

PAPER • OPEN ACCESS


X-ray source motion blur modeling and deblurring with generative diffusion for digital breast tomosynthesis

To cite this article: Mingjie Gao *et al* 2024 *Phys. Med. Biol.* **69** 115003


View the [article online](#) for updates and enhancements.

You may also like

- [Characterisation of noise and sharpness of images from four digital breast tomosynthesis systems for simulation of images for virtual clinical trials](#)
Alistair Mackenzie, Nicholas W Marshall, Andria Hadjipanteli *et al.*
- [A diffusion-based truncated projection artifact reduction method for iterative digital breast tomosynthesis reconstruction](#)
Yao Lu, Heang-Ping Chan, Jun Wei *et al.*
- [Model observer for assessing digital breast tomosynthesis for multi-lesion detection in the presence of anatomical noise](#)
Gezheng Wen, Mia K Markey, Tamara Miner Haygood *et al.*



Joining forces:
One complete
QA solution for
Dosimetry with
myQA[®], QUASAR[™]
and Radcal[®]!



The diagram is a circular graphic with a dark background and a pattern of small, colorful dots. It features four segments around a central core. The segments are labeled: 'Machine QA' (top-left, dark blue), 'Patient Specific QA' (top-right, green), 'Medical Imaging QA' (bottom, light blue), and 'Risk Management' (center, pink). The central core contains a stylized, wireframe human figure.



PAPER

OPEN ACCESS

RECEIVED
30 December 2023REVISED
27 March 2024ACCEPTED FOR PUBLICATION
19 April 2024PUBLISHED
14 May 2024

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



X-ray source motion blur modeling and deblurring with generative diffusion for digital breast tomosynthesis

Mingjie Gao^{1,2} , Jeffrey A Fessler^{1,2} and Heang-Ping Chan¹¹ Department of Radiology, University of Michigan, Ann Arbor, MI 48109, United States of America² Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, United States of AmericaE-mail: gmingjie@umich.edu**Keywords:** digital breast tomosynthesis, focal spot blur, image deblurring, image resolution, system modeling, x-ray source motion, diffusion models

Abstract

Objective. Digital breast tomosynthesis (DBT) has significantly improved the diagnosis of breast cancer due to its high sensitivity and specificity in detecting breast lesions compared to two-dimensional mammography. However, one of the primary challenges in DBT is the image blur resulting from x-ray source motion, particularly in DBT systems with a source in continuous-motion mode. This motion-induced blur can degrade the spatial resolution of DBT images, potentially affecting the visibility of subtle lesions such as microcalcifications. **Approach.** We addressed this issue by deriving an analytical in-plane source blur kernel for DBT images based on imaging geometry and proposing a post-processing image deblurring method with a generative diffusion model as an image prior. **Main results.** We showed that the source blur could be approximated by a shift-invariant kernel over the DBT slice at a given height above the detector, and we validated the accuracy of our blur kernel modeling through simulation. We also demonstrated the ability of the diffusion model to generate realistic DBT images. The proposed deblurring method successfully enhanced spatial resolution when applied to DBT images reconstructed with detector blur and correlated noise modeling. **Significance.** Our study demonstrated the advantages of modeling the imaging system components such as source motion blur for improving DBT image quality.

1. Introduction

Digital breast tomosynthesis (DBT) has significantly improved the diagnosis of breast cancer due to its high sensitivity and specificity in detecting microcalcifications (MCs), masses and architectural distortions compared to two-dimensional (2D) mammography (Chan *et al* 2014, Chong *et al* 2019, Conant *et al* 2023). In DBT, the x-ray tube moves over a limited range of angles while acquiring a small number of low-dose projection views (PVs). The PVs are subsequently reconstructed into a quasi-three-dimensional (3D) tomographic image volume with an anisotropic voxel size such that the resolution is superior in the slices parallel to the detector but is very limited in the depth direction. However, one of the primary challenges in DBT imaging is the x-ray source motion blur that can degrade the quality of DBT images, reducing their sharpness and potentially affecting the visibility of subtle lesions such as MCs (Shaheen *et al* 2011, Marshall and Bosmans 2012, Zheng *et al* 2019, Lee and Baek 2022).

The x-ray tube motion in DBT can be carried out in two modes: step-and-shoot mode and continuous-motion mode (Sechopoulos 2013, Gao *et al* 2021b). In the step-and-shoot mode, the x-ray source essentially stops at each angle, acquires a PV, then moves to the next angle. The focal spot size of this mode equals the nominal size of a stationary source. Compared with the ideal point source, the nominal source has a negligible effect on the image sharpness (Zheng *et al* 2019). In continuous-motion mode, the x-ray source moves continuously along the designated arc while capturing PVs at the respective angles by pulsing the x-rays. This approach introduces source motion blur that depends on the pulse width, causing geometric unsharpness and

substantially degrading the image resolution in the source motion direction (Marshall and Bosmans 2012, Zheng et al 2019). Very recently, Siemens introduced a flying focal spot technology for DBT to reduce the blurring effect of a continuously moving source (Michelfeit 2023). This new technology is outside the scope of our study.

X-ray source motion blur is not unique to DBT as a similar problem occurs in computed tomography (CT). Several image processing and reconstruction methods have been developed to address this issue. In CT, Tilley et al (2016) proposed to deconvolve the projection data for focal spot blur before reconstruction. In another study, Tilley et al (2018) modified the CT system forward model by incorporating focal spot blur as a shift-invariant convolution applied to the reconstructed images and used it in model-based iterative reconstruction (MBIR). Fu et al (2013) and Majee et al (2022) also modified the CT system forward model, but took a different approach by subsampling the focal spot and then averaging the projections. In DBT, Michielsen et al (2013) described the imaging process as projecting each DBT slice to the detector separately, convolving the slice projections with the slice-height-dependent source blur kernels, and then adding them up to obtain the final PV. They proposed to update the DBT slices sequentially for solving the reconstruction problem with their forward model. While their method increased the peak contrast-to-noise ratio of a simulated MC in a uniform background, the improvement was somewhat limited when applied to images with heterogeneous backgrounds.

Nonblind image deblurring is an important topic in image processing and computer vision research. Its task is to estimate sharp images from the blurred images given the blur kernel. Classic methods for nonblind image deblurring include the renowned Wiener filter (Wiener 1949) and Richardson-Lucy deconvolution (Richardson 1972). Model-based methods construct mathematical models and priors to estimate the latent sharp images using statistical methods like maximum *a posteriori* (MAP). Great efforts have been devoted to designing image priors for MAP estimation (Krishnan and Fergus 2009, Zoran and Weiss 2011, Xu et al 2013). In recent years, deep convolutional neural network (CNN) methods have emerged as powerful tools, leveraging the capacity of deep networks to learn complex mappings from blurry to sharp images (Dong et al 2022, Zhang et al 2022). More recently, denoising diffusion probabilistic models (DDPM) and score-based generative models have gained significant attention for their ability to generate high-quality samples (Song and Ermon 2019, Ho et al 2020, Song et al 2021b). Their remarkable success in image generation facilitates various inverse problems including image deblurring (Kawar et al 2022, Saharia et al 2023, Wang et al 2023) and image reconstruction (Song et al 2022, Chung et al 2023).

X-ray source blur modeling for DBT remains a challenging problem due to its shift-variant nature. Furthermore, the low-dose exposure of DBT introduces a high level of image noise. In this study, we analytically derived the in-plane blur kernel for the reconstructed DBT slices using imaging geometry and showed that it could be approximated by a shift-invariant kernel for a given slice. We proposed an effective post-processing nonblind image deblurring approach with DDPM as an image prior and applied it to the reconstructed DBT images.

2. Methods

2.1. Source blur modeling

2.1.1. DBT imaging system

Figure 1 shows a diagram of a DBT imaging system. We use x - y - z coordinates for the imaged volume and t - s coordinates for the digital detector. During the imaging process, the x-ray source moves around the compressed breast in the chest wall plane. The rotation center is denoted as O . The source rotates over a limited angular range, and a total of N_p PVs are captured. The raw projections available in clinical DBT systems usually have been preprocessed to correct for detector artifacts. We denote the source location for the central PV as S and its vertical projection onto the detector plane as D . Let $\mathbf{v} \in \mathbb{R}^{N_x N_y N_z}$ denote the (vectorized) DBT volume and $\mathbf{y}_i \in \mathbb{R}^{N_i N_s}$ denote the post-log PV at the i th scan angle. The forward imaging process can be characterized by the system matrices $\mathbf{A}_i \in \mathbb{R}^{N_i N_s \times N_x N_y N_z}$

$$\mathbf{y}_i = \mathbf{A}_i \mathbf{v} + \mathbf{n}_{i,y} \quad i = 1, \dots, N_p, \quad (1)$$

where $\mathbf{n}_{i,y} \sim \mathcal{N}(0, \sigma_{i,y}^2 \mathbf{I})$ is the additive PV noise that follows the Gaussian distribution with a mean of 0 and standard deviation of $\sigma_{i,y}$. We stack all the \mathbf{A}_i matrices to define the overall system of equations

$$\mathbf{y} = \mathbf{A} \mathbf{v} + \mathbf{n}_y \quad \text{where } \mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_{N_p} \end{bmatrix}, \mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{N_p} \end{bmatrix}, \mathbf{n}_y = \begin{bmatrix} \mathbf{n}_{1,y} \\ \vdots \\ \mathbf{n}_{N_p,y} \end{bmatrix}. \quad (2)$$

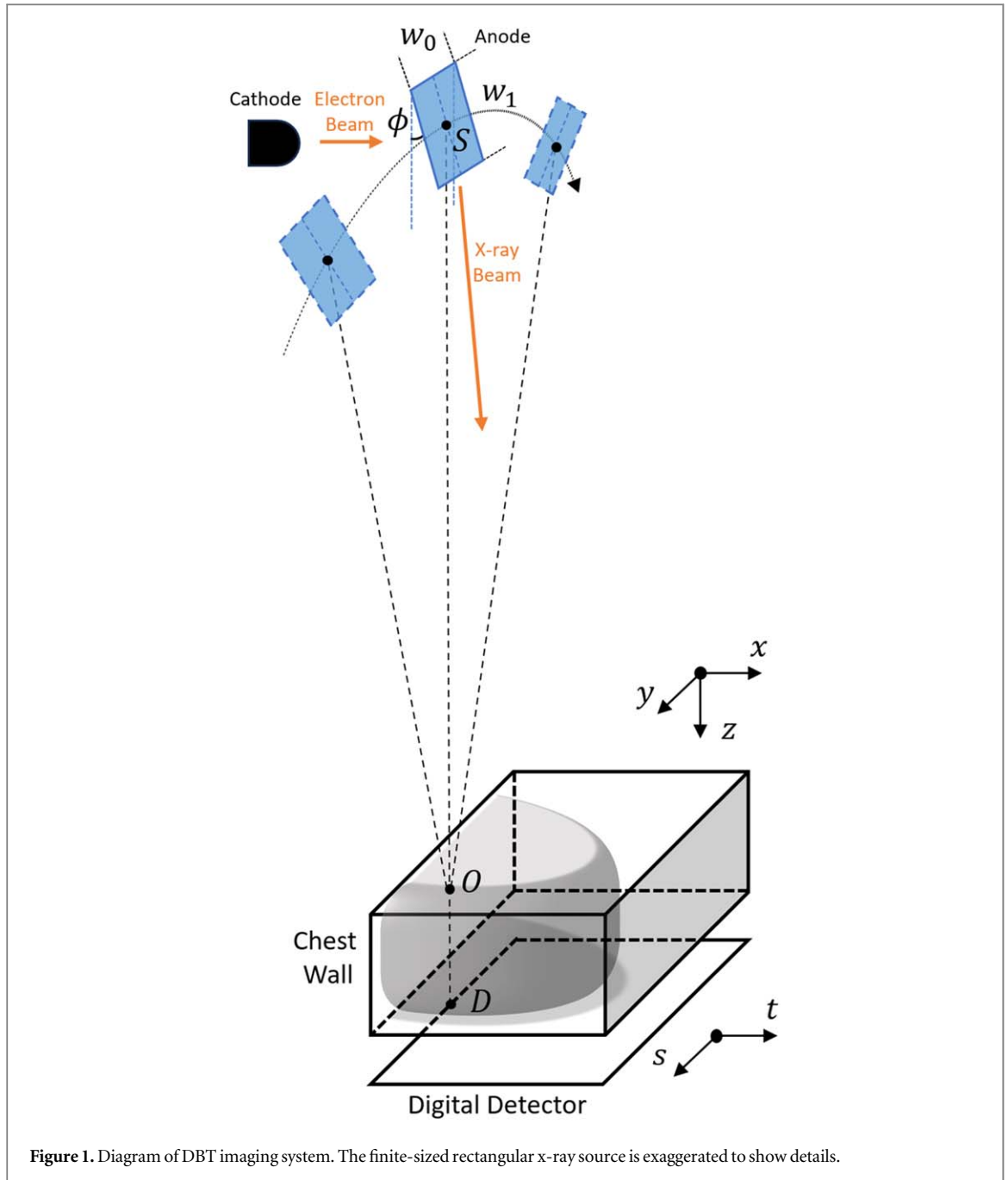


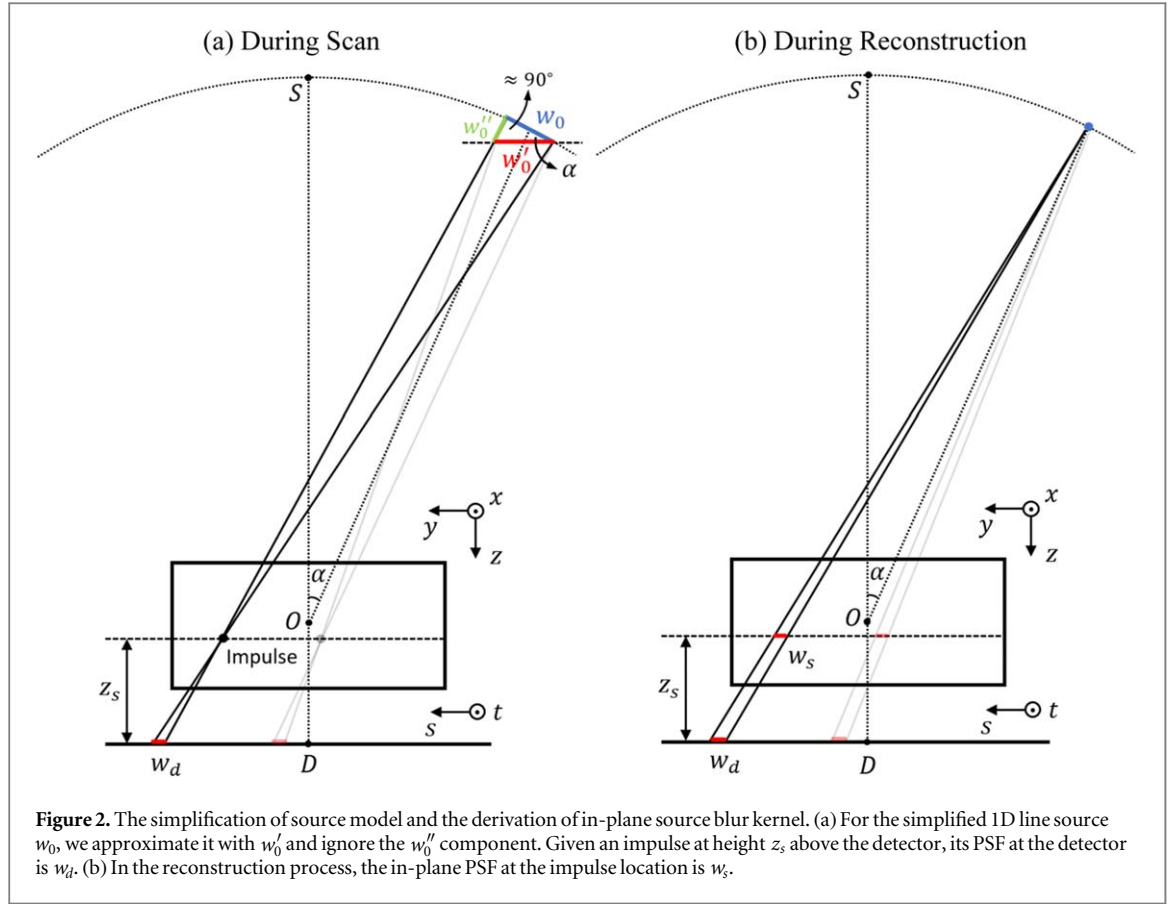
Figure 1. Diagram of DBT imaging system. The finite-sized rectangular x-ray source is exaggerated to show details.

2.1.2. In-plane source blur kernel estimation

The x-ray source is not a point source but has a finite size and shape. For simplicity, it is modeled as a rectangular source (edges w_0 and w_1 , target angle ϕ) with x-rays emitted uniformly across the anode target, as shown in figure 1. When the source rotates around O , the edge w_0 stays parallel to the tangent line of the source trajectory. In x-ray imaging, a small ϕ is often used to keep the effective focal spot size small according to the line focus principle (Bushberg *et al* 2011). This study uses a nominal focal spot size of $w_0 = w_1 \sin \phi = 0.3$ mm, which is the nominal size for most mammography systems.

For the step-and-shoot mode, the source size is the same as the stationary nominal source size. However, for the continuous-motion mode, w_0 is elongated and equal to the nominal size convolved with the source traveling distance during the x-ray pulse. For example, the Siemens Mammomat Inspiration DBT system uses a continuous-motion x-ray source with a 0.18° motion angle per pulse for most exposures (Mackenzie *et al* 2017). Given that the distance between the source and rotation center is 600 mm, the effective source size w_0 is $0.18^\circ \times (\pi/180^\circ) \times 600$ mm + 0.3 mm = 2.185 mm.

Zheng *et al* (2019) conducted a simulation study and demonstrated that a focal spot of the nominal size has negligible effect on the DBT image resolution compared with the ideal point source, while the extra blur caused by source motion leads to a substantial loss of image resolution in the source motion direction. Since the



effective source dimension $w_1 \sin \phi$ in the x -direction (the anode-cathode direction, i.e. the chestwall-anterior direction) remains at 0.3 mm at the chest wall regardless of the motion of x-ray source and decreases to less than 0.3 mm as the distance from the chest wall increases, the source blur in the x -direction has a negligible effect on image resolution (Zheng *et al* 2019). Therefore, the source model can be simplified by ignoring the x -dimension such that the rectangular source collapses to a line source. We focus on the continuous-motion x-ray source and consider a simplified 1D line source of length w_0 tangential to the source motion trajectory, as shown in figure 2(a).

During the scan process, at the scan angle α , we consider an impulse with a distance z_s from the detector (figure 2(a)). We assume that the impulse is far from the source, so the ray connecting the impulse and the end point of w_0 is perpendicular to w_0 . We further approximate w_0 with $w'_0 = w_0 / \cos \alpha$ that is parallel with the detector. Then, the point spread function (PSF) of the impulse at the detector is a line whose length $w_d(z_s, \alpha)$ can be obtained using triangle similarity

$$w_d(z_s, \alpha) = \frac{w_0}{\cos \alpha} \cdot \frac{z_s}{D_{SO} \cdot \cos \alpha + D_{OD} - z_s}, \quad (3)$$

where D_{SO} is the distance between S and O , and D_{OD} is the distance between O and D . Although the impulse is drawn to be on the source motion plane for simplicity, w_d (the PSF of the impulse from w'_0) is independent of the location of the impulse along both the x - and y -directions, i.e. shift-invariant, at a given z_s . That is, the similar triangle relationship still holds if the impulse is moved to a different y location, as depicted in gray in figure 2(a), or if the impulse is moved to a different x location, w_d remains parallel with w'_0 and maintains the same ratio. Considering w_0 and w'_0 as vectors in 3D when the impulse is moved to a general (x, y) location in figure 2(a), the approximation of w_0 with w'_0 results in the residual vector w''_0 that may cause a shift-variant blur when projected to the detector plane. However, this projected component is small and we assume it to be negligible in our modeling, as validated by our simulation study in section 4.1.

During reconstruction, the point source projector is normally used for forward and backward projections. Then, as shown in figure 2(b), the in-plane PSF of an impulse at the reconstructed DBT slice at z_s is a line whose length $w_s(z_s, \alpha)$ is

$$w_s(z_s, \alpha) = w_d(z_s, \alpha) \cdot \frac{D_{SO} \cdot \cos \alpha + D_{OD} - z_s}{D_{SO} \cdot \cos \alpha + D_{OD}} = \frac{w_0}{\cos \alpha} \cdot \frac{z_s}{D_{SO} \cdot \cos \alpha + D_{OD}}. \quad (4)$$

Note that the PSF w_s is also shift-invariant within the slice because w_d is shift-invariant and the similar triangle relationship always holds no matter where the impulse is for a given z_s .

Equation (4) is the in-plane source blur PSF from one scan angle. The source motion blur in the reconstructed DBT images is a combined effect of the blurs from all scan angles. As shown in section 4.1, the lengths w_s from different scan angles are close to their mean averaged over all angles

$$\bar{w}_s(z_s) = \frac{1}{N_p} \sum_{\alpha \in \{\alpha_1, \dots, \alpha_{N_p}\}} w_s(z_s, \alpha). \quad (5)$$

Therefore, we treat the aggregated PSF over all scan angles also as a line with a length of \bar{w}_s .

To summarize, for a continuous-motion DBT system, the in-plane blur kernel caused by the source motion in the reconstructed DBT images is a line, the size of which is shift-invariant over a DBT plane at a given slice height. Mathematically, we define a block-diagonal matrix \mathbf{B} to characterize the source motion blur and incorporate it into the DBT system of equation (2)

$$\mathbf{y} = \mathbf{A}\mathbf{B}\mathbf{v} + \mathbf{n}_y \text{ where } \mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & 0 & \dots & 0 \\ 0 & \mathbf{B}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{B}_{N_z} \end{bmatrix}. \quad (6)$$

The matrices $\mathbf{B}_z \in \mathbb{R}^{N_x N_y \times N_x N_y}$ represent the shift-invariant blur for each slice $z = 1, \dots, N_z$ and can be efficiently implemented as convolution with kernel size \bar{w}_s . Section 3.2 and section 4.1 present a simulation study to verify our derivation and justify our assumptions that simplify the DBT source blur model.

2.2. Nonblind image deblurring

Having introduced \mathbf{B} , a natural idea for addressing source motion blur is to replace \mathbf{A} with the new system matrix $\mathbf{A}\mathbf{B}$ in existing DBT reconstruction algorithms for forward and backward projections. However, the inherent limited-angle design of DBT makes the reconstruction inverse problem highly underdetermined. We found that this posed a substantial challenge, and using $\mathbf{A}\mathbf{B}$ in DBT reconstruction did not improve image sharpness. Section 5 presents a simulation study to demonstrate the limitation of this approach.

An alternative approach is to separate \mathbf{B} from \mathbf{A} and turn the source blur modeling problem into post-processing deblurring. We formulate the deblurring problem as

$$\mathbf{v}_{\text{blur}} = \mathbf{B}\mathbf{v}_{\text{true}} + \mathbf{n}_B, \quad (7)$$

where \mathbf{v}_{blur} denotes the reconstructed DBT images without any source blur modeling (i.e. by solving (2) instead of (6)), \mathbf{v}_{true} is the unknown sharp and clean images that we want to estimate, $\mathbf{n}_B \sim \mathcal{N}(0, \sigma_B^2 \mathbf{I})$ represents noise modeled as being additive Gaussian. This deblurring problem is nonblind because \mathbf{B} is known.

When deblurring DBT images, it is crucial to control the image noise level through regularization due to the low-dose exposure of DBT scans. In this work, we investigated applying DDPM as an image prior for regularizing the deblurring process. The upcoming sections first give a brief review of DDPM, and then introduce the proposed deblurring method with generative diffusion.

2.2.1. Denoising diffusion probabilistic models (DDPM)

DDPM is a class of generative models that use a diffusion process to model complex probability distributions (Sohl-Dickstein et al 2015, Ho et al 2020). These are Bayesian methods that assume that the images of interest can be represented as random vectors characterized by some probability distribution. In clinical practice, the radiologists view the DBT images as a series of in-focus planes parallel to the detector, so we consider the 2D slices $\mathbf{x}_{\text{true}} \in \mathbb{R}^{N_x N_y}$ taken from \mathbf{v}_{true} as our images of interest with the associated distribution $p(\mathbf{x}_{\text{true}})$. The diffusion process is a Markov chain that progressively adds noise to the image until a tractable distribution, such as a standard Gaussian, is achieved (Ho et al 2020). Mathematically, for a sequence of T diffusion steps, at each step $t = 1, \dots, T$,

$$\mathbf{x}_t = \sqrt{1 - \beta_t} \mathbf{x}_{t-1} + \sqrt{\beta_t} \boldsymbol{\epsilon}, \quad (8)$$

where $\beta_t \in (0,1)$ is the prescribed noise variance schedule, $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I})$ is the standard Gaussian noise. By the design of DDPM, we have $\mathbf{x}_0 = \mathbf{x}_{\text{true}}$ and $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$. Equation (8) can also be written in a non-iterative form

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}, \quad (9)$$

where $\bar{\alpha}_t = \prod_{i=1}^t (1 - \beta_{i_0})$.

To reverse the diffusion process, DDPM trains a deep neural network $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$ parameterized by θ to learn to predict the added noise from \mathbf{x}_t . The training loss is a variant of a variational lower bound, and intuitively speaking, is the mean squared error between the predicted noise and actual added noise (Ho et al 2020)

$$\operatorname{argmin}_{\theta} \mathbb{E}_{t, \mathbf{x}_0, \boldsymbol{\epsilon}} [\|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}, t)\|^2]. \quad (10)$$

Once $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$ is trained, we can generate images by randomly initializing a sample with pure Gaussian noise $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$ and then iteratively removing noise from it following the DDPM sampling procedure (Ho et al 2020). In our implementation, we use a variant of DDPM sampling called denoising diffusion implicit models (DDIM) sampling (Song et al 2021a)

$$\mathbf{x}_{t-1} = \sqrt{\bar{\alpha}_{t-1}} \left(\frac{\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)}{\sqrt{\bar{\alpha}_t}} \right) + \sqrt{1 - \bar{\alpha}_{t-1}} \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) \quad (11)$$

for $t = T, \dots, 1$.

2.2.2. Image deblurring with generative diffusion

The goal of image deblurring is to estimate the unknown sharp and clean images \mathbf{x}_{true} from the observed corrupted images $\mathbf{x}_{\text{blur}} \in \mathbb{R}^{N_x N_y}$ taken from the DBT volume \mathbf{v}_{blur} . In Bayesian image deblurring, commonly used techniques include sampling from the posterior $p(\mathbf{x}_{\text{true}} | \mathbf{x}_{\text{blur}})$ and MAP estimation. Note that the prior $p(\mathbf{x}_{\text{true}})$ is the distribution of true DBT slices, which can be effectively sampled by a well-trained DDPM using DBT images with no noise and blur.

We propose to perform posterior sampling to estimate \mathbf{x}_{true} from \mathbf{x}_{blur} . This requires us to modify the unconditional DDPM sampling to be a conditional sampling process. To do so, we exploit the connection between DDPM and score-based generative modeling following the derivation of Dhariwal and Nichol (2021). First, it has been shown that the DDPM network $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$ approximates the gradient of the log probability, also called the score function, of the distribution $p(\mathbf{x}_t)$

$$\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t) = -\frac{\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)}{\sqrt{1 - \bar{\alpha}_t}}. \quad (12)$$

To see this, recall that $p(\mathbf{x}_t | \mathbf{x}_0) \sim \mathcal{N}(\sqrt{\bar{\alpha}_t} \mathbf{x}_0, (1 - \bar{\alpha}_t) \mathbf{I})$ from (9), so

$$\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t | \mathbf{x}_0) = -\frac{\mathbf{x}_t - \sqrt{\bar{\alpha}_t} \mathbf{x}_0}{1 - \bar{\alpha}_t} = -\frac{(\mathbf{x}_t - \sqrt{\bar{\alpha}_t} \mathbf{x}_0) / \sqrt{1 - \bar{\alpha}_t}}{\sqrt{1 - \bar{\alpha}_t}} \approx -\frac{\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)}{\sqrt{1 - \bar{\alpha}_t}}.$$

and thus, using the law of iterated expectation

$$\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t) = \mathbb{E}_{p(\mathbf{x}_0)} [\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t | \mathbf{x}_0)] \approx \mathbb{E}_{p(\mathbf{x}_0)} \left[-\frac{\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)}{\sqrt{1 - \bar{\alpha}_t}} \right] = -\frac{\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)}{\sqrt{1 - \bar{\alpha}_t}},$$

which gives (12). Then, the score function of the conditional distribution $p(\mathbf{x}_t | \mathbf{x}_{\text{blur}})$ becomes

$$\begin{aligned} \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t | \mathbf{x}_{\text{blur}}) &= \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t) + \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_{\text{blur}} | \mathbf{x}_t) \\ &= -\frac{1}{\sqrt{1 - \bar{\alpha}_t}} \underbrace{(\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) - \sqrt{1 - \bar{\alpha}_t} \cdot \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_{\text{blur}} | \mathbf{x}_t))}_{\tilde{\boldsymbol{\epsilon}}_\theta(\mathbf{x}_t, t)} \end{aligned} \quad (13)$$

where Bayes rule gives $p(\mathbf{x}_t | \mathbf{x}_{\text{blur}}) = \frac{p(\mathbf{x}_t) p(\mathbf{x}_{\text{blur}} | \mathbf{x}_t)}{p(\mathbf{x}_{\text{blur}})}$ and the gradient of $\log p(\mathbf{x}_{\text{blur}})$ with respect to \mathbf{x}_t vanishes because it does not depend on \mathbf{x}_t .

When t is small, the structural details of \mathbf{x}_t are close to \mathbf{x}_0 , so we assume $\mathbf{x}_{\text{blur}} \approx \mathbf{B} \mathbf{x}_t + \mathbf{n}_{B,t}$, $\mathbf{n}_{B,t} \sim \mathcal{N}(0, \sigma_{B,t}^2 \mathbf{I})$. We omit the subscript of \mathbf{B} to simplify the notation, but it should be clear that \mathbf{B} here is the blur matrix for a slice rather than a volume. The gradient of the data-fit term is therefore

$\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_{\text{blur}} | \mathbf{x}_t) = -\frac{1}{\sigma_{B,t}^2} \mathbf{B}'(\mathbf{B}\mathbf{x}_t - \mathbf{x}_{\text{blur}})$ where $'$ denotes matrix transpose. Finally, we insert this gradient into (13) and define the modified DDPM output

$$\tilde{\epsilon}_{\theta}(\mathbf{x}_t, t) = \epsilon_{\theta}(\mathbf{x}_t, t) + \frac{\sqrt{1 - \bar{\alpha}_t}}{\sigma_{B,t}^2} \mathbf{B}'(\mathbf{B}\mathbf{x}_t - \mathbf{x}_{\text{blur}}). \quad (14)$$

We use this function in DDPM sampling to draw samples from the posterior $p(\mathbf{x}_{\text{true}} | \mathbf{x}_{\text{blur}})$ instead of $p(\mathbf{x}_{\text{true}})$. The sampling equation (11) now becomes

$$\begin{aligned} \mathbf{x}_{t-1} &= \sqrt{\bar{\alpha}_{t-1}} \left(\frac{\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t} \tilde{\epsilon}_{\theta}}{\sqrt{\bar{\alpha}_t}} \right) + \sqrt{1 - \bar{\alpha}_{t-1}} \tilde{\epsilon}_{\theta} \\ &= \sqrt{\bar{\alpha}_{t-1}} \left(\frac{\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t} \epsilon_{\theta}}{\sqrt{\bar{\alpha}_t}} \right) + \sqrt{1 - \bar{\alpha}_{t-1}} \epsilon_{\theta} - \frac{\sqrt{1 - \bar{\alpha}_t}}{\sigma_{B,t}^2} \left(\frac{\sqrt{1 - \bar{\alpha}_t}}{\sqrt{1 - \beta_t}} - \sqrt{1 - \bar{\alpha}_{t-1}} \right) \cdot \mathbf{B}'(\mathbf{B}\mathbf{x}_t - \mathbf{x}_{\text{blur}}) \\ &\approx \sqrt{\bar{\alpha}_{t-1}} \left(\frac{\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t} \epsilon_{\theta}}{\sqrt{\bar{\alpha}_t}} \right) + \sqrt{1 - \bar{\alpha}_{t-1}} \epsilon_{\theta} - \lambda \cdot \mathbf{B}'(\mathbf{B}\mathbf{x}_t - \mathbf{x}_{\text{blur}}) \end{aligned} \quad (15)$$

where we isolate $\mathbf{B}'(\mathbf{B}\mathbf{x}_t - \mathbf{x}_{\text{blur}})$ and replace its positive coefficient with a tuning parameter λ to control the balance between noise reduction and data fidelity. We drop the t -dependency of λ for easier tuning while still achieving empirically good results.

2.3. X-ray source motion deblurring for reconstructed DBTs

We have introduced a post-processing deblurring method for reconstructed DBT image slices through DDPM posterior sampling. We only run the sampling equation (15) for a small number of steps $t = \bar{T}, \dots, 1$, where $\bar{T} \ll T$ to satisfy the small t assumption. To deblur the entire DBT volume, we do slice-by-slice deblurring.

The proposed deblurring method is applicable to DBTs obtained from any reconstruction method. In this study, we investigate applying it to the model-based image reconstruction with detector blur and correlated noise modeling (DBCN) approach (Zheng *et al* 2018). By deblurring DBCN reconstructed images, the overall image reconstruction and post-processing pipeline represents a framework that employs both source motion blur and detector blur modeling.

3. Materials

3.1. DBT system configuration

We focus our study on the Siemens Mammomat Inspiration DBT system that takes 25 PVs from -25° to 25° scan angles in 2.08° intervals with $D_{SO} = 600$ mm and $D_{OD} = 50$ mm. The gap between the detector plane and the bottom of the compressed breast is 20 mm. The Siemens system uses a continuous-motion x-ray source with a typical 0.18° source motion angle (Mackenzie *et al* 2017), and we modeled it as a line source of length $w_0 = 2.185$ mm, as described in section 2.1.2. The detector pixel size is 0.085 mm \times 0.085 mm.

3.2. Verification study of blur kernel modeling

We designed a simulation study to verify the in-plane blur kernel modeling. Figure 3 shows the simulation workflow. First, we created an impulse object in the voxelized image with background value of zero. The impulse value was 0.05 mm $^{-1}$, close to the typical attenuation coefficient of breast tissues (Johns and Yaffe 1987). Then, we simulated the PVs using the point and blurry sources, denoted as \mathbf{y}_{pt} and \mathbf{y}_{blur} , respectively. The pixel values of \mathbf{y}_{pt} were generated with the segmented separable footprint (SG) projector (Zheng *et al* 2017) instead of simple ray-tracing. The generation of \mathbf{y}_{blur} used 50 SG projectors as sub-focal spots within w_0 uniformly. Because the blur occurs after the x-ray is attenuated in actual scans, we summed the sub-PVs in the pre-log domain to match the physics process and then took log to get the post-log blurry PVs. The simulation was noise-free.

Next, we reconstructed the PSFs with a point source SG projector for the following two conditions: $\text{argmin}_{\mathbf{v}} \frac{1}{2} \|\mathbf{A}\mathbf{v} - \mathbf{y}_{\text{pt}}\|^2$ and $\text{argmin}_{\mathbf{v}} \frac{1}{2} \|\mathbf{A}\mathbf{v} - \mathbf{y}_{\text{blur}}\|^2$. The reconstructed PSFs had spoke-like inter-plane artifact due to the limited-angle nature of DBT. We took the 1D PSFs through the impulse in the y -direction and took Fourier transform to obtain the impulse optical transfer functions (OTFs). Although the OTFs are complex-valued in general because the PSFs are asymmetric (except when the impulse is centered in y), their ratio, which represents the source blur OTF, is always real-valued.

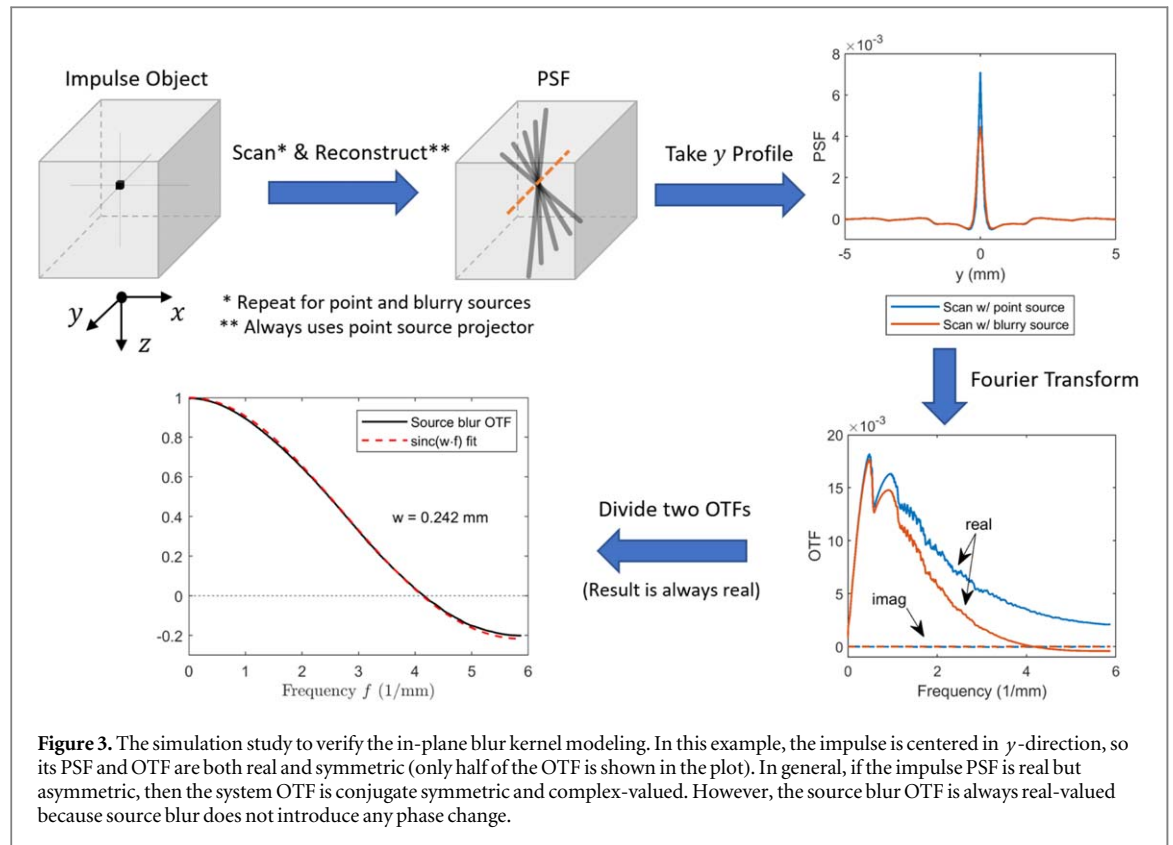


Table 1. Density and size characteristics of the virtual breast phantoms for DDPM training.

Density	Almost entirely fatty	Scattered fibroglandular dense	Heterogeneously dense	Extremely dense
GVF	5%	15%	34%	60%
No. of phantoms	10	10	25	25
Thickness after compression (mm)	52–70 (2 mm intervals)	46–64 (2 mm intervals)	36–60 (1 mm intervals)	31–55 (1 mm intervals)

Finally, recall that the Fourier transform of a 1D rectangle function $\frac{1}{w}\text{rect}\left(\frac{u}{w}\right)$ of width w and unit area is a sinc function $\text{sinc}(w \cdot f)$, where u and f are function variables. We fit the source blur OTF with $\text{sinc}(w \cdot f)$ to estimate the blur kernel length w and compared it with our analytically derived blur kernel length \bar{w}_z defined in (5). We moved the impulse object to different x , y , or z locations in the volume and repeated the experiment.

3.3. Data sets

3.3.1. VICTRE phantoms

We used the Virtual Imaging Clinical Trial for Regulatory Evaluation (VICTRE) package (Badano *et al* 2018) to create virtual phantoms with breast tissue backgrounds to train the DDPM network and test the deblurring methods. The VICTRE package can generate anthropomorphic breast phantoms that were used for virtual clinical trials. The VICTRE virtual phantoms were defined on a 3D fine grid where each voxel was assigned with a label indicating its material. The voxel size was $0.05 \text{ mm} \times 0.05 \text{ mm} \times 0.05 \text{ mm}$. The PVs of the virtual phantoms were simulated by MC-GPU, a Monte Carlo x-ray imaging simulator in the package. MC-GPU was configured for the Siemens Mammomat Inspiration DBT system, and its simulation accuracy had been validated in terms of noise and resolution (Badal *et al* 2021). It also provided an option of either using an ideal point source or a blurry source with 0.3 mm nominal size and 0.18° motion angle for the scans. The x-ray exposure was adaptively determined for each phantom by first running a quick scan with a small exposure, and then the full scan with a scaled exposure so that the mean glandular dose matched that of a real scan under automatic exposure control (AEC) (PHE 2018). MC-GPU assumed constant tube current for each PV. Scatter was simulated by MC-GPU, but we did not correct for scatter in this study.

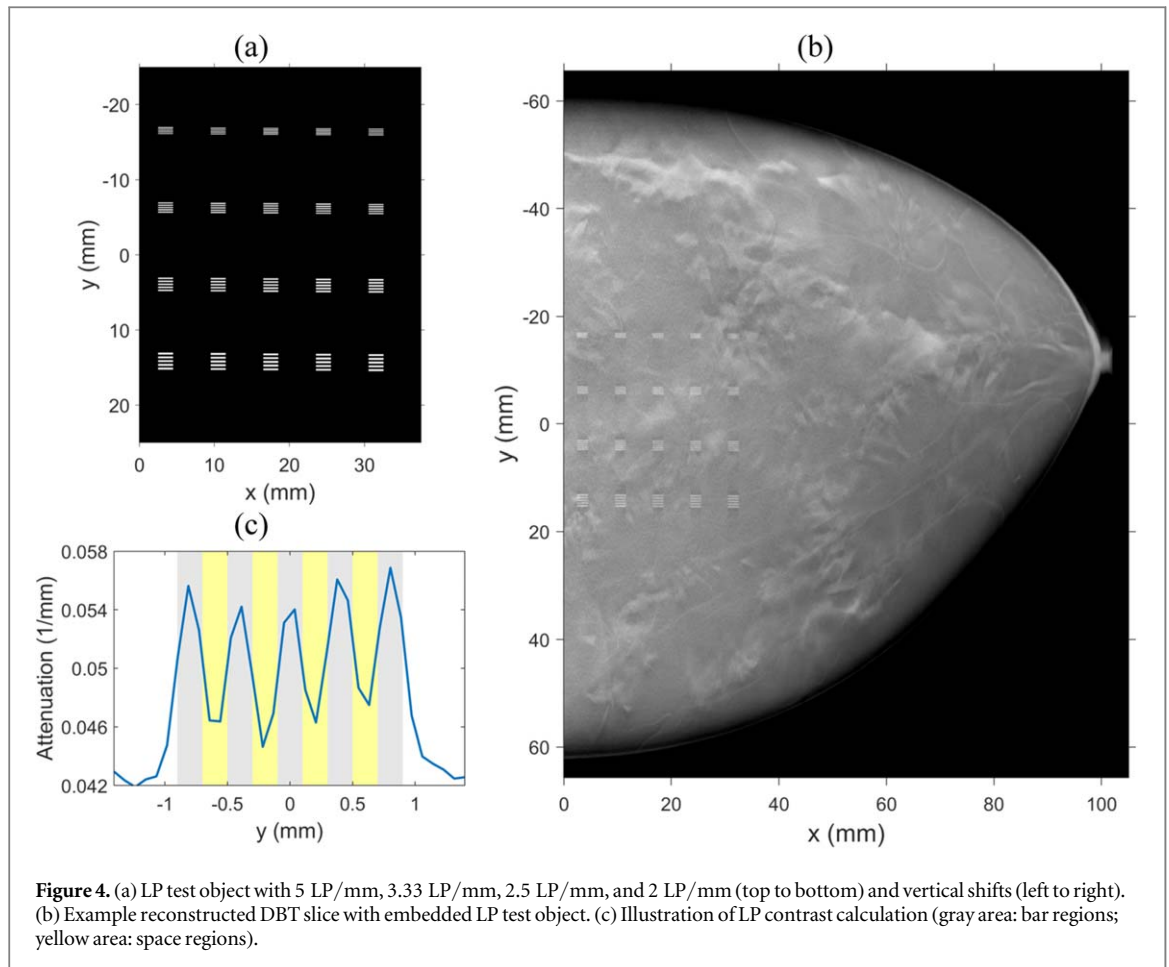


Figure 4. (a) LP test object with 5 LP/mm, 3.33 LP/mm, 2.5 LP/mm, and 2 LP/mm (top to bottom) and vertical shifts (left to right). (b) Example reconstructed DBT slice with embedded LP test object. (c) Illustration of LP contrast calculation (gray area: bar regions; yellow area: space regions).

For DDPM training, we created 70 virtual phantoms whose density and size characteristics are shown in table 1. The glandular volume fraction (GVF) setting followed that of the VICTRE study (Badano *et al* 2018). According to the formulation of the deblurring problem (7), \mathbf{x}_{true} (or \mathbf{v}_{true}) represented the sharp and noiseless images. Hence, we used the point source in MC-GPU, and increased the exposure to be 5 times the AEC to better represent the ‘noiseless’ prior DBT image distribution $p(\mathbf{x}_{\text{true}})$. We reconstructed the DBTs using DBCN (5 iterations, regularization strength = 70) and the SG projector (Zheng *et al* 2018). The reconstructed image voxel size was $0.085 \text{ mm} \times 0.085 \text{ mm} \times 1.0 \text{ mm}$. Due to the large sizes of DBT images, we trained the DDPM network using image patches instead of full DBT slices to reduce memory cost. We previously investigated the patch sizes of 32×32 pixels, 64×64 pixels, 128×128 pixels, and 256×256 pixels. All the patch sizes were found to work well for training; the network trained with 64×64 -pixel patches generated more realistic looking images with a reasonable amount of training time compared to other patch sizes. We randomly extracted 128,401 64×64 -pixel non-overlapping 2D slice patches from the reconstructed DBT images to form the DDPM training set.

To test the deblurring methods, we created a separate set of virtual test phantoms. We considered two breast sizes: an average breast, and a large breast that maximized the source blur effect. The average breast had a diameter of 105 mm at the chest wall before compression (with a cone-like shape) and a thickness of 60 mm after compression. The large breast had a diameter of 120 mm before compression and a thickness of 80 mm after compression. We created 4 phantoms for both sizes, one at each breast density, resulting in a total of 8 test phantoms. The phantoms were embedded with a line pair test object discussed next. We scanned the phantoms twice under AEC exposure in MC-GPU, first using the blurry source, and then using the ideal point source to serve as a reference standard. The DBT images of the test phantoms were also reconstructed using DBCN (5 iterations, regularization strength = 70) and the SG projector.

3.3.2. Line pair test object and image quality metrics

To quantitatively evaluate the image resolution, we designed a test object consisting of line pairs (LPs) with a range of spatial frequencies, as shown in figure 4(a). The test object was $35 \text{ mm} \times 35 \text{ mm} \times 0.2 \text{ mm}$ in size and voxelized with size of $0.05 \text{ mm} \times 0.05 \text{ mm} \times 0.05 \text{ mm}$ (same as the virtual phantoms). The LPs were 0.2 mm in thickness and 2 mm in length and were made of calcium oxalate to mimic the attenuation of small MCs. Each LP

group contained five horizontally placed bars with equal width and spacing. From top to bottom of the test object, the four bar width settings were 0.1 mm, 0.15 mm, 0.2 mm, and 0.25 mm. These were equivalent to the spatial frequencies of 5 LP/mm, 3.33 LP/mm, 2.5 LP/mm, and 2 LP/mm, respectively. The line pair groups were placed 10 mm apart in the vertical directions. Since the alignment between the test object and reconstruction voxel grid could affect the resolution of the reconstructed LPs, we duplicated the LPs in five columns from left to right of the test object with an accumulation of 0.05 mm vertical shift, simulating the possible random alignments of small object with the detector pixel array during imaging. The columns of the line pair groups were 7 mm apart in the horizontal direction.

To insert the LP test object into the virtual test phantoms, we assigned the corresponding phantom voxels with the label of calcium oxalate. The test object was parallel to the detector at a specified height from the detector and the bars were perpendicular to the source motion direction. The test object was close to the chest wall and centered in the y -direction, and was inserted well within the breast. Then, we simulated the PVs and reconstructed the images, as described in section 3.3.1. Figure 4(b) shows an example of a reconstructed DBT slice containing the test object.

We calculated the LP contrast and image noise of the reconstructed DBT images as image quality metrics. For each reconstructed LP, we averaged the central 1.5 mm bar region along x -direction to get the LP profile in y -direction. Then, we overlapped the LP profile with the ground truth locations of the five bars and four spaces on the continuous y coordinate, as illustrated in figure 4(c). The profile of the reconstructed LP was linearly interpolated (lines connecting adjacent pixel values). If the reconstructed LP was well-resolved, it should have peaks and valleys matching the corresponding locations in the ideal profile. We followed Zheng *et al* (2019) and defined the contrast of the reconstructed LP as the difference between the mean value of the five ground truth bar regions (gray area in figure 4(c)) and the mean value of the four space regions (yellow area in figure 4(c)), normalized to the contrast of the input ideal profile. Since the input ideal profile had the same contrast for all LP frequencies, the relative contrast of the LP frequencies would be the same with or without normalization. The final contrast of each LP frequency was averaged over the five shifted instances in the test object. To quantify image noise, we took 20 10×10 -pixel LP-free regions of interest (ROIs) near the LPs and calculated the root-mean-square pixel variation of the ROIs after background removal using quadratic fitting (Gao *et al* 2023). The overall image noise level was the average over all noise ROIs.

3.4. DDPM implementation

The structure of the DDPM network was a modified U-Net (Ronneberger *et al* 2015) as described in Ho *et al* (2020). The U-Net had four downsampling scales, each with three ResNet blocks (He *et al* 2016). The numbers of 3×3 convolutional kernels for the downsampling scales were 64, 128, 128, 128, respectively. Following Ho *et al* (2020), we trained the DDPM using 60 training epochs, a batch size of 256, and a learning rate of 10^{-4} with Adam optimizer (Kingma and Ba 2015). The noise variance schedule β_t was evenly spaced between $\beta_0 = 10^{-4}$ and $\beta_T = 0.02$ with $T = 1000$ (Ho *et al* 2020). The diffusion steps t were encoded with sinusoidal positional encoding (Vaswani *et al* 2017) and then added to the feature maps of the ResNet blocks. We removed all the attention layers, so the network contained only convolution and up/downsampling layers, and thus could process images of arbitrary sizes. Section 4.3.1 discusses the parameter selection of \bar{T} and λ in the DDPM posterior sampling for deblurring.

3.5. Comparison methods

We compared the proposed deblurring method with the following nonblind deblurring methods: Tikhonov regularized deblurring (Tikhonov 1963, Gunturk and Li 2013), total variation (TV) regularized deblurring, and the unfolding super-resolution network (USRNet) (Zhang *et al* 2020). Tikhonov regularized deblurring was formulated as $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \frac{1}{2} \|\mathbf{B}\mathbf{v} - \mathbf{v}_{\text{blur}}\|^2 + \frac{\mu_{\text{Tik}}}{2} \|\mathbf{v} - \mathbf{v}_{\text{blur}}\|^2$, which has an analytical solution that uses an inverse filter: $\hat{\mathbf{v}} = (\mathbf{B}'\mathbf{B} + \mu_{\text{Tik}}\mathbf{I})^{-1}(\mathbf{B}' + \mu_{\text{Tik}}\mathbf{I})\mathbf{v}_{\text{blur}}$. TV regularized deblurring was formulated as $\hat{\mathbf{v}} = \arg\min_{\mathbf{v}} \frac{1}{2} \|\mathbf{B}\mathbf{v} - \mathbf{v}_{\text{blur}}\|^2 + \mu_{\text{TV}} \|\mathbf{D}\mathbf{v}\|_1$ where \mathbf{D} contains the finite difference operators in x - and y -directions. USRNet was an end-to-end trainable CNN, and we trained it using the paired virtual phantom images with low-quality images being the simulated $1 \times$ AEC exposure using a blurry source and high-quality images being the simulated $5 \times$ AEC exposure using a point source.

4. Results

4.1. Verification study of blur kernel modeling

Figure 5(a) shows the example impulse OTFs and source blur OTFs in the simulation study of the in-plane blur kernel modeling. The impulse was placed close to the chest wall at different heights z_s . The impulse was centered

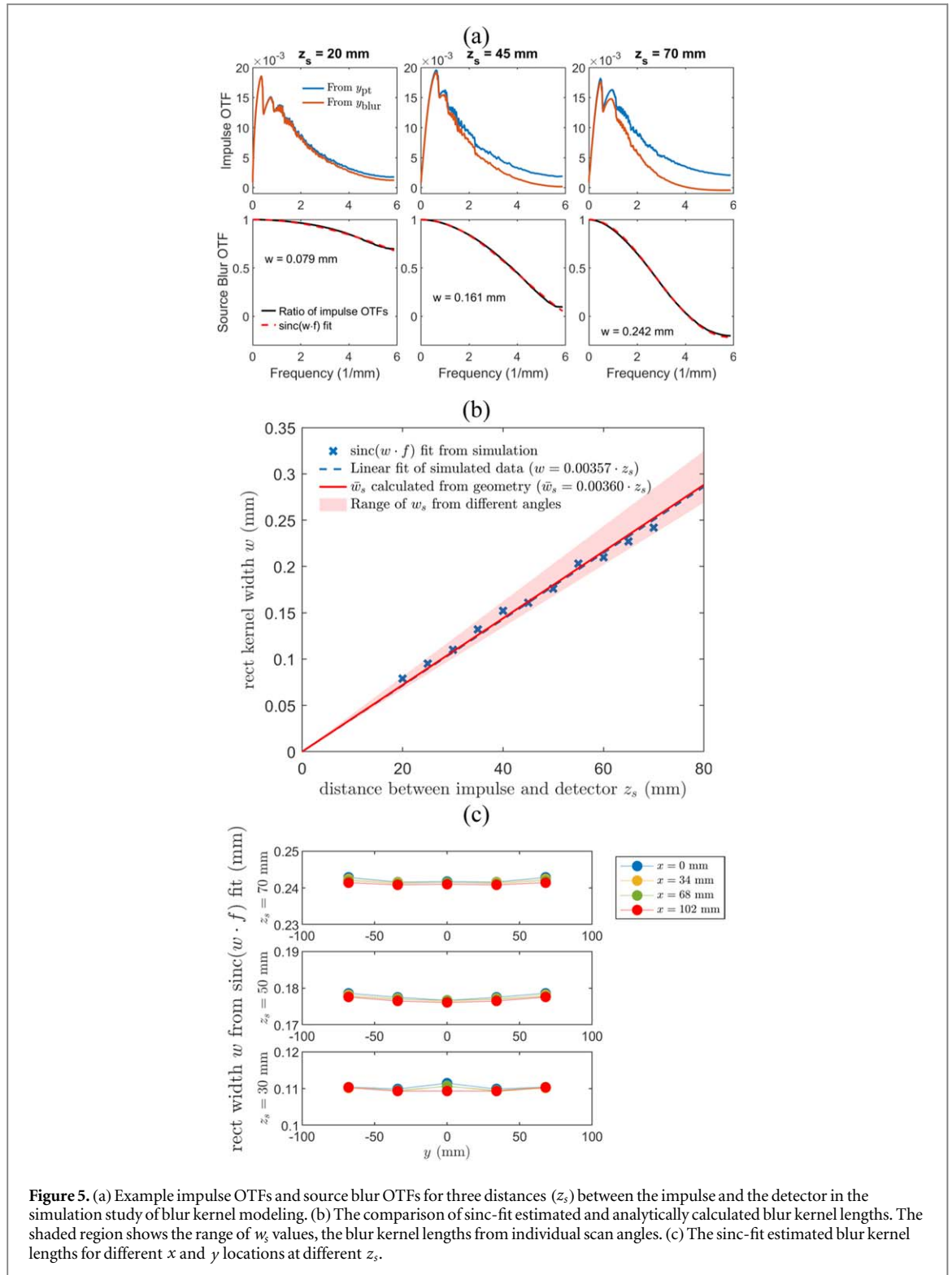


Figure 5. (a) Example impulse OTFs and source blur OTFs for three distances (z_s) between the impulse and the detector in the simulation study of blur kernel modeling. (b) The comparison of sinc-fit estimated and analytically calculated blur kernel lengths. The shaded region shows the range of w_s values, the blur kernel lengths from individual scan angles. (c) The sinc-fit estimated blur kernel lengths for different x and y locations at different z_s .

in y -direction, so its PSF and OTF were both real and symmetric. Figure 5(b) shows the scatter plot of the sinc-fit estimated blur kernel length versus z_s . The data points exhibited a good linear relationship (linear fitting result: $w = 0.00357 \cdot z_s$, correlation coefficient = 0.998, $p < 0.0001$). The linear fit of the data points had an almost perfect alignment with the analytically calculated kernel length $\bar{w}_s = 0.00360 \cdot z_s$.

The shaded region in figure 5(b) shows the range of w_s , the lengths of blur kernels from individual scan angles. For the Siemens Mammomat Inspiration DBT system that acquired 25 PVs from -25° to 25° , w_s differed from \bar{w}_s by -6.7% at 0° (lower bound in figure 5(b)) and 12.7% at 25° (upper bound in figure 5(b)). In other words, the variation of w_s was small compared to the mean \bar{w}_s . We also moved the impulse to different x , y , and z locations in the simulation. As shown in figure 5(c), the sinc-fit kernel sizes were very close with differences less

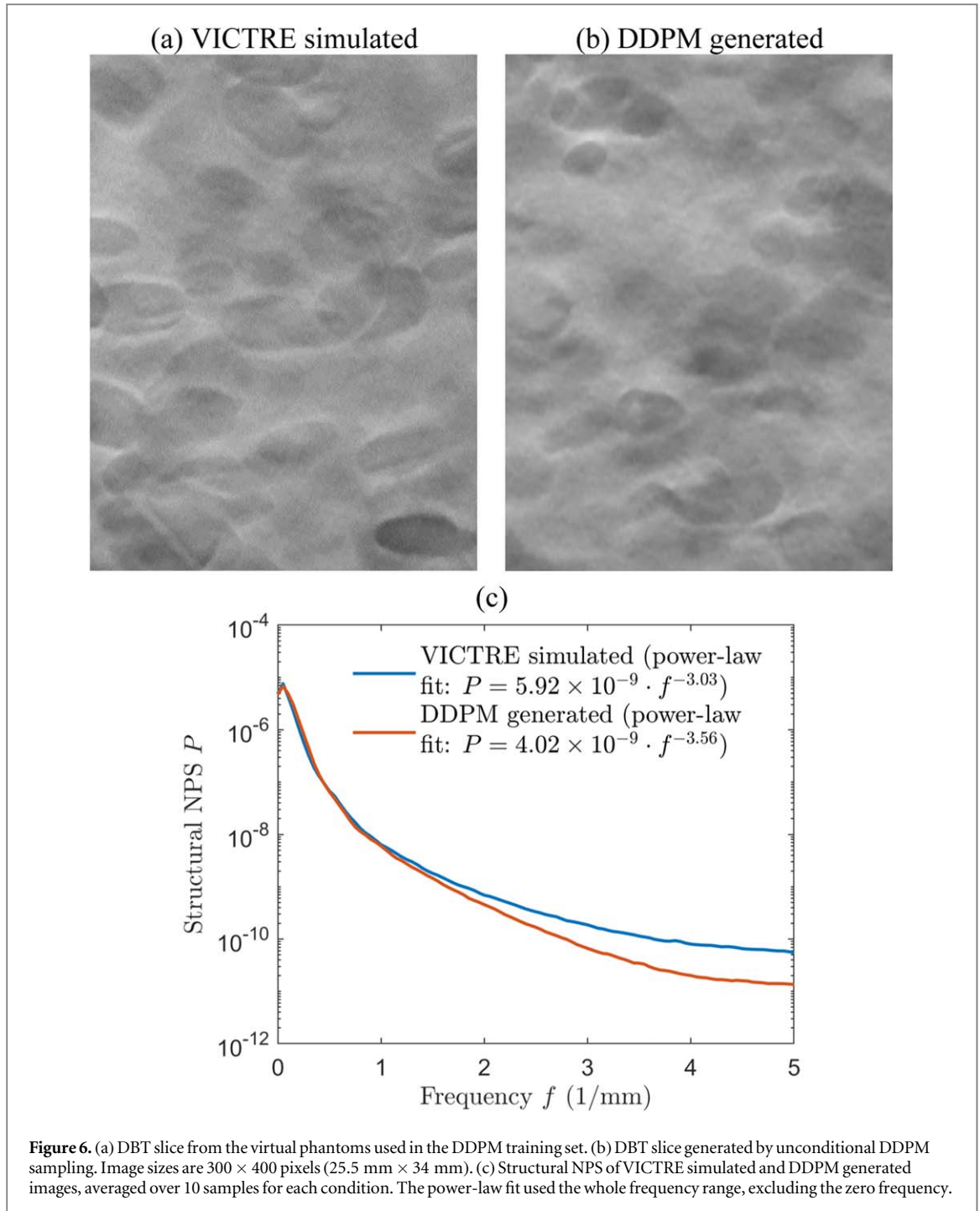
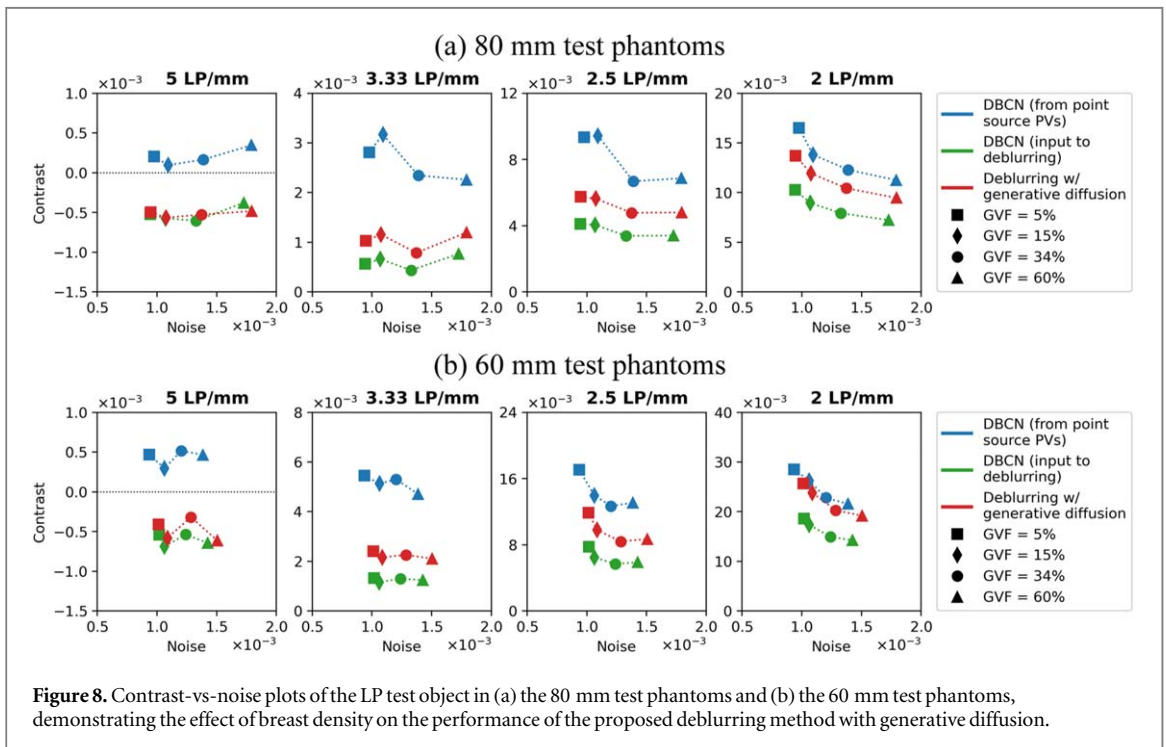
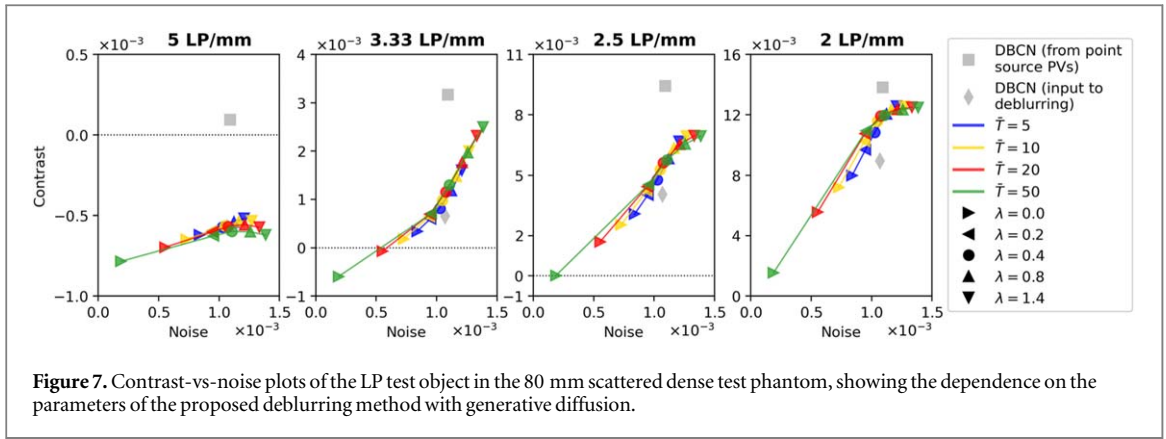


Figure 6. (a) DBT slice from the virtual phantoms used in the DDPM training set. (b) DBT slice generated by unconditional DDPM sampling. Image sizes are 300×400 pixels ($25.5 \text{ mm} \times 34 \text{ mm}$). (c) Structural NPS of VICTRE simulated and DDPM generated images, averaged over 10 samples for each condition. The power-law fit used the whole frequency range, excluding the zero frequency.

than 1.3% from the mean and were almost shift-invariant for a given z_s . These observations justified our simplification of source model of ignoring w_0'' and averaging the kernel lengths over all scan angles, resulting in a shift-invariant 1D line kernel over the area of a reconstructed DBT plane at a given z_s .

4.2. DDPM unconditional image generation

To demonstrate the ability of DDPM to produce high-quality DBT images, we ran unconditional DDPM sampling to draw samples from the prior distribution $p(\mathbf{x}_{\text{true}})$. Figure 6(a) shows an example DBT slice from the DDPM training set. Figure 6(b) shows an example of DDPM generated sample. The DDPM generated images had natural heterogeneous background textures resembling the characteristics of the training images and were free from artifacts. The generated DBT images could be considered to be 1.0 mm slices, similar to the training samples. The structural noise power spectrum (NPS) (Gao et al 2021a) of the DDPM generated images exhibited a power-law form as for mammograms (Burgess 1999) and was close to the NPS of the VICTRE simulated



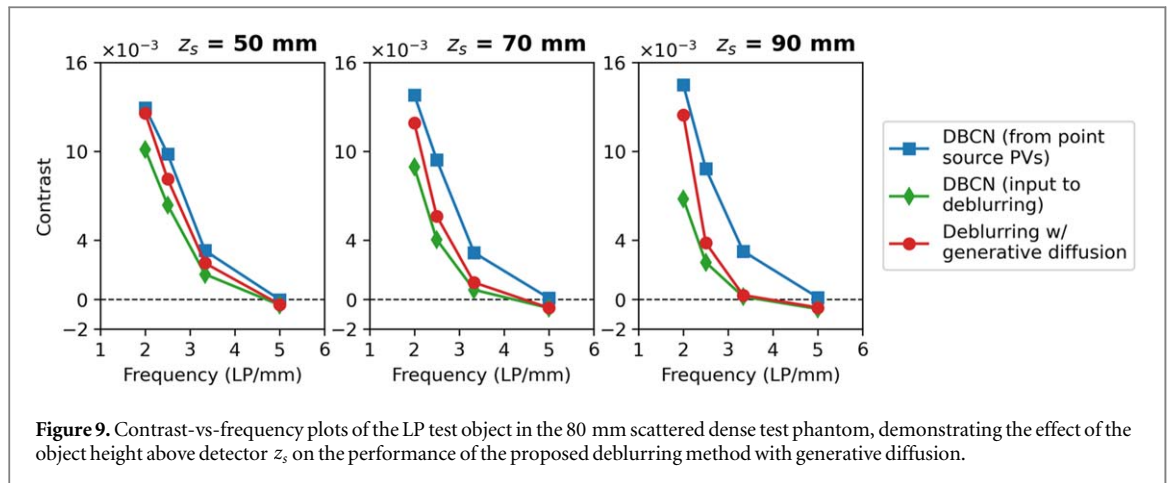
training images, as shown in figure 6(c). The exponent values of the power-law fitting to the NPS curves for the VICTRE simulated images and the DDPM generated images were 3.03 and 3.56, respectively.

4.3. Image deblurring with generative diffusion

4.3.1. Parameter selection of DDPM posterior sampling

To select the number of sampling steps \bar{T} and the weight parameter λ in DDPM posterior sampling, we did a grid search by varying $\bar{T} = 5, 10, 20, 50$, and $\lambda = 0.0, 0.2, 0.4, 0.8, 1.4$. We positioned the LP test object at $z_s = 70$ mm in the 80 mm scattered dense test phantom (note the 20 mm gap between the detector and the bottom of the compressed breast) and deblurred its DBCN reconstructed image using these parameter settings. Figure 7 shows the contrast-vs-noise plots of the LPs. The LPs with 5 LP/mm had severe blurring due to their narrow spacing, and the frequency was close to the Nyquist frequency associated with the voxel size (5.88 LP/mm). Deblurring could not recover their resolution, resulting in always negative contrasts. Therefore, we focused our attention on LP frequencies lower than 5 LP/mm.

When $\lambda = 0.0$, the deblurring method simplified to unconditional DDPM sampling. In this situation, the blur of LPs became more severe as \bar{T} increased. As we increased λ , the LP contrast improved due to the high frequency boosting from the gradient of the data-fit term. However, this enhancement also amplified background noise. Additionally, the enforcement of data fidelity at each sampling step became stronger, making the impact of \bar{T} less apparent. To balance between contrast enhancement and noise control, we selected $\lambda = 0.4$



and $\bar{T} = 20$ for the subsequent sections of this study. This parameter setting improved the image resolution after deblurring while maintaining the same image noise level as the blurry input.

4.3.2. Effect of breast densities on deblurring

To investigate the effect of breast density on the deblurring performance, we applied the proposed deblurring method to the 8 test phantoms that were 80 mm and 60 mm thick with 4 breast density settings: 5%, 15%, 34%, and 60% GVF. We placed the LP test object at $z_s = 70$ mm in these phantoms, which was 30 mm and 10 mm, respectively, deep inside the breasts from the compression paddle. Figure 8 shows the contrast-vs-noise plots of the LPs for the DBCN images and the deblurred images. The LP contrasts in the 80 mm phantoms were generally lower than those in the 60 mm phantoms because more scatter in thicker phantoms and more severe inter-plane artifacts caused a loss of object contrast. For either the 80 mm or the 60 mm phantoms, as GVF increased, the image noise increased due to the AEC mechanism in the MC-GPU simulation. In particular, the same breast thickness used the same mean glandular dose adjusted by the AEC. Hence, dense breasts absorbed more x-rays and had fewer transmitted x-rays, leading to higher image noise. Regardless, the proposed deblurring method with generative diffusion consistently improved the LP contrasts, demonstrating its robustness and flexibility to handle various image noise levels and breast densities. The relative trends of the improvement for the 60 mm and 80 mm phantoms were similar. While we mainly used the scattered dense test phantoms (GVF = 15%) in other sections of this study, our findings affirmed the applicability of the proposed method to a broader range of breast densities.

4.3.3. Effect of test object heights above detector

Due to the increased geometric unsharpness, the x-ray source blur increased as the DBT slice became closer to the x-ray source. To assess its impact on the proposed deblurring method, we placed the LP test object at $z_s = 50$ mm, 70 mm, and 90 mm in the 80 mm scattered dense test phantom, which was 50 mm, 30 mm, and 10 mm, respectively, deep inside the breast from the compression paddle. The magnification factors of these slices were 1.08, 1.12, and 1.16, respectively. Figure 9 shows the contrasts versus LP frequency for the DBCN images and the deblurred images. These contrast-vs-frequency plots resembled the modulation transfer function (MTF) curves commonly used in assessing radiographic systems, albeit with the signals represented by rectangular LPs instead of sine waves. Also, our LPs were made with calcium oxalate instead of lead in the MC-GPU simulation.

The DBCN images reconstructed from point source PVs served as the reference standard, where the LP contrasts remained almost the same irrespective of z_s due to the absence of source blur. The DBCN images reconstructed from blurry source PVs had a reduction in LP contrasts compared to the reference standard. This reduction was more pronounced when z_s increased. The proposed deblurring method with generative diffusion successfully enhanced the LP contrasts and improved the image spatial resolution across all three conditions. Nevertheless, there remained room for improvement with respect to the reference standard, especially for the challenging scenarios where the test object was closer to the x-ray source. The deblurring method was not effective for LPs that were entirely blurred and had nearly zero or negative contrasts.

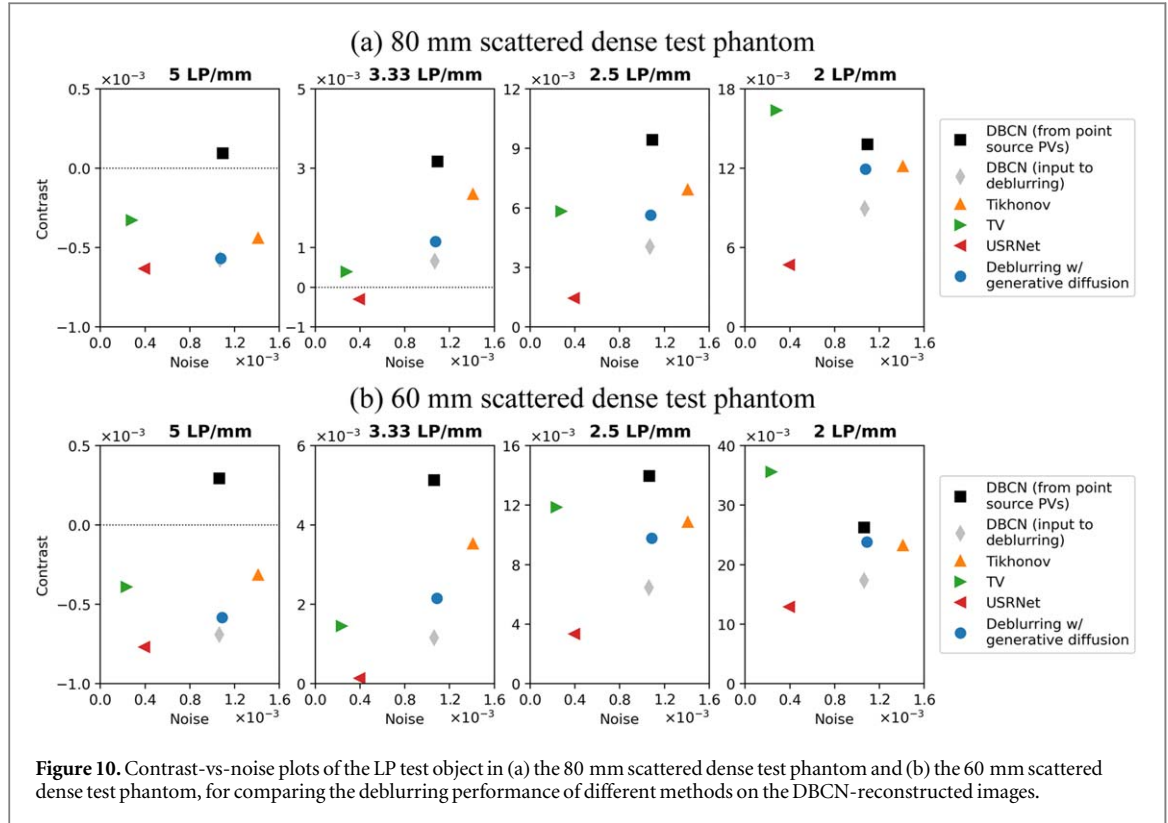


Figure 10. Contrast-vs-noise plots of the LP test object in (a) the 80 mm scattered dense test phantom and (b) the 60 mm scattered dense test phantom, for comparing the deblurring performance of different methods on the DBCN-reconstructed images.

4.3.4. Comparison of deblurring methods

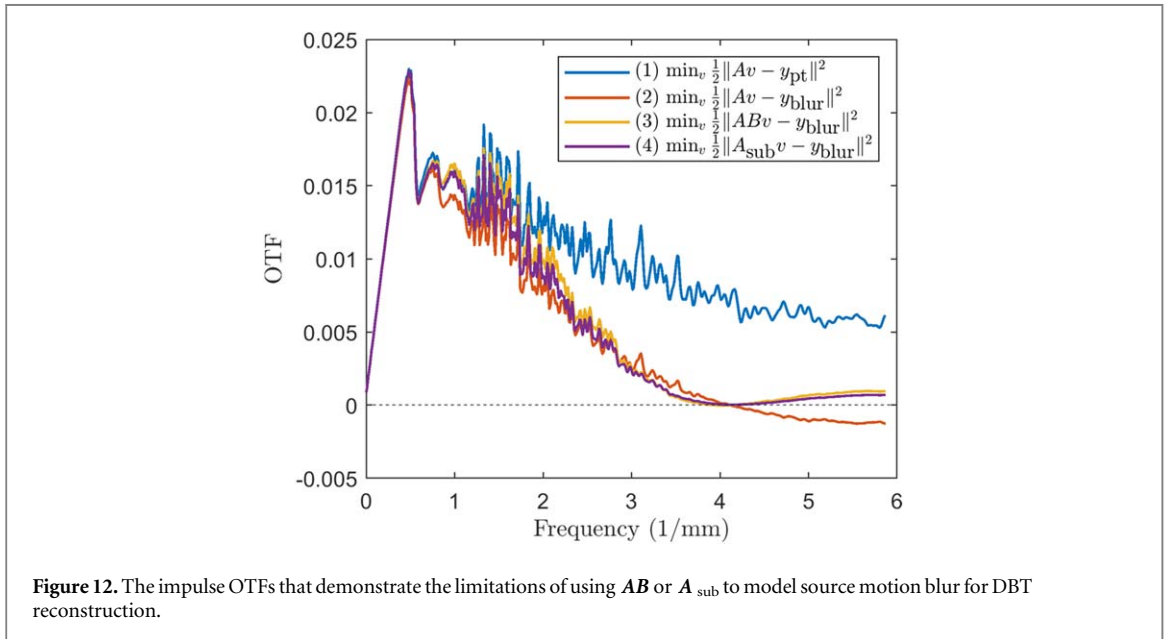
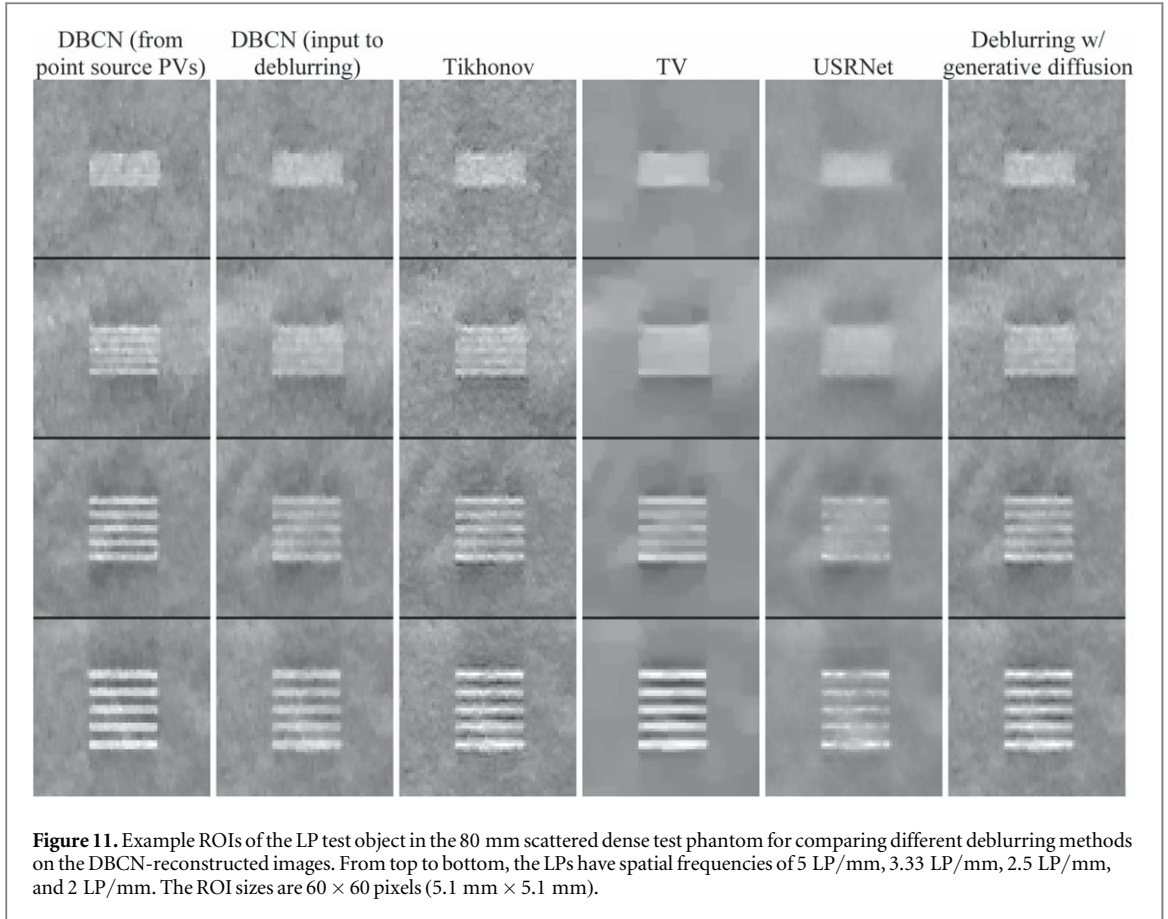
We evaluated the performance of different deblurring methods on the DBCN-reconstructed images using the 80 mm and the 60 mm scattered dense test phantoms with the LP test object placed at $z_s = 70$ mm. Figure 10 shows the contrast-vs-noise plots. Figure 11 displays the example LP ROIs from the 80 mm phantom.

Comparing the results from the 80 mm and the 60 mm phantoms, the overall trends were similar, as observed in section 4.3.2. Tikhonov regularized deblurring involved an inverse filter that inevitably enhanced the LP contrast and image noise at the same time. We tried a range of μ_{Tik} values and set it to 0.05. TV regularization ($\mu_{\text{TV}} = 0.0005$) performed very poorly on breast images, mainly because TV caused piecewise-constant patchy artifacts and could not characterize the ill-defined boundaries of soft tissue. While USRNet effectively smoothed the images, it failed to preserve the LP signals in the images. The proposed deblurring method with generative diffusion achieved an improvement in the LP contrast while maintaining a similar image noise level as the DBCN images. According to our visual judgment, it also retained the natural appearance of the tissue background without introducing artifacts.

5. Discussion

As mentioned in section 2.2, it is challenging to address source motion blur by directly using the modified system matrix \mathbf{AB} for forward and backward projections given the limited-angle and underdetermined nature of the DBT reconstruction problem. We conducted a simulation study to demonstrate the limitations of that approach. The simulation setup was the same as section 3.2, where we created an impulse object and generated noise-free \mathbf{y}_{pt} and \mathbf{y}_{blur} PVs. The impulse was placed close to the chest wall and centered in y at $z_s = 70$ mm. We reconstructed the impulse using gradient descent for the following formulations: (1) $\arg\min_{\mathbf{v}} \frac{1}{2} \|\mathbf{A}\mathbf{v} - \mathbf{y}_{\text{pt}}\|^2$, (2) $\arg\min_{\mathbf{v}} \frac{1}{2} \|\mathbf{A}\mathbf{v} - \mathbf{y}_{\text{blur}}\|^2$, (3) $\arg\min_{\mathbf{v}} \frac{1}{2} \|\mathbf{AB}\mathbf{v} - \mathbf{y}_{\text{blur}}\|^2$. We also investigated a fourth condition for the idea of introducing sub-focal spots within the blurry source: (4) $\arg\min_{\mathbf{v}} \frac{1}{2} \|\mathbf{A}_{\text{sub}}\mathbf{v} - \mathbf{y}_{\text{blur}}\|^2$. This idea was similar to the use of \mathbf{AB} and was shown to be effective for source blur modeling in CT (Fu et al 2013, Majee et al 2022). The modified system matrix \mathbf{A}_{sub} was defined as

$$\mathbf{A}_{\text{sub}} = \frac{1}{N_{\text{sub}}} \sum_{i_{\text{sub}}=1}^{N_{\text{sub}}} \mathbf{A}(\delta_{i_{\text{sub}}}), \quad (16)$$



where N_{sub} is the number of sub-focal spots, $A(\delta)$ is the perturbed system matrix of A obtained by offsetting the scan angles by δ , and $\delta_{i_{\text{sub}}} = \frac{w_0}{D_{\text{SO}}} \frac{1}{N_{\text{sub}}} \left(i_{\text{sub}} - \frac{N_{\text{sub}} + 1}{2} \right)$ in radians for $i_{\text{sub}} = 1, \dots, N_{\text{sub}}$. Figure 12 shows the OTFs of the extracted impulse PSF profiles in the y -direction. Although condition (1) was from the point source PV using the point source projector, it still largely deviated from the ideal OTF (a horizontal line with a value of one) with large oscillations even in the absence of noise. The decrease in OTF from condition (1) to condition (2) was due to the source motion blur. Their ratio corresponded to the smooth source blur OTF that was fitted by sinc functions in section 4.1. Condition (3) and (4) demonstrated that both AB and A_{sub} were able to correct the

negative phases at high frequency bands caused by the blur. However, the OTF magnitudes were not significantly improved and remained considerably lower than condition (1). This result suggested that directly using \mathbf{AB} or \mathbf{A}_{sub} for (unregularized) reconstruction cannot recover the loss in resolution caused by source motion blur in DBT.

The DDPM network uses an unsupervised training approach that solely requires high-quality images. To apply DDPM for deblurring, we simply need to integrate the gradient of the data-fit term into the DDPM sampling process, requiring no re-training or fine-tuning of the DDPM network. This unique feature endows DDPM much flexibility in terms of training data preparation because there is no need for paired low-quality and high-quality images. Moreover, it also makes the DDPM regularization very robust in that a single trained network can be applied to not only deblurring, but also other image restoration tasks as long as the specific task can be defined by a degradation operator like the blur matrix \mathbf{B} in (7).

Although the simulated training images do not contain LPs, the proposed deblurring method with the trained diffusion network is able to preserve the LP test objects and enhance their contrasts. This advantage is crucial, especially considering that MC signals in DBT images are sparse and small so they may be difficult for a network to learn. Deblurring by DDPM posterior sampling may help preserve the signals of interest such as MCs when a discrepancy exists between the training data and test data. The use of multiple diffusion steps in each deblurring process also ensures more gradual alterations of image content. In contrast, the end-to-end trained USRNet processes the images in a single step, resulting in an abrupt change in image content and a failure to retain the LPs.

To account for the possible alignments of LPs with the detector pixel array, we created shifted LPs in the test phantom. Another idea is to create LPs with an orientation slightly tilted with respect to the detector pixels and produce an oversampling of LPs. This idea proves to be challenging under the current settings because the virtual phantoms are defined on discrete voxels instead of continuous space and it is difficult to create smooth tilted LPs in the phantoms given the finite voxel resolution. In our preliminary study, we also investigated other designs of test objects, such as closely spaced bead pairs mimicking MCs. Compared with LPs, bead pairs were less discernable in the images and their quantitative metrics suffered from large variations due to the small sizes and noise. Therefore, we did not use bead pairs as our test objects.

We made a compromise by deblurring the reconstructed images and achieved moderate improvement in image sharpness. Post-processing deblurring has the advantage that it is applicable to DBT obtained from any reconstruction techniques. Nonetheless, post-processing deblurring may not be the optimal solution because the measured PVs are not exploited in the deblurring step. Future research is required to further improve the image resolution, especially for the image slices closer to the x-ray source where the blur is more severe.

We trained and tested the deblurring method using VICTRE simulated images. Besides the realism of the VICTRE phantoms, one of the main advantages of using simulation data is the availability of DBT scans from an ideal x-ray source. The point source DBT images can be used as ground truth, which are otherwise impossible to obtain from real scans, for either network training or algorithm evaluation. Our deblurring method has not been tested with real patient images due to the unavailability of data. Future work should apply this method to real patient DBT images with x-ray source motion blur to evaluate its effectiveness in clinical scenarios. We also acknowledge the importance of evaluating the proposed deblurring method using commercial phantoms with typical objects that simulate MCs. However, we do not have access to real data acquired with a continuous-motion DBT system at this time, and the use of commercial phantoms is not feasible within the scope of our current investigation. We will consider this aspect as part of our future work.

6. Conclusion

In this study, we introduced a new approach for modeling x-ray source motion blur in DBT imaging. We derived the in-plane source blur kernel for the reconstructed DBT slices based on imaging geometry and showed that it could be approximated by a shift-invariant kernel over the DBT plane at a given depth. We conducted a simulation study to validate its accuracy. Our simulation also underscored the limitations of modifying the system matrix to model source blur in DBT reconstruction, whether by incorporating the source blur matrix or introducing subsampling focal spots. In view of these limitations, we proposed a post-processing deblurring method with generative diffusion for the reconstructed DBT images using the known blur kernel. The quantitative results demonstrated that our deblurring method improved spatial resolution while maintaining the same level of image noise when applied to DBT images reconstructed with detector blur and correlated noise modeling. Future research can explore further refinements of the deblurring technique and investigate its application to human subject data for improving the diagnostic accuracy of DBT imaging.

Acknowledgments

This work is supported by the National Institutes of Health under Award Number R01 CA214981.

Data availability statement

The data cannot be made publicly available upon publication because the cost of preparing, depositing and hosting the data would be prohibitive within the terms of this research project. The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Mingjie Gao  <https://orcid.org/0000-0002-9129-0060>

References

- Badal A, Sharma D, Graff C G, Zeng R and Badano A 2021 Mammography and breast tomosynthesis simulator for virtual clinical trials *Comput. Phys. Commun.* **261** 107779
- Badano A, Graff C G, Badal A, Sharma D, Zeng R, Samuelson F W, Glick S J and Myers K J 2018 Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an *in silico* imaging trial *JAMA Netw. Open* **1** e185474
- Burgess A E 1999 Mammographic structure: data preparation and spatial statistics analysis *Proc. SPIE* **3661**
- Bushberg J T, Seibert J A, Leidholdt E M and Boone J M 2011 Diagnostic radiology *The Essential Physics of Medical Imaging* (Lippincott Williams and Wilkins) 2011, 169–576
- Chan H-P et al 2014 Digital breast tomosynthesis: observer performance of clustered microcalcification detection on breast phantom images acquired with an experimental system using variable scan angles, angular increments, and number of projection views *Radiology* **273** 675–85
- Chong A, Weinstein S P, McDonald E S and Conant E F 2019 Digital breast tomosynthesis: concepts and clinical practice *Radiology* **292** 1–14
- Chung H, Ryu D, McCann M T, Klasky M L and Ye J C 2023 Solving 3D inverse problems using pre-trained 2D diffusion models *IEEE Conf. on Computer Vision and Pattern Recognition* 2023, 22542–51
- Conant E F et al 2023 Mammographic screening in routine practice: multisite study of digital breast tomosynthesis and digital mammography screenings *Radiology* **307** e221571
- Dhariwal P and Nichol A 2021 Diffusion models beat GANs on image synthesis *Adv. Neural Inf. Process. Syst.* 8780–94
- Dong J, Roth S and Schiele B 2022 DWDN: deep Wiener deconvolution network for non-blind image deblurring *IEEE Trans. Pattern Anal. Mach. Intell.* **44** 9960–76
- Fu L, Wang J, Rui X, Thibault J B and De Man B 2013 Modeling and estimation of detector response and focal spot profile for high-resolution iterative CT reconstruction *IEEE Nuclear Science Symp. and Medical Imaging Conf.* 1–5
- Gao M, Fessler J A and Chan H-P 2021a Deep convolutional neural network with adversarial training for denoising digital breast tomosynthesis images *IEEE Trans. Med. Imaging* **40** 1805–16
- Gao M, Fessler J A and Chan H-P 2023 Model-based deep CNN-regularized reconstruction for digital breast tomosynthesis with a task-based CNN image assessment approach *Phys. Med. Biol.* **68** 245024
- Gao Y, Moy L and Heller S L 2021b Digital breast tomosynthesis: update on technology, evidence, and clinical practice *RadioGraphics* **41** 321–37
- Gunturk B K and Li X 2013 Fundamentals of image restoration *Image Restoration: Fundamentals and Advances* (Taylor and Francis Group) 2013, 25–61
- He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *IEEE Conf. on Computer Vision and Pattern Recognition* 770–8
- Ho J, Jain A and Abbeel P 2020 Denoising diffusion probabilistic models *Adv. Neural Inf. Process. Syst.* 6840–51
- Johns P C and Yaffe M J 1987 X-ray characterisation of normal and neoplastic breast tissues *Phys. Med. Biol.* **32** 675–95
- Kawar B, Elad M, Ermon S and Song J 2022 Denoising diffusion restoration models *Adv. Neural Inf. Process. Syst.* 23593–606
- Kingma D and Ba J 2015 Adam: a method for stochastic optimization *Int. Conf. on Learning Representations*
- Krishnan D and Fergus R 2009 Fast image deconvolution using hyper-Laplacian priors *Adv. Neural Inf. Process. Syst.* 1033–41
- Lee C and Baek J 2022 Effect of optical blurring of x-ray source on breast tomosynthesis image quality: modulation transfer function, anatomical noise power spectrum, and signal detectability perspectives *PLoS One* **17** e0267850
- Mackenzie A, Marshall N W, Hadjipanteli A, Dance D R, Bosmans H and Young K C 2017 Characterisation of noise and sharpness of images from four digital breast tomosynthesis systems for simulation of images for virtual clinical trials *Phys. Med. Biol.* **62** 2376–97
- Majee S, Aslan S, Gursoy D and Bouman C A 2022 CodEx: a modular framework for joint temporal de-blurring and tomographic reconstruction *IEEE Trans. Comput. Imaging* **8** 666–78
- Marshall N W and Bosmans H 2012 Measurements of system sharpness for two digital breast tomosynthesis systems *Phys. Med. Biol.* **57** 7629–50
- Michelfeit F 2023 Press release: Siemens Healthineers presents mammography system with groundbreaking new imaging technology
- Michielsen K, Van Slambrouck K, Jerebko A and Nuyts J 2013 Patchwork reconstruction with resolution modeling for digital breast tomosynthesis *Med. Phys.* **40** 031105
- PHE 2018 Technical evaluation of Siemens Mammomat Inspiration digital breast tomosynthesis system *NHS Breast Screening Programme Equipment Report* Public Health England
- Richardson W H 1972 Bayesian-based iterative method of image restoration *J. Opt. Soc. Am.* **62** 55–9
- Ronneberger O, Fischer P and Brox T 2015 U-Net: convolutional networks for biomedical image segmentation *Medical Image Computing and Computer-assisted Intervention* 234–41

- Saharia C, Ho J, Chan W, Salimans T, Fleet D J and Norouzi M 2023 Image super-resolution via iterative refinement *IEEE Trans. Pattern Anal. Mach. Intell.* **45** 4713–26
- Sechopoulos I 2013 A review of breast tomosynthesis: I. The image acquisition process *Med. Phys.* **40** 014301
- Shaheen E, Marshall N and Bosmans H 2011 Investigation of the effect of tube motion in breast tomosynthesis: continuous or step and shoot? *Proc. SPIE* **7961** 1E
- Sohl-Dickstein J, Weiss E A, Maheswaranathan N and Ganguli S 2015 Deep unsupervised learning using nonequilibrium thermodynamics *Int. Conf. on Machine Learning* 2246–55
- Song J, Meng C and Ermon S 2021a Denoising diffusion implicit models *Int. Conf. on Learning Representations*
- Song Y and Ermon S 2019 Generative modeling by estimating gradients of the data distribution *Adv. Neural Inf. Process. Syst.* 11918–30
- Song Y, Shen L, Xing L and Ermon S 2022 Solving inverse problems in medical imaging with score-based generative models *Int. Conf. on Learning Representations*
- Song Y, Sohl-Dickstein J, Kingma D P, Kumar A, Ermon S and Poole B 2021b Score-based generative modeling through stochastic differential equations *Int. Conf. on Learning Representations*
- Tikhonov A N 1963 Solution of incorrectly formulated problems and the regularization method *Soviet Math. Doklady* **4** 1035–8
- Tilley S, Jacobson M, Cao Q, Brehler M, Sisniega A, Zbijewski W and Stayman J W 2018 Penalized-likelihood reconstruction with high-fidelity measurement models for high-resolution cone-beam imaging *IEEE Trans. Med. Imaging* **37** 988–99
- Tilley S, Siewerdsen J H and Stayman J W 2016 Model-based iterative reconstruction for flat-panel cone-beam CT with focal spot blur, detector blur, and correlated noise *Phys. Med. Biol.* **61** 296–319
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser L and Polosukhin I 2017 Attention is all you need *Adv. Neural Inf. Process. Syst.*
- Wang Y, Yu J and Zhang J 2023 Zero-shot image restoration using denoising diffusion null-space model *Int. Conf. on Learning Representations*
- Wiener N 1949 *Extrapolation, Interpolation, and Smoothing of Stationary Time Series* (The MIT Press)
- Xu L, Zheng S and Jia J 2013 Unnatural L0 sparse representation for natural image deblurring *IEEE Conf. on Computer Vision and Pattern Recognition* 1107–14
- Zhang K, Van Gool L and Timofte R 2020 Deep unfolding network for image super-resolution *IEEE Conference on Computer Vision and Pattern Recognition* 3217–26
- Zhang K, Ren W, Luo W, Lai W-S, Stenger B, Yang M-H and Li H 2022 Deep image deblurring: a survey *Int. J. Comput. Vis.* **130** 2103–30
- Zheng J, Fessler J A and Chan H-P 2017 Segmented separable footprint projector for digital breast tomosynthesis and its application for subpixel reconstruction *Med. Phys.* **44** 986–1001
- Zheng J, Fessler J A and Chan H-P 2018 Detector blur and correlated noise modeling for digital breast tomosynthesis reconstruction *IEEE Trans. Med. Imaging* **37** 116–27
- Zheng J, Fessler J A and Chan H-P 2019 Effect of source blur on digital breast tomosynthesis reconstruction *Med. Phys.* **46** 5572–92
- Zoran D and Weiss Y 2011 From learning models of natural image patches to whole image restoration *Int. Conf. on Computer Vision* 479–86