undersampled knee MRI, using the HFS-SDE (HFS = high-frequency space) methodology proposed by Cao et al.[50] The conventional SENSE[51] parallel imaging method (column 2) yields a reconstruction with artefacts. In contrast, the deep learning methods (columns 3 and 4) yield higher-quality reconstructions, with HFS-SDE (column 4) achieving higher reconstruction accuracy than the supervised deep learning method (column 3). Figure courtesy of Yanjie Zhu.[50] SDE = stochastic differential equation.
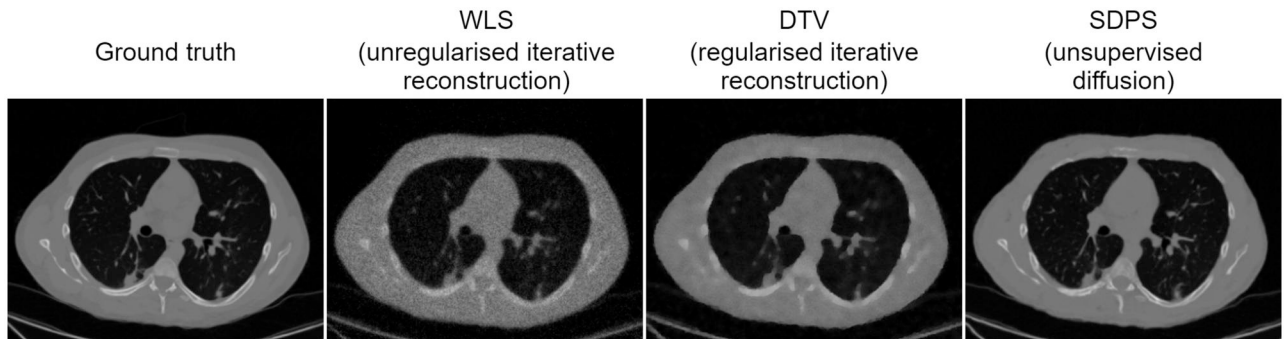
Figure 7. Reconstruction results of a multiple energy bin chest CT scan (80 keV energy bin shown) with fan beam geometry and 120 angles. Column 1: ground truth reference image. Column 2: unregularised iterative reconstruction with weighted least-squares (WLS). Column 3: state-of-the-art synergistic reconstruction with directional total variation prior (DTV). Column 4: unsupervised diffusion model adapted to a synergistic reconstruction setting, using Vazia et al's[68] spectral diffusion posterior sampling (SDPS) method. Figure courtesy of Corentin Vazia.[68]

Noise-level-aware diffusion models are an unexplored alternative seen in a denoising context, as done by Xie et al.[71]

Another PET-specific challenge is that the Poisson noise model introduces non-negativity constraints when calculating the likelihood $L(x; m)$. Existing work has clamped negative values to be zero in the likelihood calculation,[21] although this risks inadequately guiding the reconstruction at early iterations when noise is still dominant; replacing the usual Gaussian noising process with a non-negative noising process has been suggested.[21]

Other avenues of exploration include MRI-guided PET,[21] joint PET-MRI reconstruction,[72] and ultra-low-dose PET reconstruction.[73,74]

PET case study

Singh et al[21] propose memory-efficient methods for full 3D PET reconstruction. They show comparable image quality to existing state-of-the-art supervised methods while showing improved contrast recovery coefficient values for simulated lesions (eg, 98% vs 87% for supervised methods). Representative reconstructions versus other unsupervised methods are shown in Figure 8.

Disambiguation

A common misconception is that PET's Poisson noise model is incompatible with the artificial Gaussian noise used in the diffusion process. This is not an issue. In general, the diffusion process maps between a useful distribution of medical images (with some inherent noise) and a known distribution. This known distribution is chosen to be Gaussian for its nice

mathematical properties, though other options are possible. This Gaussian choice of known distribution is separate from the *inherent* Gaussian noise in some medical image modalities (eg, MRI, CT).

Ultrasound

Ultrasound imaging faces the challenge of poor signal-to-noise ratio, as well as highly structured and correlated noise.

Some unsupervised diffusion model approaches (eg, Lan et al,[75] Zhang et al[76] and Merino et al[77]) have reported significant improvements over the conventional delay-and-sum reconstruction technique and comparable results to the state-of-the-art for plane wave ultrasound imaging.

In the radiofrequency domain, ultrasound data can have a high dynamic range, making it difficult to learn priors. Stevens et al[78] explore solutions to this problem for the use case of dehazing cardiac ultrasound images.

Ultrasound case study

Zhang et al[79] show sampling multiple possible reconstructions with a diffusion model and computing the variance image to be helpful for despeckling images, as Figure 9 shows. On phantom datasets, this technique shows an improved contrast-to-noise ratio against the gold standard (18.4 dB vs 6.7 dB delay-and-sum with 75 plane wave transmissions) but does not improve axial resolution as much as supervised reconstruction techniques (eg, 0.29 mm full-width half-maximum resolution vs 0.22 mm for supervised vs 0.38 mm for delay-and-sum).
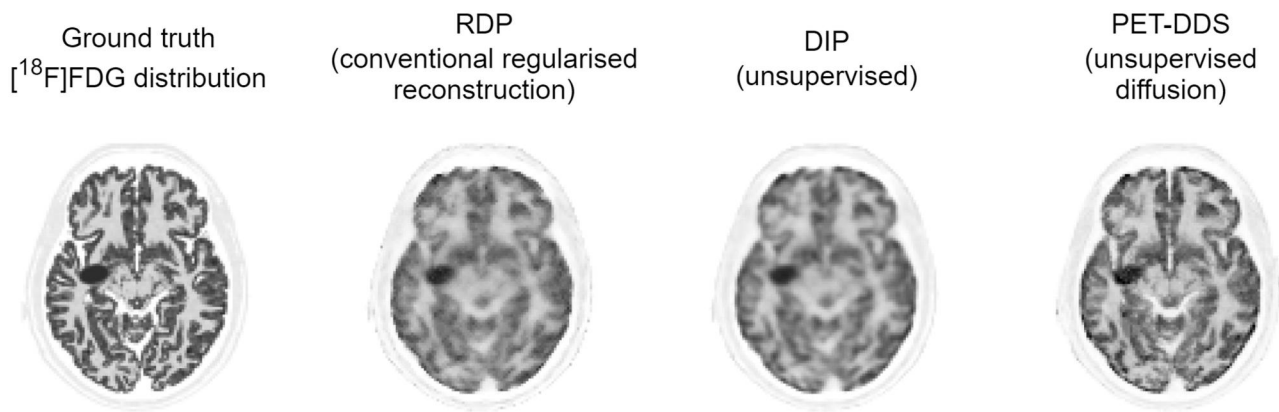
Figure 8. Example diffusion-model-based reconstruction of high-quality simulated [^{18}F]FDG-PET brain scans with an artificial lesion. Column 1: ground truth image. Column 2: reconstruction with a conventional maximum likelihood algorithm with relative difference prior (RDP) regularisation. Column 3: reconstruction with the unsupervised deep image prior (DIP) algorithm. Column 4: reconstruction with Singh et al's[21] unsupervised diffusion method PET-DDS. Figure courtesy of Imraj Singh.[21]
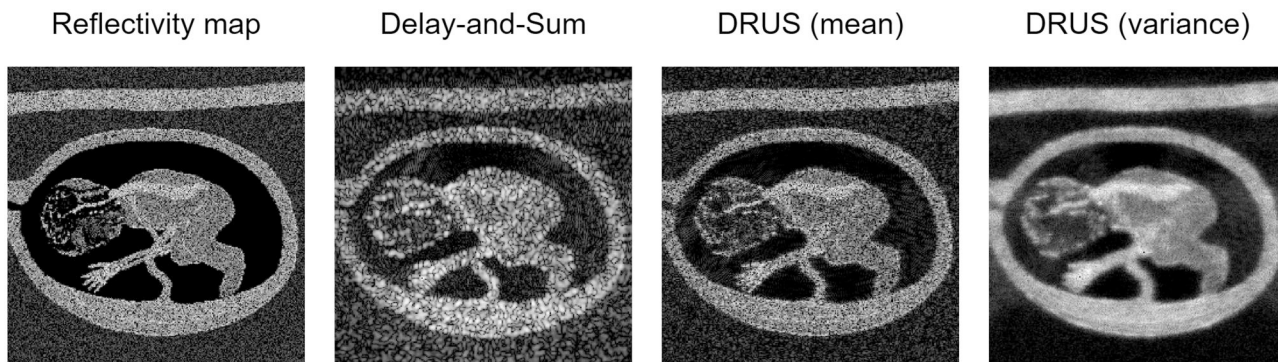
Figure 9. Example diffusion-model-based reconstruction of a simulated foetal ultrasound scan, using Zhang et al's[79] DRUS (diffusion reconstruction of ultrasound images) method. The diffusion-based reconstruction (column 3) shows improved image resolution and similarity to the true echogenicity map (column 1) compared to the standard delay-and-sum technique (with 1 plane wave transmission) (column 2). The variance of the reconstruction is also computed (column 4), highlighting increased uncertainty at tissue boundaries. Figure courtesy of Yuxin Zhang.[79]

Other modalities

To date, diffusion model reconstruction has also been explored in photoacoustic tomography,[80,81] electrical impedance tomography,[82] and electroencephalography,[83] where it has generally shown greater reductions in noise and improved generalisation to new measurement processes than conventional reconstruction techniques.

Key research directions

Image reconstruction with diffusion models is a highly active research field; last year alone saw an approximately 3-fold year-on-year increase in publications. In this section, we synthesise the field's key research directions, with a focus on unsupervised diffusion models.

Resolving mismatches between prior and measurements (avoiding hallucinations)

Often, training images have systematic differences from the image we seek to reconstruct,[84] for example, if training images are from a different scanner or hospital. In this case, steps to promote measurement consistency can hinder agreement with the diffusion prior and vice-versa (see Figure 10 for an extreme example). This can lead to hallucinations, where a high-quality reconstruction that is inconsistent with reality is produced.[85]

Some approaches attempt to resolve the mismatch via better measurement-conditioning schemes.[18,53,86] Chung et al[18] proposed resolving the mismatch by projecting the likelihood gradient onto the manifold of the diffusion prior, but this does not account for the noise in measurement data. DPS[86] is an improvement in this regard but is slow as it requires backpropagation through the score network s_θ.

Other approaches attempt to adapt the diffusion prior to the current reconstruction task via self-supervision. Güngör et al[87] fine-tune deep learning parameters controlling the prior to minimise measurement-consistency loss, outperforming other methods on out-of-distribution data. However, this approach is specific to their reconstruction architecture. Barbano et al[88] propose steerable conditional diffusion (SCD), a more general learning-free approach based on low-rank adaption (LoRA)[89] that drastically reduces hallucinatory features on out-of-distribution data. Chung and Ye[90] recently extended this work via a connection with the DIP, proposing a more efficient adaptation method for 3D inverse problem-solving in challenging imaging applications. This line of work is promising, not least due to the success of fine-tuning in the large language model field.[89]
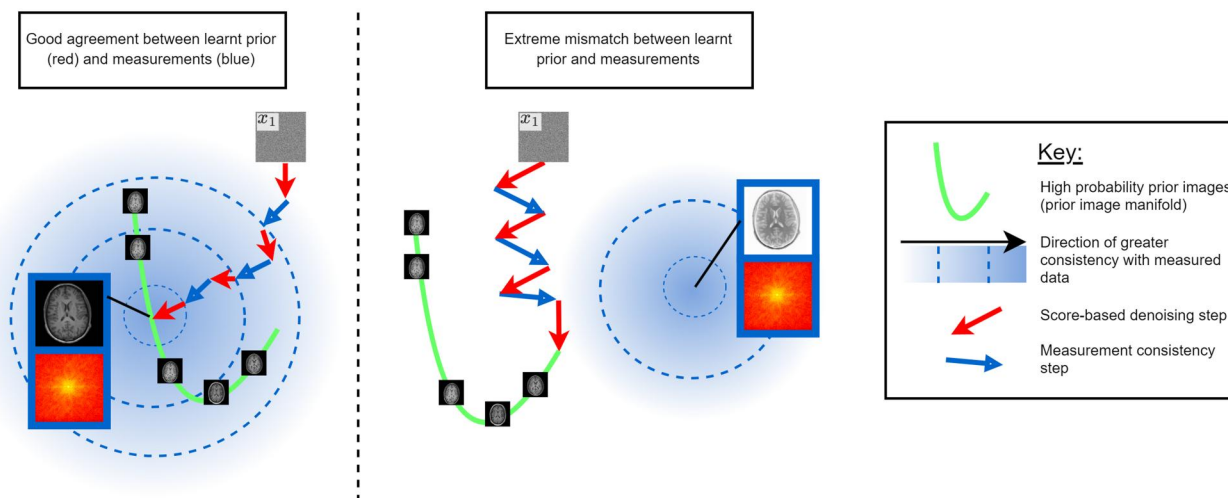
Figure 10. Visualisations of unsupervised diffusion model reconstruction, in cases of good (left) and poor (right) agreement between the prior and measurement likelihood. A diffusion model trained on irrelevant or poor-quality data can hinder rather than help reconstruction, or worse, lead to high-quality reconstructions inconsistent with reality (a so-called hallucination). Resolving small mismatches between the measurement consistency steps and the score-based denoising steps (due to noise or differences between training images and images to be reconstructed) is a subject of research interest.

Accelerating reconstruction times

Diffusion models have slow image generation, often taking thousands of iterations to generate high-quality images.[39] This is a key barrier to clinical use. Some efforts have integrated advances in diffusion model theory (eg, DDIMs) to yield faster image reconstruction algorithms.[21,26] Particularly of interest are Zhao et al[23] and Güngör et al's[87] use of adversarial training to enable generation with fewer timesteps, due to the success of GANs for supervised image reconstruction.

Other approaches are specific to image reconstruction, where we have the advantage of easily obtaining a cheap image estimate. Some authors have attempted "hot starts", that is, running the diffusion process starting from a conventional image estimate.[26] However, this estimate is unlikely to lie on an artificially noisy image manifold where the diffusion model was trained. Chung et al[91] suggest a neural network adaptation to overcome this, claiming stable state-of-the-art results on knee MRI with $10–50\times$ fewer iterations. A slower, but more widely explored, alternative is "coarse-to-fine sampling", where a coarse estimate is first generated from the diffusion model, followed by a second process of fine denoising.[32,87,92]

A separate approach has been to integrate reconstruction with acceleration schemes for existing iterative reconstruction algorithms (eg, Nesterov momentum, alternating direction method of multipliers algorithm).[39,63]

Learning from imperfect training data

It can be impossible to obtain high-quality training images ideal for training a diffusion model. For example, cardiac MRI acquisitions are undersampled to fit within a breath hold and CT acquisitions have limited intensity to reduce a patient's radiation dose. Naïvely training on imperfect images yields an imperfect prior. As a result, another research challenge is accurately learning to model clean images from noisy training images.

Liu et al[93] and Cui et al[30] consider Bayesian reconstructions as training data for score learning, to avoid using a point estimate of a noisy image. Considering the whole distribution of possible noisy reconstructions provides additional data from which to infer the clean score function. These self-supervised methods are shown to be competitive with supervised learning approaches, albeit orders of magnitude slower at inference time.

Other approaches by Aali et al[94,95] learn an approximate clean score function by introducing further known corruptions to be removed[94] (an approach best suited to a simple measurement noise model) or by leveraging Stein's unbiased risk estimate to jointly denoise data and learning its score function.[95] However, these approaches appear better suited to heavily corrupted image data than the low-level corruption in real medical images.

A different self-supervised approach is taken by Wu et al,[96] who propose the integration of an adaptive wavelet sub-network in training for denoising and in reconstruction for sparsity regularisation. The wavelet transform is used to preserve image information while reducing noise, but the integration of an additional sparsity prior may interfere with the score-based prior.

3D reconstruction

Using a single diffusion model to naïvely generate a full 3D medical volume is computationally infeasible (due to memory limitations). Most authors follow Chung et al[97] (see Figure 11) by training 1 diffusion model to generate varied 2D parallel slices from a 3D volume, and then reconstructing 2D slices separately with a hand-crafted prior to promoting similarity between neighbouring slices. This is a computationally efficient approach but can lead to slice inconsistency artefacts on slices orthogonal to the diffusion model slice direction.

Lee et al[98] proposed training multiple diffusion models on perpendicular slice directions, and iterating between the models for reconstruction. This approach does not compromise reconstruction speed while improving slice consistency, only requiring $2–3\times$ memory and training time.

Li et al[99] instead assume slice similarity across dimensions, proposing evaluating a single trained diffusion model on
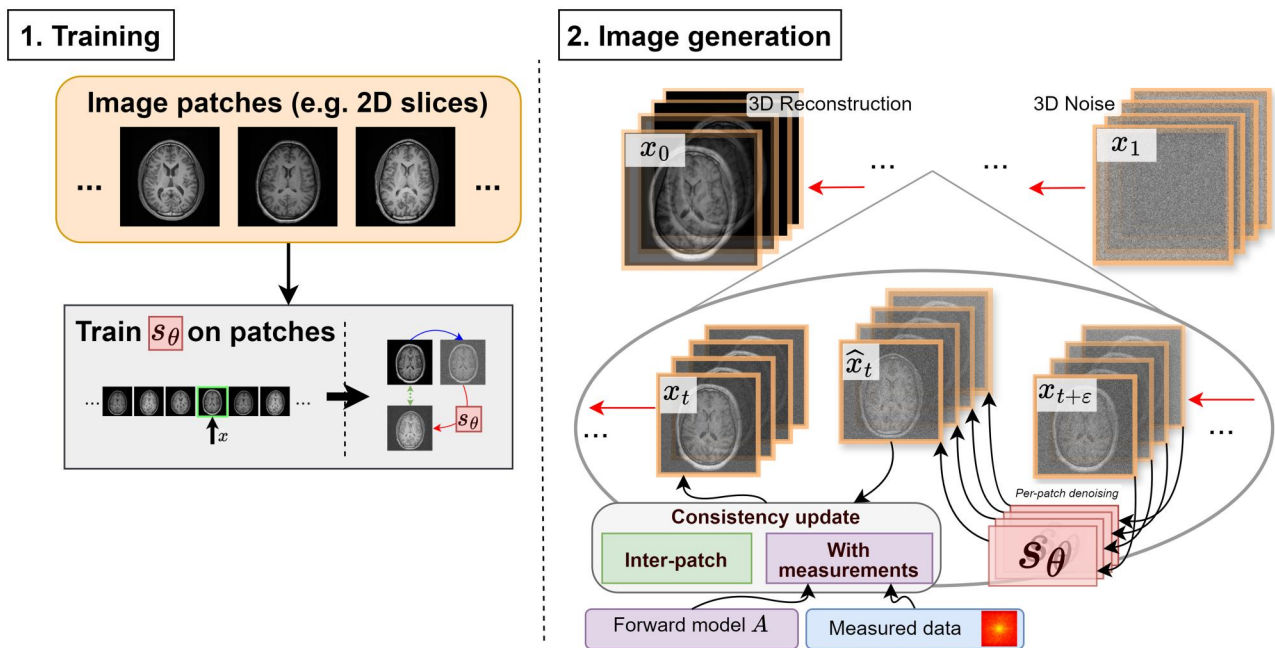
Figure 11. A simple approach to 3D reconstruction with a diffusion model.[97] A single diffusion model is trained to generate image patches (eg, whole 2D transverse slices). Then, a single diffusion step consists of patch-wise denoising, measurement consistency, and inter-patch consistency updates.

differently oriented slices to compute a 3D score. This is a strong assumption, and in many cases, it would not be expected to hold in practice. However, computing a 3D score directly rather than iterating between perpendicular scores may reduce the number of iterations required for image generation.

Song et al[20] instead propose performing the diffusion process in the latent space of a pre-trained variational autoencoder (VAE).[20,58] A small latent space could make this approach feasible for 3D reconstruction. However, due to known issues with VAEs such as image compression loss, Chen et al[53] suggest additional strategies are required to match the fine detail output by pixel-space diffusion models. Furthermore, careful validation of latent models will be required, as encouraging measurement consistency via highly non-linear encoder and decoder functions could yield unexpected effects.

Alternative diffusion model formulations

Latent diffusion models (see 3D reconstruction) are a promising alternative formulation to pixel-space diffusion models. Latent spaces other than pixel space have the potential for better modelling long-distance dependencies in an image, encouraging emphasis on image details, speeding up image generation, or alleviating memory demands.

As well as learnt latent spaces, pre-defined latent spaces have been proposed, based on the wavelet decomposition (eg, Xu et al[100]) or stacking an image onto itself (eg, Quan et al[25]) which may provide a safer and easier-to-work-with alternative to learnt spaces. In particular, several efforts have focused on the high spatial frequencies corresponding to fine details within an image.[22,50,101]

A similarly fine-detail-oriented alternative is learning a patch-based representation of the image (eg, Xia et al[66]) This approach has low memory requirements, allows parallelisation of the image reconstruction process, and can lead to very robust training. However, the focus on textural details within

patches may come at the expense of an image's longer-range spatial dependencies.

Another formulation that various authors[102-104] have instead proposed is generalised diffusion processes, which replace Gaussian noise with other degradation processes (eg, blurring, spatial frequency removal). Such approaches offer speed increases and may have better alignment with likelihood-based reconstruction processes, but risk losing the superior image quality of standard diffusion models.

Finally, instead of modelling the prior of an image, it is common to model the prior of full acquisition data from limited acquisition data (either solely[24,54,62,105] or additionally[49,55,56] to the image domain). This approach involves modelling in the measurement domain (k-space for MRI, sinograms for CT or PET), and hence risks losing some of the domain shift resilience of unsupervised diffusion. However, incorporating additional problem-specific information can result in highly successful, albeit less general, approaches.

Opportunities, challenges, and summary

Diffusion-model-based reconstruction holds the potential to improve image quality in different modalities across different hospital sites and varying scanner configurations. The ability to learn from high-quality images from different scanners (without requiring corresponding measurement data) makes these deep learning models easier to train on clinical data than previous supervised learning methods, raising the eventual possibility of "out-of-the-box" deep learning solutions for reconstruction. Many studies have already shown results for real data demonstrating improvements over the previous state-of-the-art.

Challenges remain for the clinical usage of the algorithms discussed in this article. Researchers are actively engaging with themes including how to avoid hallucinations, accelerate reconstruction times, learn from noisy data, and overcome memory limitations for the reconstruction of large volumes.

Few studies have yet been published assessing the clinical impact of images reconstructed by a diffusion model. For breast MRI, Okolie et al[106] showed at acceleration factor 2 diffusion models produce images almost indistinguishable from the original, as rated by radiologists (for 99% of cases). In the related context of denoising, Xie et al[107] show significant clinical potential for producing high-quality PET images from low-count data. Integrated reconstruction techniques would be expected to match or surpass post-processing denoising, although similar multi-institutional verification studies will be required to validate this.

In particular, in a clinical setting without ground truth data, tuning the relative strengths of the prior and the likelihood remains a difficult challenge. If the likelihood predominantly guides the reconstruction, in the worst case a noisy or artefact-dominated reconstruction is the output. If instead the diffusion prior predominantly guides the reconstruction, in the worst case a seemingly high-quality reconstruction may be obtained that has hallucinations (image structures or details that are not consistent with reality)—a much more concerning failure mode, as such hallucinations are difficult for clinicians to identify.

Self-supervised techniques to automate the selection of hyperparameters[108] and tune image generation[87,88] could mitigate such concerns.

Researchers can reduce their risk of producing hallucinatory images by robustly verifying their diffusion model has good generalisation properties, by comparing reconstructions to conventional methods, and by assessing images with appropriate quantitative metrics (rather than visually).

Uncertainty quantification may also help alleviate the issue of producing deceptively high-quality but inaccurate reconstructions. By performing multiple reconstructions initialised with different random noise images, we may quantify reconstruction uncertainty at the voxel level (eg, Luo et al[109]) This information has the potential to improve clinical analysis of reconstructed images, although dialogue with clinicians will be essential to establish useful ways to present such information. Uncertainty information could also prove useful to downstream image analysis tasks such as segmentation and classification.

In summary, diffusion models have set new state-of-the-art results for reconstruction tasks across many key medical imaging modalities. The improved robustness to domain shift of unsupervised diffusion models relative to supervised deep learning is a particular strength that will give these models a greater chance of clinical translation. Further work is needed, to speed up reconstruction, reconcile trained priors with measured data, and validate the robustness of reconstructions. Fortunately, the rate of progress in the field is fast, with many avenues remaining for further improvements.

Acknowledgements

Funding

Conflicts of interest

No competing interest is declared.

References

1. Ye JC, Eldar YC, Unser M, eds. *Deep Learning for Biomedical Image Reconstruction*. Cambridge University Press; 2023.
2. Reader AJ, Pan B. AI for PET image reconstruction. *Br J Radiol.* 2023;96(1150):20230292.
3. Lin DJ, Johnson PM, Knoll F, Lui YW. Artificial intelligence for MR image reconstruction: an overview for clinicians. *J Magn Reson Imaging.* 2021;53(4):1015-1028.
4. Sohl-Dickstein J, Weiss E, Maheswaranathan N, Ganguli S. Deep unsupervised learning using nonequilibrium thermodynamics. In: *Proceedings of the 32nd International Conference on Machine Learning.* Lille, France: PMLR; Vol. 37. 2015:2256-2265.
5. Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. In: Larochelle H, Ranzato M, Hadsell R, Balcan MF, Lin H, eds. *Advances in Neural Information Processing Systems.* Vol. 33. Curran Associates, Inc.; 2020:6840-6851.
6. Chung H, Ye JC. Score-based diffusion models for accelerated MRI. *Med Image Anal.* 2022;80:102479.
7. Srinram A, Zbontar J, Murrell T, et al. End-to-end variational networks for accelerated MRI reconstruction. In: Martel AL, Abolmaesumi P, Stoyanov D, et al., eds. *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020.* Springer International Publishing; 2020:64-73.
8. Hashimoto F, Ote K, Onishi Y. PET image reconstruction incorporating deep image prior and a forward projection model. *IEEE Trans Radiat Plasma Med Sci.* 2022;6(8):841-846.
9. Dhariwal P, Nichol A. Diffusion models beat GANs on image synthesis. In: Ranzato M, Beygelzimer A, Dauphin Y, Liang PS, Vaughan JW, eds. *Advances in Neural Information Processing Systems.* Vol. 34. Curran Associates, Inc.; 2021:8780-8794.
10. Croitoru FA, Hondru V, Ionescu RT, Shah M. Diffusion models in vision: a survey. *IEEE Trans Pattern Anal Mach Intell.* 2023; 45(9):10850-10869.
11. Song J, Meng C, Ermon S. Denoising diffusion implicit models. Poster presented at: *The Ninth International Conference on Learning Representations*, May 3-7, 2021; Virtual Event, Austria.
12. Song Y, Ermon S. Improved techniques for training score-based generative models. In: Larochelle H, Ranzato M, Hadsell R, Balcan MF, Lin H, eds. *Advances in Neural Information Processing Systems.* Vol 33. Curran Associates, Inc.; 2020:12438-12448.
13. Vincent P. A connection between score matching and denoising autoencoders. *Neural Comput.* 2011;23(7):1661-1674.
14. Song Y, Sohl-Dickstein J, Kingma DP, Kumar A, Ermon S, Poole B. Score-based generative modeling through stochastic differential equations. Poster presented at: *The Ninth International Conference on Learning Representations*; May 3-7, 2021; Virtual Event, Austria.
15. Jalal A, Arvinte M, Daras G, Price E, Dimakis AG, Tamir J. Robust compressed sensing MRI with deep generative priors. In: Ranzato M, Beygelzimer A, Dauphin Y, Liang PS, Vaughan JW, eds. *Advances in Neural Information Processing Systems.* Vol. 34. Curran Associates, Inc.; 2021:14938-14954.
16. He Z, Zhang Y, Guan Y, et al. Iterative reconstruction for low-dose CT using deep gradient priors of generative model. *IEEE Trans Radiat Plasma Med Sci.* 2022;6(7):741-754.
17. Ramzi Z, Remy B, Lanusse F, Starck JL, Ciuciu P. Denoising score-matching for uncertainty quantification in inverse

problems. Poster presented at: *NeurIPS 2020 Workshop on Deep Learning and Inverse Problems*; Dec 6-12, 2020; Virtual Event.

18. Chung H, Sim B, Ryu D, Ye JC. Improving diffusion models for inverse problems using manifold constraints. In: Koyejo S, Mohamed S, Agarwal A, Belgrave D, Cho K, Oh A, eds. *Advances in Neural Information Processing Systems*. Vol. 35. Curran Associates, Inc.; 2022:25683-25696.

19. Chung H, Lee S, Ye JC. Decomposed diffusion sampler for accelerating large-scale inverse problems. Poster presented at: *The Twelfth International Conference on Learning Representations*, May 7-11, 2024; Vienna, Austria.

20. Song B, Kwon SM, Zhang Z, Hu X, Qu Q, Shen L. Solving inverse problems with latent diffusion models via hard data consistency. In: *The Twelfth International Conference on Learning Representations*, May 7-11, 2024; Vienna, Austria.

21. Singh IRD, Denker A, Barbano R, et al. Score-based generative models for PET image reconstruction. *MELBA*. 2024;2(Generative Models):547-585.

22. Li Y, Sun X, Wang S, Qin Y, Pan J, Chen P. Generation model meets Swin transformer for unsupervised low-dose CT reconstruction. *Mach Learn Sci Technol*. 2024;5(2):025005.

23. Zhao X, Yang T, Li B, Yang A, Yan Y, Jiao C. DiffGAN: an adversarial diffusion model with local transformer for MRI reconstruction. *Magn Reson Imaging*. 2024;109:108-119.

24. Korkmaz Y, Cukur T, Patel VM. Self-supervised MRI reconstruction with unrolled diffusion models. In: Greenspan H, Madabhushi A, Mousavi P, et al., eds. *Medical Image Computing and Computer Assisted Intervention—MICCAI 2023*. Lecture Notes in Computer Science. Springer Nature Switzerland; 2023:491-501.

25. Quan C, Zhou J, Zhu Y, et al. Homotopic gradients of generative density priors for MR image reconstruction. *IEEE Trans Med Imaging*. 2021;40(12):3265-3278.

26. Jiang W, Xiong Z, Liu F, Ye N, Sun H. Fast controllable diffusion models for undersampled MRI reconstruction. In: *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, Athens, Greece. 2024:1–5.

27. Bian W, Jang A, Liu F. Diffusion modeling with domain-conditioned prior guidance for accelerated MRI and qMRI reconstruction. *IEEE Transactions on Medical Imaging*. 2024;PP. Epub ahead of print.

28. Erlacher M, Zach M. Joint non-linear MRI inversion with diffusion priors. arXiv, arXiv:2310.14842, 2023, preprint: not peer reviewed.

29. Cui ZX, Liu C, Cao C, et al. Meta-learning enabled score-based generative model for 1.5T-like image reconstruction from 0.5T MRI. arXiv, arXiv:2305.02509, 2023, preprint: not peer reviewed.

30. Cui ZX, Cao C, Liu S, et al. Self-score: self-supervised learning on score-based models for MRI reconstruction. arXiv, arXiv:2209.00835, 2022, preprint: not peer reviewed.

31. Gao Z, Zhou SK. U^{2}MRPD: unsupervised undersampled MRI reconstruction by prompting a large latent diffusion model. arXiv, arXiv:2402.10609, 2024, preprint: not peer reviewed.

32. Peng C, Guo P, Zhou SK, Patel VM, Chellappa R. Towards performant and reliable undersampled MR reconstruction via diffusion model sampling. In: Wang L, Dou Q, Fletcher PT, Speidel S, Li S, eds. *Medical Image Computing and Computer Assisted Intervention—MICCAI 2022*. Lecture Notes in Computer Science. Springer Nature Switzerland; 2022:623-633.

33. Song Y, Shen L, Xing L, Ermon S. Solving inverse problems in medical imaging with score-based generative models. In: *The Ninth International Conference on Learning Representations*, May 3-7, 2021; Virtual Event, Austria.

34. Nichol AQ, Dhariwal P. Improved denoising diffusion probabilistic models. In: *Proceedings of the 38th International Conference on Machine Learning*. PMLR; 2021:8162-8171.

35. Safari M, Yang X, Fatemi A, Archambault L. MRI motion artifact reduction using a conditional diffusion probabilistic model (MAR-CDPM). *Med Phys*. 2023;51(4):2598-2610.

36. Cao Y, Wang L, Zhang J, Xia H, Yang F, Zhu Y. Accelerating multi-echo MRI in k-space with complex-valued diffusion probabilistic model. In: *2022 16th IEEE International Conference on Signal Processing*. Vol. 1. Curran Associates, Inc.; 2022:479-484.

37. Tu Z, Liu D, Wang X, et al. WKGM: weighted k-space generative model for parallel imaging reconstruction. *NMR Biomed*. 2023; 36(11):e5005.

38. Zhang W, Xiao Z, Tao H, Zhang M, Xu X, Liu Q. Low-rank tensor assisted K-space generative model for parallel imaging reconstruction. *Magn Reson Imaging*. 2023;103:198-207.

39. Hou R, Li F, Zeng T. Fast and reliable score-based generative model for parallel MRI. *IEEE Trans Neural Netw Learn Syst*. 2024:1-14.

40. Chen L, Tian X, Wu J, et al. JSMoCo: joint coil sensitivity and motion correction in parallel MRI with a self-calibrating score-based diffusion model. arXiv, arXiv:2310.09625, 2023, preprint: not peer reviewed.

41. Cui ZX, Cao C, Cheng J, et al. SPIRiT-diffusion: self-consistency driven diffusion model for accelerated MRI. arXiv, arXiv:2304.05060, 2023, preprint: not peer reviewed.

42. Chan TJ, Rajapakse CS. Learning the domain specific inverse NUFFT for accelerated spiral MRI using diffusion models. arXiv, arXiv:2404.12361, 2024, preprint: not peer reviewed.

43. Tan J, Zhang X, Lv Y, Xu X, Li G. Self-supervised fetal MRI 3D reconstruction based on radiation diffusion generation model. arXiv, arXiv:2310.10209, 2023, preprint: not peer reviewed.

44. Levac B, Jalal A, Tamir JI. Accelerated motion correction for MRI using score-based generative models. In: *2023 IEEE 20th International Symposium on Biomedical Imaging*, Cartagena, Colombia. IEEE; 2023:1-5.

45. Levac B, Jalal A, Ramchandran K, Tamir JI. MRI reconstruction with side information using diffusion models. In: *57th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA. IEEE; 2023:1436-1442.

46. Li Y, Zhao L, Tian Y, Zhao S. T1 and T2 mapping reconstruction based on conditional DDPM. In: Camara O, Puyol-Antón E, Sermesant M, et al., eds. *Statistical Atlases and Computational Models of the Heart. Regular and CMRxRecon Challenge Papers*. Lecture Notes in Computer Science. Springer Nature Switzerland; 2024:303-313.

47. Qiu S, Pan S, Liu Y, et al. Spatiotemporal diffusion model with paired sampling for accelerated cardiac cine MRI. arXiv, arXiv:2403.08758, 2024, preprint: not peer reviewed.

48. Xiang T, Yue W, Lin Y, Yang J, Wang Z, Li X. DiffCMR: fast cardiac MRI reconstruction with diffusion probabilistic models. In: Camara O, Puyol-Antón E, Sermesant M, et al., eds. *Statistical Atlases and Computational Models of the Heart. Regular and CMRxRecon Challenge Papers*. Lecture Notes in Computer Science. Springer Nature Switzerland; 2024:380-389.

49. Yu C, Guan Y, Ke Z, Lei K, Liang D, Liu Q. Universal generative modeling in dual domains for dynamic MRI. *NMR Biomed*. 2023;36(12):e5011.

50. Cao C, Cui Z-X, Wang Y, et al. High-frequency space diffusion model for accelerated MRI. *IEEE Trans Med Imaging*. 2024;43 (5):1853-1865.

51. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. *Magn Reson Med*. 1999;42 (5):952-962.

52. Braure T, Ginsburger K. Conditioning generative latent optimization to solve imaging inverse problems. arXiv, arXiv:2307.16670, 2023, preprint: not peer reviewed.

53. Chen H, Hao Z, Guo L, Xiao L. Mitigating data consistency induced discrepancy in cascaded diffusion models for sparse-view CT reconstruction. arXiv, arXiv:2403.09355, 2024, preprint: not peer reviewed.

54. Guan B, Yang C, Zhang L, et al. Generative modeling in sinogram domain for sparse-view CT reconstruction. *IEEE Trans Radiat Plasma Med Sci*. 2024;8(2):195-207.

55. Pan S, Lo SY, Chang CW, et al. Patient-specific 3D volumetric CBCT image reconstruction with single x-ray projection using denoising diffusion probabilistic model. In: *Medical Imaging 2024: Imaging Informatics for Healthcare, Research, and Applications*. Vol. 12931. SPIE; 2024:136-143.

56. Xia W, Tseng HW, Niu C, et al. Parallel diffusion model-based sparse-view cone-beam breast CT. arXiv, arXiv:2303.12861, 2024, preprint: not peer reviewed.

57. Lopez-Montes A, McSkimming T, Skeats A, et al. Stationary CT imaging of intracranial hemorrhage with diffusion posterior sampling reconstruction. arXiv, arXiv:2407.11196, 2024, preprint: not peer reviewed.

58. He L, Yan H, Luo M, et al. Iterative reconstruction based on latent diffusion model for sparse data reconstruction. arXiv, arXiv:2307.12070, 2023, preprint: not peer reviewed.

59. Liu J, Anirudh R, Thiagarajan JJ, et al. DOLCE: a model-based probabilistic diffusion framework for limited-angle CT reconstruction. In: *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. Curran Associates, Inc.; 2023:10464-10474.

60. Xu K, Krishnan A, Li TZ, et al. Zero-shot CT field-of-view completion with unconditional generative diffusion prior. In: *Medical Imaging with Deep Learning*, hort Paper Track. July 10-12, 2023; Nashville, TN, USA.

61. Wang Y, Li Z, Wu W. Time-reversion fast-sampling score-based model for limited-angle CT reconstruction. *IEEE Trans Med Imaging*. 2024;PP. Epub ahead of print.

62. Huang B, Lu S, Zhang L, Lin B, Wu W, Liu Q. One sample diffusion modeling in projection domain for low-dose CT imaging. *IEEE Trans Radiat Plasma Med Sci*. 2024;PP. Epub ahead of print.

63. Xia W, Shi Y, Niu C, Cong W, Wang G. Diffusion prior regularized iterative reconstruction for low-dose CT. arXiv, arXiv:2310.06949, 2023, preprint: not peer reviewed.

64. Du W, Cui H, He L, Chen H, Zhang Y, Yang H. Structure-aware diffusion for low-dose CT imaging. *Phys Med Biol*. 2024;69 (15):155008.

65. Li Z, Chang D, Zhang Z, et al. Dual-domain collaborative diffusion sampling for multi-source stationary computed tomography reconstruction. *IEEE Trans Med Imaging*. 2024;PP. Epub ahead of print.

66. Xia W, Cong W, Wang G. Patch-based denoising diffusion probabilistic model for sparse-view CT reconstruction. arXiv, arXiv:2211.10388, 2022, preprint: not peer reviewed.

67. Vazia C, Bousse A, Froment J, et al. Spectral CT two-step and one-step material decomposition using diffusion posterior sampling. In: *2024 32nd European Signal Processing Conference (EUSIPCO)*, Lyon, France, 2024.

68. Vazia C, Bousse A, Vedel B, et al. Diffusion posterior sampling for synergistic reconstruction in spectral computed tomography. In: *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, Athens, Greece, 2024:1–5.

69. Liu Y, Zhou X, Wei C, Xu Q. Sparse-view spectral CT reconstruction and material decomposition based on multi-channel SGM. *IEEE Trans Med Imaging*. 2024:PP. Epub ahead of print.

70. Zhou X, Xu Q, Liu Y, Wei C, Wei L. SGM-based sparsity reconstruction under non-standard geometry of robotic CT. In: *Medical Imaging 2024: Physics of Medical Imaging*. SPIE. Vol. 12925. 2024:385-392.

71. Xie H, Gan W, Zhou B, et al. DDPET-3D: dose-aware diffusion model for 3D ultra low-dose PET imaging. *J Nucl Med*. 2024;65 (Suppl 2):241797.

72. Xie T, Cui Z-X, Luo C, et al. Joint diffusion: mutual consistency-driven diffusion model for PET-MRI co-reconstruction. *Phys Med Biol*. 2024;69(15).

73. Hu R, Wu D, Tivnan M, et al. Unsupervised low-dose PET image reconstruction based on pre-trained denoising diffusion probabilistic prior. *J Nucl Med*. 2024;65(Suppl 2):241109.

74. Webber G, Mizuno Y, Howes OD, Hammers A, King AP, Andrew J. Reader generative-model-based fully 3D PET image reconstruction by conditional diffusion sampling. In: *2024 IEEE Nuclear Science Symposium, Medical Imaging Conference and International Symposium on Room-Temperature Semiconductor Detectors (NSS MIC RTSD)*, Tampa, FL, USA. 2024. To appear.

75. Lan H, Li Z, He Q, Luo J. Fast sampling generative model for ultrasound image reconstruction. arXiv, arXiv:2312.09510, 2023, preprint: not peer reviewed.

76. Zhang Y, Huneau C, Idier J, Mateus D. Ultrasound imaging based on the variance of a diffusion restoration model. In: *2024 32nd European Signal Processing Conference (EUSIPCO)*, Lyon, France, 2024. To appear.

77. Merino S, Salazar I, Lavarello R. Generative models for ultrasound image reconstruction from single plane-wave simulated data. In: *2024 IEEE UFFC Latin America Ultrasonics Symposium (LAUS)*, Montevideo, Uruguay. IEEE; 2024:1-4.

78. Stevens TSW, Meral FC, Yu J, Apostolakis IZ, Robert JL, Van Sloun RJG. Dehazing ultrasound using diffusion models. *IEEE Trans Med Imaging*. 2024:PP. Epub ahead of print.

79. Zhang Y, Huneau C, Idier J, Mateus D. Diffusion reconstruction of ultrasound images with informative uncertainty. arXiv, arXiv:2310.20618, 2023, preprint: not peer reviewed.

80. Song X, Wang G, Zhong W, et al. Sparse-view reconstruction for photoacoustic tomography combining diffusion model with model-based iteration. *Photoacoustics*. 2023;33:100558.

81. Tong S, Lan H, Nie L, Luo J, Gao F. Score-based generative models for photoacoustic image reconstruction with rotation consistency constraints. arXiv, arXiv:2306.13843, 2023, preprint: not peer reviewed.

82. Wang H, Xu G, Zhou Q. A comparative study of variational autoencoders, normalizing flows, and score-based diffusion models for electrical impedance tomography. *J Inverse Ill-Pose P*. 2024;32 (4):795–813.

83. Zeng H, Xia N, Qian D, Hattori M, Wang C, Kong W. DM-RE2I: a framework based on diffusion model for the reconstruction from EEG to image. *Biomed Signal Process Control*. 2023; 86(A):105125.

84. Darestani MZ, Liu J, Heckel R. Test-time training can close the natural distribution shift performance gap in deep learning based compressed sensing. In: *Proceedings of the 39th International Conference on Machine Learning*. PMLR; 2022;162:4754-4776.

85. Bhadra S, Kelkar VA, Brooks FJ, Anastasio MA. On hallucinations in tomographic image reconstruction. *IEEE Trans Med Imaging*. 2021;40(11):3249-3260.

86. Chung H, Kim J, Mccann MT, Klasky ML, Ye JC. Diffusion posterior sampling for general noisy inverse problems. In: *The Eleventh International Conference on Learning Representations*, May 1-5, 2023; Kigali, Rwanda.

87. Güngör A, Dar SU, Öztürk Ş, et al. Adaptive diffusion priors for accelerated MRI reconstruction. *Med Image Anal*. 2023;88:102872.

88. Barbano R, Denker A, Chung H, et al. Steerable conditional diffusion for out-of-distribution adaptation in imaging inverse problems. arXiv, arXiv:2308.14409, 2023, preprint: not peer reviewed.

89. Hu EJ, Shen Y, Wallis P, et al. LoRA: low-rank adaptation of large language models. In: *The Tenth International Conference on Learning Representations*, April 25-29, 2022; Virtual Event.

90. Chung H, Ye JC. Deep diffusion image prior for efficient OOD adaptation in 3D inverse problems. arXiv, arXiv:2407.10641, 2024, preprint: not peer reviewed.

91. Chung H, Sim B, Ye JC. Come-closer-diffuse-faster: accelerating conditional diffusion models for inverse problems through stochastic contraction. In: *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, LA, USA. Curran Associates Inc.; 2022:12403-12412.

92. Zhang C, Chen Y, Fan Z, et al. TF-DiffRecon: texture coordination MRI reconstruction method based on diffusion model and modified MF-UNet method. In: *2024 IEEE International*

Symposium on Biomedical Imaging (ISBI), Athens, Greece, 2024:1–5.

93. Liu Y, Cui ZX, Liu C, et al. Score-based diffusion models with self-supervised learning for accelerated 3D multi-contrast cardiac magnetic resonance imaging. arXiv, arXiv:2310.04669, 2023, preprint: not peer reviewed.

94. Aali A, Daras G, Levac B, Kumar S, Dimakis AG, Tamir JI. Ambient diffusion posterior sampling: solving inverse problems with diffusion models trained on corrupted data. Poster presented at: *NeurIPS 2023 Workshop on Deep Learning and Inverse Problems,* December 16, 2024; New Orleans, USA.

95. Aali A, Arvinte M, Kumar S, Tamir JI. Solving inverse problems with score-based generative priors learned from noisy data. In: *57th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA. IEEE; 2023:837–843.

96. Wu W, Wang Y, Liu Q, Wang G, Zhang J. Wavelet-improved score-based generative model for medical imaging. *IEEE Trans Med Imaging*. 2024;43(3):966-979.

97. Chung H, Ryu D, Mccann MT, Klasky ML, Ye JC. Solving 3D inverse problems using pre-trained 2D diffusion models. In: *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, Canada. IEEE; 2023:22542-22551.

98. Lee S, Chung H, Park M, Park J, Ryu W, Ye J. Improving 3D imaging with pre-trained perpendicular 2D diffusion models. In: *2023 IEEE/CVF International Conference on Computer Vision*, Vancouver, Canada. IEEE; 2023:10676-10686.

99. Li Z, Wang Y, Zhang J, Wu W, Yu H. Two-and-a-half order score-based model for solving 3D ill-posed inverse problems. *Comput Biol Med*. 2024;168:107819.

100. Xu K, Lu S, Huang B, Wu W, Liu Q. Stage-by-stage wavelet optimization refinement diffusion model for sparse-view CT reconstruction. *IEEE Trans Med Imaging*. 2024;PP. Epub ahead of print.

101. Guan Y, Yu C, Cui Z, Zhou H, Liu Q. Correlated and multifrequency diffusion modeling for highly under-sampled MRI reconstruction. *IEEE Trans Med Imaging*. 2024;PP. Epub ahead of print.

102. Mirza MU, Dalmaz O, Bedel HA, et al. Learning Fourier-constrained diffusion bridges for MRI reconstruction. arXiv, arXiv:2308.01096, 2023, preprint: not peer reviewed.

103. Huang J, Aviles-Rivero AI, Schönlieb CB, Yang G. CDiffMR: can we replace the Gaussian noise with K-space undersampling for fast MRI? In: Greenspan H, Madabhushi A, Mousavi P, et al., eds. *Medical Image Computing and Computer Assisted Intervention—MICCAI 2023*. Lecture Notes in Computer Science. Springer Nature Switzerland; 2023:3-12.

104. Shen G, Li M, Farris CW, Anderson S, Zhang X. K-space cold diffusion: learning to reconstruct accelerated MRI without noise. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023: 26th International Conference, Vancouver, BC, Canada, October 8-12, 2023, Proceedings, Part X*. Springer-Verlag; 2023:3-12.

105. Xie Y, Li Q. Measurement-conditioned denoising diffusion probabilistic model for under-sampled medical image reconstruction. In: Wang L, Dou Q, Fletcher PT, Speidel S, Li S, eds. *Medical Image Computing and Computer Assisted Intervention—MICCAI 2022*. Lecture Notes in Computer Science. Springer Nature Switzerland; 2022:655-664.

106. Okolie A, Dirrichs T, Huck LC, et al. Accelerating breast MRI acquisition with generative AI models. *Eur Radiol*. 2024. Epub ahead of print.

107. Xie H, Gan W, Zhou B, et al. Dose-aware diffusion model for 3D low-dose PET: multi-institutional validation with reader study and real low-dose data. *J Nucl Med*. 2024;65(Suppl 2):241797-241797.

108. Ozturkler B, Liu C, Eckart B, Mardani M, Song J, Kautz J. SMRD: SURE-based robust MRI reconstruction with diffusion models. In: Greenspan H, Madabhushi A, Mousavi P, et al., eds. *Medical Image Computing and Computer Assisted Intervention—MICCAI 2023*. Lecture Notes in Computer Science. Springer Nature Switzerland; 2023:199-209.

109. Luo G, Blumenthal M, Heide M, Uecker M. Bayesian MRI reconstruction with joint uncertainty estimation using diffusion models. *Magn Reson Med*. 2023;90(1):295-311.